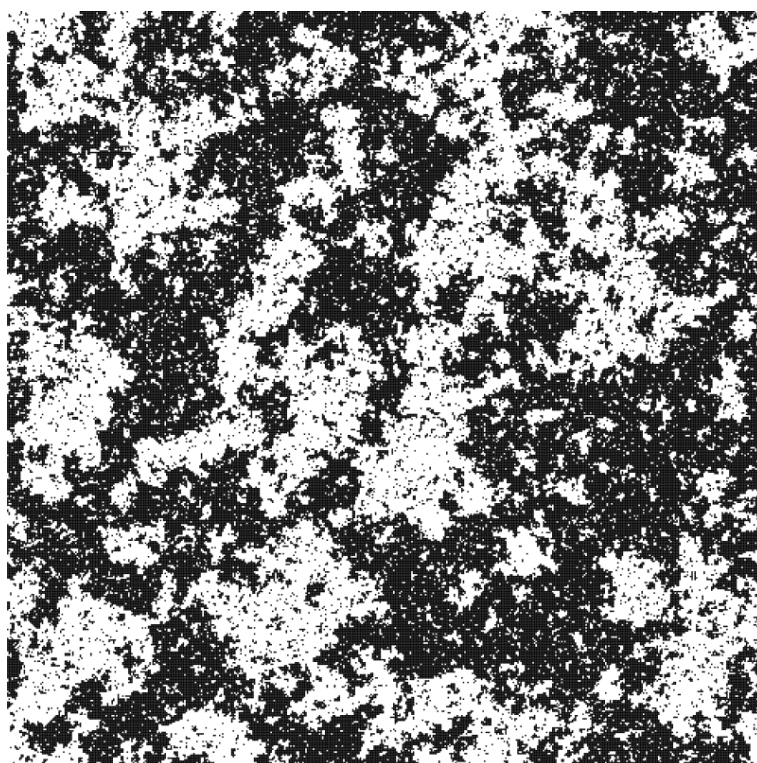


Introduction aux champs aléatoires markoviens et gibbsiens

Y. Velenik



— Version du 25 mars 2013 —

Dernière version téléchargeable à l'adresse
<http://www.unige.ch/math/folks/velenik/cours.html>

2006-2007

Table des matières

1	Introduction	7
1.1	Chaînes de Markov et champs markoviens	8
1.2	Le modèle d'Ising	9
1.2.1	Définition	9
1.2.2	Transition de phase	10
1.2.3	Champ magnétique	14
1.2.4	Conditions au bord	15
1.3	Quelques ouvrages de référence	16
2	Le modèle d'Ising	19
2.1	Définition du modèle	19
2.2	Quelques interprétations	20
2.2.1	Ferro-aimant	20
2.2.2	Gaz réticulaire	21
2.2.3	Alliage binaire	21
2.2.4	Quelques autres interprétations et applications	22
2.3	Les inégalités de corrélation	22
2.3.1	Inégalités GKS	22
2.3.2	Inégalités FKG	22
2.4	Limite thermodynamique	23
2.5	Critères d'unicité/non-unicité	27
2.6	Aimantation et transition de phase	28
2.6.1	Quelques propriétés élémentaires de l'aimantation	28
2.6.2	Énergie libre	29
2.7	Transition de phase à basse température	33
2.8	Unicité à haute température	36
2.9	Unicité en champ magnétique non nul	39
2.10	Preuve des inégalités de corrélation	41
2.10.1	Inégalités GKS	41
2.10.2	Inégalités FKG	43

2.11	Fonctions convexes	44
3	Champs markoviens et mesures de Gibbs	47
3.1	Définitions	47
3.1.1	Champs aléatoires markoviens	47
3.1.2	Mesures de Gibbs	48
3.1.3	Potentiels canoniques	48
3.2	Équivalence	49
4	Simulation de champs markoviens	51
4.1	Méthode de Monte-Carlo	51
4.2	Simulation parfaite	52
4.3	Algorithmes par amas	57
4.3.1	FK percolation	57
4.3.2	Dynamique de Swendsen-Wang	59
5	Modèles avec symétrie continue	61
5.1	Symétrie et mesures de Gibbs	61
5.2	Définition du modèle	62
5.3	Absence de brisure de symétrie	63
5.4	Transitions de phases	68
5.4.1	Brisure de symétrie	68
5.4.2	Transition de Kosterlitz-Thouless	68
6	Modèles avec spin non-borné	71
6.1	Le modèle gaussien	71
6.2	La représentation en marche aléatoire	72
6.3	Mesures en volume infini	74
6.3.1	Le cas massif	74
6.3.2	Le cas non-massif	75
6.4	Modèles effectifs de surfaces aléatoires	78
7	Positivité sous réflexion	79
7.1	Mesures positives sous réflexion	80
7.1.1	Définition et propriétés élémentaires	80
7.1.2	Construction de mesures positives sous réflexion	81
7.2	L'estimation de l'échiquier	82
7.3	Domination gaussienne et borne infrarouge	86
7.4	Applications	89
7.4.1	Condensation d'ondes de spin et transitions de phase	89
7.4.2	Transition de phase dans les modèles $O(N)$ anisotropes	94
7.4.3	Transition de phase dans un modèle gaussien avec double-puits	100

8	Les champs gibbsiens	107
8.1	Cadre mathématique	108
8.1.1	Espace de configurations.	108
8.1.2	Structure d'espace de probabilité.	108
8.1.3	Structure topologique et métrique.	109
8.1.4	Convergence des champs aléatoires	111
8.2	Probabilités conditionnelles et spécifications	112
8.2.1	Noyaux de probabilité	112
8.2.2	Probabilités conditionnelles	112
8.2.3	Spécifications	114
8.3	Potentiels et mesures de Gibbs	115
8.4	Propriétés des mesures de Gibbs	117
8.4.1	Structure de $\mathcal{G}(\Pi)$	117
8.4.2	Propriétés globales et tribu asymptotique	118
8.4.3	Propriétés des mesures extrémales	120
8.4.4	Mesures de Gibbs et limite thermodynamique	122
	Bibliographie	124

Chapitre 1

Introduction

De très nombreuses situations sont naturellement modélisées par une famille de variables aléatoires $(X_i)_{i \in V}$, indexées par les sommets d'un graphe $G = (V, E)$, où V dénote l'ensemble des sommets de G , et E l'ensemble des ses arêtes.

Exemple 1.0.1.

- Physique : $V = \mathbb{Z}^d$, et chaque sommet correspond à un atome du réseau cristallin correspondant. X_i peut alors représenter une quantité associée à l'atome placé en i , par exemple son spin, ou son état d'excitation, etc.
- Écologie : on décompose un territoire en petites parcelles, et on place un sommet dans chacune d'entre elles; dans ce cas on peut prendre $V = \mathbb{Z}^2$. X_i peut alors représenter le type de plantes se trouvant dans la parcelle i .
- Épidémiologie : chaque sommet correspond à un individu, et deux sommets sont reliés si ces individus ont des contacts fréquents (par exemple, membres de la même famille, partenaires sexuels, etc.). X_i peut alors indiquer si l'individu i est malade ou en bonne santé.
- Sociologie : chaque sommet correspond à un individu, et X_i à l'opinion de l'individu i sur un certain sujet.
- Imagerie : $G = \mathbb{Z}^2$, et chaque sommet correspond à un pixel de l'image. X_i peut alors représenter la couleur du pixel correspondant.

On peut évidemment en trouver beaucoup d'autres. Un point commun à tous ces exemples est que les variables ne sont en général pas indépendantes, mais que leur dépendance doit refléter la structure locale du graphe. Par exemple, la loi de X_i ne devrait dépendre principalement que des variables X_j dans un voisinage de i . Supposons pour le moment que G soit un graphe fini, et associons à chaque sommet $i \in V$ son voisinage

$$\mathcal{N}_i \stackrel{\text{déf}}{=} \{j \in V : (i, j) \in E\} \equiv \{j \sim i\}.$$

Une formalisation de la remarque précédente revient à exiger que

$$\text{Prob}(X_i = a \mid X_j, j \neq i) = \text{Prob}(X_i = a \mid X_j, j \in \mathcal{N}_i); \quad (1.1)$$

c'est ce qu'on appelle la propriété de Markov. Les lois de probabilité sur G possédant cette propriété sont appelées champs aléatoires markoviens sur G . La définition précise dans le cas d'un graphe infini est plus délicate et sera discutée plus loin dans le cours.

1.1 Chaînes de Markov et champs markoviens

Soit $(X_t)_{t \in \mathbb{N}}$ une chaîne de Markov à temps discret sur un espace S avec probabilités de transition $p_t(i, j)$, $i, j \in S$, $t \in \mathbb{N}$. Pour simplifier, considérons simplement une fenêtre de temps finie, c'est-à-dire restreignons-nous à la famille (X_0, \dots, X_T) , $0 < T < \infty$, et supposons l'espace d'état S fini. Par définition, la chaîne satisfait la propriété de Markov suivante :

$$\begin{aligned} \text{Prob}(X_t = x_t \mid X_0 = x_0, \dots, X_{t-1} = x_{t-1}) &= \text{Prob}(X_t = x_t \mid X_{t-1} = x_{t-1}) \\ &= p_{t-1}(x_{t-1}, x_t), \end{aligned}$$

pour tout $0 < t \leq T$ et $x_0, \dots, x_t \in S$. Cette propriété, reposant sur l'ordre total propre à \mathbb{N} , n'est pas aussi symétrique que (1.1). Il est cependant aisé de vérifier qu'elle implique la propriété suivante :

$$\text{Prob}(X_t = x_t \mid X_s = x_s, s \neq t) = \text{Prob}(X_t = x_t \mid X_{t-1} = x_{t-1}, X_{t+1} = x_{t+1}),$$

pour tout $0 < t < T$ et $x_0, \dots, x_T \in S$, ce qui est exactement (1.1). En effet, en notant $\pi(\cdot)$ la loi de X_0 , on a

$$\text{Prob}(X_0 = x_0, X_1 = x_1, \dots, X_T = x_T) = \pi(x_0)p_0(x_0, x_1) \cdots p_{T-1}(x_{T-1}, x_T),$$

et donc

$$\begin{aligned} &\text{Prob}(X_t = x_t \mid X_s = x_s, s \neq t) \\ &= \frac{\pi(x_0)p_0(x_0, x_1) \cdots p_{T-1}(x_{T-1}, x_T)}{\sum_{y \in S} \pi(x_0)p_0(x_0, x_1) \cdots p_{t-1}(x_{t-1}, y)p_t(y, x_{t+1}) \cdots p_{T-1}(x_{T-1}, x_T)} \\ &= \frac{p_{t-1}(x_{t-1}, x_t)p_t(x_t, x_{t+1})}{\sum_{y \in S} p_{t-1}(x_{t-1}, y)p_t(y, x_{t+1})}, \end{aligned}$$

ce qui montre que le membre de gauche n'est fonction que de x_{t-1} et x_{t+1} (et x_t évidemment).

Les chaînes de Markov fournissent donc un exemple de champs aléatoires markoviens (sur le graphe avec sommets $V = \mathbb{N}$ et arêtes $E = \{(i, j) : i, j \in \mathbb{N}, |i - j| = 1\}$). La réciproque n'est pas vraie : on peut construire des champs markoviens sur \mathbb{Z} qui ne sont pas des chaînes de Markov ; ainsi, même sur \mathbb{N} , cette notion est plus générale. C'est bien sûr l'extension à des graphes plus généraux que \mathbb{N} (ou \mathbb{Z}) qui est l'application la plus intéressante et utile des champs markoviens.

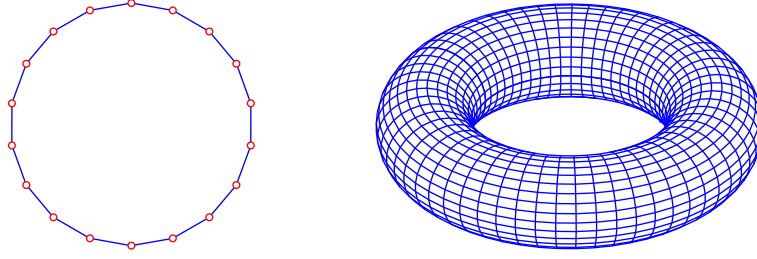


FIGURE 1.1 – Graphes avec condition au bord périodique, en dimension 1 et 2.

1.2 Le modèle d'Ising

1.2.1 Définition

Dans cette section, nous allons discuter de façon informelle un exemple particulièrement important de champ markovien : le modèle d'Ising. Son analyse rigoureuse nous occupera une bonne partie du cours (*cf.* Chapitre 2).

Le modèle d'Ising a été introduit en 1925 par le physicien Ernst Ising dans sa thèse de doctorat, suivant la suggestion de son directeur de thèse Wilhelm Lenz. Son but était de décrire la transition entre les régimes ferromagnétique et paramagnétique des matériaux ferromagnétiques, mais il possède de très nombreuses autres interprétations, *cf.* la discussion plus bas.

Ce modèle peut être défini sur un graphe arbitraire, mais il est usuel de le définir sur le graphe avec ensemble de sommets $V_N = \{1, \dots, N\}^d$ et avec une arête connectant chaque paire de sommets plus-proches-voisins (c'est-à-dire tels que $\|j - i\|_2 = 1$). Afin de ne pas introduire de distinction entre les sommets se trouvant au bord de V_N et ceux se trouvant à l'intérieur, nous ajoutons également un lien entre chaque paire de sommets i et j tels qu'il existe $k \in \{1, \dots, d\}$ avec $i_k = 1$, $j_k = N$ et $i_{k'} = j_{k'}$ pour tout $k' \neq k$ (*cf.* figure 1.1). On dit alors que le modèle est à condition au bord périodique ; notons $G_N^{\text{per}} \stackrel{\text{déf}}{=} (V_N, E_N^{\text{per}})$ le graphe correspondant.

On appelle configuration du modèle d'Ising un élément $\omega \in \Omega_N \stackrel{\text{déf}}{=} \{-1, 1\}^{V_N}$; Ω_N est l'espace des configurations. Le rôle des variables aléatoires X_i décrites ci-dessus est tenu par les variables aléatoires $\sigma_i : \Omega_N \rightarrow \{-1, 1\}$, $\sigma_i(\omega) \stackrel{\text{déf}}{=} \omega_i$; σ_i est appelée le **spin** au sommet i . À chaque configuration ω est associée son énergie

$$\mathcal{H}_{N;\beta}(\omega) \stackrel{\text{déf}}{=} -\beta \sum_{(i,j) \in E_N^{\text{per}}} \sigma_i(\omega) \sigma_j(\omega),$$

où le paramètre $\beta \in \mathbb{R}^+$ est appelé **température inverse**. On introduit alors la mesure de probabilité suivante sur $(\Omega_N, \mathcal{F}_N)$ (la tribu \mathcal{F}_N étant ici simplement l'ensemble des parties de Ω_N),

$$\mu_{N;\beta}(\omega) \stackrel{\text{déf}}{=} \frac{1}{Z_{N,\beta}} \exp(-\mathcal{H}_N(\omega)).$$

La constante de normalisation

$$Z_{N;\beta} \stackrel{\text{d\'ef}}{=} \sum_{\omega \in \Omega_N} \exp(-\beta \mathcal{H}_N(\omega))$$

est appelée **fonction de partition** et jouera un rôle important dans l'étude du chapitre 2. On voit que cette mesure favorise les configurations ayant une énergie basse.

Il est immédiat de vérifier (exercice) que l'on a bien la propriété de Markov : soit $s \in \Omega_N$, alors la probabilité que σ_i prenne la valeur s_i étant donné que le reste de la configuration est donné par s est

$$\mu_{N;\beta}(\sigma_i = s_i \mid \sigma_j = s_j, \forall j \neq i) = \frac{\exp(\beta s_i \sum_{j \sim i} s_j)}{\exp(\beta \sum_{j \sim i} s_j) + \exp(-\beta \sum_{j \sim i} s_j)},$$

ce qui ne dépend bien que de $\{s_j : j \sim i\}$.

1.2.2 Transition de phase

On voit que, lorsque $\beta \neq 0$, un spin va avoir tendance à prendre la même valeur que la majorité de ses voisins (puisque $s_i \sum_{j \sim i} s_j$ est maximal lorsque les signes de s_i et $\sum_{j \sim i} s_j$ coïncident). On peut alors se demander si cette interaction entre spins voisins va conduire à un ordre dans tout le système. Commençons par considérer deux cas limites.

1. $\beta = 0$. Dans ce cas, les spins forment une famille de variables aléatoires indépendantes suivant chacune une loi de Bernoulli de paramètre $\frac{1}{2}$. La loi des grands nombres implique que lorsque N est très grand, $N^{-d} \sum_{i \in V_N} \sigma_i \approx 0$, et le théorème central limite montre que cette dernière quantité est en fait de l'ordre de $N^{-d/2}$. En particulier, une réalisation typique du champ aura approximativement la même densité de spins prenant valeur 1 et -1 .
2. $\beta = \infty$. Dans ce cas, les spins sont très fortement dépendants : seules deux configurations ont probabilité strictement positive (égale à $1/2$ par symétrie), $\omega_+ \equiv 1$ et $\omega_- \equiv -1$. En effet, toute autre configuration ω' a une énergie plus élevée, et donc $\lim_{\beta \rightarrow \infty} \mu_{N;\beta}(\omega') / \mu_{N;\beta}(\omega_+) = 0$.

Manifestement, aucune information ne peut être transmise lorsque $\beta = 0$: la connaissance de la valeur d'un spin ne fournit aucune information sur l'état des autres spins. La situation est opposée lorsque $\beta = \infty$: la connaissance de la valeur du spin en un sommet donné du graphe détermine complètement la configuration lorsque $\beta = \infty$, et donc que l'information est transmise arbitrairement loin dans ce cas. Le problème à présent est de déterminer quel type de comportement a lieu pour les valeurs intermédiaires (plus intéressantes) de β .

Les figures 1.2 et 1.3 contiennent des configurations typiques du modèle d'Ising avec condition au bord périodique en dimension, respectivement, 1 et 2, pour diverses valeurs du paramètre β (en fait, en fonction du paramètre $p_\beta \stackrel{\text{d\'ef}}{=} 1 - e^{-2\beta} \in [0, 1]$). Dans les deux cas, on constate (comme on pouvait s'y attendre) une augmentation de la taille des amas de spins de même valeur lorsque p_β augmente. En dimension 1, la symétrie entre les deux valeurs possibles du spin semble maintenue, en tout cas jusqu'à des valeurs de ce paramètre

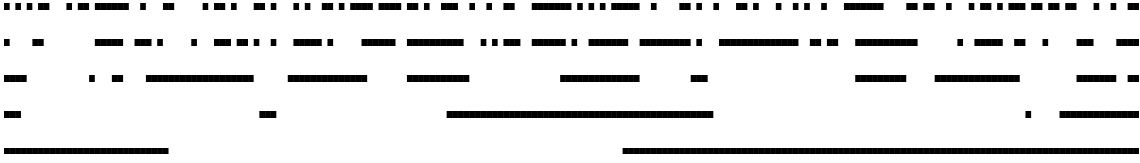


FIGURE 1.2 – Configurations typiques du modèle d’Ising en dimension 1 avec condition au bord périodique ($N = 500$), pour différentes valeurs du paramètre $p_\beta \stackrel{\text{déf}}{=} 1 - e^{-2\beta}$: (de haut en bas) de 0 à 0,5, 0,9, 0,95, 0,99.

très proches de 1 ; l’ordre apparent pour $p_\beta \approx 1$ semble bien n’être qu’un phénomène de taille finie, et il semble raisonnable d’imaginer que la symétrie serait encore présente, pour toute valeur de $p_\beta < 1$, si on considérait un système suffisamment grand. Par contre, en dimension 2, la situation semble très différente : si les amas grandissent tout en maintenant la symétrie entre les deux types de spins jusqu’à $p_\beta \approx 0.58$, un changement abrupt a lieu près de cette valeur, et à partir de $p_\beta \approx 0.59$, les densités des deux types de spins ne sont plus égales, le système semblant avoir spontanément fait un choix en faveur d’un des deux types de spin.

Afin de faire une analyse un peu plus convaincante, il est utile de faire des observations plus quantitatives. Les graphiques présentés sur la figure 1.4 montrent le comportement approché (moyenné sur un échantillon) de la fonction suivante,

$$\langle |N^{-d} \sum_{i \in V_N} \sigma_i| \rangle_{N;\beta} \equiv \langle |m_{N;\beta}| \rangle_{N;\beta},$$

où $\langle \cdot \rangle_{N;\beta}$ est une notation standard pour l’espérance sous la mesure $\mu_{N;\beta}$. Cette quantité mesure bien la différence entre les densités des deux espèces de spins. (La raison pour laquelle on prend la valeur absolue est que sinon l’espérance est nulle par symétrie.) La variable aléatoire $m_{N;\beta}$ est appelée **aimantation**.

Ces graphes confirment l’analyse précédente : en dimension 1, il semble bien que la fonction limite, que l’on notera $m^*(\beta)$, soit identiquement nulle, sauf en $p_\beta = 1$ (c’est-à-dire $\beta = \infty$) où elle vaut 1, alors qu’en dimension 2, la fonction limite semble non triviale, devenant strictement positive à partir d’une valeur $p \approx 0.58$ (la courbe limite est aussi représentée).

Nous verrons dans le chapitre 2 que tout ceci est effectivement correct. On appelle **transition de phase** le phénomène abrupt observé en dimension 2 (et qui est en fait aussi présent en dimensions supérieures) ; nous en verrons une définition précise plus tard. La perte de symétrie qui l’accompagne est appelée **brisure spontanée de symétrie**. L’absence de transition de phase en dimension 1 est le résultat principal de la thèse d’Ising [12], et est en fait très simple ; nous démontrerons un résultat plus général dans la section 2.8 (et en verrons des preuves alternatives élémentaires en exercices). La preuve de l’existence d’une transition de phase en dimension 2 et plus est due à un autre physicien, Rudolf Peierls, qui introduisit en 1936 un argument qui a depuis lors été généralisé à une classe immense de systèmes [20]. Nous l’étudierons dans la section 2.7. En dimension 2, il est en fait possible de calculer explicitement la fonction limite $m^*(\beta)$ apparaissant dans le

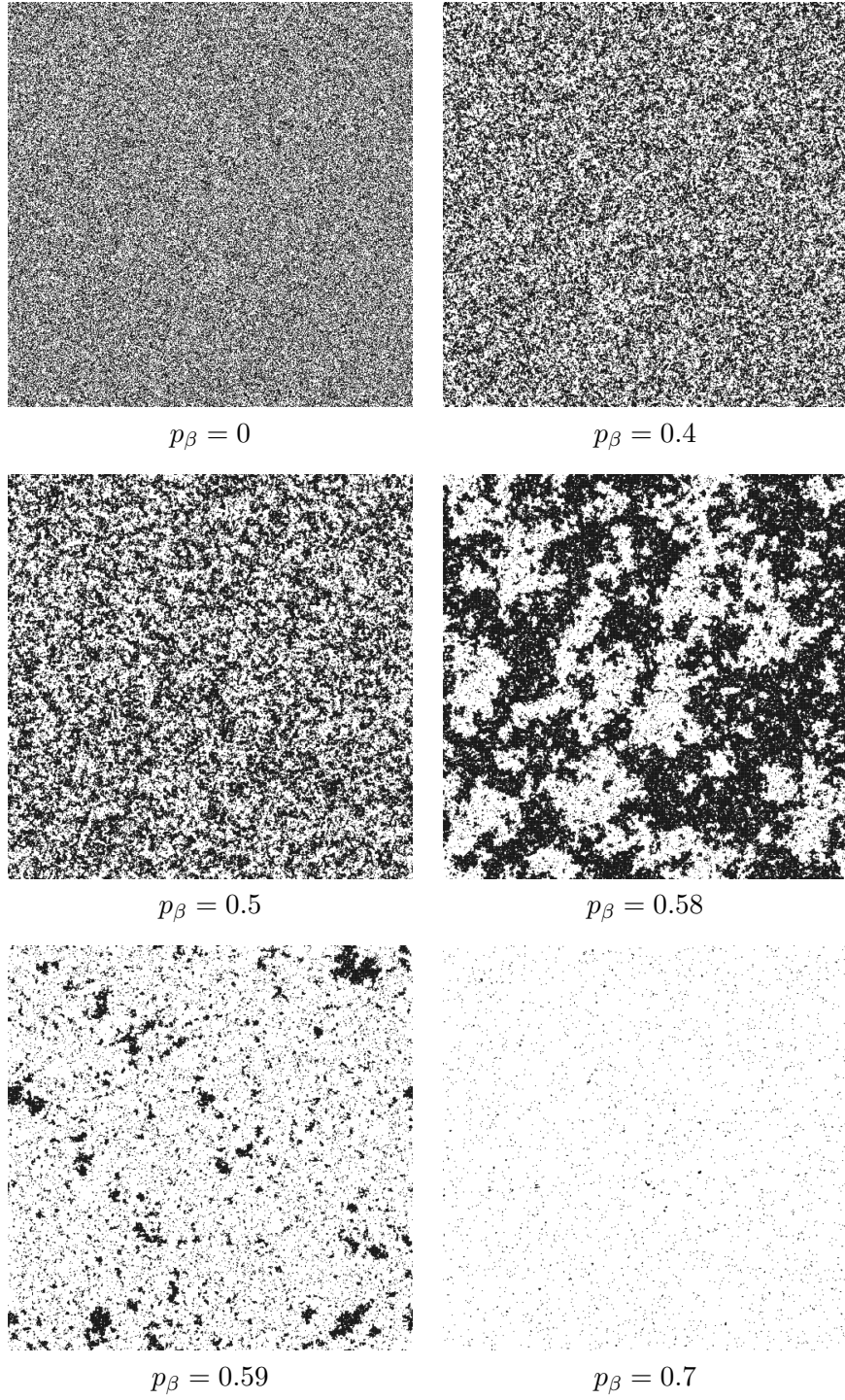


FIGURE 1.3 – Configurations typiques du modèle d'Ising en dimension 2 avec condition au bord périodique ($N = 500$), pour différentes valeurs du paramètre $p_\beta \stackrel{\text{déf}}{=} 1 - e^{-2\beta}$.

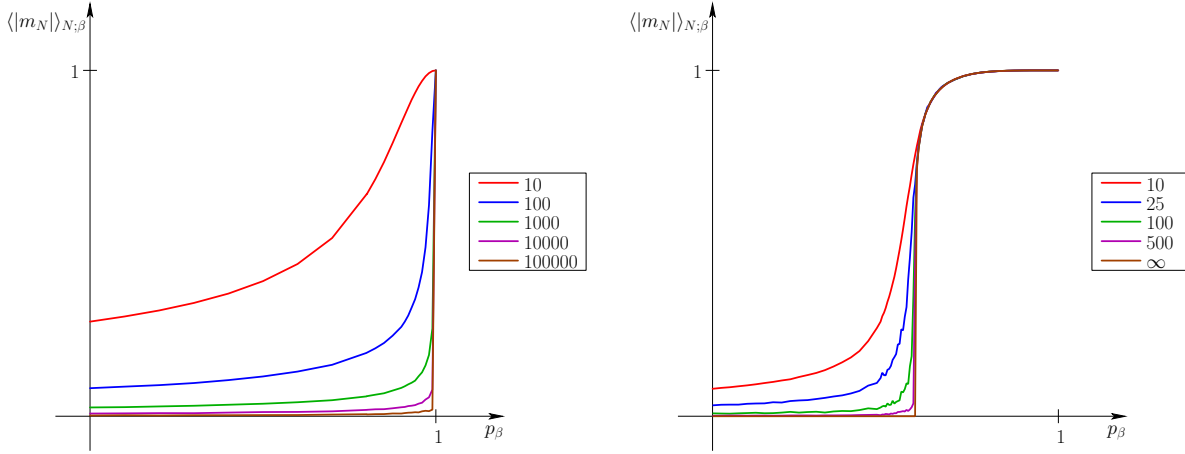


FIGURE 1.4 – Espérance de $|m_{N;\beta}|$ en fonction de p_β pour le modèle d’Ising en dimension 1 (gauche) et 2 (droite) avec condition au bord périodique, pour différentes valeurs de N . (Les oscillations présentes sur certaines courbes sont dues aux fluctuations statistiques.)

graphe de la figure 1.4 ; ce résultat est dû au chimiste (prix Nobel de chimie, mais aussi physicien et mathématicien virtuose à ses heures) Lars Onsager [19, 29]. On trouve qu’elle est donnée par la fonction

$$m^*(p) \stackrel{\text{déf}}{=} \left[\max \left(1 - \left(\frac{2(1-p)}{p(2-p)} \right)^4, 0 \right) \right]^{1/8}, \quad (1.2)$$

ce qui implique en particulier que la valeur de β à laquelle la transition a lieu, appelée *valeur critique* et notée β_c , est

$$\beta_c = \frac{1}{2} \operatorname{asinh}(1) \cong 0.4406867935097715239 \dots,$$

ce qui correspond à $p_{\beta_c} = \sqrt{2}/(1 + \sqrt{2}) \cong 0.5857864376269049655 \dots$. Ces formules, dont la preuve est de nature plus combinatoire, ne seront pas démontrées dans ce cours.

On voit que le phénomène de transition de phase est d’autant plus marqué que la taille du système est importante. Il est donc mathématiquement raisonnable d’approximer les très grands systèmes par des systèmes infinis (par la suite, il faudra bien sûr estimer les corrections dues aux effets de taille finie). Au vu des exemples précédents, on constate que des systèmes de taille modérée sont déjà raisonnablement approchés par le cas limite¹. Un tel passage à des systèmes infinis est appelé *passage à la limite thermodynamique*. Dans cette limite, il va être possible de donner des définitions précises des transitions de phases et d’autres concepts associés.

Le passage à la limite pose plusieurs problèmes. En particulier, celui de définir de façon précise les champs de Markov sur des graphes infinis. Il existe plusieurs moyens de procéder, et nous en étudierons deux dans ce cours. La première méthode, particulièrement

1. Surtout si l’on pense que pour les applications à la physique, on a souvent N^d de l’ordre de 10^{20} !

appropriée pour étudier les propriétés générales de telles mesures, est de donner un sens à la propriété de Markov énoncée plus haut : on peut montrer que, pour un graphe fini, la donnée d'un tel système de probabilités conditionnelles (ceci sous-entend certaines propriétés de consistance) permet de reconstruire une unique mesure de probabilité. L'extension de ce résultat à des graphes infinis n'est pas vraie : il existe en général une infinité de mesures de probabilité compatibles avec un tel système de probabilités conditionnelles. L'absence d'unicité correspondra précisément au régime où il y a transition de phase (intuitivement, dans le cas du modèle d'Ising, on obtiendra, par exemple, une mesure avec une densité supérieure à 1/2 de spin +1, et une autre avec une densité supérieure à 1/2 de spin -1, et également toutes les combinaisons convexes de ces deux mesures).

La seconde approche est plus intuitive, mais moins adaptée à la construction d'une théorie générale des champs markoviens. Dans cette approche, on considère des suites de mesures de probabilité sur des graphes de plus en plus grands, et on définit les mesures sur le graphe limite comme étant l'ensemble des points d'accumulation de ces suites. Pour cela, il faudra évidemment introduire une topologie appropriée sur l'espace des mesures de probabilité. L'ensemble de mesures obtenu de cette façon est le même (modulo quelques subtilités que l'on discutera plus loin) que celui obtenu avec l'approche précédente. En particulier, l'existence de plusieurs points d'accumulation correspondra à nouveau à la présence d'une transition de phase.

1.2.3 Champ magnétique

Nous venons de voir que la présence d'une transition de phase dans ce modèle se traduit par une brisure de la symétrie entre les deux espèces de spins, les configurations typiques possédant des densités différentes de chacun des deux types de spins. Il est donc naturel de généraliser le modèle d'Ising en introduisant un paramètre supplémentaire permettant de jouer sur la symétrie entre les spins. Ce paramètre a aussi une interprétation naturelle en terme de la modélisation originelle d'Ising, et était déjà présent dans son analyse : le **champ magnétique** (extérieur) h . On associe donc à chaque configuration l'énergie suivante,

$$\mathcal{H}_{N;\beta,h}(\omega) \stackrel{\text{déf}}{=} -\beta \sum_{(i,j) \in E_N^{\text{per}}} \sigma_i(\omega) \sigma_j(\omega) - h \sum_{i \in V_N} \sigma_i(\omega),$$

et la mesure de probabilité correspondante,

$$\mu_{N;\beta,h}(\omega) \stackrel{\text{déf}}{=} \frac{1}{Z_{N;\beta,h}} \exp(-\mathcal{H}_{N;\beta,h}(\omega)).$$

La figure 1.5 montre l'effet du paramètre h sur le comportement de la fonction limite $m(\beta, h)$, définie par

$$m(\beta, h) \stackrel{\text{déf}}{=} \lim_{N \rightarrow \infty} \langle N^{-d} \sum_{i \in V_N} \sigma_i \rangle_{N;\beta,h} \equiv \langle m_{N;\beta,h} \rangle_{N;\beta,h}.$$

Le graphe de gauche montre que l'asymétrie induite par la présence d'un champ magnétique non nul fait disparaître la transition abrupte observée dans le cas $h = 0$: $m(\beta, h)$

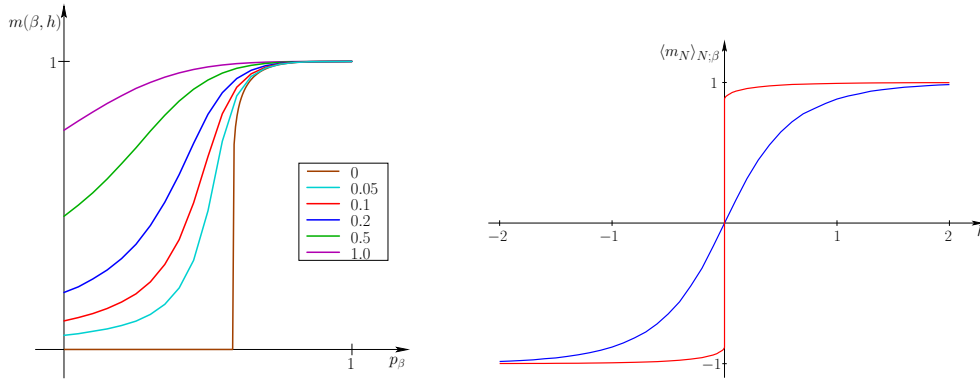


FIGURE 1.5 – Gauche : l’aimantation $m(\beta, h)$ en fonction de p_β pour le modèle d’Ising en dimension 2 avec condition au bord périodique, pour diverses valeurs de h . Droite : l’aimantation $m(\beta, h)$ en fonction du champ magnétique h pour le modèle d’Ising en dimension 2 avec condition au bord périodique, lorsque $p_\beta = 0.2$ et $p_\beta = 0.6$.

dépend de manière lisse de p_β . La dépendance de $m(\beta, h)$ par rapport au paramètre h (graphe de droite) rend par contre très manifeste la transition de phase : si $m(\beta, h)$ est une fonction lisse de h lorsque p_β est inférieur à la valeur critique, elle devient discontinue en $h = 0$ (où elle prend la valeur 0, alors que les limites $\lim_{h \downarrow 0} m(\beta, h) = -\lim_{h \uparrow 0} m(\beta, h)$ sont non nulles) lorsque p est supérieur à la valeur critique. En fait cette dernière limite coïncide avec la valeur obtenue pour $m^*(\beta)$ lorsque $h = 0$.

La présence des deux comportements typiques (aimantations positive et négative) lorsque $h = 0$ et $p > p_{\beta_c}$ peut ainsi être vue comme une trace de la sensibilité à la perturbation par un champ extérieur : pour $p < p_{\beta_c}$, l’introduction d’un petit champ magnétique $h > 0$ produit une aimantation positive approximativement proportionnelle (la “réponse” du système à la perturbation est linéaire pour petit h), alors que pour $p > p_{\beta_c}$, l’introduction d’un champ magnétique infinitésimal $h > 0$ produit une aimantation d’ordre 1 ! Intuitivement, on peut dire, dans ce dernier cas, qu’en champ magnétique nul le système “hésite” entre deux comportements différents, et que l’introduction d’un champ magnétique non nul arbitraire suffit à faire pencher la balance dans la direction correspondante.

1.2.4 Conditions au bord

Il y a une autre façon, très utile, de comprendre la présence de cette transition de phase comme résultant d’une sensibilité extrême du système à certaines perturbations : au lieu (ou en plus) de briser partout la symétrie entre les deux espèces de spin par l’introduction d’un champ magnétique, on peut également considérer l’effet de différentes conditions au bord. On ne considère alors plus le modèle sur un graphe fini, avec condition au bord périodique, mais sur un graphe infini (ici \mathbb{Z}^d) avec une configuration de spins *gelée* à l’extérieur de la boîte. Plus précisément, on associe à chaque configuration $\omega \in \Omega$ son

1.3. QUELQUES OUVRAGES DE RÉFÉRENCE

énergie à l'intérieur d'un sous-ensemble fini $V \Subset \mathbb{Z}^d$,

$$\mathcal{H}_{V;\beta,h}(\omega) \stackrel{\text{déf}}{=} -\beta \sum_{\substack{\{i,j\} \cap V \neq \emptyset \\ i \sim j}} \sigma_i(\omega) \sigma_j(\omega) - h \sum_{i \in V} \sigma_i(\omega).$$

Remarquez que cette énergie prend également en compte l'interaction entre les spins de V et ceux hors de V (grâce à la première somme).

Notons $\Omega \stackrel{\text{déf}}{=} \{-1, 1\}^{\mathbb{Z}^d}$ et $\Omega_V \stackrel{\text{déf}}{=} \{-1, 1\}^V$. Soit $\bar{\omega} \in \Omega$; étant donnée une configuration $\omega \in \Omega_V$, on construit une nouvelle configuration $\omega\bar{\omega} \in \Omega$ par

$$(\omega\bar{\omega})_i \stackrel{\text{déf}}{=} \begin{cases} \omega_i & \text{si } i \in V, \\ \bar{\omega}_i & \text{sinon.} \end{cases}$$

On introduit alors une nouvelle la mesure probabilité sur $(\Omega_V, \mathcal{F}_V)$,

$$\mu_{V;\beta,h}^{\bar{\omega}}(\omega) \stackrel{\text{déf}}{=} \frac{1}{Z_{V;\beta,h}^{\bar{\omega}}} \exp(-\mathcal{H}_{V;\beta,h}(\omega\bar{\omega})).$$

$\bar{\omega}$ est appelée la **condition au bord**. Cette dernière peut favoriser un type de spin au voisinage du bord de V . La question est de déterminer si cette information peut se propager dans V tout entier. Il n'est pas difficile à imaginer, et nous le démontrerons au chapitre 2, que deux configurations jouent un rôle extrême (dans le sens qu'elles favorisent au maximum la valeur 1, resp. -1) : $\omega_+ \equiv 1$ et $\omega_- \equiv -1$. On notera les mesures correspondantes simplement $\mu_{V;\beta,h}^+$ et $\mu_{V;\beta,h}^-$.

Soit $(V_N)_{N \geq 1}$ une suite croissante (pour l'inclusion) de sous-ensembles finis de \mathbb{Z}^d tels que $\bigcup_{N \geq 1} V_N = \mathbb{Z}^d$, ce que l'on notera $V_N \uparrow \mathbb{Z}$. On introduit alors les aimantations moyennes²

$$m^+(\beta) \stackrel{\text{déf}}{=} \lim_{N \rightarrow \infty} \langle |V_N|^{-1} \sum_{i \in V_N} \sigma_i \rangle_{V_N;\beta,0}^+,$$

$$m^-(\beta) \stackrel{\text{déf}}{=} \lim_{N \rightarrow \infty} \langle |V_N|^{-1} \sum_{i \in V_N} \sigma_i \rangle_{V_N;\beta,0}^-.$$

Nous verrons au chapitre 2 que $m^+(\beta) = \lim_{h \downarrow 0} m(\beta, h)$ et $m^-(\beta) = \lim_{h \uparrow 0} m(\beta, h) = -m^+(\beta)$. En particulier, $m^+(\beta) = m^-(\beta)$ si $\beta < \beta_c$, alors que $m^+(\beta) \neq m^-(\beta)$ si $\beta > \beta_c$.

1.3 Quelques ouvrages de référence

Le livre de Georgii [8] est probablement le livre le plus proche du cours; il est malheureusement plutôt difficile d'accès. Le livre de Simon [26] est plus facile d'accès, mais a un point de vue assez différent de celui du cours. En particulier, les techniques employées sont du type analyse fonctionnelle plutôt que probabilistes; il contient aussi une discussion détaillée du cas quantique.

2. Pour tout $C \subset \mathbb{Z}^d$, $|C| \stackrel{\text{déf}}{=} \#\{i \in C\}$.

Les livres de Ruelle [24], Sinai [27] et Israel [13] sont tous des classiques. Ils ont certes un peu vieilli, mais restent une excellente source d'informations. La longue préface du livre d'Israel est une très belle introduction à la thermodynamique.

Les livres de Prum [23] et Kindermann et Snell [15] sont plus introductifs. Le premier est en français, et le second est à présent disponible gratuitement sur internet.

1.3. QUELQUES OUVRAGES DE RÉFÉRENCE

Chapitre 2

Le modèle d'Ising

Notre but dans ce chapitre est de déterminer sous quelles conditions sur ses paramètres β et h , le modèle d'Ising possède une transition de phase, et de confirmer ainsi rigoureusement les observations faites dans l'introduction. Au passage, nous aurons l'occasion de découvrir quelques-unes des méthodes très variées (issues de la combinatoire, de l'algèbre, de l'analyse complexe, de l'analyse convexe, de la théorie des probabilités, etc.) utilisées dans l'étude mathématique des champs aléatoires.

2.1 Définition du modèle

Nous commençons par quelques définitions et notations, dont certaines diffèrent légèrement de celles utilisées dans l'introduction.

Configurations. Une configuration du modèle d'Ising est une application $\omega : \mathbb{Z}^d \rightarrow \{-1, 1\}$, $i \mapsto \omega(i) \equiv \omega_i$. L'ensemble de toutes les configurations, appelé **espace des configurations**, est noté Ω . On note $\omega|_\Lambda$ la restriction de la configuration $\omega \in \Omega$ à $\Lambda \subseteq \mathbb{Z}^d$, c'est-à-dire la famille $(\omega_i)_{i \in \Lambda}$. L'ensemble des configurations restreintes à $\Lambda \subseteq \mathbb{Z}^d$ est noté Ω_Λ . Étant donné $\omega \in \Omega_\Lambda$, $\Lambda \subseteq \mathbb{Z}^d$, et $\bar{\omega} \in \Omega$, on construit la configuration $\omega\bar{\omega} \in \Omega$ par $(\omega\bar{\omega})_i = \omega_i$ si $i \in \Lambda$, et $(\omega\bar{\omega})_i = \bar{\omega}_i$ sinon.

Spin. À chaque sommet $i \in \mathbb{Z}^d$, nous attachons une variable aléatoire $\sigma_i : \Omega \rightarrow \{-1, 1\}$, $\omega \mapsto \sigma_i(\omega) \stackrel{\text{déf}}{=} \omega_i$, le **spin** en i .

Hamiltonien. On associe à chaque configuration $\omega \in \Omega$ son énergie dans un domaine $\Lambda \Subset \mathbb{Z}^d$. Celle-ci dépend de deux paramètres : la **température inverse** $\beta \in \mathbb{R}$, que l'on suppose positive, et le **champ magnétique** $h \in \mathbb{R}$. Elle est donnée par l'**Hamiltonien** suivant

$$\mathcal{H}_{\Lambda; \beta, h}(\omega) \stackrel{\text{déf}}{=} -\beta \sum_{\substack{\{i, j\} \cap \Lambda \neq \emptyset \\ i \sim j}} \sigma_i(\omega) \sigma_j(\omega) - h \sum_{i \in \Lambda} \sigma_i(\omega).$$

Mesure de Gibbs. Soit $\Lambda \Subset \mathbb{Z}^d$. Nous allons à présent introduire une mesure de probabilité sur $(\Omega_\Lambda, \mathcal{F}_\Lambda)$, où \mathcal{F}_Λ est la tribu produit (et donc simplement l'ensemble des parties

2.2. QUELQUES INTERPRÉTATIONS

de Ω_Λ ici). On fixe également une configuration $\bar{\omega} \in \Omega$, la **condition au bord**. La **mesure de Gibbs** du modèle d'Ising avec interaction entre plus-proches-voisins, paramètres β et h , et condition au bord $\bar{\omega}$, est la mesure de probabilité sur $(\Omega_\Lambda, \mathcal{F}_\Lambda)$ définie par

$$\mu_{\Lambda;\beta,h}^{\bar{\omega}}(\omega) \stackrel{\text{déf}}{=} \frac{1}{Z_{\Lambda;\beta,h}^{\bar{\omega}}} \exp(-\mathcal{H}_{\Lambda;\beta,h}(\omega\bar{\omega})).$$

La constante de normalisation $Z_{\Lambda;\beta,h}^{\bar{\omega}}$ est appelée **fonction de partition**.

Par la suite, il sera souvent utile de voir cette mesure de probabilité comme étant définie sur (Ω, \mathcal{F}) , où \mathcal{F} est la tribu produit, c'est-à-dire la tribu engendrée par les **cylindres** $\mathcal{C}_{\eta,\Lambda'} \stackrel{\text{déf}}{=} \{\omega \in \Omega : \omega_{\Lambda'} = \eta\}$, $\Lambda' \subseteq \mathbb{Z}^d$, $\eta \in \Omega_{\Lambda'}$. Ceci peut se faire naturellement en posant, pour $\omega \in \Omega$,

$$\mu_{\Lambda;\beta,h}^{\bar{\omega}}(\omega) \stackrel{\text{déf}}{=} \begin{cases} \frac{1}{Z_{\Lambda;\beta,h}^{\bar{\omega}}} \exp(-\mathcal{H}_{\Lambda;\beta,h}(\omega)) & \text{si } \omega_{\Lambda^c} = \bar{\omega}_\Lambda \\ 0 & \text{sinon.} \end{cases}$$

En d'autres termes, on ne donne de poids qu'aux configurations coïncidant avec la condition au bord $\bar{\omega}$ hors de Λ . Nous ne ferons pas de distinction de notation entre ces deux mesures, et utiliserons l'une ou l'autre interprétation lorsque cela se révèlera avantageux.

Conditions au bord. Parmi les différentes conditions au bord, deux sont particulièrement intéressantes : $\bar{\omega} \equiv 1$, appelée condition au bord $+$, et $\bar{\omega} \equiv -1$, appelée condition au bord $-$. Les mesures de Gibbs correspondantes sont notées simplement $\mu_{\Lambda;\beta,h}^+$ et $\mu_{\Lambda;\beta,h}^-$. Il sera aussi parfois utile de considérer un autre type de condition au bord : la condition au bord libre. Il s'agit d'un système sans interaction avec l'extérieur. La mesure est définie comme précédemment, mais en restreignant, dans l'Hamiltonien, la somme aux couples $\{i, j\} \subseteq \Lambda$:

$$\mu_{\Lambda;\beta,h}^\emptyset(\omega) \stackrel{\text{déf}}{=} \frac{1}{Z_{\Lambda;\beta,h}^\emptyset} \exp\left(\beta \sum_{\substack{\{i,j\} \subseteq \Lambda \\ i \sim j}} \sigma_i(\omega)\sigma_j(\omega) + h \sum_{i \in \Lambda} \sigma_i(\omega)\right).$$

On note la mesure correspondante $\mu_{\Lambda;\beta,h}^\emptyset$.

2.2 Quelques interprétations

Du point de vue mathématique, l'intérêt de ce modèle est évident, puisqu'il s'agit d'un des modèles les plus simples de champ aléatoire (variables aléatoires binaires, interactions quadratiques, etc.), et on a déjà eu, dans l'introduction, un aperçu de la richesse de son comportement. Ce modèle est cependant aussi intéressant pour décrire qualitativement (et parfois quantitativement) une grande variété de phénomènes, dont on liste ici quelques exemples, principalement tirés de la physique.

2.2.1 Ferro-aimant

Il s'agit de l'interprétation originelle du modèle d'Ising. Le paramètre β correspond à l'inverse de la température. Les sommets de \mathbb{Z}^d correspondent aux positions des atomes

d'un réseau cristallin. Chaque atome possède un moment magnétique (le *spin*) qui est supposé ne pouvoir prendre que deux orientations, indexées par $+1$ ou -1 . Ce modèle associe donc une énergie plus basse lorsque les spins sont alignés entre eux, et alignés avec le champ magnétique extérieur h . Le but est d'expliquer et de décrire la transition entre les comportements paramagnétique et ferromagnétique du système lorsque sa température est changée. Quelle que soit sa température, un morceau de fer placé dans un champ magnétique va développer une aimantation en réponse à ce dernier (les spins s'alignant avec le champ magnétique). À haute température (au-dessus de 1043°K , la température de Curie du fer), cette aimantation disparaît lorsque le champ magnétique est enlevé ; on parle de comportement paramagnétique. À basse température (c'est-à-dire en-dessous de la température de Curie), l'aimantation développée en réponse au champ magnétique extérieur ne disparaît plus lorsque celui-ci est enlevé : il reste une aimantation résiduelle, appelée aimantation spontanée ; on parle de comportement ferromagnétique.

2.2.2 Gaz réticulaire

Dans cette interprétation, on suppose que \mathbb{R}^d a été partitionné en cellules cubiques de côté 1. Chaque cellule est soit vide ($\omega_i = -1$), soit occupée par une molécule ($\omega_i = +1$). Il est impossible d'avoir deux molécules dans la même cellule. Il est alors naturel de faire le changement de variables $n_i \stackrel{\text{def}}{=} (1 + \omega_i)/2$, n_i représentant alors le nombre de molécules (0 ou 1) dans la cellule i , et l'Hamiltonien prend la forme (à une constante additive triviale près) : $-\frac{1}{2}\beta' \sum_{i \sim j} n_i n_j - \mu \sum_i n_i$, où $\beta' \stackrel{\text{def}}{=} 4\beta$ à l'interprétation d'une température inverse, et $\mu \stackrel{\text{def}}{=} (\frac{1}{2}h - 2d)$ est le potentiel chimique (quantité qui mesure le coût énergétique lié à l'addition d'une molécule supplémentaire dans le système). L'énergie d'interaction favorise donc la condensation des molécules, puisque l'énergie décroît lorsque deux cellules voisines sont toutes deux occupées.

Ce que l'on désire comprendre dans ce cas-là, c'est la transition liquide/vapeur, c'est-à-dire la transition entre une phase dense et une phase diluée.

2.2.3 Alliage binaire

Dans cette interprétation, chaque sommet du réseau est soit occupé par un atome de type A ($\omega_i = -1$), soit un atome de type B ($\omega_i = +1$). Le terme d'interaction favorise le regroupement d'atomes de la même espèce, alors que le champ magnétique permet de contrôler les densités respectives des deux espèces (différence de potentiels chimiques), similairement à ce que l'on a vu pour le cas de l'interprétation en termes de gaz sur réseau.

On désire alors comprendre la transition de *démixtion* dans de tels alliages binaires : lorsque plusieurs composés dans un alliage sont partiellement miscibles et qu'un composé en solution a dépassé sa limite de solubilité, plusieurs phases de compositions différentes se forment.

2.2.4 Quelques autres interprétations et applications

Il y a de très nombreuses autres interprétations possibles (et utilisées) : agents économiques, modèle écologiques, etc.

Du côté des applications plus pratiques, ce modèle est utilisé en analyse et traitement d'images. Dans ce cas, il faut imaginer chaque sommet comme un pixel de l'image, et ω_i comme son état (allumé ou éteint). Divers algorithmes permettent alors d'effectuer des tâches aussi diverses que la restauration d'images bruitées, la segmentation, ou la détection des contours. Bien sûr, il est en général intéressant d'avoir plus de deux états par pixel (ce qui ne permet de décrire que des images monochromes), et on peut alors utiliser diverses généralisations du modèle d'Ising (par exemple le modèle de Potts, qui autorise $q \geq 2$ états, et se réduit au modèle d'Ising lorsque $q = 2$).

2.3 Les inégalités de corrélation

Le but de cette section est d'introduire un outil particulièrement important dans l'étude du modèle d'Ising : les inégalités de corrélations. La preuve de ces inégalités est reléguée à la fin du chapitre, *cf.* la section 2.10.

Parmi la multitude de telles inégalités, deux jouent un rôle prépondérant : les inégalités GKS et FKG.

2.3.1 Inégalités GKS

Les inégalités GKS (pour Griffiths, Kelly et Sherman [9, 14]) sont limitées aux conditions au bord + et libre et aux champs magnétiques positifs. Elles s'appliquent aux produits de spins, c'est-à-dire aux fonctions de la forme $\sigma_A \stackrel{\text{déf}}{=} \prod_{i \in A} \sigma_i$.

Théorème 2.3.1 (Inégalités GKS). *Soit $\Lambda \Subset \mathbb{Z}^d$, et $h \geq 0$. Quels que soient $A, B \subseteq \Lambda$ et $\beta \geq 0$,*

$$\langle \sigma_A \rangle_{\Lambda; \beta, h}^+ \geq 0, \quad (2.1)$$

$$\langle \sigma_A \sigma_B \rangle_{\Lambda; \beta, h}^+ \geq \langle \sigma_A \rangle_{\Lambda; \beta, h}^+ \langle \sigma_B \rangle_{\Lambda; \beta, h}^+. \quad (2.2)$$

Ces inégalités restent vraies pour la mesure $\mu_{\Lambda; \beta, h}^\emptyset$.

2.3.2 Inégalités FKG

Les inégalités FKG (pour Fortuin, Kasteleyn et Ginibre [7]) sont une extension de l'inégalité suivante sur les fonctions réelles : soit f et g deux fonctions croissantes de \mathbb{R} dans \mathbb{R} et μ une mesure de probabilité sur \mathbb{R} ; alors

$$\langle fg \rangle_\mu \geq \langle f \rangle_\mu \langle g \rangle_\mu.$$

La preuve est élémentaire, puisqu'il suffit de dupliquer le système :

$$\langle fg \rangle_\mu - \langle f \rangle_\mu \langle g \rangle_\mu = \frac{1}{2} \langle (f(x) - f(x'))(g(x) - g(x')) \rangle_{\mu \otimes \mu}$$

et d'observer que $f(x) - f(x')$ et $g(x) - g(x')$ ont nécessairement le même signe puisque f et g sont toutes deux croissantes.

Soit $\Lambda \Subset \mathbb{Z}^d$. L'ensemble $\{-1, 1\}$ étant totalement ordonné, on peut définir un ordre partiel sur Ω_Λ : $\omega \leq \omega'$ si et seulement si $\omega_i \leq \omega'_i$ pour tout $i \in \Lambda$. Ceci permet de définir la notion de fonction croissante : $f : \Omega_\Lambda \rightarrow \mathbb{R}$ est croissante si et seulement si $\omega \leq \omega' \implies f(\omega) \leq f(\omega')$. Une classe importante de fonctions croissantes est formée des produits des variables d'occupation $n_i \stackrel{\text{déf}}{=} \frac{1}{2}(1 + \sigma_i) : n_A \stackrel{\text{déf}}{=} \prod_{i \in A} n_i$.

Les inégalités FKG affirment que les fonctions croissantes sont positivement corrélées. Leur grand avantage est d'être applicables quelle que soit la condition au bord, et pour toute valeur (pas seulement positive) du champ magnétique.

Théorème 2.3.2 (Inégalités FKG). *Soit $\Lambda \Subset \mathbb{Z}^d$ et $\bar{\omega}$ une condition au bord arbitraire. Alors, quels que soient $\beta \geq 0$ et $h \in \mathbb{R}$, pour toute paire de fonctions croissantes f et g ,*

$$\langle fg \rangle_{\Lambda; \beta, h}^{\bar{\omega}} \geq \langle f \rangle_{\Lambda; \beta, h}^{\bar{\omega}} \langle g \rangle_{\Lambda; \beta, h}^{\bar{\omega}}. \quad (2.3)$$

2.4 Limite thermodynamique

On n'a pour l'instant défini le modèle d'Ising que pour des sous-ensembles finis de \mathbb{Z}^d . Une façon naturelle de définir le modèle sur \mathbb{Z}^d tout entier (on dit parfois "en volume infini") est de considérer une suite de parties finies $\Lambda_n \uparrow \mathbb{Z}^d$, et une suite de conditions au bord $(\bar{\omega}_n)_{n \geq 1}$. On aimerait à présent prendre la limite de la suite de mesures $(\mu_{\Lambda_n; \beta, h}^{\bar{\omega}_n})_{n \geq 1}$. Pour cela, il est nécessaire d'introduire une topologie appropriée.

On dit qu'une fonction $f : \Omega \rightarrow \mathbb{R}$ est locale si elle est mesurable par rapport à \mathcal{F}_Λ pour un Λ fini (c'est-à-dire qu'elle ne dépend que des spins situés à l'intérieur de Λ) ; en d'autres termes, une fonction est locale si elle ne dépend que de l'état d'un nombre fini de spins. On note $\text{supp}(f)$ le support de la fonction f , c'est-à-dire le plus petit ensemble de sommets dont les spins déterminent la valeur de f .

On dira qu'une suite de mesures μ_n sur (Ω, \mathcal{F}) converge vers la mesure μ si et seulement si

$$\lim_{n \rightarrow \infty} \langle f \rangle_{\mu_n} \rightarrow \langle f \rangle_\mu,$$

pour toute fonction locale f .

On appellera mesure de Gibbs (en volume infini) du modèle d'Ising tout point d'accumulation des suites $(\mu_{\Lambda_n; \beta, h}^{\bar{\omega}_n})_{n \geq 1}$ introduites ci-dessus.

Remarque 2.4.1. *Supposons que l'on ait montré que $\langle f \rangle_{\mu_n}$ converge vers une valeur $c(f)$ lorsque $n \rightarrow \infty$, pour toute fonction locale f . Peut-on en déduire l'existence d'une mesure de probabilité μ sur (Ω, \mathcal{F}) telle que $c(f) = \langle f \rangle_\mu$? La réponse est positive, comme le montre l'argument suivant.*

Soit $\mathcal{F}_\Lambda = \bigcup_{C \subseteq \Omega_\Lambda} \{\sigma|_C \in C\}$ la tribu des événements ne dépendant que des spins dans $\Lambda \Subset \mathbb{Z}^d$, et $\mathcal{C} = \bigcup_{\Lambda \Subset \mathbb{Z}^d} \mathcal{F}_\Lambda$ l'algèbre des cylindres fini-dimensionnels. On définit une fonction $\mu : \mathcal{C} \rightarrow [0, 1]$ par

$$\mu(A) \stackrel{\text{déf}}{=} \lim_{n \rightarrow \infty} \mu_n(A),$$

2.4. LIMITE THERMODYNAMIQUE

ce qui peut se faire puisque A est un événement local. Nous allons vérifier à présent que μ est une mesure de probabilité sur \mathcal{C} , et par conséquent peut être étendue, grâce au théorème d'extension de Carathéodory, à une mesure de probabilité sur $\mathcal{F} = \sigma(\mathcal{C})$.

Manifestement, $\mu(\Omega) = 1$. Vérifions à présent que μ est finiment additive : soient $A_1, A_2 \in \mathcal{C}$, $A_1 \cap A_2 = \emptyset$. Alors,

$$\mu(A_1 \cup A_2) = \lim_{n \rightarrow \infty} \mu_n(A_1 \cup A_2) = \lim_{n \rightarrow \infty} (\mu_n(A_1) + \mu_n(A_2)) = \mu(A_1) + \mu(A_2).$$

Il reste donc, pour établir la σ -additivité, à montrer que toute suite décroissante $A_1 \supseteq A_2 \supseteq \dots$ telle que, pour tout i , $A_i \in \mathcal{C}$ et $\mu(A_i) \geq c > 0$, satisfait $\bigcap_i A_i \neq \emptyset$. Puisque $A_i \neq \emptyset$, pour tout i , on peut trouver une suite de configurations $\omega^{(k)} \in A_k$, $k = 1, 2, \dots$. Ordonnons totalement les sites de $\mathbb{Z}^d : i_1, i_2, \dots$. Puisque $\omega_{i_1} \in \{-1, 1\}$, il existe une sous-suite $(k_j^1)_j$ le long de laquelle $\omega_{i_1}^{(k_j^1)} = \text{const} \equiv \bar{\omega}_{i_1}$. On peut alors en extraire une nouvelle sous-suite $(k_j^2)_j$ le long de laquelle $\omega_{i_2}^{(k_j^2)} = \text{const} \equiv \bar{\omega}_{i_2}$. En procédant de la sorte pour tous les sites, on obtient finalement une configuration $\bar{\omega}$ avec la propriété que $\bar{\omega} \in A_n$, pour tout $n \geq 1$. En effet, quel que soit $n \geq 1$, il existe $j_0(n)$ suffisamment grand tel que $\omega^{k_j^j} \equiv \bar{\omega}$ sur le support de A_n pour tout $j > j_0(n)$, et donc $\bar{\omega} \in A_n$. Par conséquent, $\bigcap_n A_n \neq \emptyset$.

Comme cela a déjà été mentionné dans l'introduction, il n'y a pas nécessairement unicité de la mesure limite associée à un couple (β, h) donné (l'existence d'au moins un point d'accumulation est par contre garantie par compacité de l'espace des mesures de probabilité sur (Ω, \mathcal{F}) pour la topologie que l'on vient d'introduire). On dira qu'il y a transition de phase du 1er ordre en (β, h) si et seulement si il y a non-unicité. Nous verrons plus loin que cette définition correspond précisément à la transition de phase observée dans l'introduction via l'apparition d'une discontinuité de l'aimantation comme fonction du champ magnétique.

La topologie introduite reposant sur les fonctions locales, il est utile de déterminer certaines classes de fonctions locales engendrant toutes les autres. Le lemme suivant en propose deux, qui sont particulièrement adaptées aux inégalités de corrélation présentées dans la section précédente.

Lemme 2.4.1. *Soit f une fonction locale. Alors*

$$f = \sum_{A \subseteq \text{supp}(f)} \hat{f}_A \sigma_A,$$

où $\hat{f}_A \stackrel{\text{déf}}{=} 2^{-|\text{supp}(f)|} \sum_{\tilde{\omega} \in \Omega_{\text{supp}(f)}} f(\tilde{\omega}) \sigma_A(\tilde{\omega})$. Par conséquent, on a également

$$f = \sum_{A \subseteq \text{supp}(f)} \tilde{f}_A n_A,$$

pour des coefficients \tilde{f}_A appropriés.

Démonstration. Pour démontrer la première affirmation, il suffit d'utiliser la relation

$$2^{-|B|} \sum_{A \subseteq B} \sigma_A(\tilde{\omega}) \sigma_A(\omega) = \mathbf{1}_{\{\omega \equiv \tilde{\omega} \text{ sur } B\}}.$$

Cette relation d'orthogonalité se démontre très facilement. Supposons tout d'abord que $\omega_i = \tilde{\omega}_i$, pour tout $i \in B$. Alors $\sigma_A(\tilde{\omega})\sigma_A(\omega) = \prod_{i \in A} \tilde{\omega}_i \omega_i = 1$, puisque $\tilde{\omega}_i \omega_i = 1$ pour tout $i \in B$. Supposons donc à présent qu'il existe $i \in B$ tel que $\omega_i \neq \tilde{\omega}_i$ (et donc $\omega_i \tilde{\omega}_i = -1$); on a alors

$$\begin{aligned} \sum_{A \subseteq B} \sigma_A(\tilde{\omega})\sigma_A(\omega) &= \sum_{A \subseteq B \setminus \{i\}} \left(\sigma_A(\tilde{\omega})\sigma_A(\omega) + \sigma_{A \cup \{i\}}(\tilde{\omega})\sigma_{A \cup \{i\}}(\omega) \right) \\ &= \sum_{A \subseteq B \setminus \{i\}} (\sigma_A(\tilde{\omega})\sigma_A(\omega) + \omega_i \tilde{\omega}_i \sigma_A(\tilde{\omega})\sigma_A(\omega)) \\ &= \sum_{A \subseteq B \setminus \{i\}} (\sigma_A(\omega \tilde{\omega}) + (-1)\sigma_A(\omega \tilde{\omega})) = 0. \end{aligned}$$

La seconde affirmation suit immédiatement en insérant $\sigma_A = \prod_{i \in A} (2n_i - 1)$ dans la première formule. \square

Les inégalités de corrélation introduites dans la section précédente permettent de démontrer de nombreux résultats concernant les limites possibles. Commençons par montrer que les mesures avec conditions au bord $+$ et $-$ admettent bien chacune une limite, quelle que soit la suite $\Lambda_n \uparrow \mathbb{Z}^d$ considérée.

Théorème 2.4.1. *Soit $\beta \geq 0$ et $h \in \mathbb{R}$. Pour toute suite $\Lambda_n \uparrow \mathbb{Z}^d$, la suite de mesures $(\mu_{\Lambda_n; \beta, h}^+)_n$ converge vers une mesure $\mu_{\beta, h}^+$ (indépendante de la suite $(\Lambda_n)_{n \geq 1}$). De même, la suite $(\mu_{\Lambda_n; \beta, h}^-)_n$ converge vers une mesure $\mu_{\beta, h}^-$ (indépendante de la suite $(\Lambda_n)_{n \geq 1}$).*

De plus, les mesures $\mu_{\beta, h}^+$ et $\mu_{\beta, h}^-$ sont invariantes sous l'action du groupe des translations de \mathbb{Z}^d : si θ_t est la translation par $t \in \mathbb{Z}^d$, $\langle f \circ \theta_t \rangle_{\beta, h}^+ = \langle f \rangle_{\beta, h}^+$.

Démonstration. Montrons tout d'abord la convergence de la suite de mesures $\mu_{\Lambda_n; \beta, h}^+$. Les inégalités FKG impliquent que pour toute fonction f locale et croissante,

$$\langle f \rangle_{\Lambda_n; \beta, h}^+ \geq \langle f \rangle_{\Lambda_{n+1}; \beta, h}^+. \quad (2.4)$$

En effet, on vérifie facilement que

$$\langle f \rangle_{\Lambda_n; \beta, h}^+ = \langle f \mid \sigma_i = 1, \forall i \in \Lambda_{n+1} \setminus \Lambda_n \rangle_{\Lambda_{n+1}; \beta, h}^+.$$

Comme l'indicatrice $\mathbf{1}_{\{\sigma_i=1, \forall i \in \Lambda_{n+1} \setminus \Lambda_n\}}$ est une fonction croissante, il suit bien des inégalités FKG que

$$\begin{aligned} \langle f \rangle_{\Lambda_n; \beta, h}^+ &= \frac{\langle f \mathbf{1}_{\{\sigma_i=1, \forall i \in \Lambda_{n+1} \setminus \Lambda_n\}} \rangle_{\Lambda_{n+1}; \beta, h}^+}{\langle \mathbf{1}_{\{\sigma_i=1, \forall i \in \Lambda_{n+1} \setminus \Lambda_n\}} \rangle_{\Lambda_{n+1}; \beta, h}^+} \\ &\geq \frac{\langle f \rangle_{\Lambda_{n+1}; \beta, h}^+ \langle \mathbf{1}_{\{\sigma_i=1, \forall i \in \Lambda_{n+1} \setminus \Lambda_n\}} \rangle_{\Lambda_{n+1}; \beta, h}^+}{\langle \mathbf{1}_{\{\sigma_i=1, \forall i \in \Lambda_{n+1} \setminus \Lambda_n\}} \rangle_{\Lambda_{n+1}; \beta, h}^+} = \langle f \rangle_{\Lambda_{n+1}; \beta, h}^+. \end{aligned}$$

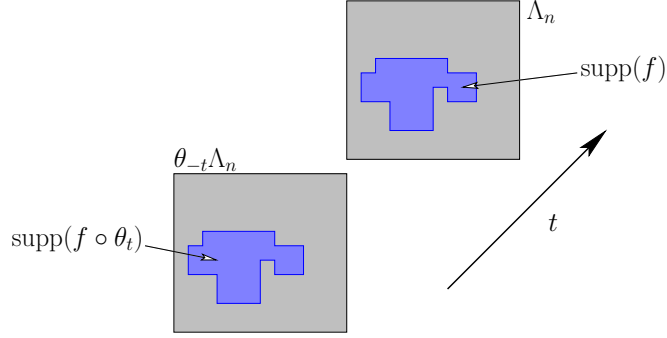


FIGURE 2.1 – Preuve de l'invariance sous les translations.

f étant nécessairement bornée, on en déduit la convergence de la suite $(\langle f \rangle_{\Lambda_n; \beta, h}^+)_n$ pour toute fonction locale croissante f . Par conséquent, une application du lemme 2.4.1 montre que, pour toute fonction locale g ,

$$\begin{aligned} \lim_{n \rightarrow \infty} \langle g \rangle_{\Lambda_n; \beta, h}^+ &= \sum_{A \subseteq \text{supp}(g)} \tilde{g}_A \lim_{n \rightarrow \infty} \langle n_A \rangle_{\Lambda_n; \beta, h}^+ \\ &= \sum_{A \subseteq \text{supp}(g)} \tilde{g}_A \langle n_A \rangle_{\beta, h}^+ = \langle g \rangle_{\beta, h}^+, \end{aligned}$$

puisque les fonctions n_A sont locales et croissantes. Par conséquent, on a bien convergence de la suite $(\mu_{\Lambda_n; \beta, h}^+)_n$ vers une mesure $\mu_{\beta, h}^+$.

Vérifions à présent que la limite ne dépend pas de la suite $\Lambda_n \uparrow \mathbb{Z}^d$ considérée. Soit $\Lambda_n^1 \uparrow \mathbb{Z}^d$ et $\Lambda_n^2 \uparrow \mathbb{Z}^d$ deux telles suites, et notons $\mu_{\beta, h}^{+,1}$ et $\mu_{\beta, h}^{+,2}$ les limites correspondantes. On construit une nouvelle suite $\Delta_n \uparrow \mathbb{Z}^d$ de la façon suivante : $\Delta_1 = \Lambda_1^1$, et, pour $k \geq 1$,

$$\begin{aligned} \Delta_{2k} &= \bigcap \left\{ \Lambda_n^2 : \Lambda_n^2 \supseteq \Delta_{2k-1} \right\}, \\ \Delta_{2k+1} &= \bigcap \left\{ \Lambda_n^1 : \Lambda_n^1 \supseteq \Delta_{2k} \right\}. \end{aligned}$$

La convergence de la suite $(\mu_{\Delta_n; \beta, h}^+)_{n \geq 1}$ implique donc que $\mu_{\beta, h}^{+,1} = \mu_{\beta, h}^{+,2}$, ces dernières correspondant, respectivement, aux limites obtenues pour les sous-suites $(\mu_{\Delta_{2n+1}; \beta, h}^+)_{n \geq 1}$ et $(\mu_{\Delta_{2n}; \beta, h}^+)_{n \geq 1}$.

La preuve pour la suite $(\mu_{\Lambda_n; \beta, h}^-)_n$ est identique.

L'invariance sous les translations est immédiate. Soit $t \in \mathbb{Z}^d$, f une fonction locale et $\Lambda_n \uparrow \mathbb{Z}^d$. Alors $f \circ \theta_t$ est également une fonction locale et $\theta_{-t}\Lambda_n \uparrow \mathbb{Z}^d$ ($\theta_{-t}A \stackrel{\text{déf}}{=} A - t$, pour tout $A \subseteq \mathbb{Z}^d$). On a donc

$$\langle f \rangle_{\Lambda_n; \beta, h}^+ \rightarrow \langle f \rangle_{\beta, h}^+ \quad \text{et} \quad \langle f \circ \theta_t \rangle_{\theta_{-t}\Lambda_n; \beta, h}^+ \rightarrow \langle f \circ \theta_t \rangle_{\beta, h}^+,$$

et la conclusion suit puisque $\langle f \circ \theta_t \rangle_{\theta_{-t}\Lambda_n; \beta, h}^+ = \langle f \rangle_{\Lambda_n; \beta, h}^+$. □

2.5 Critères d'unicité/non-unicité

L'importance des mesures $\mu_{\beta,h}^+$ et $\mu_{\beta,h}^-$ est rendue manifeste par le résultat suivant, qui permet de réduire le problème de l'unicité/non-unicité des mesures de Gibbs du modèle d'Ising à celui de comparer uniquement ces deux mesures.

Théorème 2.5.1. *Les affirmations suivantes sont équivalentes :*

1. *Il y a une unique mesure de Gibbs en volume infini.*
2. $\mu_{\beta,h}^+ = \mu_{\beta,h}^-$.
3. $\langle \sigma_0 \rangle_{\beta,h}^+ = \langle \sigma_0 \rangle_{\beta,h}^-$.

Démonstration. Une application des inégalités FKG montre que, pour toute fonction f locale et croissante, dont le support est inclus dans Λ_n ,

$$\langle f \rangle_{\Lambda_n;\beta,h}^- \leq \langle f \rangle_{\Lambda_n;\beta,h}^{\bar{\omega}} \leq \langle f \rangle_{\Lambda_n;\beta,h}^+ , \quad (2.5)$$

quelle que soit la condition au bord $\bar{\omega}$. En effet, en utilisant que, pour tout $\omega \in \Omega_{\Lambda_n}$ et $\bar{\omega}, \tilde{\omega} \in \Omega$,

$$\mathcal{H}_{\Lambda;\beta,h}(\omega\bar{\omega}) = \mathcal{H}_{\Lambda;\beta,h}(\omega\tilde{\omega}) + \beta \sum_{i \in \Lambda, j \notin \Lambda} \omega_i(\tilde{\omega}_j - \bar{\omega}_j),$$

on a par exemple

$$\langle f \rangle_{\Lambda_n;\beta,h}^+ = \frac{\langle e^{\beta \sum_{i \in \Lambda, j \notin \Lambda} (1 - \bar{\omega}_j) \sigma_i} f \rangle_{\Lambda_n;\beta,h}^{\bar{\omega}}}{\langle e^{\beta \sum_{i \in \Lambda, j \notin \Lambda} (1 - \bar{\omega}_j) \sigma_i} \rangle_{\Lambda_n;\beta,h}^{\bar{\omega}}} \geq \langle f \rangle_{\Lambda_n;\beta,h}^{\bar{\omega}},$$

puisque la fonction $\exp(\beta \sum_{i \in \Lambda, j \notin \Lambda} (1 - \bar{\omega}_j) \sigma_i)$ est croissante.

On en déduit qu'il y a une unique mesure de Gibbs si et seulement si $\mu_{\beta,h}^- = \mu_{\beta,h}^+$. Pour montrer qu'il est suffisant de vérifier $\langle \sigma_0 \rangle_{\beta,h}^- = \langle \sigma_0 \rangle_{\beta,h}^+$, on utilise une fois de plus les inégalités FKG. On observe tout d'abord que pour tout $A \Subset \mathbb{Z}^d$, la fonction $\sum_{i \in A} n_i - n_A$ est croissante. Par conséquent, (2.5) implique que

$$\langle \sum_{i \in A} n_i - n_A \rangle_{\Lambda_n;\beta,h}^- \leq \langle \sum_{i \in A} n_i - n_A \rangle_{\Lambda_n;\beta,h}^+ ,$$

ce qui nous fournit l'inégalité suivante sur les mesures limites,

$$\sum_{i \in A} (\langle n_i \rangle_{\beta,h}^+ - \langle n_i \rangle_{\beta,h}^-) \geq \langle n_A \rangle_{\beta,h}^+ - \langle n_A \rangle_{\beta,h}^- .$$

Il suit de (2.5) que le membre de droite de cette dernière inégalité est positif. Par invariance sous les translations, la condition $\langle \sigma_0 \rangle_{\beta,h}^- = \langle \sigma_0 \rangle_{\beta,h}^+$ (et donc $\langle n_0 \rangle_{\beta,h}^- = \langle n_0 \rangle_{\beta,h}^+$) implique que le membre de gauche est nul. Il suit que $\langle n_A \rangle_{\beta,h}^+ = \langle n_A \rangle_{\beta,h}^-$ pour tout A fini. Ceci conclut la démonstration grâce au Lemme 2.4.1. \square

2.6 Aimantation et transition de phase

Dans cette section, nous relierons la présence de plusieurs mesures de Gibbs en volume infini avec certaines propriétés de l'aimantation, faisant ainsi le pont entre les résultats mathématiques démontrés dans ce chapitre et les observations faites dans l'introduction. Pour ce faire, il sera nécessaire d'introduire une quantité très importante dans l'étude de ces champs aléatoires : l'énergie libre.

2.6.1 Quelques propriétés élémentaires de l'aimantation

Le critère d'unicité de la section 2.5 fait intervenir les espérances $\langle \sigma_0 \rangle_{\beta, h}^+$ et $\langle \sigma_0 \rangle_{\beta, h}^-$. La première question que l'on peut se poser est s'il y a une relation entre ces espérances et l'espérance de l'aimantation dans une boîte finie. Le lemme suivant montre qu'elles coïncident dans la limite, au moins pour de bonnes boîtes. C'est la raison pour laquelle nous appellerons également aimantation ces espérances.

Proposition 2.6.1. *Soit $\Lambda_n = \{-n, \dots, n\}^d$. Alors*

$$\langle \sigma_0 \rangle_{\beta, h}^+ = \lim_{n \rightarrow \infty} \langle |\Lambda_n|^{-1} \sum_{i \in \Lambda_n} \sigma_i \rangle_{\beta, h}^+,$$

et similairement pour $\langle \sigma_0 \rangle_{\beta, h}^-$.

Démonstration. On utilise (2.4). D'une part,

$$\min_{i \in \Lambda_n} \langle \sigma_i \rangle_{\Lambda_n, \beta, h}^+ \geq \langle \sigma_0 \rangle_{\Lambda_{2n}, \beta, h}^+,$$

et donc

$$\lim_{n \rightarrow \infty} \frac{1}{|\Lambda_n|} \sum_{i \in \Lambda_n} \langle \sigma_i \rangle_{\Lambda_n, \beta, h}^+ \geq \lim_{n \rightarrow \infty} \langle \sigma_0 \rangle_{\Lambda_{2n}, \beta, h}^+ = \langle \sigma_0 \rangle_{\beta, h}^+.$$

D'autre part, soit $R > 0$; pour tout $i \in \Lambda_N$ se trouvant à distance au moins R du bord de Λ_N , $\langle \sigma_i \rangle_{\Lambda_n, \beta, h}^+ \leq \langle \sigma_0 \rangle_{\Lambda_R, \beta, h}^+$. On a donc

$$\lim_{n \rightarrow \infty} \frac{1}{|\Lambda_n|} \sum_{i \in \Lambda_n} \langle \sigma_i \rangle_{\Lambda_n, \beta, h}^+ \leq \langle \sigma_0 \rangle_{\Lambda_R, \beta, h}^+.$$

La conclusion suit en laissant $R \rightarrow \infty$. □

Montrons à présent quelques propriétés élémentaires de ces fonctions.

Proposition 2.6.2. *1. Soit $\Lambda \in \mathbb{Z}^d$. $\langle \sigma_0 \rangle_{\Lambda; \beta, h}^+$ et $\langle \sigma_0 \rangle_{\Lambda; \beta, h}^-$ sont des fonctions croissantes de h , pour tout $\beta \geq 0$.*
2. Soit $\Lambda \in \mathbb{Z}^d$. $\langle \sigma_0 \rangle_{\Lambda; \beta, h}^+$ (resp. $\langle \sigma_0 \rangle_{\Lambda; \beta, h}^-$) est une fonction croissante (resp. décroissante) de β , pour tout $h \geq 0$ (resp. $h \leq 0$).
3. $\langle \sigma_0 \rangle_{\beta, h}^+$ (resp. $\langle \sigma_0 \rangle_{\beta, h}^-$) est une fonction continue à droite (resp. à gauche) du champ magnétique h .

Remarque 2.6.1. Évidemment, les propriétés de monotonie sont encore vérifiées dans la limite thermodynamique. Il est en fait possible de montrer que $\langle \sigma_0 \rangle_{\beta, h}^+$ est concave (et en particulier continue) pour $h \geq 0$. Nous ne le ferons pas ici, la preuve reposant sur une inégalité de corrélation, l'inégalité GHS, que nous ne discuterons pas. Par contre, nous prouverons plus loin (Théorème 2.9.1) que $\langle \sigma_0 \rangle_{\beta, h}^+$ est en fait une fonction analytique de h pour $h \neq 0$.

Par symétrie, on a des résultats analogues pour $\langle \sigma_0 \rangle_{\beta, h}^-$ et h négatif.

Démonstration. 1. C'est une conséquence immédiate de (2.3) :

$$\frac{\partial}{\partial h} \langle \sigma_0 \rangle_{\Lambda; \beta, h}^+ = \sum_{i \in \Lambda} \left(\langle \sigma_0 \sigma_i \rangle_{\Lambda; \beta, h}^+ - \langle \sigma_0 \rangle_{\Lambda; \beta, h}^+ \langle \sigma_i \rangle_{\Lambda; \beta, h}^+ \right) \geq 0.$$

2. C'est une conséquence immédiate de (2.2) :

$$\frac{\partial}{\partial \beta} \langle \sigma_0 \rangle_{\Lambda; \beta, h}^+ = \sum_{\substack{\{i, j\} \cap \Lambda \neq \emptyset \\ i \sim j}} \left(\langle \sigma_0 \sigma_i \sigma_j \rangle_{\Lambda; \beta, h}^+ - \langle \sigma_0 \rangle_{\Lambda; \beta, h}^+ \langle \sigma_i \sigma_j \rangle_{\Lambda; \beta, h}^+ \right) \geq 0.$$

3. Soit $(h_k)_{k \geq 1}$ une suite telle que $h_k \downarrow h$. D'une part, $\langle \sigma_0 \rangle_{\beta, h}^+$ est une fonction monotone de h , par le point 1, et donc la limite des $\langle \sigma_0 \rangle_{\beta, h_k}^+$ existe et satisfait

$$\lim_{h_k \downarrow h} \langle \sigma_0 \rangle_{\beta, h_k}^+ \geq \langle \sigma_0 \rangle_{\beta, h}^+.$$

D'autre part, il suit de (2.4) que, pour tout $\Lambda \Subset \mathbb{Z}^d$,

$$\lim_{h_k \downarrow h} \langle \sigma_0 \rangle_{\beta, h_k}^+ \leq \lim_{h_k \downarrow h} \langle \sigma_0 \rangle_{\Lambda; \beta, h_k}^+ = \langle \sigma_0 \rangle_{\Lambda; \beta, h}^+,$$

puisque $\langle \sigma_0 \rangle_{\Lambda; \beta, h}^+$ est évidemment une fonction continue de h . On conclut en laissant $\Lambda \uparrow \mathbb{Z}^d$. \square

2.6.2 Énergie libre

Nous allons avoir besoin d'une notion appropriée de convergence de parties finies de \mathbb{Z}^d vers \mathbb{Z}^d tout entier. L'idée est que l'on désire extraire les effets de volume, et négliger les corrections dues aux effets de surface. Nous dirons qu'une suite $(\Lambda_n)_{n \geq 1}$ telle que $\Lambda_n \uparrow \mathbb{Z}^d$ converge vers \mathbb{Z}^d au sens de van Hove, $\Lambda_n \uparrow \mathbb{Z}^d$, si et seulement si

$$\lim_{n \rightarrow \infty} |\partial \Lambda_n| / |\Lambda_n| = 0, \quad (2.6)$$

où $\partial C \stackrel{\text{def}}{=} \{i \in C : \exists j \notin C, j \sim i\}$.

Théorème 2.6.1. Soit $f_{\Lambda}^{\bar{\omega}}(\beta, h) \stackrel{\text{def}}{=} |\Lambda|^{-1} \log \mathbf{Z}_{\Lambda; \beta, h}^{\bar{\omega}}$. La limite

$$f(\beta, h) \stackrel{\text{def}}{=} \lim_{\Lambda \uparrow \mathbb{Z}^d} f_{\Lambda}^{\bar{\omega}}(\beta, h)$$

existe et est indépendante de $\bar{\omega}$ et de la suite $\Lambda \uparrow \mathbb{Z}^d$. De plus, la fonction $f(\beta, h)$ est convexe.

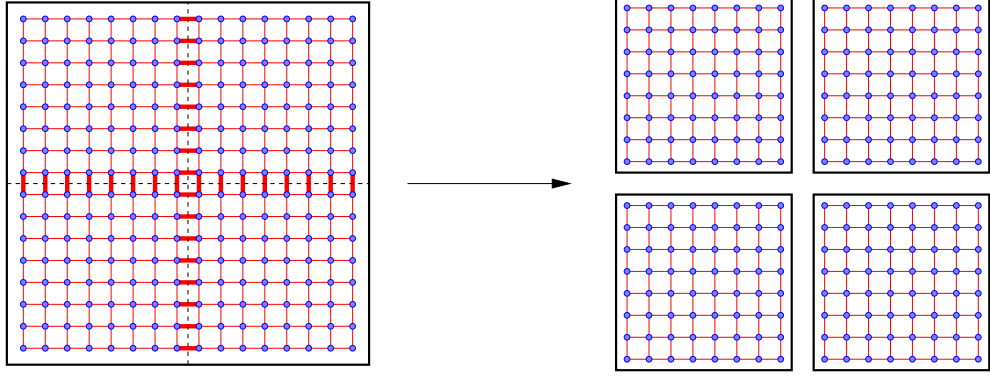


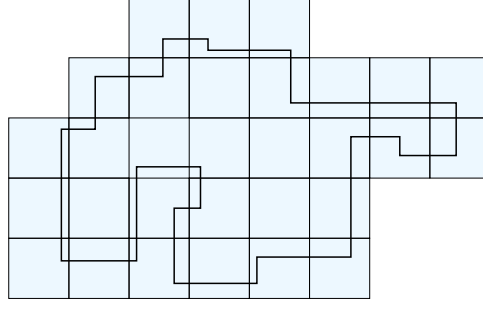
FIGURE 2.2 – Un cube B_{n+1} et sa partition en 2^d cubes B_n . Les interactions entre les différents sous-cubes sont indiquées.

La quantité $f(\beta, h)$ joue un rôle essentiel dans la relation entre physique statistique et thermodynamique; on l'appelle **énergie libre**. Nous verrons plus bas que les propriétés analytiques de cette fonction sont profondément liées à la présence de transitions de phase.

Démonstration. Existence de la limite. On commence par démontrer la convergence dans le cas de la condition au bord libre. La preuve se fait en deux étapes : tout d'abord pour une suite de boîtes cubiques, puis pour des boîtes générales.

Soit $B_n \stackrel{\text{déf}}{=} \{i = (i_1, \dots, i_d) \in \mathbb{Z}^d : 1 \leq i_k \leq 2^n\}$. L'énergie libre associée à la boîte B_{n+1} peut aisément être comparée à celle associée à la boîte B_n . En effet, si l'on décompose B_{n+1} en 2^d copies disjointes de B_n , notées $B_n^{(1)}, \dots, B_n^{(2^d)}$ (cf. Figure 2.2), alors l'énergie provenant de l'interaction entre les spins de deux sous-boîtes différentes est bornée par β multiplié par le nombre de sites dans une face de B_n , soit $2^{n(d-1)}$. Comme il y a précisément $\frac{1}{2}d2^d$ faces à considérer (chacune étant partagée entre deux cubes), on en déduit que

$$\begin{aligned}
 \mathbf{Z}_{B_{n+1}; \beta, h}^\varnothing &= \sum_{\omega \in \Omega_{B_{n+1}}} \prod_{\substack{\{i, j\} \subseteq B_{n+1} \\ i \sim j}} e^{\beta \omega_i \omega_j} \prod_{i \in B_{n+1}} e^{h \omega_i} \\
 &\geq e^{-\beta \frac{1}{2} d 2^d 2^{n(d-1)}} \sum_{\omega \in \Omega_{B_{n+1}}} \prod_{k=1}^{2^d} \prod_{\substack{\{i, j\} \subseteq B_n^{(k)} \\ i \sim j}} e^{\beta \omega_i \omega_j} \prod_{i \in B_n^{(k)}} e^{h \omega_i} \\
 &= e^{-\beta d 2^{(n+1)(d-1)}} \prod_{k=1}^{2^d} \sum_{\omega \in \Omega_{B_n^{(k)}}} \prod_{\substack{\{i, j\} \subseteq B_n^{(k)} \\ i \sim j}} e^{\beta \omega_i \omega_j} \prod_{i \in B_n^{(k)}} e^{h \omega_i} \\
 &= e^{-\beta d 2^{(n+1)(d-1)}} \left(\mathbf{Z}_{B_n; \beta, h}^\varnothing \right)^{2^d},
 \end{aligned}$$


 FIGURE 2.3 – Un ensemble Λ_n et un recouvrement possible par des translatés de B_k .

et donc que

$$\begin{aligned}
 f_{B_{n+1}}^\varnothing(\beta, h) &= 2^{-d(n+1)} \log \mathbf{Z}_{B_{n+1}; \beta, h}^\varnothing \\
 &\geq 2^{-d(n+1)} \log (\mathbf{Z}_{B_n; \beta, h}^\varnothing)^{2^d} - \frac{\beta d 2^{(n+1)(d-1)}}{2^{d(n+1)}} \\
 &= f_{B_n}^\varnothing(\beta, h) - \beta d 2^{-(n+1)}.
 \end{aligned}$$

En procédant de la même façon pour obtenir une borne dans l'autre direction, on arrive à

$$|f_{B_{n+1}}^\varnothing(\beta, h) - f_{B_n}^\varnothing(\beta, h)| \leq \beta d 2^{-(n+1)}.$$

L'existence de la limite le long de la suite de cubes B_n suit immédiatement ; on la note $f(\beta, h)$.

Considérons à présent une suite $\Lambda_n \uparrow \mathbb{Z}^d$ arbitraire. On recouvre Λ_n par des translatés disjoints de B_k , k fixé (*cf.* Figure 2.3). On note $\Lambda_n^{\text{ext}, k}$ la partie de \mathbb{Z}^d ainsi obtenue. On vérifie facilement que $\Lambda_n^{\text{ext}, k} \setminus \Lambda_n$ contient au plus $|\partial \Lambda_n| |B_k|$ sites.

En procédant comme ci-dessus par découpage de $\Lambda_n^{\text{ext}, k}$ en ses sous-boîtes (exercice !), on voit que, pour tout $\epsilon > 0$, on peut trouver $k_0(\epsilon, \beta, d)$ tel que, pour tout $k \geq k_0$,

$$|f_{\Lambda_n^{\text{ext}, k}}^\varnothing(\beta, h) - f_{B_k}^\varnothing(\beta, h)| < \epsilon/3.$$

Comme $f_{B_k}^\varnothing(\beta, h) \rightarrow f(\beta, h)$, on a donc également

$$|f_{\Lambda_n^{\text{ext}, k}}^\varnothing(\beta, h) - f(\beta, h)| < \epsilon/2$$

pour tout $k > k_1(\epsilon, \beta, d)$.

D'autre part, on peut comparer $\mathbf{Z}_{\Lambda_n; \beta, h}^\varnothing$ et $\mathbf{Z}_{\Lambda_n^{\text{ext}, k}; \beta, h}^\varnothing$ de façon similaire, en éliminant toutes les interactions liant un spin de $\Lambda_n^{\text{ext}, k} \setminus \Lambda_n$ à ses voisins. On obtient ainsi

$$|f_{\Lambda_n}^\varnothing(\beta, h) - f_{\Lambda_n^{\text{ext}, k}}^\varnothing(\beta, h)| \leq \frac{2d|B_k||\partial \Lambda_n|}{|\Lambda_n|} C(\beta, h).$$

2.6. AIMANTATION ET TRANSITION DE PHASE

Par conséquent, il suit de (2.6) qu'il existe $n_0(\epsilon, k)$ tel que, pour tout $n > n_0$,

$$|f_{\Lambda_n}^\varnothing(\beta, h) - f_{\Lambda_n^{\text{ext},k}}^\varnothing(\beta, h)| \leq \epsilon/2.$$

La conclusion suit.

Indépendance de la condition au bord. Le fait que la même limite est atteinte quelle que soit la condition au bord se vérifie de façon élémentaire, puisque pour tout $\bar{\omega}$,

$$e^{\beta d|\partial\Lambda|} \mathbf{Z}_{\Lambda;\beta,h}^\varnothing \geq \mathbf{Z}_{\Lambda;\beta,h}^{\bar{\omega}} \geq e^{-\beta d|\partial\Lambda|} \mathbf{Z}_{\Lambda;\beta,h}^\varnothing,$$

et que l'affirmation suit alors immédiatement de la propriété (2.6).

Convexité. La convexité peut se vérifier de multiples façons. Par exemple, en utilisant l'inégalité de Hölder,

$$\begin{aligned} \mathbf{Z}_{\Lambda;\alpha\beta_1+(1-\alpha)\beta_2,\alpha h_1+(1-\alpha)h_2}^\varnothing &= \sum_{\omega \in \Omega_\Lambda} e^{-\alpha \mathcal{H}_{\Lambda;\beta_1,h_1}(\omega) - (1-\alpha) \mathcal{H}_{\Lambda;\beta_2,h_2}(\omega)} \\ &\leq \left(\sum_{\omega \in \Omega_\Lambda} e^{-\mathcal{H}_{\Lambda;\beta_1,h_1}(\omega)} \right)^\alpha \left(\sum_{\omega \in \Omega_\Lambda} e^{-\mathcal{H}_{\Lambda;\beta_2,h_2}(\omega)} \right)^{(1-\alpha)}, \end{aligned}$$

et donc

$$f_\Lambda^\omega(\alpha\beta_1 + (1-\alpha)\beta_2, \alpha h_1 + (1-\alpha)h_2) \leq \alpha f_\Lambda^\omega(\beta_1, h_1) + (1-\alpha) f_\Lambda^\omega(\beta_2, h_2).$$

□

Le théorème suivant montre que la présence d'une transition de phase dans le modèle d'Ising se manifeste par une singularité de l'énergie libre : une dérivée première de l'énergie libre est discontinue ; c'est pour cette raison que l'on parle de transition de phase du *premier* ordre.

Théorème 2.6.2. *On a les identités suivantes, pour tout $\beta \geq 0$ et $h \in \mathbb{R}$,*

$$\begin{aligned} \frac{\partial}{\partial h^+} f(\beta, h) &= \langle \sigma_0 \rangle_{\beta,h}^+, \\ \frac{\partial}{\partial h^-} f(\beta, h) &= \langle \sigma_0 \rangle_{\beta,h}^-, \end{aligned}$$

où $\frac{\partial}{\partial h^+}$ et $\frac{\partial}{\partial h^-}$ représentent les dérivées à droite et à gauche respectivement. En particulier, la dérivée

$$\frac{\partial}{\partial h} f(\beta, h)$$

existe si et seulement si il y a unicité de la mesure de Gibbs en (β, h) .

Démonstration. La démonstration qui suit repose sur un certain nombre de résultats élémentaires sur les fonctions convexes, qui sont regroupés dans la section 2.11.

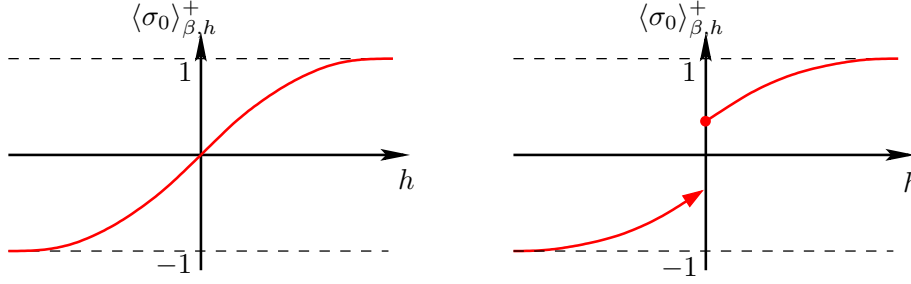


FIGURE 2.4 – Représentation schématique de $\langle \sigma_0 \rangle_{\beta, h}^+$ en fonction de h . Gauche : régime d'unicité. Droite : régime de non-unicité.

Soit Λ_n le cube de côté $2n + 1$. La convexité de $f(\beta, h)$ implique que les dérivées à gauche et à droite par rapport à h existent, sont respectivement continues à gauche et à droite, et diffèrent au plus pour un ensemble dénombrable de valeurs de h . Par conséquent, pour tout h il est possible de trouver une suite $h_k \downarrow h$ telle que f est différentiable pour tous les h_k , et donc

$$\frac{\partial}{\partial h^+} f(\beta, h) = \lim_{h_k \downarrow h} \frac{\partial}{\partial h} f(\beta, h_k).$$

En volume fini,

$$\frac{\partial}{\partial h} f_{\Lambda_n}^+(\beta, h_k) = \frac{1}{|\Lambda_n|} \sum_{i \in \Lambda_n} \langle \sigma_i \rangle_{\Lambda_n, \beta, h_k}^+,$$

et donc

$$\frac{\partial}{\partial h} f(\beta, h_k) = \lim_{n \rightarrow \infty} \frac{1}{|\Lambda_n|} \sum_{i \in \Lambda_n} \langle \sigma_i \rangle_{\Lambda_n, \beta, h_k}^+ = \langle \sigma_0 \rangle_{\beta, h_k}^+,$$

puisque que la dérivée existe. La dernière identité suit de la proposition 2.6.1. On en conclut que

$$\frac{\partial}{\partial h^+} f(\beta, h) = \lim_{h_k \downarrow h} \langle \sigma_0 \rangle_{\beta, h_k}^+ = \langle \sigma_0 \rangle_{\beta, h}^+,$$

la dernière inégalité suivant du point 3 de la proposition 2.6.2.

On montre de la même façon que $\frac{\partial}{\partial h^-} f(\beta, h) = \langle \sigma_0 \rangle_{\beta, h}^-$, ce qui prouve que l'existence de la dérivée est équivalente à $\langle \sigma_0 \rangle_{\beta, h}^+ = \langle \sigma_0 \rangle_{\beta, h}^-$, ce qui caractérise le régime d'unicité. \square

2.7 Transition de phase à basse température

Armé du critère établi dans la section 2.5, nous pouvons nous tourner vers la détermination de conditions sur les paramètres β et h (et la dimension d) assurant l'existence ou l'absence d'une transition de phase du premier ordre.

Commençons par nous intéresser au modèle sans champ magnétique, c'est-à-dire $h = 0$. Avant de procéder, il convient de faire une remarque.

2.7. TRANSITION DE PHASE À BASSE TEMPÉRATURE

Tout d'abord, $\langle \sigma_0 \rangle_{\beta,0}^+ = -\langle \sigma_0 \rangle_{\beta,0}^-$, et donc $\langle \sigma_0 \rangle_{\beta,0}^+ \neq \langle \sigma_0 \rangle_{\beta,0}^-$ si et seulement si $\langle \sigma_0 \rangle_{\beta,h}^+ > 0$. Il suit de la proposition 2.6.2 que $\langle \sigma_0 \rangle_{\beta,0}^+$ est une fonction croissante de β . Par conséquent, la température critique inverse $\beta_c(d) \in [0, \infty]$, définie par

$$\beta_c(d) \stackrel{\text{déf}}{=} \sup\{\beta \geq 0 : \langle \sigma_0 \rangle_{\beta,0}^+ = 0\},$$

sépare (pour $h = 0$) le régime d'unicité du régime de non-unicité : il y aura transition de phase du premier ordre lorsque $\beta > \beta_c$, mais pas lorsque $\beta < \beta_c$. La question est donc de déterminer si $0 < \beta_c < \infty$.

Le résultat principal de cette sous-section est le théorème suivant, qui montre que lorsque $d \geq 2$ et $h = 0$, $\beta_c < \infty$ et donc il y a transition de phase à basses températures.

Théorème 2.7.1. *Pour toute dimension $d \geq 2$, on a $\beta_c(d) < \infty$.*

Remarque 2.7.1. *Ce théorème ne dit rien de ce qui se passe lorsque $\beta = \beta_c$. Cette question est beaucoup plus délicate, et sort du cadre de ce cours. On peut montrer [2] que pour $d \geq 4$, $\langle \sigma_0 \rangle_{\beta_c,0}^+ = 0$, et donc il y a unicité en β_c . Il suit également du calcul explicite (1.2) qu'il en est de même lorsque $d = 2$. Le cas de la dimension 3 est encore ouvert aujourd'hui. Le consensus est qu'il devrait aussi y avoir unicité. Notez que ceci n'est pas générique : par exemple, il est connu que dans le cas du modèle de Potts à q états, une généralisation du modèle d'Ising dans laquelle chaque spin peut prendre q valeurs, il y a transition de phase du premier ordre à la température critique dès que q est assez grand [16].*

Remarque 2.7.2. *Ce théorème montre qu'il n'y a pas unicité lorsque $\beta > \beta_c$ et $h = 0$, mais ne décrit pas l'ensemble des mesures de Gibbs. En dimension 2, on peut en fait montrer que toutes les mesures de Gibbs en volume infini sont de la forme $\alpha \mu_{\beta,0}^+ + (1 - \alpha) \mu_{\beta,0}^-$, $1 \geq \alpha \geq 0$ (Théorème d'Aizenman-Higuchi [1, 10]). Ce n'est plus le cas en dimension 3 et plus [5].*

Démonstration. La méthode utilisée dans la preuve suivante est très importante, et peut être étendue à des situations beaucoup plus générales ; on l'appelle l'argument de Peierls.

Soit Λ_n le cube de côté $2n + 1$ centré à l'origine. Il est suffisant de montrer qu'uniformément en n , $\langle \sigma_0 \rangle_{\Lambda_n;\beta,0}^+ > c > 0$ pour tout β suffisamment grand. Nous ne traiterons que du cas de la dimension 2, puisque le cas général est similaire¹. Bien sûr, $\langle \sigma_0 \rangle_{\Lambda_n;\beta,0}^+ = 1 - 2\mu_{\Lambda_n;\beta,0}^+(\sigma_0 = -1)$, et il est donc suffisant de montrer que $\mu_{\Lambda_n;\beta,0}^+(\sigma_0 = -1) < \frac{1}{2} - c$, pour une constante $c > 0$ indépendante de n .

La preuve repose sur une représentation géométrique des configurations du modèle d'Ising. À chaque sommet $i \in \mathbb{Z}^d$, on associe le carré $\mathcal{S}_i \stackrel{\text{déf}}{=} i + [-\frac{1}{2}, \frac{1}{2}]^2$ centré en i . À une configuration ω telle que $\omega_i \equiv 1$ hors de Λ_n , on associe le sous-ensemble de \mathbb{R}^2 défini par

$$\mathcal{M}(\omega) \stackrel{\text{déf}}{=} \bigcup_{i \in \Lambda_n : \omega_i = -1} \mathcal{S}_i.$$

1. En fait, on peut utiliser l'inégalité (2.11) pour montrer que $\langle \sigma_0 \rangle_{\Lambda_n;\beta,0}^+$ est une fonction croissante de la dimension (exercice!), et le résultat général suit donc de celui en dimension 2.

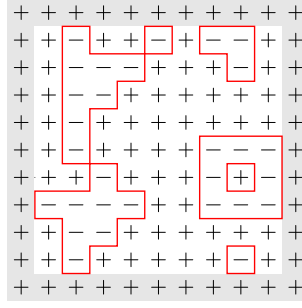


FIGURE 2.5 – Les contours d’une configuration du modèle d’Ising bidimensionnel dans une boîte finie avec condition au bord +.

Les composantes connexes maximales du bord de $\mathcal{M}(\omega)$ sont appelées les **contours** de la configuration ω , et sont notées $\Gamma = (\gamma_1, \dots, \gamma_{m(\omega)})$ (voir la figure 2.5). Il est évident qu’une telle configuration ω est entièrement déterminée par l’ensemble de ses contours.

L’énergie d’une configuration ω s’exprime de manière particulièrement simple en termes de ses contours :

$$\mathcal{H}_{\Lambda_n; \beta, 0}(\omega) = 2\beta \sum_{\gamma \in \Gamma(\omega)} |\gamma| + C,$$

où la constante $C = C(n, \beta)$ est indépendante de la configuration ω , et $|\gamma|$ représente la longueur du contour γ . En effet, il suffit de réécrire l’Hamiltonien sous la forme

$$\mathcal{H}_{\Lambda; \beta, 0}(\omega) = -\beta \sum_{\substack{\{i, j\} \cap \Lambda \neq \emptyset \\ i \sim j}} (\sigma_i(\omega) \sigma_j(\omega) - 1) + C,$$

et d’observer que $\sigma_i(\omega) \sigma_j(\omega) - 1 \neq 0$ si et seulement si $\omega_i \neq \omega_j$, c’est-à-dire si et seulement si les sommets i et j sont séparés par un segment d’un contour.

À présent, l’observation cruciale est que

$$\{\omega : \omega_0 = -1\} \subseteq \{\omega : \exists \gamma^* \in \Gamma(\omega), \gamma^* \text{ entoure } 0\}.$$

Soit γ^* un contour ; à chaque configuration ω telle que $\Gamma(\omega) \ni \gamma^*$, on peut associer la configuration $\mathcal{E}_{\gamma^*}(\omega)$ telle que $\Gamma(\mathcal{E}_{\gamma^*}(\omega)) = \Gamma(\omega) \setminus \{\gamma^*\}$ (en d’autres termes, la configuration $\mathcal{E}_{\gamma^*}(\omega)$ est la configuration qu’on obtient en partant de ω et en enlevant le contour γ^*). Introduisant l’ensemble

$$\mathfrak{C}(\gamma^*) \stackrel{\text{déf}}{=} \{\mathcal{E}_{\gamma^*}(\omega) : \Gamma(\omega) \ni \gamma^*\}$$

de toutes les configurations que l’on peut obtenir en éliminant le contour γ^* d’une autre

configuration, on peut alors écrire

$$\begin{aligned}
 \mu_{\Lambda_n; \beta, 0}^+(\sigma_0 = -1) &\leq \mu_{\Lambda_n; \beta, 0}^+(\exists \gamma^* \in \Gamma, \text{ entourant } 0) \\
 &\leq \frac{\sum_{\gamma^* \text{ entourant } 0} \sum_{\omega: \Gamma(\omega) \ni \gamma^*} \prod_{\gamma \in \Gamma(\omega)} e^{-2\beta|\gamma|}}{\sum_{\omega} \prod_{\gamma \in \Gamma(\omega)} e^{-2\beta|\gamma|}} \\
 &= \sum_{\gamma^* \text{ entourant } 0} e^{-2\beta|\gamma^*|} \frac{\sum_{\omega \in \mathfrak{C}(\gamma^*)} \prod_{\gamma \in \Gamma(\omega)} e^{-2\beta|\gamma|}}{\sum_{\omega} \prod_{\gamma \in \Gamma(\omega)} e^{-2\beta|\gamma|}} \\
 &\leq \sum_{\gamma^* \text{ entourant } 0} e^{-2\beta|\gamma^*|} \\
 &\leq \sum_{k \geq 4} e^{-2\beta k} \#\{\gamma^* \text{ entourant } 0, |\gamma^*| = k\} \\
 &\leq 4 \sum_{k \geq 4} \frac{k}{2} 3^{k-1} e^{-2\beta k},
 \end{aligned}$$

ce qui est strictement inférieur à $1/2$, dès que β est suffisamment grand. La dernière inégalité provient des observations suivantes :

- Le nombre total de contours de longueur k partant d'un point donné est au plus égal à $4 \cdot 3^{k-1}$. En effet, on a 4 directions possibles pour le premier segment, puis au plus 3 pour les $k - 1$ segments suivants (puisque le contour ne peut pas revenir sur ses pas).
- Un contour de longueur k entourant l'origine intersecte nécessairement l'ensemble $\{(u - \frac{1}{2}, 0) : u = 1, \dots, [k/2]\}$. Par conséquent, on peut choisir le point de départ dans cet ensemble et appliquer l'observation précédente.

□

2.8 Unicité à haute température

Dans cette section, nous terminons l'analyse du modèle sans champ magnétique, en montrant le résultat complémentaire que $\beta_c > 0$, et en prouvant qu'il y a toujours unicité en dimension 1.

Il existe plusieurs approches pour montrer ce type de résultat : par exemple, on peut démontrer que l'énergie libre est analytique pour β suffisamment petit, en utilisant un développement perturbatif (le développement en amas, *cluster expansion* en anglais) ; une autre approche très générale repose sur le *critère d'unicité de Dobrushin*. Nous utiliserons une méthode plus probabiliste, aussi très générale : une technique de couplage. Un couplage de deux mesures de probabilité P_1 et P_2 sur un espace probabilisable commun (Ω, \mathcal{F}) est une mesure de probabilité Q sur $(\Omega \times \Omega, \mathcal{F} \times \mathcal{F})$ ayant les propriétés suivantes :

$$Q(A \times \Omega) = P_1(A), \quad Q(\Omega \times A) = P_2(A),$$

pour tout $A \in \mathcal{F}$, c'est-à-dire que les marginales de Q par rapport à chacune des deux variables redonnent les mesures P_1 et P_2 .

Théorème 2.8.1. *Pour tout $d \geq 1$, $\beta_c(d) > 0$. De plus, $\beta_c(1) = \infty$.*

Démonstration. Soit Λ_n le cube de côté $2n + 1$ centré à l'origine. On désire construire un couplage P des deux mesures $\mu_{\Lambda_n; \beta, 0}^+$ et $\mu_{\Lambda_n; \beta, 0}^-$. P doit donc être une mesure de probabilité sur $(\Omega \times \Omega, \mathcal{F} \times \mathcal{F})$ telle que

$$\sum_{\omega' \in \Omega_{\Lambda_n}} P(\omega, \omega') = \mu_{\Lambda_n; \beta, 0}^+(\omega), \quad \sum_{\omega \in \Omega_{\Lambda_n}} P(\omega, \omega') = \mu_{\Lambda_n; \beta, 0}^-(\omega').$$

On désire que ce couplage possède également la propriété suivante :

$$0 \leq \langle \sigma_0 \rangle_{\Lambda_n; \beta, 0}^+ - \langle \sigma_0 \rangle_{\Lambda_n; \beta, 0}^- \leq 2P(0 \xleftrightarrow{\neq} \Lambda_n^c), \quad (2.7)$$

où $\{0 \xleftrightarrow{\neq} \Lambda_n^c\}$ représente l'évènement "il existe un chemin reliant 0 à l'extérieur de Λ_n le long duquel ω et ω' diffèrent en chaque site".

Le couplage est construit à l'aide de l'algorithme suivant. On munit Λ_n d'un ordre total arbitraire, et si $\Delta \subseteq \mathbb{Z}^d$, on note $\partial^{\text{ext}} \Delta \stackrel{\text{déf}}{=} \{i \in \mathbb{Z}^d \setminus \Delta : \exists j \in \Delta, j \sim i\}$.

ÉTAPE 0. On pose $\Delta = \Lambda_n$, $\omega_i = 1$ pour tout $i \notin \Lambda_n$ et $\omega'_i = -1$ pour tout $i \notin \Lambda_n$.

ÉTAPE 1. On suppose que l'on a déjà construit ω et ω' pour tout $i \notin \Delta$.

- Si $\omega \equiv \omega'$ sur $\partial^{\text{ext}} \Delta$, alors on étend ω et ω' à Λ_n tout entier en choisissant $\omega \equiv \omega'$ sur Δ avec la mesure $\mu_{\Delta; \beta, 0}^\omega = \mu_{\Delta; \beta, 0}^{\omega'}$ et le couplage est terminé.
- Si ça n'est pas le cas, soit i le premier site de Δ dont un voisin j au moins est tel que $\omega_j \neq \omega'_j$. On étend la construction de ω et ω' à $(\mathbb{Z}^d \setminus \Delta) \cup \{i\}$ de la façon suivante : Soit $\rho = \mu_{\Delta; \beta, 0}^\omega(\sigma_i = 1)$ et $\rho' = \mu_{\Delta; \beta, 0}^{\omega'}(\sigma'_i = 1)$. Alors on pose

$$(\omega_i, \omega'_i) = \begin{cases} (1, 1) & \text{avec probabilité } \rho \wedge \rho', \\ (-1, -1) & \text{avec probabilité } (1 - \rho) \wedge (1 - \rho'), \\ (1, -1) & \text{avec probabilité } \rho - (\rho \wedge \rho'), \\ (-1, 1) & \text{avec probabilité } \rho' - (\rho \wedge \rho'). \end{cases}$$

Ayant défini le couplage au site i , on remplace Δ par $\Delta \setminus \{i\}$ et on retourne à l'étape 1.

Il est clair que cette construction s'achève en un nombre fini de pas. Vérifions qu'elle fournit bien un couplage de $\mu_{\Lambda_n; \beta, 0}^+$ et $\mu_{\Lambda_n; \beta, 0}^-$.

On va vérifier qu'à chaque étape de la construction on a l'identité (l'identité analogue pour l'autre marginale suit par symétrie) :

$$\sum_{\bar{\omega}' \in \Omega_{\Lambda_n \setminus \Delta}} P(\sigma_k = \bar{\omega}_k, \sigma'_k = \bar{\omega}'_k, \forall k \notin \Delta) = \mu_{\Lambda_n}^+(\sigma_k = \bar{\omega}_k, \forall k \notin \Delta), \quad (2.8)$$

2.8. UNICITÉ À HAUTE TEMPÉRATURE

pour tout $\bar{\omega} \in \Omega_{\Lambda_n} \setminus \Delta$. Supposons qu'on l'ait vérifiée pour Δ et montrons que cette propriété est encore vérifiée à l'étape suivante. Il y a deux cas à considérer, selon que $\bar{\omega}$ et $\bar{\omega}'$ coïncident ou non sur $\partial^{\text{ext}} \Delta$. Dans le premier cas, on a

$$\begin{aligned} P(\sigma_i = \bar{\omega}_i, \forall i \in \Delta \mid \sigma_k = \bar{\omega}_k, \sigma'_k = \bar{\omega}'_k, \forall k \notin \Delta) \\ = \mu_{\Lambda_n}^+(\sigma_i = \bar{\omega}_i, \forall i \in \Delta \mid \sigma_k = \bar{\omega}_k, \forall k \notin \Delta), \end{aligned}$$

pour tout choix des $\bar{\omega}_i$, $i \in \Delta$, et le résultat suit de (2.8), puisque

$$\begin{aligned} P(\sigma_i = \bar{\omega}_i, \forall i \in \Lambda_n) &= P(\sigma_i = \bar{\omega}_i, \forall i \in \Delta \mid \sigma_k = \bar{\omega}_k, \forall k \notin \Delta) \\ &\quad \times P(\sigma_k = \bar{\omega}_k, \forall k \notin \Delta) \\ &= \mu_{\Lambda_n}^+(\sigma_i = \bar{\omega}_i, \forall i \in \Delta \mid \sigma_k = \bar{\omega}_k, \forall k \notin \Delta) \\ &\quad \times \mu_{\Lambda_n}^+(\sigma_k = \bar{\omega}_k, \forall k \notin \Delta) \\ &= \mu_{\Lambda_n}^+(\sigma_i = \bar{\omega}_i, \forall i \in \Lambda_n). \end{aligned}$$

Dans le second cas, soit j le site mis à jour lors de cette étape de la procédure. Alors

$$\begin{aligned} \sum_{\bar{\omega}'_j = \pm 1} P(\sigma_j = \bar{\omega}_j, \sigma'_j = \bar{\omega}'_j \mid \sigma_k = \bar{\omega}_k, \sigma'_k = \bar{\omega}'_k, \forall k \notin \Delta) \\ = \mu_{\Lambda_n}^+(\sigma_j = \bar{\omega}_j \mid \sigma_k = \bar{\omega}_k, \forall k \notin \Delta), \end{aligned}$$

car $(\rho \wedge \rho') + (\rho - (\rho \wedge \rho')) = \rho$ et $((1 - \rho) \wedge (1 - \rho')) + (\rho' - (\rho \wedge \rho')) = 1 - \rho$. On a donc bien

$$\begin{aligned} \sum_{\bar{\omega}' \in \Omega_{\Delta \cup \{j\}}} P(\sigma_k = \bar{\omega}_k, \sigma'_k = \bar{\omega}'_k, \forall k \notin \Delta \setminus \{j\}) \\ = \mu_{\Lambda_n}^+(\sigma_j = \bar{\omega}_j \mid \sigma_k = \bar{\omega}_k, \forall k \notin \Delta) \mu_{\Lambda_n}^+(\sigma_k = \bar{\omega}_k, \forall k \notin \Delta) \\ = \mu_{\Lambda_n}^+(\sigma_k = \bar{\omega}_k, \forall k \notin \Delta \setminus \{j\}). \end{aligned}$$

Vérifions à présent que le couplage ainsi construit possède également la propriété (2.8) : puisque

$$\begin{aligned} \langle \sigma_0 \rangle_{\Lambda_n; \beta, 0}^+ - \langle \sigma_0 \rangle_{\Lambda_n; \beta, 0}^- &= 2\mu_{\Lambda_n; \beta, 0}^+(\sigma_0 = 1) - \mu_{\Lambda_n; \beta, 0}^-(\sigma_0 = 1) \\ &= 2\langle \mathbf{1}_{\{\sigma_0=1\}} - \mathbf{1}_{\{\sigma'_0=1\}} \rangle_P \\ &= 2\langle \mathbf{1}_{\{\sigma_0=1, \sigma'_0=-1\}} - \mathbf{1}_{\{\sigma_0=-1, \sigma'_0=1\}} \rangle_P \\ &\leq 2P(\sigma_0 = 1, \sigma'_0 = -1) \\ &\leq 2P(\sigma_0 \neq \sigma'_0), \end{aligned}$$

et que la seule façon d'avoir $\sigma_0 \neq \sigma'_0$ sous P est qu'il y ait un chemin connectant 0 à $\partial \Lambda_n$ le long duquel $\omega \neq \omega'$. Pour voir cela, supposons que $\omega_0 \neq \omega'_0$ et considérons l'étape de la construction ci-dessus lors de laquelle ω_0 a été mis à jour. Manifestement celui-ci

n'a pas pu être mis à jour simultanément avec d'autres sommets à l'aide de la procédure appliquée lorsque $\omega \equiv \omega'$ sur $\partial^{\text{ext}}\Delta$, car sinon $\omega_0 = \omega'_0$ par construction. Donc i a été mis à jour individuellement. Mais 0 n'a pu être choisi lors de cette étape que si l'un au moins de ses voisins hors de Δ , appelons-le j , satisfaisait $\omega_j \neq \omega'_j$. En répétant cet argument, on construit de proche en proche un chemin reliant 0 et $\partial\Lambda_n$ le long duquel ω et ω' ne prennent jamais la même valeur.

L'observation suivante est qu'à chaque étape de l'algorithme précédent, la probabilité que $\omega_i \neq \omega'_i$ est au plus de $|\rho - \rho'| \leq \mu_{\{i\};\beta,0}^+(\sigma_i = 1) - \mu_{\{i\};\beta,0}^-(\sigma_i = 1) = \tanh(2d\beta)$. On en conclut que

$$\begin{aligned} P(0 \not\longleftrightarrow \Lambda_n^c) &\leq \sum_{w: 0 \longleftrightarrow \Lambda_n^c} P(\sigma_i \neq \sigma'_i, \forall i \in w) \\ &\leq \sum_{k \geq n} (\tanh(2d\beta))^k \#\{w : 0 \longleftrightarrow \Lambda_n^c, |w| = k\} \\ &\leq (2d/(2d-1)) \sum_{k \geq n} (\tanh(2d\beta))^k (2d-1)^k, \end{aligned}$$

et cette dernière expression tend vers 0, lorsque $n \rightarrow \infty$, pour tout β suffisamment petit. (Pour obtenir la borne $2d(2d-1)^{k-1}$ sur le nombre de chemins auto-évitant de longueur k partant de 0, il suffit d'observer qu'à chaque pas, sauf le premier, il y a au plus $2d-1$ voisins encore inoccupés.) Ceci prouve le résultat pour $d \geq 2$. Pour $d = 1$, on a bien entendu un résultat plus fort :

$$P(0 \not\longleftrightarrow \Lambda_n^c) \leq 2P(\sigma_i \neq \sigma'_i, \forall i \in \Lambda_n, i \leq 0) \leq 2(\tanh(2\beta))^{n+1},$$

qui tend vers 0 pour tout $\beta \geq 0$. □

2.9 Unicité en champ magnétique non nul

Il nous reste donc à étudier l'effet du champ magnétique h . Dans cette section, nous allons montrer qu'il y a toujours unicité de la mesure de Gibbs du modèle d'Ising lorsque le champ magnétique $h \neq 0$. Il existe deux façons de démontrer ce résultat : la première passe par une autre inégalité de corrélation (inégalité GHS) qui implique que $\langle \sigma_0 \rangle_{\beta,h}^+$ est une fonction concave (et donc continue) de $h \geq 0$. Ici, nous utiliserons un autre type d'argument, montrant que l'énergie libre est analytique lorsque $h \neq 0$. Bien entendu, la fonction de partition en volume fini est analytique en h , et ne s'annule pas (c'est une somme de termes strictement positifs). Ce qu'il convient donc d'exclure, c'est qu'un zéro complexe approche l'axe réel dans la limite thermodynamique.

Théorème 2.9.1. *L'énergie libre $f(\beta, h)$, vue comme fonction de la variable complexe h , est analytique dans le domaine $|\Re h| > |\Im h|$. En particulier, il y a une unique mesure de Gibbs dès que $h \in \mathbb{R} \setminus \{0\}$.*

Remarque 2.9.1. *Il est en fait possible démontrer que $f(\beta, h)$ est analytique dans tout le domaine $\Re h \neq 0$; c'est le célèbre Théorème du cercle de Lee-Yang [18].*

Remarque 2.9.2. Il suit du théorème 2.6.2 que l'aimantation est également une fonction analytique de h lorsque $\Re h \neq 0$.

Démonstration. Puisque l'énergie libre ne dépend pas de la condition au bord, il suffit de considérer le modèle avec condition au bord libre. Par symétrie, on peut aussi se restreindre au cas $\Re h > 0$, c'est-à-dire se restreindre au domaine $\mathcal{D} \stackrel{\text{déf}}{=} \{\Re h > |\Im h|\}$. On verra plus bas que, pour tout $\Lambda \in \mathbb{Z}^d$,

$$|\mathbf{Z}_{\Lambda;\beta,h}^\varnothing| \geq \mathbf{Z}_{\Lambda;\beta,\Re h-|\Im h|}^\varnothing \geq \mathbf{Z}_{\Lambda;\beta,0}^\varnothing > 0, \quad (2.9)$$

dans le domaine \mathcal{D} .

Soit $\Lambda_n \uparrow \mathbb{Z}^d$ une suite croissante de cubes. On introduit la suite de fonctions $g_n(h) \stackrel{\text{déf}}{=} (\mathbf{Z}_{\Lambda_n;\beta,h}^\varnothing)^{1/|\Lambda_n|}$. On a les propriétés suivantes :

- $g_n(h)$ est analytique dans le domaine \mathcal{D} , puisque $|\mathbf{Z}_{\Lambda_n;\beta,h}^\varnothing| > 0$ (on choisit la détermination qui est réelle pour $h \in \mathbb{R}_+$.)
- $|g_n(h)| \leq g_n(\Re h)$, et donc g_n est localement uniformément bornée dans \mathcal{D} .
- La suite de fonctions g_n converge vers une fonction g sur l'axe réel (Théorème 2.6.1).

On rappelle deux théorèmes classiques d'analyse complexe :

Théorème de convergence de Vitali. Soit D un domaine ouvert de \mathbb{C} , et (f_n) une suite de fonctions analytiques, localement uniformément bornées, dans D , convergeant sur un ensemble ayant un point d'accumulation dans D . Alors f_n converge localement uniformément dans D , la limite étant par conséquent une fonction analytique.

Théorème de Hurwitz. Soit D un ouvert de \mathbb{C} et (f_n) une suite de fonction analytiques convergeant localement uniformément dans D vers une fonction analytique f . Si $f_n(z) \neq 0$, pour tout $z \in D$ et pour tout n , alors f est soit identiquement nulle soit jamais nulle dans D .

Il suit du Théorème de convergence de Vitali que la suite de fonctions g_n converge localement uniformément vers une fonction g analytique dans le domaine \mathcal{D} . De plus, le théorème de Hurwitz montre que g ne s'annule pas dans \mathcal{D} ; par conséquent, $\log g$ est aussi analytique dans \mathcal{D} , ce qui démontre le théorème.

Il reste à montrer (2.9). On utilise un argument de duplication du système : on écrit

$$|\mathbf{Z}_{\Lambda;\beta,h}^\varnothing|^2 = \sum_{\omega, \omega'} e^{\beta \sum_{\{i,j\} \subset \Lambda, i \sim j} (\omega_i \omega_j + \omega'_i \omega'_j) + \sum_{i \in \Lambda} (h \omega_i + \bar{h} \omega'_i)}.$$

On fait le changement de variables suivant : $\cos \theta_i \stackrel{\text{déf}}{=} \frac{1}{2}(\omega_i + \omega'_i)$ et $\sin \theta_i \stackrel{\text{déf}}{=} \frac{1}{2}(\omega_i - \omega'_i)$; manifestement $\theta_i \in \{0, \pi/2, \pi, 3\pi/2\}$. On vérifie aisément que

$$\begin{aligned} \omega_i \omega_j + \omega'_i \omega'_j &= 2 \cos(\theta_i - \theta_j) = e^{i(\theta_i - \theta_j)} + e^{-i(\theta_i - \theta_j)}, \\ h \omega_i + \bar{h} \omega'_i &= 2 \Re h \cos(\theta_i) + 2i \Im h \sin(\theta_i) = (\Re h + \Im h) e^{i\theta_i} + (\Re h - \Im h) e^{-i\theta_i}. \end{aligned}$$

Par conséquent, en développant l'exponentielle, on voit que

$$|\mathbf{Z}_{\Lambda;\beta,h}^\varnothing|^2 = \sum_{(\theta_i)_{i \in \Lambda}} \sum_{\substack{\underline{m} = (m_i)_{i \in \Lambda} \\ m_i \in \{0,1,2,3\}}} \alpha_{\underline{m}} e^{i \sum_{i \in \Lambda} m_i \theta_i},$$

avec des coefficients $\alpha_{\underline{m}}$ positifs, et croissants en $\Re h + \Im h$ et en $\Re h - \Im h$. À présent, on observe que

$$\sum_{(\theta_i)_{i \in \Lambda}} e^{i \sum_{i \in \Lambda} m_i \theta_i} = \prod_{i \in \Lambda} \sum_{\theta_i} e^{i m_i \theta_i} = \begin{cases} 4^{|\Lambda|} & \text{si } m_i = 0, \forall i \in \Lambda \\ 0 & \text{sinon.} \end{cases}$$

On en déduit que $|\mathbf{Z}_{\Lambda; \beta, h}^{\sigma}|^2 = 4^{|\Lambda|} \alpha_{(0,0,\dots,0)}$, et donc que $|\mathbf{Z}_{\Lambda; \beta, h}^{\sigma}|^2$ est croissant en $\Re h + \Im h$ et en $\Re h - \Im h$. Ceci prouve (2.9), puisque $\Re h - |\Im h| = \min(\Re h + \Im h, \Re h - \Im h)$ (et évidemment, pour $h \in \mathbb{R}$, la fonction de partition est strictement positive puisqu'il s'agit d'une somme d'exponentielles). \square

2.10 Preuve des inégalités de corrélation

Dans cette section, nous démontrons les inégalités de corrélation utilisées dans les sections précédentes.

2.10.1 Inégalités GKS

Il est plus naturel de les énoncer dans un cadre un peu plus général que celui introduit précédemment.

Soit Λ un sous-ensemble fini de \mathbb{Z}^d , et $\mathbf{J} = (J_A)_{A \subseteq \Lambda}$, une famille de nombres réels. On considère la mesure de probabilité suivante sur $(\Omega_{\Lambda}, \mathcal{F}_{\Lambda})$:

$$\nu_{\Lambda; \mathbf{J}}(\omega) \stackrel{\text{déf}}{=} \frac{1}{Z_{\Lambda; \mathbf{J}}} \exp\left\{ \sum_{C \subseteq \Lambda} J_C \omega_C \right\},$$

où l'on a utilisé la notation $\omega_C = \prod_{i \in C} \omega_i$.

Remarque 2.10.1. Les mesures $\mu_{\Lambda; \beta, h}^+$, $\mu_{\Lambda; \beta, h}^{\sigma}$ et $\mu_{\Lambda; \beta, h}^{\text{per}}$ peuvent être mises sous la forme ci-dessus, avec $J_C \geq 0$, $\forall C \subseteq \Lambda$, si $h \geq 0$. Par exemple, pour la condition au bord +, il suffit de prendre

$$J_C = \begin{cases} h + \beta \cdot \#\{j \notin \Lambda : j \sim i\} & \text{si } C = \{i\} \subseteq \Lambda, \\ \beta & \text{si } C = \{i, j\} \subseteq \Lambda, i \sim j \\ 0 & \text{sinon.} \end{cases}$$

Théorème 2.10.1. Soient $\mathbf{J} = (J_C)_{C \subseteq \Lambda}$ et $\mathbf{J}' = (J'_C)_{C \subseteq \Lambda}$ tels que $J_C \geq |J'_C|$ (en particulier $J_C \geq 0$), pour tout $C \subseteq \Lambda$. Alors, pour tout $A, B \subseteq \Lambda$,

$$\langle \sigma_A \rangle_{\Lambda; \mathbf{J}} \geq 0, \tag{2.10}$$

$$\langle \sigma_A \sigma_B \rangle_{\Lambda; \mathbf{J}} \geq \langle \sigma_A \rangle_{\Lambda; \mathbf{J}} \langle \sigma_B \rangle_{\Lambda; \mathbf{J}}, \tag{2.11}$$

$$\langle \sigma_A \rangle_{\Lambda; \mathbf{J}} \geq \langle \sigma_A \rangle_{\Lambda; \mathbf{J}'}, \tag{2.12}$$

2.10. PREUVE DES INÉGALITÉS DE CORRÉLATION

Démonstration. Manifestement la fonction de partition $\mathbf{Z}_{\Lambda;\mathbf{J}}$ est strictement positive. Il suffit donc de démontrer les résultats correspondants pour les numérateurs.

$$\begin{aligned} \mathbf{Z}_{\Lambda;\mathbf{J}} \langle \sigma_A \rangle_{\Lambda;\mathbf{J}} &= \sum_{\omega} \omega_A \prod_{C \subseteq \Lambda} e^{J_C \omega_C} \\ &= \sum_{\omega} \omega_A \prod_{C \subseteq \Lambda} \sum_{n \geq 0} \frac{J_C^n}{n!} \omega_C^n \\ &= \sum_{\omega} \omega_A \sum_{\substack{(n_C)_{C \subseteq \Lambda} \\ n_C \geq 0}} \prod_{C \subseteq \Lambda} \frac{J_C^{n_C}}{n_C!} \omega_C^{n_C} \\ &= \sum_{\substack{(n_C)_{C \subseteq \Lambda} \\ n_C \geq 0}} \prod_{C \subseteq \Lambda} \frac{J_C^{n_C}}{n_C!} \sum_{\omega} \omega_A \prod_{C \subseteq \Lambda} \omega_C^{n_C}. \end{aligned}$$

La sommation sur ω dans la dernière expression est de la forme

$$\sum_{\omega} \prod_{i \in \Lambda} \omega_i^{m_i},$$

où les m_i dépendent des n_C et de A . On voit alors aisément que

$$\sum_{\omega} \prod_{i \in \Lambda} \omega_i^{m_i} = \prod_{i \in \Lambda} \sum_{\omega_i = \pm 1} \omega_i^{m_i} \geq 0,$$

puisque $\sum_{\omega_i = \pm 1} \omega_i^{m_i}$ vaut 2 lorsque m_i est pair, et 0 lorsque m_i est impair. Ceci prouve l'affirmation (2.10).

Afin de démontrer (2.11), il est utile de dupliquer le système, c'est-à-dire d'écrire la covariance de la façon suivante :

$$\langle \sigma_A \sigma_B \rangle_{\Lambda;\mathbf{J}} - \langle \sigma_A \rangle_{\Lambda;\mathbf{J}} \langle \sigma_B \rangle_{\Lambda;\mathbf{J}} = \langle \sigma_A (\sigma_B - \sigma'_B) \rangle_{\nu_{\Lambda;\mathbf{J}} \otimes \nu_{\Lambda;\mathbf{J}}},$$

la première mesure agissant sur ω , et la seconde sur ω' , et où l'on a introduit $\sigma_i(\omega, \omega') = \omega_i$ et $\sigma'_i(\omega, \omega') = \omega'_i$. On peut donc écrire

$$(\mathbf{Z}_{\Lambda;\mathbf{J}})^2 \langle \sigma_A (\sigma_B - \sigma'_B) \rangle_{\nu_{\Lambda;\mathbf{J}} \otimes \nu_{\Lambda;\mathbf{J}}} = \sum_{\omega, \omega'} \omega_A (\omega_B - \omega'_B) \prod_{C \subseteq \Lambda} e^{J_C (\omega_C + \omega'_C)}.$$

Introduisant les variables $\omega''_i \stackrel{\text{déf}}{=} \omega_i \omega'_i = \omega'_i / \omega_i$, on obtient

$$\begin{aligned} \sum_{\omega, \omega'} \omega_A (\omega_B - \omega'_B) \prod_{C \subseteq \Lambda} e^{J_C (\omega_C + \omega'_C)} &= \sum_{\omega, \omega''} \omega_A \omega_B (1 - \omega''_B) \prod_{C \subseteq \Lambda} e^{J_C (1 + \omega''_C) \omega_C} \\ &= \sum_{\omega''} (1 - \omega''_B) \sum_{\omega} \omega_A \omega_B \prod_{C \subseteq \Lambda} e^{J_C (1 + \omega''_C) \omega_C}. \end{aligned}$$

Puisque $1 - \omega''_B \geq 0$, il suffit d'appliquer (2.10) à la somme sur ω (avec constantes de couplages $J_C (1 + \omega''_C) \geq 0$) pour obtenir (2.11).

La preuve de (2.12) est essentiellement identique. On écrit

$$\begin{aligned} (\mathbf{Z}_{\Lambda; \mathbf{J}} \mathbf{Z}_{\Lambda; \mathbf{J}'}) \langle \sigma_A - \sigma'_A \rangle_{\nu_{\Lambda; \mathbf{J}} \otimes \nu_{\Lambda; \mathbf{J}'}} &= \sum_{\omega, \omega'} (\omega_A - \omega'_A) \prod_{C \subseteq \Lambda} e^{J_C \omega_C + J'_C \omega'_C} \\ &= \sum_{\omega''} (1 - \omega''_A) \sum_{\omega} \omega_A \prod_{C \subseteq \Lambda} e^{(J_C + J'_C \omega''_C) \omega_C}, \end{aligned}$$

et on conclut comme précédemment puisque $J_C + J'_C \omega''_C \geq 0$. \square

2.10.2 Inégalités FKG

À nouveau, il est utile de considérer une situation substantiellement plus générale.

Soit $x = (x_1, \dots, x_n), y = (y_1, \dots, y_n) \in \Xi^n$, $\Xi \subseteq \mathbb{R}$, $n \geq 1$. On définit² $x \vee y = (x_1 \vee y_1, \dots, x_n \vee y_n)$ et $x \wedge y = (x_1 \wedge y_1, \dots, x_n \wedge y_n)$. On écrit $x \leq y$ lorsque $x \vee y = y$, et on dit qu'une fonction $f : \mathbb{R}^n \rightarrow \mathbb{R}$ est croissante si $x \leq y \implies f(x) \leq f(y)$.

On a alors le résultat suivant :

Théorème 2.10.2. *Soient f_1, \dots, f_4 des fonctions boréliennes, positives, sur Ξ^n , telles que*

$$f_1(x)f_2(y) \leq f_3(x \wedge y)f_4(x \vee y), \quad \forall x, y \in \Xi^n.$$

Soit $\mu = \mu_1 \otimes \dots \otimes \mu_n$ une mesure-produit σ -finie sur $\mathcal{B}(\Xi^n)$. Alors,

$$\langle f_1 \rangle_{\mu} \langle f_2 \rangle_{\mu} \leq \langle f_3 \rangle_{\mu} \langle f_4 \rangle_{\mu}.$$

Pour le modèle d'Ising, $\Xi = \{-1, 1\}$. En choisissant $\mu_i(\sigma_i = s) = e^{hs + \beta \sum_{j \in \Lambda, j \sim i} \bar{\omega}_j s}$, $s = \pm 1$, pour tout $i \in \Lambda$, et notant p la densité de $\mu_{\Lambda; \beta, h}^{\bar{\omega}}$ par rapport à $\mu = \bigotimes_{i \in \Lambda} \mu_i$, on voit que le théorème précédent appliqué aux fonctions

$$f_1(\omega) = p(\omega)f(\omega), f_2(\omega) = p(\omega)g(\omega), f_3(\omega) = p(\omega), f_4(\omega) = p(\omega)f(\omega)g(\omega),$$

(f et g peuvent être supposées positives sans perte de généralité, sinon il suffit de leur ajouter une constante, ce qui ne change pas leur covariance) implique (2.3) pourvu que

$$p(\omega)p(\omega') \leq p(\omega \vee \omega')p(\omega \wedge \omega').$$

Or, puisque

$$p(\omega) \propto \exp\left(\beta \sum_{\substack{i, j \in \Lambda \\ i \sim j}} \omega_i \omega_j\right),$$

il suffit de vérifier que

$$\omega_i \omega_j + \omega'_i \omega'_j \leq (\omega_i \vee \omega'_i)(\omega_j \vee \omega'_j) + (\omega_i \wedge \omega'_i)(\omega_j \wedge \omega'_j),$$

ce qui se voit immédiatement sur les 16 cas possibles (en fait, moins en utilisant les symétries présentes).

Il reste donc à démontrer le Théorème 2.10.2.

2. On utilise les notations standards $a \vee b \stackrel{\text{déf}}{=} \max(a, b)$ et $a \wedge b \stackrel{\text{déf}}{=} \min(a, b)$.

Démonstration du Théorème 2.10.2. L'idée est la suivante. Écrivons $x = (X, u)$ et $y = (Y, v)$, où $X = (x_1, \dots, x_{n-1})$ et $Y = (y_1, \dots, y_{n-1})$. On va montrer que l'inégalité

$$f_1(x)f_2(y) \leq f_3(x \wedge y)f_4(x \vee y). \quad (2.13)$$

implique l'inégalité

$$\tilde{f}_1(X)\tilde{f}_2(Y) \leq \tilde{f}_3(X \wedge Y)\tilde{f}_4(X \vee Y), \quad (2.14)$$

où $\tilde{f}_i = \langle f_i(X, \cdot) \rangle_{\mu_n}$, la moyenne étant prise sur la dernière coordonnée. Par conséquent, en appliquant cet argument n fois, on obtient l'inégalité recherchée.

Le membre de gauche de (2.14) peut s'écrire

$$\begin{aligned} \langle f_1(X, u)f_2(Y, v) \rangle_{\mu_n \otimes \mu_n} &= \langle \mathbf{1}_{\{u=v\}} f_1(X, u)f_2(Y, v) \rangle_{\mu_n \otimes \mu_n} \\ &\quad + \langle \mathbf{1}_{\{u < v\}} (f_1(X, u)f_2(Y, v) + f_1(X, v)f_2(Y, u)) \rangle_{\mu_n \otimes \mu_n}. \end{aligned}$$

De même, le membre de droite de (2.14) s'écrit

$$\begin{aligned} \langle f_3(X \wedge Y, u)f_4(X \vee Y, v) \rangle_{\mu_n \otimes \mu_n} &= \\ &\quad \langle \mathbf{1}_{\{u=v\}} f_3(X \wedge Y, u)f_4(X \vee Y, v) \rangle_{\mu_n \otimes \mu_n} \\ &\quad + \langle \mathbf{1}_{\{u < v\}} (f_3(X \wedge Y, u)f_4(X \vee Y, v) + f_3(X \wedge Y, v)f_4(X \vee Y, u)) \rangle_{\mu_n \otimes \mu_n}. \end{aligned}$$

On voit donc que lorsque $u = v$, l'inégalité (2.13) donne immédiatement le résultat désiré. Lorsque $u < v$, on pose $A \stackrel{\text{déf}}{=} f_1(X, u)f_2(Y, v)$, $B \stackrel{\text{déf}}{=} f_1(X, v)f_2(Y, u)$, $C \stackrel{\text{déf}}{=} f_3(X \wedge Y, u)f_4(X \vee Y, v)$ et $D \stackrel{\text{déf}}{=} f_3(X \wedge Y, v)f_4(X \vee Y, u)$. Il suffit alors d'observer que (2.13) implique que $A \leq C$, $B \leq D$ et

$$\begin{aligned} AB &= f_1(X, u)f_2(Y, u)f_1(X, v)f_2(Y, v) \\ &\leq f_3(X \wedge Y, u)f_4(X \vee Y, u)f_3(X \wedge Y, v)f_4(X \vee Y, v) = CD. \end{aligned}$$

(2.14) sera donc établie si l'on peut montrer que $A + B \leq C + D$. Mais ceci est évident puisque

$$(C + D - A - B)/C \geq 1 - (A + B)/C + AB/C^2 = (1 - A/C)(1 - B/C) \geq 0.$$

(Observer que l'inégalité désirée est trivialement satisfaite si $C = 0$.) □

2.11 Résultats élémentaires sur les fonctions convexes

Dans cette section sont regroupés quelques résultats élémentaires sur les fonctions convexes d'une variable réelle.

Théorème 2.11.1. *Soit $I \subseteq \mathbb{R}$ un ouvert, et $f : I \rightarrow \mathbb{R}$ une fonction convexe. Alors*

1. *Les dérivées à droite et à gauche, $\partial^+ f(x)$ et $\partial^- f(x)$ existent en tout point $x \in I$.*
2. *$\partial^+ f(x) \geq \partial^- f(x)$, pour tout $x \in I$.*

3. $\partial^+ f$ et $\partial^- f$ sont croissantes.
4. $\partial^+ f$ est continue à droite, $\partial^- f$ est continue à gauche.
5. L'ensemble $\{x : \partial^+ f(x) \neq \partial^- f(x)\}$ est au plus dénombrable.
6. Si $(g_n)_{n \geq 1}$ est une suite de fonctions convexes de $I \rightarrow \mathbb{R}$ convergeant ponctuellement vers une fonction g , alors g est également convexe. De plus, si g est dérivable en x ,
 $\lim_{n \rightarrow \infty} \partial^+ g_n(x) = \lim_{n \rightarrow \infty} \partial^- g_n(x) = g'(x)$.

Démonstration. Laissée en exercice. □

Chapitre 3

Champs aléatoires markoviens et mesures de Gibbs

Dans ce chapitre, nous revenons à la problématique générale. Nous allons commencer par définir précisément la notion de champ aléatoire markovien sur un graphe, fini, et celle de mesure de Gibbs avec potentiel plus-proches-voisins. Nous montrerons ensuite l'équivalence de ces deux notions de champs aléatoires.

3.1 Définitions

Soit $G = (V, E)$ un graphe fini simple (c'est-à-dire sans boucles, ni arêtes multiples), et S un ensemble dénombrable non-vidé. On note $i \sim j$ la relation “ i et j sont connectés par une arête de E ”. Une clique de G est un sous-ensemble $K \subseteq V$ tel que $i \sim j, \forall i, j \in K$.

On note $\Omega \stackrel{\text{d\'ef}}{=} S^V$ l'espace des configurations, et $\mathcal{F} = \mathcal{P}(\Omega)$ l'ensemble des parties de Ω . On note $\sigma_i, i \in V$, les variables aléatoires définies par $\sigma_i(\omega) = \omega_i$.

L'ensemble des mesures de probabilité sur (Ω, \mathcal{F}) est noté $\mathcal{M}_1(\Omega, \mathcal{F})$; on appelle de telles mesures de probabilité des **champs aléatoires**. Afin d'alléger l'écriture, on utilise la même notation pour le champ aléatoire et sa densité, c'est-à-dire qu'on écrit $\mathbb{P}(\omega)$ pour $\mathbb{P}(\{\omega\})$.

On dira qu'un champ aléatoire est **strictement positif** si $\mathbb{P}(\omega) > 0$ pour toute configuration $\omega \in \Omega$.

3.1.1 Champs aléatoires markoviens

Un champ aléatoire $\mathbb{P} \in \mathcal{M}_1(\Omega, \mathcal{F})$ est **markovien** s'il est strictement positif et possède la propriété de Markov suivante : pour toute configuration $\omega \in \Omega$ et tout $i \in V$,

$$\mathbb{P}(\sigma_i = \omega_i \mid \sigma_k = \omega_k, k \neq i) = \mathbb{P}(\sigma_i = \omega_i \mid \sigma_k = \omega_k, k \sim i).$$

Remarque 3.1.1. *Tout champ aléatoire sur G peut être vu comme un champ aléatoire markovien sur un graphe auquel on aura ajouté suffisamment d'arêtes; en particulier,*

3.1. DÉFINITIONS

tout champ aléatoire sur G est un champ aléatoire markovien sur le graphe complet avec sommets V (c'est-à-dire le graphe $G = (V, E)$ tel que $i \sim j, \forall i, j \in V$ tels que $i \neq j$).

3.1.2 Mesures de Gibbs

On appelle **potentiel** une collection $(\Phi_A)_{A \subseteq V}$ de fonctions $\Phi_A : \Omega \rightarrow \mathbb{R}$ telles que chaque Φ_A ne dépende que des valeurs $(\omega_i)_{i \in A}$ prise par ω sur A .

Un potentiel $(\Phi_A)_{A \subseteq V}$ est dit **plus-proches-voisins** (p.p.v.) si $\Phi_A \equiv 0$ lorsque A n'est pas une clique de G .

Étant donné un potentiel $(\Phi_A)_{A \subseteq V}$, on définit l'énergie de la configuration ω par

$$\mathcal{H}(\omega) \stackrel{\text{déf}}{=} \sum_{A \subseteq V} \Phi_A(\omega).$$

\mathcal{H} est appelé l'Hamiltonien.

La mesure de Gibbs associée au potentiel $(\Phi_A)_{A \subseteq V}$ est le champ aléatoire $\mathbb{P} \in \mathcal{M}_1(\Omega, \mathcal{F})$ défini par

$$\mathbb{P}(\omega) = \frac{1}{\mathbf{Z}} e^{-\mathcal{H}(\omega)},$$

où la fonction de partition $\mathbf{Z} = \sum_{\omega \in \Omega} e^{-\mathcal{H}(\omega)}$.

3.1.3 Potentiels canoniques

Manifestement, il existe une infinité de potentiels donnant lieu à la même mesure de Gibbs. On parle de potentiels **physiquement équivalents**. Il est utile de pouvoir associer à une mesure de Gibbs un unique potentiel de façon canonique.

Soit $\mathbb{P} \in \mathcal{M}_1(\Omega, \mathcal{F})$ un champ aléatoire strictement positif, et soit $\mathbf{0}$ un des éléments de S . Le potentiel **canonique** $(\tilde{\Phi}_A)_{A \subseteq V}$ associé à \mathbb{P} est défini par

$$\tilde{\Phi}_A(\omega) \stackrel{\text{déf}}{=} \sum_{B \subseteq A} (-1)^{|A \setminus B|} \Psi(\omega^B),$$

où $\Psi(\omega) \stackrel{\text{déf}}{=} -(\log \mathbb{P}(\omega) - \log \mathbb{P}(\mathbf{0}))$, avec $\mathbf{0} \in \Omega$, $\sigma_i(\mathbf{0}) = 0, \forall i \in V$, et $\omega^B \in \Omega$ est définie par

$$\omega_i^B \stackrel{\text{déf}}{=} \begin{cases} \omega_i & \text{si } i \in B, \\ 0 & \text{sinon.} \end{cases}$$

On vérifie immédiatement que ceci définit bien un potentiel. Le lemme suivant implique que si \mathbb{P} est une mesure de Gibbs associée à un potentiel $(\Phi_A)_{A \subseteq V}$, et si $(\tilde{\Phi}_A)_{A \subseteq V}$ est le potentiel canonique associé à \mathbb{P} , alors $(\Phi_A)_{A \subseteq V}$ et $(\tilde{\Phi}_A)_{A \subseteq V}$ sont physiquement équivalents ; par conséquent, on peut associer de manière unique un potentiel canonique à une mesure de Gibbs.

Lemme 3.1.1. *Soit \mathbb{P} un champ aléatoire strictement positif et $(\tilde{\Phi}_A)_{A \subseteq V}$ le potentiel canonique qui lui est associé. Alors \mathbb{P} est la mesure de Gibbs associée à ce potentiel.*

Démonstration. La mesure de Gibbs associée au potentiel $(\tilde{\Phi}_A)_{A \subseteq V}$ est donnée par

$$\begin{aligned}
 \tilde{\mathbb{P}}(\omega) &= \frac{1}{\tilde{\mathbf{Z}}} \exp \left\{ - \sum_{A \subseteq V} \tilde{\Phi}_A(\omega) \right\} \\
 &= \frac{1}{\tilde{\mathbf{Z}}} \exp \left\{ - \sum_{A \subseteq V} \sum_{B \subseteq A} (-1)^{|A \setminus B|} \Psi(\omega^B) \right\} \\
 &= \frac{1}{\tilde{\mathbf{Z}}} \exp \left\{ - \sum_{B \subseteq V} \Psi(\omega^B) \sum_{C \subseteq V \setminus B} (-1)^{|C|} \right\} \\
 &= \frac{1}{\tilde{\mathbf{Z}}} \exp \{ -\Psi(\omega) \} \\
 &= \frac{1}{\tilde{\mathbf{Z}} \mathbb{P}(\mathbf{0})} \mathbb{P}(\omega) \\
 &= \mathbb{P}(\omega).
 \end{aligned}$$

La quatrième identité suit du fait que $\sum_{C \subseteq V \setminus B} (-1)^{|C|} = (1 - 1)^{|V \setminus B|}$ est nulle pour tout $B \neq V$, et la dernière du fait que \mathbb{P} et $\tilde{\mathbb{P}}$ sont des mesures de probabilité (en particulier, on montre par la même occasion que $\tilde{\mathbf{Z}} = 1/\mathbb{P}(\mathbf{0})$). \square

3.2 Équivalence

Dans les sous-sections précédentes, on a défini deux types de champs aléatoires sur G . Le théorème suivant, habituellement appelé théorème de Hammersley-Clifford (mais sur lequel de nombreux autres probabilistes ont travaillé, parmi lesquels Averintsev, Spitzer, Besag, Preston, Grimmett, Kozlov, Sullivan), montre que ces deux notions coïncident lorsque le potentiel associé est p.p.v..

Théorème 3.2.1. *Un champ aléatoire $\mathbb{P} \in \mathcal{M}_1(\Omega, \mathcal{F})$ est markovien si et seulement si \mathbb{P} est une mesure de Gibbs avec potentiel p.p.v..*

Démonstration. Soit \mathbb{P} une mesure de Gibbs avec potentiel p.p.v., et soit $\omega \in \Omega$. Alors

$$\begin{aligned}
 \mathbb{P}(\sigma_i = \omega_i \mid \sigma_k = \omega_k, \forall k \neq i) &= \frac{\prod_{A \subseteq V} e^{-\Phi_A(\omega)}}{\sum_{\substack{\eta \in \Omega \\ \eta_j = \omega_j, \forall j \neq i}} \prod_{A \subseteq V} e^{-\Phi_A(\eta)}} \\
 &= \frac{\prod_{\substack{A \subseteq V \\ A \ni i}} e^{-\Phi_A(\omega)}}{\sum_{\substack{\eta \in \Omega \\ \eta_j = \omega_j, \forall j \neq i}} \prod_{\substack{A \subseteq V \\ A \ni i}} e^{-\Phi_A(\eta)}}.
 \end{aligned}$$

3.2. ÉQUIVALENCE

Le membre de droite ne dépendant que de ω_i et des $(\omega_j)_{j \sim i}$, on en déduit que

$$\mathbb{P}(\sigma_i = \omega_i \mid \sigma_k = \omega_k, \forall k \neq i) = \mathbb{P}(\sigma_i = \omega_i \mid \sigma_k = \omega_k, \forall k \sim i),$$

et donc que \mathbb{P} est markovien.

Réciproquement, supposons que \mathbb{P} soit un champ aléatoire markovien et $(\tilde{\Phi}_A)_{A \subseteq V}$ son potentiel canonique. Le Lemme 3.1.1 montre que \mathbb{P} est la mesure de Gibbs associée au potentiel $(\tilde{\Phi}_A)_{A \subseteq V}$. Il ne reste plus qu'à vérifier que celui-ci est p.p.v., c'est-à-dire que $\tilde{\Phi}_A \equiv 0$ si A n'est pas une clique. Soit $A \subseteq V$, et $i, j \in A$ tels que $i \not\sim j$. On a alors

$$\begin{aligned} \tilde{\Phi}_A(\omega) &= \sum_{B \subseteq A} (-1)^{|A \setminus B|} \Psi(\omega^B) \\ &= \sum_{B \subseteq A \setminus \{i, j\}} (-1)^{|A \setminus B|} \left\{ \Psi(\omega^B) - \Psi(\omega^{B \cup \{i\}}) - \Psi(\omega^{B \cup \{j\}}) + \Psi(\omega^{B \cup \{i, j\}}) \right\}. \end{aligned}$$

Il suffit donc de montrer que

$$\log \mathbb{P}(\omega^B) - \log \mathbb{P}(\omega^{B \cup \{i\}}) - \log \mathbb{P}(\omega^{B \cup \{j\}}) + \log \mathbb{P}(\omega^{B \cup \{i, j\}}) = 0,$$

or ceci est équivalent à

$$\frac{\mathbb{P}(\omega^{B \cup \{i\}})}{\mathbb{P}(\omega^B)} = \frac{\mathbb{P}(\omega^{B \cup \{i, j\}})}{\mathbb{P}(\omega^{B \cup \{j\}})},$$

c'est-à-dire

$$\frac{\mathbb{P}(\sigma_i = \omega_i \mid \sigma_k = \omega_k^{B \cup \{j\}}, \forall k \neq i)}{\mathbb{P}(\sigma_i = 0 \mid \sigma_k = \omega_k^{B \cup \{j\}}, \forall k \neq i)} = \frac{\mathbb{P}(\sigma_i = \omega_i \mid \sigma_k = \omega_k^B, \forall k \neq i)}{\mathbb{P}(\sigma_i = 0 \mid \sigma_k = \omega_k^B, \forall k \neq i)}.$$

Puisque \mathbb{P} est markovien, cette dernière identité peut également s'écrire

$$\frac{\mathbb{P}(\sigma_i = \omega_i \mid \sigma_k = \omega_k^{B \cup \{j\}}, \forall k \sim i)}{\mathbb{P}(\sigma_i = 0 \mid \sigma_k = \omega_k^{B \cup \{j\}}, \forall k \sim i)} = \frac{\mathbb{P}(\sigma_i = \omega_i \mid \sigma_k = \omega_k^B, \forall k \sim i)}{\mathbb{P}(\sigma_i = 0 \mid \sigma_k = \omega_k^B, \forall k \sim i)}.$$

et la conclusion suit puisque $\omega_k^{B \cup \{j\}} = \omega_k^B$ pour $k \sim i$. □

Chapitre 4

Simulation de champs markoviens

Dans ce chapitre, nous allons décrire quelques méthodes permettant d'obtenir des échantillons d'un champ markovien en volume fini. La simulation de champs aléatoires est un domaine très actif, et on pourrait facilement y consacrer un cours entier. Ici, nous nous contenterons donc d'un bref survol.

4.1 Méthode de Monte-Carlo

Soit $\mathbb{P} \in \mathcal{M}_1(\Omega, \mathcal{F})$ un champ markovien sur le graphe $G = (V, E)$, à valeur dans un ensemble S que l'on supposera fini, $S = \{1, \dots, |S|\}$. On désire tirer une configuration de Ω avec probabilité \mathbb{P} . Une façon naturelle de procéder est d'introduire une chaîne de Markov irréductible et apériodique sur l'ensemble fini Ω , choisie de sorte à ce que son unique mesure invariante soit \mathbb{P} . Nous allons discuter le cas standard de l'algorithme du "réservoir thermique" (*heat bath*, en anglais). Dans celui-ci, on passe d'une configuration ω à une autre configuration ω' de la façon suivante : on tire un nombre $u \in [0, 1]$ avec la mesure uniforme, et un sommet $i \in V$, également de façon uniforme. On partitionne (à un ensemble de mesure 0 près) l'intervalle $[0, 1] = \bigvee_{s \in S} \mathcal{I}_s$ en $|S|$ sous-intervalles ouverts disjoints de longueur $|\mathcal{I}_s| = \mathbb{P}(\sigma_i = s \mid \sigma_j = \omega_j, \forall j \sim i)$. On pose alors

$$\omega'_j = \begin{cases} \omega_j & \text{si } j \neq i, \\ s & \text{si } j = i \text{ et } u \in \mathcal{I}_s. \end{cases} \quad (4.1)$$

En d'autres termes, après avoir choisi un sommet au hasard, on fixe la valeur de la configuration en i à $s \in S$ avec probabilité

$$\mathbb{P}(\sigma_i = s \mid \sigma_j = \omega_j, \forall j \sim i).$$

Notons $\mathbb{T}(\omega \rightarrow \omega')$ les probabilités de transition correspondantes.

Lemme 4.1.1. *L'unique distribution stationnaire de la chaîne de Markov construite ci-dessus est \mathbb{P} .*

4.2. SIMULATION PARFAITE

Démonstration. Tout d’abord, la chaîne construite est manifestement irréductible et apériodique, et par conséquent possède une unique mesure stationnaire. Soit $\omega' \in \Omega$. On a

$$\begin{aligned} \sum_{\omega \in \Omega} \mathbb{P}(\omega) \mathbb{T}(\omega \rightarrow \omega') &= \sum_{i \in V} \sum_{\substack{\omega: \\ \omega_j = \omega'_j, \forall j \neq i}} \mathbb{P}(\omega) \mathbb{T}(\omega \rightarrow \omega') \\ &= |V|^{-1} \sum_{i \in V} \mathbb{P}(\sigma_i = \omega'_i \mid \sigma_k = \omega'_k, \forall k \sim i) \sum_{\substack{\omega: \\ \omega_j = \omega'_j, \forall j \neq i}} \mathbb{P}(\omega) \\ &= |V|^{-1} \sum_{i \in V} \mathbb{P}(\omega') \\ &= \mathbb{P}(\omega'), \end{aligned}$$

ce qui montre que la distribution stationnaire de la chaîne est bien \mathbb{P} . \square

On est alors confronté à un problème d’ordre pratique : on ne peut évidemment laisser cet algorithme tourner infiniment longtemps, et il faut donc l’interrompre après un nombre fini M d’itérations. Comment doit-on choisir M si l’on veut être assuré d’être proche de la distribution stationnaire ? Une approche possible est d’étudier la vitesse de convergence de la chaîne. Malheureusement, une telle approche ne fournit en général que des bornes trop grossières pour être réellement utiles en pratique, même si elles peuvent donner des indications¹. Par chance, il existe une variante de l’algorithme ci-dessus, ne nécessitant qu’un nombre fini (mais aléatoire !) de pas, et garantissant que la distribution obtenue est *exactement* la distribution stationnaire : on parle alors de simulation parfaite.

4.2 Simulation parfaite

Nous avons vu qu’il est aisé de construire des chaînes de Markov dont la mesure stationnaire est le champ markovien que l’on désire simuler. Nous allons à présent en donner une variante appelée “couplage depuis le passé” (*coupling from the past*, en anglais) qui permet de générer des configurations distribuées exactement selon la mesure stationnaire tout en ne nécessitant qu’un nombre fini d’itérations. Cette variante est due à Propp et Wilson [22].

Soient $q_{\omega, \omega'}$, $\omega, \omega' \in \Omega$, les probabilités de transition d’une chaîne de Markov X irréductible et apériodique de mesure stationnaire \mathbb{P} . Afin d’analyser la dépendance en l’état initial de la chaîne, on démarre, au temps t_0 , $|\Omega|$ copies de celle-ci, chacune partant d’un état différent. L’évolution se fait de façon indépendante, tant que les trajectoires ne coïncident pas. Si deux (ou plus) trajectoires coïncident à un temps t , alors on les fait évoluer ensemble pour tous les temps ultérieurs, comme représenté sur la Figure 4.1. On

1. Pour nuancer un peu ces propos, de tels algorithmes de simulations sont utilisés régulièrement, par exemple par les physiciens, même en l’absence d’information sur la vitesse de convergence ; pragmatiques, ils se contentent souvent de lancer la simulation à partir de différentes configurations initiales, et estiment avoir attendu suffisamment longtemps si les résultats obtenus ne dépendent pas (trop) de la configuration initiale.

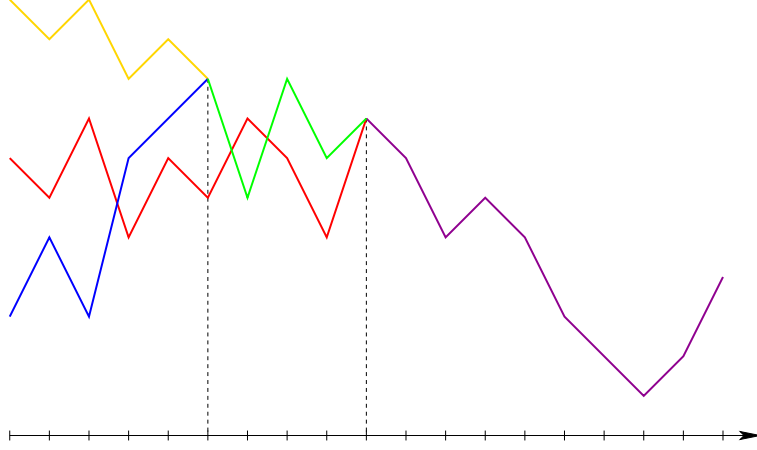


FIGURE 4.1 – Trois copies d’une chaîne partant de trois états différents. Les (ici, deux) instants de coalescence sont indiqués.

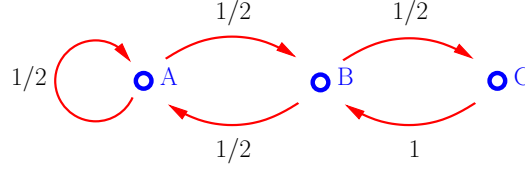


FIGURE 4.2 – Un exemple montrant qu’à l’instant de coalescence, on n’a pas la distribution stationnaire : l’état C ne pouvant être atteint qu’en venant de l’état B, il n’est pas possible que les chaînes soient dans l’état C au moment de la coalescence.

réalise ainsi un couplage de ces $|\Omega|$ chaînes. Une façon équivalente de présenter ce couplage est de considérer une famille $(f_k)_{k \in \mathbb{Z}}$ de fonctions aléatoires i.i.d. $f_k : \Omega \rightarrow \Omega$ de loi \mathbb{Q} telles que, pour tout $\omega' \in \Omega$, $\mathbb{Q}(f_0(\omega) = \omega') = q_{\omega, \omega'}$ indépendamment pour chaque $\omega \in \Omega$. La distribution au temps $t \in \mathbb{N}$ de la chaîne partant de l’état ω au temps $t_0 < t$ est donc identique à celle de $F_{t_0}^t(\omega) \stackrel{\text{déf}}{=} f_{t-1} \circ f_{t-2} \circ \cdots \circ f_{t_0+1} \circ f_{t_0}(\omega)$.

Lemme 4.2.1. *Il existe \mathbb{Q} -presque-sûrement un temps $N < \infty$ tel que F_1^N est une constante.*

Démonstration. Par irréductibilité et apériodicité, il existe $K < \infty$ tel que la probabilité d’aller de ω à ω' en K pas est strictement positive pour toute paire $\omega, \omega' \in \Omega$. Par conséquent, $\mathbb{Q}(F_1^K = \text{const})$ est également strictement positive. Les événements $\{F_{nK+1}^{(n+1)K} = \text{const}\}$, $n \geq 0$, étant i.i.d., on peut trouver, avec \mathbb{Q} -probabilité 1, un n tel que $F_{nK+1}^{(n+1)K}$ soit constante.

Soit $t_2 > t_1 \geq 0$; observant que $F_{t_1}^{t_2} = \text{const}$ implique $F_0^{t'} = \text{const}$ pour tout $t' > t_2$, on en déduit que F_0^t sera constante pour tout t suffisamment grand. \square

Manifestement, une fois que les $|\Omega|$ trajectoires ont coalescé, toute information sur l’état de départ est perdue, et on pourrait donc penser que l’on obtient à cet instant

4.2. SIMULATION PARFAITE

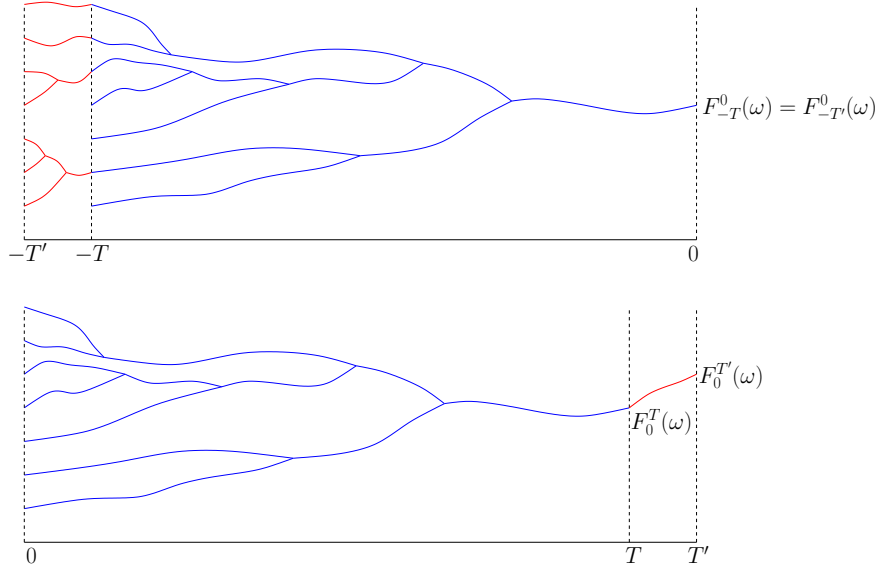


FIGURE 4.3 – Représentation schématique de l'évolution. Haut : couplage depuis le passé ; F^0_{-T} étant constante, c'est également le cas de $F^0_{-T'}$, et leur unique image est identique. Bas : évolution vers le futur ; F^T_0 étant constante, $F^{T'}_0$ l'est aussi, mais il n'est plus vrai en général que leur unique image est identique.

un échantillon distribué selon la loi stationnaire. Mais c'est faux, comme on peut le voir facilement sur l'exemple de la Figure 4.2 ; le problème est que l'observation n'est pas faite en un temps déterministe, ce qui conduit à un biais. Cette idée n'est cependant pas à rejeter complètement, et il se trouve qu'une modification très simple permet de la faire fonctionner.

L'idée est de ne pas chercher à coupler “dans le futur” comme on vient de le faire, mais “depuis le passé” : on va démarrer les $|\Omega|$ chaînes à un temps suffisamment reculé dans le passé, et observer le résultat au temps 0. Plus précisément, on sait que $F^0_{-t} = f_{-1} \circ f_{-2} \circ \dots \circ f_{-t+1} \circ f_{-t}$ est constante pour tout t suffisamment grand (remarquez que l'on n'exige pas que le temps auquel la coalescence de toutes les copies a lieu soit égal à 0, mais seulement que cela ait eu lieu *avant* 0 ; cette différence est cruciale).

Lemme 4.2.2. *Soit M une variable aléatoire telle que F^0_{-M} soit constante (M peut être choisie finie, \mathbb{Q} -presque sûrement). Alors la distribution de l'unique image de F^0_{-M} est précisément \mathbb{P} .*

Démonstration. Soit $\omega \in \Omega$. Comme F^0_{-T} est \mathbb{Q} -presque sûrement constante pour tous les T suffisamment grands, la limite

$$F^0_{-\infty}(\omega) \stackrel{\text{déf}}{=} \lim_{T \rightarrow \infty} F^0_{-T}(\omega)$$

existe \mathbb{Q} -presque sûrement, et est \mathbb{Q} -presque sûrement indépendante de ω (voir la Fi-

gure 4.3). De plus, on a l'égalité suivante (presque sûrement)

$$F_{-\infty}^0 = F_{-M}^0.$$

La conclusion suit, puisque, par ergodicité de la chaîne de Markov, $F_{-\infty}^0$ suit la distribution invariante,

$$F_{-\infty}^0 \stackrel{d}{=} \mathbb{P}.$$

□

Remarque 4.2.1. *Il peut être utile de réaliser ce qui ne marcherait pas si l'on avait procédé "vers le futur" : dans ce cas, il n'est pas vrai que la limite*

$$\lim_{T \rightarrow \infty} F_0^T(\omega)$$

existe. En effet, il est même faux que $F_0^T(\omega) = F_0^{T+1}(\omega)$ en général ! Un coup d'oeil à la Figure 4.3 peut peut-être aider.

En résumé : si `FctAleatoire()` est une routine qui retourne une fonction f distribuée selon \mathbb{Q} , alors l'algorithme

```

t ← 0
F_t^0 ← identité
répéter
|   t ← t - 1
|   f ← FctAleatoire()
|   F_t^0 ← F_{t+1}^0 ∘ f
jusqu'à F_t^0 est constante
retourner l'unique valeur dans l'image de F_t^0
    
```

renvoie un élément de Ω distribué selon \mathbb{P} après un temps presque-sûrement fini.

Faisons à présent quelques commentaires à propos de l'implémentation d'un tel algorithme.

Temps d'échantillonnage. Il n'est pas nécessaire (et pas désirable du tout !) d'appliquer l'algorithme ci-dessus pour chaque $t < 0$. Il suffit bien sûr de choisir une suite décroissante de temps $T_k < 0$, $T_k \downarrow -\infty$, et de vérifier successivement, pour $k = 1, 2, \dots$, si $F_{T_k}^0$ est constante. On peut montrer que le choix $T_k = -2^k$ est proche du choix optimal.

La dynamique sous-jacente. Bien entendu, un pas de temps dans l'algorithme ci-dessus ne correspond pas nécessairement à l'application d'un pas de la chaîne de Markov sous-jacente. par exemple, dans le cas de la dynamique du bain thermique, un pas de la chaîne de Markov ne modifie qu'au plus un spin, et ne favorise donc guère la coalescence des trajectoires. Il est beaucoup plus judicieux pour chaque pas de l'algorithme ci-dessus, d'effectuer un nombre suffisant de pas de la chaîne de Markov.

Il est aussi important de bien choisir cette dynamique sous-jacente. Plus elle converge rapidement, plus l'algorithme de couplage depuis le passé s'arrêtera rapidement. Par exemple, appliquée au modèle d'Ising, la dynamique de bain thermique introduite précédemment converge rapidement dans le régime d'unicité, mais sa convergence est catastrophique dans le régime de coexistence des phases. Heureusement, il en existe de meilleures, au moins dans certaines situations, *cf.* la section 4.3.

Couplage indépendant. Le couplage utilisé ci-dessus est tel que $f_0(\omega)$ est choisie comme étant égale à ω' avec probabilité $q_{\omega, \omega'}$ indépendamment pour chaque $\omega \in \Omega$. Le choix d'un tel couplage indépendant entre les différentes trajectoires jusqu'à leur rencontre n'est bien entendu pas le seul possible. En fait, on peut en général faire beaucoup mieux, en choisissant un couplage qui favorise une coalescence plus rapide du processus. On en verra un exemple ci-dessous.

Couplage monotone. Une faiblesse de l'approche esquissée ci-dessus devrait être évidente : la nécessité de considérer des chaînes partant de chaque état de Ω peut sembler rendre cette approche inutile pour les champs aléatoires. Après tout, pour le modèle d'Ising dans une boîte carrée de 500 sites sur 500 sites (la taille utilisée pour les simulations de la Figure 1.3), le nombre de configurations est déjà de $2^{250000} \simeq 10^{75257}$! Heureusement, dans de nombreuses situations importantes il est possible d'utiliser des propriétés d'ordre afin d'éliminer cette difficulté. Décrivons le principe dans le cas du modèle d'Ising. Nous avons vu, dans la sous-section 2.10.2, que Ω_{Λ_n} peut être muni d'un ordre partiel naturel. On introduit le couplage suivant : on tire un nombre u uniformément dans $[0, 1]$ et un site i uniformément dans $\Lambda_n = \{-n, \dots, n\}^d$ et on définit $f_0(\omega)$ comme dans (4.1), c'est-à-dire qu'on pose $f_0(\omega) = \omega'$, où $\omega'_j = \omega_j$ pour tout $j \neq i$, et ω'_i prend valeur 1 avec probabilité

$$\mu_{\Lambda_n; \beta, h}^{\bar{\omega}}(\sigma_i = 1 \mid \sigma_k = \omega_k, \forall k \sim i) = \left(1 + \exp(-2\beta \sum_{j \sim i} \omega_j - 2h)\right)^{-1},$$

et valeur -1 sinon. Remarquez que ce couplage n'est pas du tout indépendant, puisque l'on utilise les mêmes i et u pour toute configuration initiale ω . En observant que

$$\left(1 + \exp(-2\beta \sum_{j \sim i} \omega_j - 2h)\right)^{-1}$$

est une fonction croissante de ω , on constate que ce couplage possède la propriété suivante : si $\omega \leq \omega'$, alors $f_0(\omega) \leq f_0(\omega')$. Un tel couplage est dit **monotone**.

L'intérêt de ce couplage est qu'il suffit de comparer *deux* trajectoires : celles partant des configurations $\omega \equiv 1$ et $\omega \equiv -1$ dans Λ_n . En effet, comme toute trajectoire partant d'une autre configuration va toujours être prise en sandwich entre ces deux trajectoires-ci, ce seront les dernières à coalescer.

En outre, ce couplage est plus efficace que le couplage indépendant, dans le sens qu'il couple plus rapidement des trajectoires distinctes.

4.3 Algorithmes par amas

Dans le cas du modèle d'Ising, la dynamique de bain thermique converge très lentement dans le régime de coexistence des phases. En effet, la dynamique va rapidement converger vers un état où la boîte se retrouve partagée en régions occupées soit par la phase $+$, soit par la phase $-$, séparées par des interfaces. La relaxation de ces interfaces est très lente, car il n'y a pas de biais en faveur d'une des deux phases (sauf éventuellement au bord du système). Ceci donne lieu à des temps de convergence explosant comme une exponentielle étirée avec la taille du système.

Nous allons présenter dans cette section une dynamique alternative, moins locale, qui ne souffre pas de ces défauts (du moins sur \mathbb{Z}^d et pour des conditions au bord appropriées; remarquez aussi que d'un point de vue théorique, il y a très peu de résultats sur les vitesses de convergence de cet algorithme, bien que d'un point de vue pratique il fonctionne si bien qu'il est, avec d'autres algorithmes analogues, presque toujours utilisé aujourd'hui). Cette dynamique repose sur une représentation géométrique du modèle d'Ising connue sous le nom de **représentation de Fortuin-Kasteleyn**, ou **représentation en amas aléatoires** (*random cluster representation*, en anglais). Celle-ci est aussi extrêmement utile pour l'étude mathématique du modèle d'Ising et de certaines de ses généralisations.

4.3.1 FK percolation

L'espace de configuration du modèle est $\Xi \stackrel{\text{déf}}{=} \{0, 1\}^{\mathcal{L}(\mathbb{Z}^d)}$, où $\mathcal{L}(\mathbb{Z}^d) \stackrel{\text{déf}}{=} \{\{i, j\} \subset \mathbb{Z}^d : i \sim j\}$ est l'ensemble de toutes les paires (non-ordonnées) de plus-proches-voisins de \mathbb{Z}^d , que l'on appellera les **liens**. Étant donné un élément $\omega \in \Xi$ et un lien e , on introduit la variable aléatoire $n_e(\omega) = \omega_e$; un lien $e \in \mathcal{L}(\mathbb{Z}^d)$ sera dit **ouvert** ou **fermé** dans la configuration ω selon que $n_e(\omega) = 1$ ou 0 . Les composantes connexes maximales de liens ouverts sont appelés les **amas** de ω , chaque sommet isolé de \mathbb{Z}^d étant considéré comme un amas.

Soit $\Lambda \Subset \mathbb{Z}^d$ et $\bar{n} \in \Xi$. La mesure de FK dans Λ , avec condition au bord \bar{n} et paramètres $p \in [0, 1]$ et $q > 0$, est la mesure de probabilité suivante sur la tribu-produit associée à Ξ

$$\nu_{\Lambda; p, q}^{\bar{n}}(\omega) \stackrel{\text{déf}}{=} \begin{cases} (\mathbf{Z}_{\Lambda; p, q}^{\text{FK}, \bar{n}})^{-1} \left(\frac{p}{1-p} \right)^{\sum_{e \cap \Lambda \neq \emptyset} n_e(\omega)} q^{\mathfrak{N}(\omega)} & \text{si } n_e(\omega) = \bar{n}_e, \forall e \cap \Lambda = \emptyset, \\ 0 & \text{sinon,} \end{cases}$$

où la variable aléatoire $\mathfrak{N}(\omega)$ compte le nombre d'amas disjoints intersectant Λ , et la fonction de partition $\mathbf{Z}_{\Lambda; p, q}^{\text{FK}, \bar{n}}$ sert de constante de normalisation.

Dans cette section, on va s'intéresser tout particulièrement au cas de la mesure avec condition au bord ouverte (*wired boundary condition*, en anglais), $\bar{n} \equiv 1$, que l'on notera $\nu_{\Lambda; p, q}^1$.

Le cas $q = 1$ est connu sous le nom de **modèle de percolation**; il s'agit d'un modèle extrêmement étudié. Dans ce cas, on voit que la mesure introduite ci-dessus n'est autre que le produit sur $\mathcal{L}(\mathbb{Z}^d)$ de mesures de Bernoulli de paramètre p (c'est-à-dire que chaque

4.3. ALGORITHMES PAR AMAS

lien est ouvert avec probabilité p et fermé avec probabilité $1 - p$, indépendamment de l'état des autres liens).

Le facteur $q^{\mathfrak{N}(\omega)}$ introduit une dépendance entre ces variables aléatoires, et pour cette raison le modèle de Fortuin-Kasteleyn est souvent présenté comme un modèle de *percolation dépendante*.

Il se trouve que le modèle de Fortuin-Kasteleyn se réduit à de nombreux autres modèles d'intérêt (pas tous décrits dans ce cours) lorsque q varie : au modèle d'Ising lorsque $q = 2$, au modèle de Potts à q états, lorsque $q \in \mathbb{N}$, $q \geq 3$, à la mesure uniforme sur les arbres couvrants de \mathbb{Z}^d dans la limite $q \downarrow 0$ et $q/p \rightarrow 0$.

Nous ne décrivons ici que le cas $q = 2$, et sa relation avec le modèle d'Ising. Soit $\mu_{\Lambda;\beta,0}^+$ la mesure de Gibbs en volume fini du modèle d'Ising dans la boîte Λ , avec condition au bord $+$, et paramètres β et $h = 0$.

Considérons les algorithmes suivants :

```

Entrées : une configuration  $\omega$  tirée selon la mesure  $\mu_{\Lambda;\beta,0}^+$ 
pour chaque lien  $e = \{i, j\}$  de  $\mathcal{L}(\mathbb{Z}^d)$  faire
    si  $\omega_i \neq \omega_j$  alors
         $\eta_e \leftarrow 0$ 
    sinon
         $u \leftarrow \text{Uniforme}([0, 1])$ 
        si  $u < p$  alors
             $\eta_e \leftarrow 1$ 
        sinon
             $\eta_e \leftarrow 0$ 
        fin
    fin
fin
retourner la configuration  $\eta \in \Xi$ 

```

```

Entrées : une configuration  $\eta$  tirée selon la mesure  $\nu_{\Lambda;p,2}^1$ 
pour chaque amas fini de la configuration  $\eta$  faire
     $s \leftarrow \text{Bernoulli}(\{-1, 1\}, 1/2)$ 
    pour chaque site  $i$  de l'amas faire  $\omega_i \leftarrow s$ 
fin
pour chaque site  $i$  de l'amas infini faire  $\omega_i \leftarrow 1$ 
retourner la configuration  $\omega \in \{-1, 1\}^{\mathbb{Z}^d}$ 

```

Lemme 4.3.1. *Supposons la relation $p = 1 - e^{2\beta}$ satisfaite. Alors la configuration η retournée par le premier algorithme est distribuée selon $\nu_{\Lambda;p,2}^1$, et la configuration ω retournée par le second algorithme est distribuée selon $\mu_{\Lambda;\beta,0}^+$.*

Démonstration. Le lemme est une conséquence immédiate de l'observation suivante : on obtient un couplage $P_{\Lambda,p}$ des mesures $\mu_{\Lambda;\beta,0}^+$ et $\nu_{\Lambda;p,2}^1$ en associant au couple $(\omega, \eta) \in \Omega \times \Xi$

une probabilité proportionnelle à

$$\prod_{e=\{i,j\} \in \mathcal{L}(\mathbb{Z}^d)} \left(p^{\eta_e} (1-p)^{1-\eta_e} \mathbf{1}_{\{(\omega_i - \omega_j)\eta_e = 0\}} \right)$$

si $\omega_i = 1$ pour tout $i \notin \Lambda$ et $\eta_e = 1$ pour tout $e \cap \Lambda = \emptyset$. En d'autres termes, on choisit indépendamment et uniformément le spin en chaque site et l'état de chaque lien, puis on conditionne sur le fait que si deux sites sont reliés par un lien ouvert alors les spins correspondants coïncident. Vérifions que ceci fournit bien un couplage : en notant $P_{\Lambda;p}^{\text{spins}}$ la marginale sur Ω , et $P_{\Lambda;p}^{\text{liens}}$ la marginale sur Ξ , on a

$$\begin{aligned} P_{\Lambda;p}^{\text{spins}}(\omega) &\propto \sum_{\eta} \prod_{e=\{i,j\}, e \cap \Lambda \neq \emptyset} \left(p^{\eta_e} (1-p)^{1-\eta_e} \mathbf{1}_{\{(\omega_i - \omega_j)\eta_e = 0\}} \right) \\ &= \prod_{e=\{i,j\}, e \cap \Lambda \neq \emptyset} (1-p)^{\mathbf{1}_{\{\omega_i \neq \omega_j\}}} \\ &= e^{-2\beta \sum_{e=\{i,j\}, e \cap \Lambda \neq \emptyset} \mathbf{1}_{\{\omega_i \neq \omega_j\}}} \\ &\propto \mu_{\Lambda;\beta,0}^+(\omega), \end{aligned}$$

où la somme est prise sur toutes les configurations $\eta \in \Xi$ telles que $\eta_e = 1$ pour tout $e \cap \Lambda = \emptyset$. Similairement,

$$\begin{aligned} P_{\Lambda;p}^{\text{liens}}(\eta) &\propto \sum_{\omega} \prod_{e=\{i,j\}} \left(p^{\eta_e} (1-p)^{1-\eta_e} \mathbf{1}_{\{(\omega_i - \omega_j)\eta_e = 0\}} \right) \\ &= 2^{\mathfrak{N}(\eta)} \prod_{e=\{i,j\}} \left(p^{\eta_e} (1-p)^{1-\eta_e} \right) \\ &\propto \nu_{\Lambda;p,2}^1(\eta), \end{aligned}$$

où la somme est prise sur toutes les configurations $\omega \in \Omega$ telles que $\omega_i = 1$ pour tout $i \notin \Lambda$. \square

4.3.2 Dynamique de Swendsen-Wang

Il est à présent très facile de construire une nouvelle chaîne de Markov dont la mesure stationnaire est $\mu_{\Lambda;\beta,0}^+$: on passe d'une configuration ω à une configuration ω' en appliquant successivement (et dans l'ordre) chacun des deux algorithmes présentés avant le lemme 4.3.1. La première application génère (aléatoirement) une configuration η de liens, compatible avec la configuration de spin ω ; la seconde application retourne une nouvelle configuration (aléatoire) de spins ω' , compatible avec la configuration de liens intermédiaire η .

Lemme 4.3.2. *La chaîne de Markov introduite ci-dessus converge vers $\mu_{\Lambda;\beta,0}^+$.*

Démonstration. Le lemme 4.3.1 montre que si la configuration ω est distribuée selon $\mu_{\Lambda;\beta,0}^+$, alors c'est encore le cas pour la configuration ω' , donc $\mu_{\Lambda;\beta,0}^+$ est bien une mesure stationnaire de la chaîne. Puisque n'importe quelle configuration ω' (compatible avec la condition

au bord) peut être atteinte en une étape (ceci est possible, par exemple, si tous les liens de la configuration intermédiaire η sont fermés), la chaîne est irréductible et apériodique, et donc ergodique. \square

Les grands avantages de cette dynamique par rapport à celle du bain thermique sont sa nature non-locale, qui lui permet de se déplacer plus rapidement dans l'espace des configurations, et le fait qu'elle ne distingue pas les différentes phases d'équilibre et fonctionne ainsi très efficacement même dans le régime de coexistence des phases.

Remarque 4.3.1. *Une limitation de cette approche est liée à la condition au bord. S'il est bien entendu possible de traiter également les conditions au bord –, libre et périodique, il n'en est pas de même pour des conditions au bord générales $\bar{\omega}$, car celles-ci induisent en général des contraintes sur les connections autorisées à l'intérieur de Λ (si $\bar{\omega}_i \neq \bar{\omega}_j$, alors il ne peut pas y avoir d'amas connectant i à j).*

Remarque 4.3.2. *Malgré le caractère non-monotone de cette dynamique, il est tout de même possible de l'utiliser dans une approche “couplage depuis le passé”, en utilisant une technique dite de *chaîne dominante* (bounding chain, en anglais) afin de détecter le moment de la coalescence. Nous ne discuterons pas de cette technique ici.*

Chapitre 5

Modèles avec symétrie continue

Dans le chapitre 2, nous avons étudié de façon assez détaillée la transition de phase du modèle d'Ising. En particulier, nous avons vu que la transition de phase en température (qui a lieu lorsque $h = 0$ et $d \geq 2$) correspond à une brisure spontanée de symétrie : lorsque $\beta < \beta_c$, il y a une unique mesure de Gibbs en volume infini, invariante sous l'échange des spins $+$ et $-$; au contraire, lorsque $\beta > \beta_c$, il n'y a plus unicité, et l'on a montré l'existence de deux mesures distinctes $\mu_{\beta,0}^+$ et $\mu_{\beta,0}^-$ n'étant plus invariantes sous l'échange des spins (puisque, par exemple, $\langle \sigma_0 \rangle_{\beta,0}^+ > 0$). Notons toutefois que l'image de $\mu_{\beta,0}^+$ sous échange des spins est précisément $\mu_{\beta,0}^-$, et réciproquement.

Dans le présent chapitre, nous allons étudier l'effet de la présence d'une symétrie continue (c'est-à-dire l'invariance sous l'action d'un groupe de Lie) sur les transitions de phases. Nous verrons en particulier que celles-ci sont beaucoup plus difficiles à briser : il n'y a généralement pas de transition de phase pour de tels modèles markoviens en dimension 2.

Avant de faire cela, nous allons brièvement discuter de façon générale la relation entre les symétries présentes au niveau de l'Hamiltonien, et celles présentes au niveau des mesures de Gibbs (en volume infini).

5.1 Symétrie et mesures de Gibbs

On considère un champ markovien dont l'espace des configurations est $\Omega = S^{\mathbb{Z}^d}$. Soit $(\mu_{\Lambda_n}^{\omega_n})_{n \geq 1}$ une suite de mesures de Gibbs en volume fini, convergeant vers une mesure de Gibbs en volume infini μ . Soit G un groupe de transformations $g : S \rightarrow S$; on note $g(\omega)$ la configuration telle que $\sigma_i(g(\omega)) = g(\omega_i)$, pour tout $i \in \mathbb{Z}^d$. On suppose que l'Hamiltonien du système est invariant sous l'action de G , c'est-à-dire que $H_\Lambda(g(\omega)) = H_\Lambda(\omega)$ pour tout $g \in G$, $\omega \in \Omega$ et $\Lambda \Subset \mathbb{Z}^d$.

Lemme 5.1.1. *Sous les hypothèses décrites ci-dessus, les mesures $g(\mu) \stackrel{\text{def}}{=} \mu \circ g^{-1}$, $g \in G$, sont aussi des mesures de Gibbs en volume infini (pouvant éventuellement coïncider avec*

5.2. DÉFINITION DU MODÈLE

μ). En particulier, s'il y a unicité de la mesure de Gibbs en volume infini, on doit avoir $\mathbf{g}(\mu) = \mu$, pour tout $\mathbf{g} \in \mathbf{G}$.

Démonstration. Il nous suffit de construire une suite de mesures de Gibbs en volume fini convergeant vers $g(\mu)$. Étant donné que, pour toute fonction f , tout sous-ensemble $\Lambda \Subset \mathbb{Z}^d$ et toute configuration $\bar{\omega}$,

$$\begin{aligned} \sum_{\substack{\omega: \\ \omega_i = \bar{\omega}_i, \forall i \notin \Lambda}} e^{-\mathcal{H}_\Lambda(\omega)} f(\omega) &= \sum_{\substack{\omega: \\ \omega_i = \mathbf{g}^{-1}(\bar{\omega}_i), \forall i \notin \Lambda}} e^{-\mathcal{H}_\Lambda(\mathbf{g}(\omega))} f(\mathbf{g}(\omega)) \\ &= \sum_{\substack{\omega: \\ \omega_i = \mathbf{g}^{-1}(\bar{\omega}_i), \forall i \notin \Lambda}} e^{-\mathcal{H}_\Lambda(\omega)} f(\mathbf{g}(\omega)), \end{aligned}$$

on voit immédiatement que

$$\langle f \rangle_{\Lambda_n}^{\mathbf{g}(\bar{\omega}_n)} = \langle f \circ \mathbf{g} \rangle_{\Lambda_n}^{\bar{\omega}_n}.$$

Or, le membre de droite converge, lorsque $n \rightarrow \infty$, vers

$$\langle f \circ \mathbf{g} \rangle_\mu = \langle f \rangle_{\mu \circ \mathbf{g}^{-1}} = \langle f \rangle_{\mathbf{g}(\mu)},$$

et la suite $(\mu_{\Lambda_n}^{\mathbf{g}(\bar{\omega}_n)})_{n \geq 1}$ convient donc. \square

Remarque 5.1.1. Nous avons supposé, pour simplifier les notations, que le groupe de symétrie agissait directement sur S . On aurait pu supposer que son action n'était définie que sur Ω , et qu'en particulier le groupe puisse agir sur les coordonnées $i \in \mathbb{Z}^d$. Ceci permet par exemple de discuter la symétrie sous les translations. Il est facile de vérifier que l'argument ci-dessus s'applique également à cette situation plus générale.

Dans le cas du modèle d'Ising avec champ magnétique nul, l'Hamiltonien est invariant sous l'action du groupe de symétrie $\mathbf{G} = \{\mathbf{e}, \mathbf{f}\}$ sur $\{-1, 1\}$, où \mathbf{e} est l'identité et $\mathbf{f}(s) \stackrel{\text{déf}}{=} -s$. Nous avons vu que cette symétrie est préservée lorsque $d = 1$ ou lorsque $d \geq 2$ et $\beta < \beta_c(d)$, puisqu'il y a unicité de la mesure en volume infini dans ces cas-là, alors qu'elle est brisée en dimension $d \geq 2$ lorsque $\beta > \beta_c$, puisque $\mu_{\beta,0}^+ \neq \mu_{\beta,0}^- = \mathbf{f}(\mu_{\beta,0}^+)$.

5.2 Définition du modèle

Une configuration du modèle est une application $\omega : \mathbb{Z}^d \rightarrow \mathbb{S}^1$, où \mathbb{S}^1 est le cercle unité ; les éléments de \mathbb{S}^1 sont paramétrés par l'angle correspondant. L'ensemble des configurations est noté Ω . À chaque site $i \in \mathbb{Z}^d$, nous attachons une variable aléatoire $\vartheta_i : \Omega \rightarrow \mathbb{S}^1$ associant à une configuration ω l'angle correspondant à ω_i .

Pour $\Lambda \Subset \mathbb{Z}^d$, l'Hamiltonien est défini par

$$\mathcal{H}_\Lambda(\omega) \stackrel{\text{déf}}{=} \sum_{\substack{\{i,j\} \cap \Lambda \neq \emptyset \\ i \sim j}} V(\vartheta_j(\omega) - \vartheta_i(\omega)),$$

où $V : \mathbb{S}^1 \rightarrow \mathbb{R}$ est une fonction paire de classe C^2 et telle que $V(0) = 0$ (pensez à $V(x) = 1 - \cos x$; dans ce cas $V(\vartheta_j(\omega) - \vartheta_i(\omega))$ correspond, à une constante sans importance près, au produit scalaire entre les spins en i et j).

La mesure de Gibbs dans Λ , avec condition au bord $\bar{\vartheta} \in \Omega$, est la mesure de probabilité $\mu_{\Lambda}^{\bar{\vartheta}}$ sur (Ω, \mathcal{F}) définie par

$$\mu_{\Lambda}^{\bar{\vartheta}}(d\omega) \stackrel{\text{déf}}{=} \begin{cases} (\mathbf{Z}_{\Lambda}^{\bar{\vartheta}})^{-1} e^{-\mathcal{H}_{\Lambda}(\omega)} \lambda^{\Lambda}(d\omega) & \text{si } \omega \equiv \bar{\vartheta} \text{ hors de } \Lambda, \\ 0 & \text{sinon,} \end{cases}$$

où λ est la mesure uniforme (mesure de Haar) sur \mathbb{S}^1 .

Manifestement, l'Hamiltonien est invariant sous l'action du groupe $\text{SO}(2)$ des rotations du cercle. Nous allons voir qu'il est plus difficile de briser une telle symétrie continue qu'une symétrie discrète.

5.3 Absence de brisure de symétrie

Le résultat principal de cette section est le théorème suivant, habituellement appelé Théorème de Mermin-Wagner (bien que ces derniers n'en aient démontré qu'une version beaucoup plus faible) ; la version que nous en donnons ici est essentiellement due à Pfister [21]. Notons $r_{\psi} : \Omega \rightarrow \Omega$ la rotation simultanée par un angle $\psi : \omega_i \mapsto \omega_i + \psi, \forall i \in \mathbb{Z}^d$; on utilise la même notation pour son action sur les fonctions et sur les mesures : $r_{\psi} f \stackrel{\text{déf}}{=} f \circ r_{-\psi}$, $\langle f \rangle_{r_{\psi} \mu_{\Lambda}^{\bar{\vartheta}}} \stackrel{\text{déf}}{=} \langle r_{-\psi} f \rangle_{\mu_{\Lambda}^{\bar{\vartheta}}}$.

Théorème 5.3.1. *En dimensions 1 et 2, toutes les mesures de Gibbs (en volume infini) sont invariantes sous l'action de $\text{SO}(2)$: pour toute fonction locale f ,*

$$\lim_{\Lambda \nearrow \mathbb{Z}^d} |\langle f \rangle_{\Lambda}^{\bar{\vartheta}} - \langle r_{\psi} f \rangle_{\Lambda}^{\bar{\vartheta}}| = 0,$$

pour tout choix de condition au bord $\bar{\vartheta}$, et pour tout angle ψ .

Remarque 5.3.1. *Le résultat ci-dessus peut être étendu à des systèmes invariants sous l'action d'un groupe de Lie compact connexe arbitraire [6] (les spins prenant valeur dans une variété appropriée). L'hypothèse de régularité C^2 de la fonction V peut également être affaiblie [11] (la continuité suffit).*

Avant de passer à la preuve, il est utile d'expliquer heuristiquement la différence essentielle avec le cas du modèle d'Ising. Considérons le modèle dans la boîte $\Lambda_n = \{-n, \dots, n\}^2$, avec condition au bord $\bar{\vartheta} \equiv 0$. Quelle est l'énergie minimale d'une configuration avec $\theta_0(\omega) = \pi$? Dans le cas du modèle d'Ising, la présence d'un spin $-$ dans une telle boîte avec condition au bord $+$ nécessite la formation d'un contour entourant l'origine. En particulier, le coût énergétique pour mettre un spin -1 au centre de la boîte est d'au moins 8β (si tous les autres spins sont dans l'état $+1$). La situation est ici radicalement différente :

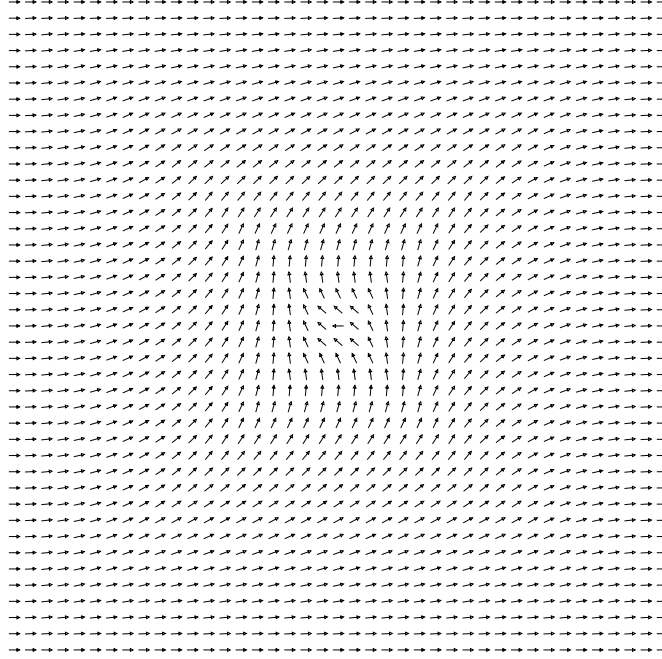


FIGURE 5.1 – Une onde de spin retournant le spin situé au centre de la boîte.

le coût minimal pour avoir $\theta_0(\omega) = \pi$ tend vers 0 lorsque $n \rightarrow \infty$! En effet, considérons la configuration (cf. Figure 5.1)

$$\omega_i^{\text{sw}} \stackrel{\text{déf}}{=} \left(1 - \frac{\log(1 + \|i\|_\infty)}{\log(1 + n)}\right) \pi,$$

si $i \in \Lambda_n$ et $\omega_i^{\text{sw}} = 0$ sinon. Son énergie peut facilement être estimée : les seules contributions non-nulles à l'énergie (se souvenir que $V(0) = 0$) sont celles provenant des paires de sites i, j voisins avec $\|i\|_\infty = \|j\|_\infty - 1$. Dans ce cas, en faisant un développement en série de Taylor,

$$\begin{aligned} V(\omega_i^{\text{sw}} - \omega_j^{\text{sw}}) &= V\left(\pi \frac{\log(1 + \frac{1}{\|j\|_\infty})}{\log(1 + n)}\right) \simeq V\left(\frac{\pi}{\|j\|_\infty \log(1 + n)}\right) \\ &\simeq \frac{\pi^2 V''(0)}{2 (\log(1 + n))^2} \frac{1}{\|j\|_\infty^2}. \end{aligned}$$

On a donc

$$\mathcal{H}_{\Lambda_n}(\omega^{\text{sw}}) \simeq \frac{\pi^2 V''(0)}{2 (\log(1 + n))^2} \sum_{r=1}^{n+1} 4(2r-1)r^{-2} \leq \frac{C}{\log(1 + n)},$$

qui tend bien vers 0 lorsque $n \rightarrow \infty$. C'est l'existence de telles excitations d'énergie arbitrairement faible, appelées **ondes de spin**, qui rend impossible l'application d'un argument de Peierls. Un calcul similaire peut être fait en dimension 1 (en prenant $\omega_i^{\text{sw}} =$

$(1 - (\|i\|_\infty/n)\pi)$, mais pas en dimensions $d \geq 3$, l'énergie minimale ne tendant plus vers 0 (essayez !). On verra plus bas que dans ce dernier cas, il peut y avoir brisure de la symétrie continue à basse température.

La preuve du théorème 5.3.1 rend rigoureux cet argument heuristique.

Démonstration du théorème 5.3.1. Soit $\Lambda_n = \{-n, \dots, n\}^d$. On introduit

$$R \stackrel{\text{déf}}{=} \min\{n : \text{supp}(f) \subset \Lambda_n\},$$

$$N \stackrel{\text{déf}}{=} \max\{n : \Lambda_n \subset \Lambda\}.$$

Comme $N \rightarrow \infty$ lorsque $\Lambda \nearrow \mathbb{Z}^d$, on peut, sans perte de généralité, supposer Λ suffisamment grand pour que $N > R$.

Soit $\Psi \in \Omega$ une configuration telle que $\Psi_i = \psi$ pour tout $i \in \Lambda_R$, et $\Psi_i = 0$ pour tout $i \notin \Lambda_N$; on spécifiera cette configuration plus tard. Soit $t_\Psi : \Omega \rightarrow \Omega$ la transformation $\omega \mapsto (\omega_i - \Psi_i)_{i \in \mathbb{Z}^d}$. On considère la mesure de probabilité suivante :

$$\mu_{\Lambda}^{\bar{\vartheta}; \Psi}(d\omega) \stackrel{\text{déf}}{=} \begin{cases} (\mathbf{Z}_{\Lambda}^{\bar{\vartheta}})^{-1} e^{-\mathcal{H}_{\Lambda}(t_{\Psi}(\omega))} \lambda^{\Lambda}(d\omega) & \text{si } \omega \equiv \bar{\vartheta} \text{ hors de } \Lambda, \\ 0 & \text{sinon.} \end{cases}$$

Notez que la fonction de partition est inchangée. L'intérêt de cette nouvelle mesure est que l'on a manifestement

$$\langle r_{\psi} f \rangle_{\Lambda}^{\bar{\vartheta}} = \langle f \rangle_{\Lambda}^{\bar{\vartheta}; \Psi}, \quad (5.1)$$

et qu'on se ramène donc à comparer l'espérance de f sous deux mesures différentes. Pour ce faire, on va utiliser l'inégalité de Pinsker.

Lemme 5.3.1 (Inégalité de Pinsker). *Soit μ et ν deux mesures de probabilité sur un même espace, avec μ absolument continue par rapport à ν . Alors,*

$$\mathbf{H}(\mu | \nu) \geq \frac{1}{2} (\|\mu - \nu\|_{\text{TV}})^2,$$

où $\mathbf{H}(\mu | \nu) \stackrel{\text{déf}}{=} \langle \log \frac{d\mu}{d\nu} \rangle_{\mu}$ est l'entropie relative de μ par rapport à ν et $\|\mu - \nu\|_{\text{TV}} \stackrel{\text{déf}}{=} \langle |1 - \frac{d\mu}{d\nu}| \rangle_{\nu}$ est la distance en variation totale entre les mesures μ et ν .

Démonstration du lemme 5.3.1. Notons $m = \frac{d\mu}{d\nu} - 1$. On a évidemment $\langle m \rangle_{\nu} = 0$. On peut donc écrire

$$\begin{aligned} \mathbf{H}(\mu | \nu) &= \langle (1 + m) \log(1 + m) \rangle_{\nu} \\ &= \langle (1 + m) \log(1 + m) - m \rangle_{\nu} && (\langle m \rangle_{\nu} = 0) \\ &\geq \langle \frac{m^2}{2(1 + \frac{m}{3})} \rangle_{\nu} && (\text{voir plus bas}) \\ &= \langle \frac{m^2}{2(1 + \frac{m}{3})} \rangle_{\nu} \langle 1 + \frac{m}{3} \rangle_{\nu} && (\langle m \rangle_{\nu} = 0) \\ &\geq \frac{1}{2} (\langle |m| \rangle_{\nu})^2 && (\text{Cauchy-Schwarz}) \\ &= \frac{1}{2} (\|\mu - \nu\|_{\text{TV}})^2. \end{aligned}$$

La première inégalité résulte d'une application de

$$(1+x)\log(1+x) - x \geq \frac{1}{2}x^2 \left(1 + \frac{x}{3}\right)^{-1},$$

elle-même conséquence de l'inégalité de Jensen,

$$\begin{aligned} (1+x)\log(1+x) - x &= x^2 \int_0^1 dt \int_0^t ds \frac{1}{1+xs} \\ &\geq \frac{1}{2}x^2 \frac{1}{1+x \int_0^1 dt \int_0^t ds 2s} = \frac{1}{2}x^2 \left(1 + \frac{x}{3}\right)^{-1}. \end{aligned}$$

□

Revenons à la preuve du théorème 5.3.1. En utilisant (5.1) et en appliquant l'inégalité de Pinsker, on obtient

$$|\langle f \rangle_{\Lambda}^{\bar{\vartheta}} - \langle r_{\psi} f \rangle_{\Lambda}^{\bar{\vartheta}}| \leq \|f\|_{\infty} \sqrt{2\mathbf{H}(\mu_{\Lambda}^{\bar{\vartheta}} | \mu_{\Lambda}^{\bar{\vartheta};\Psi})}.$$

Il reste à contrôler l'entropie relative entre les deux mesures. Un développement de Taylor au second ordre nous donne l'expression suivante

$$\begin{aligned} \mathbf{H}(\mu_{\Lambda}^{\bar{\vartheta}} | \mu_{\Lambda}^{\bar{\vartheta};\Psi}) &= \langle \mathcal{H}_{\Lambda}(\omega) - \mathcal{H}_{\Lambda}(\mathbf{t}_{\Psi}(\omega)) \rangle_{\Lambda}^{\bar{\vartheta}} \\ &= \sum_{\substack{\{i,j\} \cap \Lambda \neq \emptyset \\ i \sim j}} \langle V(\vartheta_j(\omega) - \vartheta_i(\omega)) - V(\vartheta_j(\omega) - \Psi_j - \vartheta_i(\omega) + \Psi_i) \rangle_{\Lambda}^{\bar{\vartheta}} \\ &= \sum_{\substack{\{i,j\} \cap \Lambda \neq \emptyset \\ i \sim j}} \left[\langle V'(\vartheta_j(\omega) - \vartheta_i(\omega)) \rangle_{\Lambda}^{\bar{\vartheta}} (\Psi_i - \Psi_j) \right. \\ &\quad \left. + \langle V''(t_{ij}) \rangle_{\Lambda}^{\bar{\vartheta}} \frac{1}{2} (\Psi_i - \Psi_j)^2 \right], \end{aligned}$$

pour certaines fonctions $t_{ij} : \Omega \rightarrow \mathbb{S}^1$. Puisque V est deux fois continûment différentiable sur \mathbb{S}^1 , on sait que $V''(t) \leq C$ uniformément pour tout $t \in \mathbb{S}^1$. Afin d'éliminer les termes de premier ordre, on utilise la méthode suivante : puisque l'entropie relative est toujours positive (il suffit d'appliquer l'inégalité de Jensen), on peut écrire

$$\mathbf{H}(\mu_{\Lambda}^{\bar{\vartheta}} | \mu_{\Lambda}^{\bar{\vartheta};\Psi}) \leq \mathbf{H}(\mu_{\Lambda}^{\bar{\vartheta}} | \mu_{\Lambda}^{\bar{\vartheta};\Psi}) + \mathbf{H}(\mu_{\Lambda}^{\bar{\vartheta}} | \mu_{\Lambda}^{\bar{\vartheta};-\Psi}).$$

Le second terme du membre de droite de cette dernière expression peut être traité comme ci-dessus, et donne lieu aux mêmes termes de premier ordre *mais avec le signe opposé*. On obtient donc

$$\mathbf{H}(\mu_{\Lambda}^{\bar{\vartheta}} | \mu_{\Lambda}^{\bar{\vartheta};\Psi}) \leq C \sum_{\substack{\{i,j\} \cap \Lambda \neq \emptyset \\ i \sim j}} (\Psi_i - \Psi_j)^2.$$

La conclusion du théorème est donc immédiate si on choisit pour Ψ la configuration utilisée dans l'argument heuristique donné plus haut. Il est cependant instructif de déterminer la configuration optimale ; c'est le contenu du lemme suivant.

Lemme 5.3.2. *La fonctionnelle*

$$\sum_{\substack{\{i,j\} \cap \Lambda \neq \emptyset \\ i \sim j}} (\Psi_i - \Psi_j)^2,$$

définie sur les fonctions $\Psi : \mathbb{Z}^d \rightarrow \mathbb{R}$ avec $\Psi_i = 0$ pour tout $i \notin \Lambda_N$ et $\Psi_i = 1$ pour tout $i \in \Lambda_R$, atteint son minimum pour la fonction $\Psi_k^{R,N} \stackrel{\text{déf}}{=} \mathbb{P}_k(X \text{ visite } \Lambda_R \text{ avant de quitter } \Lambda_N)$, où \mathbb{P}_k est la loi de la marche aléatoire simple sur \mathbb{Z}^d avec condition initiale $X_0 = k$.

De plus,

$$\lim_{N \rightarrow \infty} \sum_{\substack{\{i,j\} \cap \Lambda \neq \emptyset \\ i \sim j}} (\Psi_i^{R,N} - \Psi_j^{R,N})^2 = 0,$$

si et seulement si $d = 1$ ou 2 .

Démonstration du Lemme 5.3.2. Soit $\Delta_{N,R} \stackrel{\text{déf}}{=} \Lambda_N \setminus \Lambda_R$. Il convient tout d'abord de remarquer que la propriété de Markov pour la marche aléatoire simple implique que $\Psi^{R,N}$ est harmonique dans $\Delta_{N,R}$, c'est-à-dire que, pour tout $i \in \Delta_{N,R}$,

$$\Psi_i^{R,N} = \frac{1}{2d} \sum_{j \sim i} \Psi_j^{R,N}.$$

Soit à présent $a \in \mathbb{R}$ et $\delta \in \mathbb{R}^{\mathbb{Z}^d}$ telle que $\delta_i = 0$ pour tout $i \notin \Lambda_N$ et $\delta_i = a$ pour tout $i \in \Lambda_R$; on a

$$\begin{aligned} \sum_{\substack{\{i,j\} \cap \Lambda \neq \emptyset \\ i \sim j}} (\Psi_i^{R,N} - \Psi_j^{R,N})(\delta_i - \delta_j) &= \sum_{i \in \Lambda} \delta_i \sum_{j \sim i} (\Psi_i^{R,N} - \Psi_j^{R,N}) \\ &= a \sum_{i \in \Lambda_R} \sum_{j \sim i} (1 - \Psi_j^{R,N}), \end{aligned} \quad (5.2)$$

puisque, par harmonicité,

$$\sum_{i \in \Delta_{N,R}} \delta_i \sum_{j \sim i} (\Psi_i^{R,N} - \Psi_j^{R,N}) = \sum_{i \in \Delta_{N,R}} \delta_i (2d \Psi_i^{R,N} - \sum_{j \sim i} \Psi_j^{R,N}) = 0.$$

On en déduit que, pour tout $\delta \in \mathbb{R}^{\mathbb{Z}^d}$ comme ci-dessus avec $a = 0$,

$$\sum_{\substack{\{i,j\} \cap \Lambda \neq \emptyset \\ i \sim j}} (\Psi_i^{R,N} + \delta_i - \Psi_j^{R,N} - \delta_j)^2 = \sum_{\substack{\{i,j\} \cap \Lambda \neq \emptyset \\ i \sim j}} (\Psi_i^{R,N} - \Psi_j^{R,N})^2 + \sum_{\substack{\{i,j\} \cap \Lambda \neq \emptyset \\ i \sim j}} (\delta_i - \delta_j)^2,$$

et la première affirmation du lemme est démontrée.

Une seconde application de (5.2), avec $\delta = \Psi^{R,N}$ (et donc $a = 1$), et la propriété de Markov montrent que

$$\begin{aligned} \sum_{\substack{\{i,j\} \cap \Lambda \neq \emptyset \\ i \sim j}} (\Psi_i^{R,N} - \Psi_j^{R,N})^2 &= \sum_{i \in \Lambda_R} \sum_{j \sim i} \mathbb{P}_j(X \text{ quitte } \Lambda_N \text{ avant de visiter } \Lambda_R) \\ &= 2d \sum_{i \in \Lambda_R} \mathbb{P}_i(X \text{ quitte } \Lambda_N \text{ avant de revisiter } \Lambda_R). \end{aligned}$$

Or $\lim_{N \rightarrow \infty} \mathbb{P}_i(X \text{ quitte } \Lambda_N \text{ avant de revisiter } \Lambda_R) = 0$ si et seulement si la marche aléatoire simple est récurrente, c'est-à-dire si et seulement si $d = 1$ ou 2 . \square

\square

5.4 Transitions de phases

Dans cette section, nous énonçons sans preuve¹ deux résultats très importants concernant les mesures de Gibbs avec symétrie continue. Pour simplifier, nous nous restreignons à un modèle bien particulier, le modèle XY , pour lequel $V(x) = -\beta \cos(x)$ (c'est-à-dire que l'interaction est proportionnelle au produit scalaire entre les spins). $\beta \geq 0$ possède à nouveau l'interprétation d'une température inverse.

5.4.1 Brisure de symétrie

Le premier résultat montre que le théorème 5.3.1 est optimal pour ce modèle, en ce sens qu'à basse température il y a brisure de l'invariance sous $SO(2)$ dès que $d \geq 3$.

Théorème 5.4.1. *En dimension $d \geq 3$, il existe une famille de mesures de Gibbs $(\mu_\beta^\psi)_{\psi \in \mathbb{S}^1}$ telles que, pour tout β suffisamment grand,*

$$\langle \cos(\theta_0 - \psi) \rangle_\beta^\psi > 0.$$

Démonstration. Ce sera le contenu du Théorème 7.4.3. \square

Remarque 5.4.1. *Bien sûr, ceci implique que les mesures μ_β^ψ ne sont pas invariantes sous $SO(2)$, car dans ce cas $\langle \cos(\theta_0 - \tilde{\psi}) \rangle_\beta^\psi$ est indépendant de $\tilde{\psi}$ et donc*

$$\langle \cos(\theta_0 - \psi) \rangle_\beta^\psi = \frac{1}{2\pi} \langle \int_0^{2\pi} \cos(\theta_0 - \tilde{\psi}) d\tilde{\psi} \rangle_\beta^\psi = 0.$$

5.4.2 Transition de Kosterlitz-Thouless

Considérons le modèle XY en dimension 2. Le Théorème 5.3.1 montre que toutes les mesures de Gibbs sont invariantes sous $SO(2)$, mais ne suffit pas à impliquer l'unicité. On peut cependant montrer, à l'aide d'inégalités de corrélation adaptées, qu'il implique qu'il y a une unique mesure de Gibbs μ_β invariante sous les translations, pour tout $\beta \geq 0$ [4]. Comme l'on s'attend en fait à ce qu'il n'y ait aucune mesure de Gibbs non-invariante sous les translations pour ce modèle lorsque $d = 2$ (mais aucune preuve de cela n'est connue à l'heure actuelle), il ne devrait pas y avoir de transition de phase dans le sens discuté dans les chapitres précédents.

Le remarquable théorème suivant, dû à Fröhlich et Spencer prouve toutefois qu'il y a une autre sorte de transition de phase dans ce modèle, ce qui montre que la non-unicité des mesures de Gibbs ne représente qu'un type (certes important) de transition de phases.

1. La preuve du théorème 5.4.1 sera donnée au chapitre 7, Théorème 7.4.3 ; la preuve du théorème 5.4.2 est quant à elle beaucoup trop longue et difficile pour ce cours.

Théorème 5.4.2. *Il existe $0 < \beta_{\text{KT}} < \infty$ tel que*

— *pour tout $\beta < \beta_{\text{KT}}$, il existe $C(\beta)$ et $m(\beta) > 0$ telles que*

$$\langle \cos(\vartheta_i - \vartheta_j) \rangle_\beta \leq C(\beta) \exp(-m(\beta) \|j - i\|_2),$$

uniformément en $i, j \in \mathbb{Z}^d$;

— *pour tout $\beta > \beta_{\text{KT}}$, il existe $c(\beta) > 0$ et D telles que*

$$\langle \cos(\vartheta_i - \vartheta_j) \rangle_\beta \geq c(\beta) \|j - i\|_2^{-D/\beta},$$

uniformément en $i, j \in \mathbb{Z}^d$.

Remarque 5.4.2. *Contrairement au théorème de Mermin-Wagner, qui est valide pour tout groupe de Lie, ce dernier résultat semble n'être vrai que pour les systèmes dont le groupe de symétrie est abélien. En particulier, les corrélations du modèle de Heisenberg classique, défini comme le modèle XY, mais dans lequel les spins prennent valeur dans la sphère \mathbb{S}^2 (possédant ainsi un Hamiltonien invariant sous l'action de $\text{SO}(3)$), devraient décroître exponentiellement à toute température. Ceci reste un des problèmes ouverts majeurs en physique statistique.*

Modèles avec spin non-borné

Jusqu'à présent, dans tous les modèles considérés, les spins étaient toujours à valeurs dans un ensemble compact. Une importante conséquence est la trivialité du problème de l'*existence* d'une mesure de Gibbs en volume infini, l'espace des mesures de probabilité sur (Ω, \mathcal{F}) étant compact (ceci suit, par exemple, des théorèmes de Tychonov et de Banach-Alaoglu). Dans cette section, nous allons considérer un modèle très simple, mais intéressant, de système dans lequel les spins prennent valeur dans un espace non-compact, en l'occurrence \mathbb{R} . Nous verrons que le problème de l'existence devient non-trivial, et que les cas de non-existence correspondent à des phénomènes intéressants.

6.1 Le modèle gaussien

Les champs markoviens suivant une loi gaussiennes jouent un rôle important dans les applications, car ils se prêtent à de nombreux calculs explicites. Afin de simplifier l'analyse, nous nous contenterons d'étudier un modèle très simple, généralement appelé **cristal harmonique** ou **champ libre discret**, pour des raisons que nous n'aborderons pas ici. La motivation pour étudier cet exemple particulier de champ markovien est qu'il permet d'étudier de manière détaillée le problème de l'existence d'une mesure en volume infini, et qu'il a suscité de nombreux travaux en théorie des probabilités ces dernières années, dans le cadre des modèles effectifs d'interfaces (*cf.* Section 6.4). Il est également très similaire aux modèles étudiés dans le chapitre précédent.

Une configuration de ce modèle est une application $\omega : \mathbb{Z}^d \rightarrow \mathbb{R}$. L'ensemble des configurations est comme d'habitude noté Ω . À chaque site $i \in \mathbb{Z}^d$, nous attachons une variable aléatoire $\varphi_i : \Omega \rightarrow \mathbb{R}$, $\omega \mapsto \omega_i$.

Soit Λ une partie finie de \mathbb{Z}^d . L'Hamiltonien est défini par

$$\mathcal{H}_\Lambda(\omega) \stackrel{\text{déf}}{=} \frac{\beta}{4d} \sum_{\substack{\{i,j\} \cap \Lambda \neq \emptyset \\ i \sim j}} (\varphi_i(\omega) - \varphi_j(\omega))^2 + \frac{m^2}{2} \sum_{i \in \Lambda} \varphi_i^2,$$

6.2. LA REPRÉSENTATION EN MARCHE ALÉATOIRE

où $\beta \geq 0$ est la température inverse et $m \geq 0$ est la masse.

La mesure de Gibbs dans Λ , avec condition au bord $\bar{\varphi} \in \Omega$, est la mesure de probabilité $\mu_{\Lambda;\beta,m}^{\bar{\varphi}}$ sur (Ω, \mathcal{F}) définie par

$$\mu_{\Lambda;\beta,m}^{\bar{\varphi}}(d\omega) \stackrel{\text{déf}}{=} \begin{cases} \frac{1}{Z_{\Lambda;\beta,m}^{\bar{\varphi}}} e^{-\mathcal{H}_{\Lambda}(\omega)} \ell^{\Lambda}(d\omega) & \text{si } \omega \equiv \bar{\varphi} \text{ hors de } \Lambda, \\ 0 & \text{sinon,} \end{cases}$$

où ℓ est la mesure de Lebesgue sur \mathbb{R} .

Remarque 6.1.1. *On peut se débarrasser du paramètre β en faisant le changement de variables $\varphi \mapsto \varphi/\sqrt{\beta}$; on supposera donc dorénavant que $\beta = 1$ et on l'omettra des notations.*

Ce qui fait l'intérêt de ce modèle, c'est que la mesure $\mu_{\Lambda;m}^{\bar{\varphi}}$ est une mesure gaussienne, et qu'il est donc possible de calculer explicitement un grand nombre de quantités. Nous verrons quelques exemples dans la suite.

Une conséquence du caractère gaussien de ces mesures est qu'il est aisé de vérifier la convergence d'une suite de mesures de Gibbs vers une mesure en volume infini : en effet, celle-ci sera nécessairement gaussienne, et il suffit donc de vérifier la convergence de la moyenne et des covariances, quantités facilement calculables dans ce modèle. Pour voir ça, il suffit d'observer que la marginale de $\mu_{\Lambda;m}^{\bar{\varphi}}$ sur un ensemble $A \subseteq \mathbb{Z}^d$ est une mesure gaussienne sur $\mathbb{R}^{|A|}$.

6.2 La représentation en marche aléatoire

Nous allons à présent dériver des formules très utiles permettant d'exprimer la moyenne et les covariances de $\mu_{\Lambda;m}^{\bar{\varphi}}$ en termes de la fonction de Green de la marche aléatoire simple X sur \mathbb{Z}^d .

Soit Z le processus de Markov sur $\mathbb{Z}^d \cup \{\dagger\}$, où \dagger est un nouvel état (que l'on appellera la mort), dont les probabilités de transition sont données par la matrice

$$P(i,j) \stackrel{\text{déf}}{=} \begin{cases} \frac{1}{1+m^2} \frac{1}{2d} & \text{si } i,j \in \mathbb{Z}^d, i \sim j, \\ 1 - \frac{1}{1+m^2} & \text{si } i \in \mathbb{Z}^d \text{ et } j = \dagger, \\ 1 & \text{si } i = j = \dagger, \\ 0 & \text{sinon.} \end{cases}$$

Manifestement, \dagger est un état absorbant pour ce processus. Remarquez que le processus ne peut mourir (c'est-à-dire commencer en $i \in \mathbb{Z}^d$ et être absorbé en \dagger) que si $m \neq 0$. De façon informelle, on peut dire que ce processus décrit la marche aléatoire simple X , jouant à la roulette russe à chaque pas, avec une probabilité de survie égale à $1/(1+m^2)$. On note \mathbb{E}_i l'espérance du processus Z partant de i au temps 0.

Théorème 6.2.1. *Posons $\bar{\varphi}_\dagger \stackrel{\text{déf}}{=} 0$. Vue comme une mesure sur \mathbb{R}^Λ , la mesure de Gibbs $\mu_{\Lambda;m}^{\bar{\varphi}}$ dans Λ avec condition au bord $\bar{\varphi}$ est une mesure Gaussienne de moyenne $u_i^{\bar{\varphi},\Lambda,m}$ donnée par*

$$u_i^{\bar{\varphi},\Lambda,m} = \mathbb{E}_i(\bar{\varphi}_{Z_{\tau_\Lambda}}),$$

où $\tau_\Lambda = \min\{n \geq 0 : X_n \notin \Lambda\}$, et de matrice de covariance $G^\Lambda(i, j)$ donnée par

$$G^\Lambda(i, j) = (1 + m^2)^{-1} \mathbb{E}_i \left(\sum_{n=0}^{\tau_\Lambda-1} \mathbf{1}_{\{Z_n=j\}} \right).$$

Démonstration. Notons simplement $u = u^{\bar{\varphi},\Lambda,m}$, $\Lambda^c \stackrel{\text{déf}}{=} \mathbb{Z}^d \setminus \Lambda$ et $\mathbb{Z}_\dagger^d \stackrel{\text{déf}}{=} \mathbb{Z}^d \cup \{\dagger\}$. On vérifie aisément que u est l'unique solution de

$$\begin{cases} \sum_{j \in \mathbb{Z}_\dagger^d} (P(i, j) - \delta_{i,j}) u_j = 0 & \forall i \in \Lambda, \\ u_i = \bar{\varphi}_i & \forall i \in \Lambda^c \cup \{\dagger\}. \end{cases}$$

En effet, on a d'une part, par la propriété de Markov que, si $i \in \Lambda$,

$$\sum_{j \in \mathbb{Z}_\dagger^d} P(i, j) \mathbb{E}_j(\bar{\varphi}_{Z_{\tau_\Lambda}}) = \mathbb{E}_i(\bar{\varphi}_{Z_{\tau_\Lambda}}),$$

et d'autre part, si $i \notin \Lambda$, $\mathbb{E}_i(\bar{\varphi}_{Z_{\tau_\Lambda}}) = \bar{\varphi}_i$ puisque dans ce cas $\tau_\Lambda = 0$.

Pour toute configuration ω telle que $\varphi(\omega) \equiv \bar{\varphi}$ hors de Λ , on peut réécrire

$$\mathcal{H}_\Lambda(\omega) = \frac{1+m^2}{2} \sum_{i,j \in \Lambda} \varphi_i(\omega) (\delta_{ij} - P(i, j)) \varphi_j(\omega) - \sum_{i \in \Lambda} \varphi_i(\omega) \frac{1}{2d} \sum_{\substack{j \notin \Lambda \\ i \sim j}} \bar{\varphi}_j + C(\bar{\varphi}),$$

où $C(\bar{\varphi})$ est indépendant de ω . Nous allons exprimer la seconde somme du membre de droite en termes de la fonction u . Puisque, pour tout $i \in \Lambda$,

$$0 = \sum_{j \in \mathbb{Z}_\dagger^d} (P(i, j) - \delta_{i,j}) u_j = \sum_{j \sim i} P(i, j) u_j - u_i = \frac{1}{1+m^2} \frac{1}{2d} \sum_{j \sim i} u_j - u_i,$$

on obtient la relation

$$\frac{1}{2d} \sum_{\substack{j \notin \Lambda \\ j \sim i}} \bar{\varphi}_j = \frac{1}{2d} \sum_{\substack{j \notin \Lambda \\ j \sim i}} u_j = (1+m^2) u_i - \sum_{\substack{j \in \Lambda \\ j \sim i}} \frac{1}{2d} u_j = (1+m^2) \sum_{j \in \Lambda} (\delta_{ij} - P(i, j)) u_j.$$

Par conséquent, l'Hamiltonien peut se réécrire sous la forme

$$\mathcal{H}_\Lambda(\omega) = \frac{1+m^2}{2} \sum_{\substack{i,j \in \Lambda \\ i \sim j}} (\varphi_i(\omega) - u_i) (\delta_{ij} - P(i, j)) (\varphi_j(\omega) - u_j) + C(\bar{\varphi}),$$

le dernier terme étant indépendant de ω , et n'intervenant donc pas dans la mesure de Gibbs. On en déduit que cette dernière mesure (vue comme une mesure sur \mathbb{R}^Λ) est bien une mesure gaussienne de moyenne u , et de matrice de covariance donnée par l'inverse de la matrice $(1 + m^2)(\mathcal{I}_\Lambda - \mathcal{P}_\Lambda)$, où \mathcal{I}_Λ est l'identité et les éléments de matrices de \mathcal{P}_Λ sont donnés par $P(i, j)$, $i, j \in \Lambda$. Il reste à vérifier que l'expression annoncée dans l'énoncé du théorème correspond bien à cet inverse. Mais cela suit immédiatement de la réinterprétation de la série de Neumann,

$$(\mathcal{I}_\Lambda - \mathcal{P}_\Lambda)^{-1} = \sum_{k \geq 0} \mathcal{P}_\Lambda^k,$$

puisque

$$(\mathcal{P}_\Lambda^k)_{ij} = \mathbb{P}_i(X_k = j, \tau_\Lambda > k).$$

□

6.3 Mesures en volume infini

Grâce au Théorème 6.2.1, il est possible de caractériser complètement l'ensemble des mesures de Gibbs en volume infini. Nous ne le ferons pas ici, et nous contenterons d'étudier quelques cas particuliers, notre principal intérêt étant d'exhiber des conditions sous lesquelles aucune limite n'existe.

6.3.1 Le cas massif

Nous nous intéressons tout d'abord au cas massif, $m > 0$. Dans ce cas, le modèle se comporte essentiellement comme un modèle avec spin compact. On a en particulier le théorème suivant.

Théorème 6.3.1. *Soit $\Lambda_k \nearrow \mathbb{Z}^d$ et $(\bar{\varphi}_k)_{k \geq 1}$ une suite de conditions au bord bornées. Alors, toutes les suites $(\mu_{\Lambda_k; m}^{\bar{\varphi}_k})_{k \geq 1}$ de ce type convergent vers la même limite μ_m . De plus,*

$$\langle \varphi_i \rangle_m = 0, \quad \langle \varphi_i \varphi_j \rangle_m \leq c_1 e^{-c_2 \|j-i\|},$$

pour des constantes $c_1(m)$ et $c_2(m)$, uniformément en $i, j \in \mathbb{Z}^d$.

Remarque 6.3.1. *Si on permet des conditions au bord non-bornées, alors on peut montrer qu'il existe un ensemble non-dénombrable de mesures de Gibbs en volume infini.*

Remarque 6.3.2. *En travaillant plus dur, le comportement asymptotique des covariances peut être déterminé précisément. En particulier, on obtient que $c_2(m) = m + O(m^2)$ lorsque la masse m est petite. Ceci est la raison de l'introduction de la présence du carré dans l'Hamiltonien.*

Démonstration. Vérifions tout d'abord que pour tout sous-ensemble $A \Subset \mathbb{Z}^d$, la moyenne et les covariances de la marginale de $\mu_{\Lambda_k; m}^{\bar{\varphi}_k}$ sur A convergent vers une limite indépendante des

conditions au bord et de la suite Λ_k . Soit $M = \sup_{j \in \mathbb{Z}^d} |\bar{\varphi}_k| < \infty$. D'après le théorème 6.2.1, on a

$$\begin{aligned} \left| \langle \varphi_i \rangle_{\Lambda_k; m}^{\bar{\varphi}_k} \right| &= |\mathbb{E}_i(\bar{\varphi}_k(X_{\tau_\Lambda}))| \\ &\leq M \mathbb{P}_i(\tau_\Lambda > d(\{i\}, \Lambda_k^c)) \\ &\leq M(1 + m^2)^{-d(\{i\}, \Lambda_k^c)}, \end{aligned}$$

et cette dernière expression tend vers 0 lorsque $k \rightarrow \infty$. On a utilisé ici le fait que la marche doit survivre un temps au moins égal à $d(\{i\}, \Lambda_k^c)$ pour sortir de Λ_k en partant de i .

Similairement,

$$\langle \varphi_i \varphi_j \rangle_{\Lambda_k; m}^{\bar{\varphi}_k} = (1 + m^2)^{-1} \sum_{n \geq 0} \mathbb{P}_i(Z_n = j, \tau_{\Lambda_k} > n) \geq 0.$$

On en déduit que $\langle \varphi_i \varphi_j \rangle_{\Lambda_k; m}^{\bar{\varphi}_k}$ est indépendant de $\bar{\varphi}$ et croissant en Λ_k . Il suffit donc de vérifier que la limite est finie; on va vérifier qu'elle est non seulement finie, mais exponentiellement décroissante en $\|j - i\|$. En introduisant la variable aléatoire $\Upsilon_{\dagger} \stackrel{\text{déf}}{=} \min\{n \geq 0 : Z_n = \dagger\}$ ("heure du décès"), on peut écrire la limite sous la forme

$$\begin{aligned} \sum_{n \geq 0} \mathbb{P}_i(Z_n = j, \Upsilon_{\dagger} > n) &\leq \sum_{n \geq \|j - i\|} \mathbb{P}_i(\Upsilon_{\dagger} > n) \\ &\leq \sum_{n \geq \|j - i\|} n \mathbb{P}_i(\Upsilon_{\dagger} = n) \\ &\leq (1 + m^2)^{-\|j - i\|} \sum_{n \geq 1} n (1 + m^2)^{-n}. \end{aligned}$$

□

6.3.2 Le cas non-massif

Nous allons maintenant considérer le cas non-massif, $m = 0$. Dans ce cas, comme on l'a déjà mentionné, le processus Z ne meurt plus et se réduit donc à la marche aléatoire simple X . Le comportement de ce modèle est beaucoup plus intéressant dans ce régime.

Théorème 6.3.2. *Il existe des constantes $c(d) > 0$ telles que la variance du champ dans une boîte $\Lambda = \{-N, \dots, N\}^d$ satisfasse*

$$\langle \varphi_i^2 \rangle_{\Lambda; 0}^{\bar{\varphi}} - \left(\langle \varphi_i \rangle_{\Lambda; 0}^{\bar{\varphi}} \right)^2 = \begin{cases} (c(1) + o(1))N & \text{si } d = 1 \\ (c(2) + o(1)) \log N & \text{si } d = 2 \\ c(d) + o(1) & \text{si } d \geq 3, \end{cases}$$

lorsque $N \rightarrow \infty$, indépendamment de la condition au bord $\bar{\varphi}$. En particulier, il n'y a pas de mesures de Gibbs en volume infini lorsque $d = 1$ ou 2.

6.3. MESURES EN VOLUME INFINI

En plus hautes dimensions, $d \geq 3$, il y a un ensemble non-dénombrable de mesures de Gibbs, même pour des conditions au bord bornées. De plus, pour toute mesure limite μ ,

$$\langle \varphi_i \varphi_j \rangle_\mu - \langle \varphi_i \rangle_\mu \langle \varphi_j \rangle_\mu = (r(d) + o(1)) \|j - i\|^{2-d},$$

lorsque $\|j - i\| \rightarrow \infty$.

Démonstration. L'indépendance des (co)variances de la condition au bord $\bar{\varphi}$ suit directement de la représentation en marche aléatoire. L'existence d'un ensemble non-dénombrable de mesures de Gibbs en volume infini lorsque $d \geq 3$ est évidente, puisque chaque condition au bord de la forme $\bar{\varphi} \equiv z$, $z \in \mathbb{R}$, donne lieu à une mesure limite différente, de moyenne $\mu_m(\varphi_i) = z$, pour tout i . Ces mesures ne forment bien sûr qu'une partie des mesures de Gibbs en volume infini : il y en a beaucoup d'autres lorsque $\bar{\varphi}$ n'est pas supposée bornée.

Les comportements asymptotiques donnés ici suivent d'une analyse détaillée de la fonction de Green de la marche aléatoire X , que l'on peut trouver dans de nombreux ouvrages, par exemple [28] et [17]. Nous ne le ferons pas ici. \square

La divergence des variances en dimensions 1 et 2 peut aussi être vue comme résultant de la symétrie continue de l'Hamiltonien : on voit en effet que l'Hamiltonien ne dépend que des gradients du champ (c'est pourquoi on parle parfois pour les modèles de ce type de **modèles gradients**). Par conséquent, en introduisant, pour $t \in \mathbb{R}$, la transformation $v_t : \Omega \rightarrow \Omega$, $(v_t \omega)_i = \omega_i - t$, on voit que

$$\mathcal{H}_\Lambda(v_t \omega) = \mathcal{H}_\Lambda(\omega), \quad \forall t \in \mathbb{R}.$$

On a montré dans le chapitre 5 qu'une symétrie continue ne pouvait pas être brisée lorsque $d = 1$ ou 2 . Bien sûr, les résultats du chapitre 5 ne s'appliquent pas directement ici, puisque le groupe de symétrie n'est plus compact (il est isomorphe à \mathbb{R}). Cependant, on va voir qu'il est possible d'adapter la preuve du Théorème 5.3.1 de façon à obtenir des bornes inférieures pour la variance qui ont la bonne divergence, et sont valides beaucoup plus généralement. On considère la classe de mesures définies comme précédemment, mais à partir d'un Hamiltonien de la forme

$$\mathcal{H}_\Lambda(\omega) = \sum_{\substack{\{i,j\} \cap \Lambda \neq \emptyset \\ i \sim j}} V(\varphi_j(\omega) - \varphi_i(\omega)),$$

avec $V : \mathbb{R} \rightarrow \mathbb{R}$ paire, et telle que $V(0) = 0$. Le modèle gaussien correspond ainsi au cas $V(x) = \frac{1}{4d}x^2$.

Théorème 6.3.3. *On considère la classe de modèles introduits ci-dessus, avec V de classe C^2 , satisfaisant $V(0) = 0$ et $\sup_{x \in \mathbb{R}} V''(x) < \bar{c} < \infty$. Soit $d = 1$ ou 2 , $\Lambda_n = \{-n, \dots, n\}^d$, et posons $T_n(1) = \sqrt{n}$ et $T_n(2) = \sqrt{\log n}$. Il existe une constante c telle que, pour tout $\alpha > 0$,*

$$\mu_{\Lambda_n}^{\bar{\varphi}} \left(\left| \varphi_0 - \mu_{\Lambda_n,0}^{\bar{\varphi}}(\varphi_0) \right| > \alpha T_n(d) \right) \geq \frac{1}{c} e^{-c\alpha^2}.$$

Démonstration. Comme annoncé, la preuve est très proche de celle du Théorème 5.3.1. Quitte à traduire la condition au bord, on peut supposer que $\langle \varphi_0 \rangle_{\Lambda_n;0}^{\bar{\varphi}} = 0$. De plus, on peut supposer, sans perte de généralité, que $\mu_{\Lambda_n;0}^{\bar{\varphi}}(\varphi_0 > 0) \geq \frac{1}{2}$. Soit $\Psi : \mathbb{Z}^d \rightarrow \mathbb{R}$ tel que $\Psi \equiv 0$ hors de Λ_n , et $\Psi_0 = \alpha T_n(d)$, et soit $\mathbf{v}_\Psi : \Omega \rightarrow \Omega$, $(\mathbf{v}_\Psi \omega)_i = \omega_i - \Psi_i$. On introduit une mesure déformée définie de la façon suivante

$$\mu_{\Lambda_n;0}^{\bar{\varphi};\Psi}(\mathcal{A}) \stackrel{\text{déf}}{=} \mu_{\Lambda_n;0}^{\bar{\varphi}}(\mathbf{v}_\Psi \mathcal{A}), \quad \forall \mathcal{A} \in \mathcal{F}.$$

On utilise à présent, en remplacement de l'inégalité de Pinsker, l'inégalité d'entropie suivante.

Lemme 6.3.1. *Soit μ et ν deux mesures de probabilité sur un même espace, absolument continues l'une par rapport à l'autre, et soit \mathcal{E} un événement tel que $\nu(\mathcal{E}) > 0$. Alors,*

$$\mu(\mathcal{E}) \geq \exp \left(- \frac{\mathbf{H}(\nu | \mu) + e^{-1}}{\nu(\mathcal{E})} \right) \nu(\mathcal{E}).$$

Démonstration du Lemme 6.3.1. En utilisant l'inégalité de Jensen, et l'inégalité $x \log x \geq -e^{-1}$, valide pour tout $x > 0$, on peut écrire

$$\log \frac{\mu(\mathcal{E})}{\nu(\mathcal{E})} = \log \nu \left(\frac{d\mu}{d\nu} \middle| \mathcal{E} \right) \geq \nu \left(\log \frac{d\mu}{d\nu} \middle| \mathcal{E} \right) = - \frac{\mu \left(\frac{d\nu}{d\mu} \log \frac{d\nu}{d\mu} \mathbf{1}_{\mathcal{E}} \right)}{\nu(\mathcal{E})} \geq - \frac{\mu \left(\frac{d\nu}{d\mu} \log \frac{d\nu}{d\mu} \right) + e^{-1}}{\nu(\mathcal{E})}.$$

□

Le lemme précédent implique dans notre cas que

$$\mu_{\Lambda_n;0}^{\bar{\varphi}}(|\varphi_0| > \alpha T_n(d)) \geq \frac{1}{2} \exp \left\{ -2 \left(\mathbf{H} \left(\mu_{\Lambda_n;0}^{\bar{\varphi};\Psi} \middle| \mu_{\Lambda_n;0}^{\bar{\varphi}} \right) + e^{-1} \right) \right\},$$

puisque $\{|\varphi_0| > \alpha T_n(d)\} \supset \{\varphi_0 > \Psi_0\}$, et donc

$$\mu_{\Lambda_n;0}^{\bar{\varphi};\Psi}(|\varphi_0| > \alpha T_n(d)) = \mu_{\Lambda_n;0}^{\bar{\varphi}}(|\varphi_0| > 0) \geq \frac{1}{2}.$$

La preuve se ramène donc, comme dans le cas du Théorème 5.3.1, à un contrôle de l'entropie relative. En procédant exactement de la même façon que lors de la démonstration de ce théorème, on obtient

$$\mathbf{H} \left(\mu_{\Lambda_n;0}^{\bar{\varphi};\Psi} \middle| \mu_{\Lambda_n;0}^{\bar{\varphi}} \right) \leq c \sum_{\substack{\{i,j\} \cap \Lambda_n \neq \emptyset \\ i \sim j}} (\Psi_j - \Psi_i)^2,$$

et donc finalement, en optimisant sur Ψ ,

$$\mathbf{H} \left(\mu_{\Lambda_n;0}^{\bar{\varphi};\Psi} \middle| \mu_{\Lambda_n;0}^{\bar{\varphi}} \right) \leq c' \alpha^2 T_n(d)^2 \mathbb{P}_0(X \text{ quitte } \Lambda_n \text{ sans revisiter } 0).$$

Cette dernière probabilité peut être estimée (par exemple en utilisant le fait qu'elle est égale à l'inverse de $G_{\Lambda_n}(0,0) = \sum_{k \geq 0} \mathbb{P}_0(X_k = 0, \tau_{\Lambda_n} > k)$, ce qui est une conséquence immédiate de la propriété de Markov, et des estimations standards de cette fonction de Green [28, 17]); on obtient ainsi qu'elle est d'ordre $T_n(d)^{-2}$, ce qui démontre le théorème. □

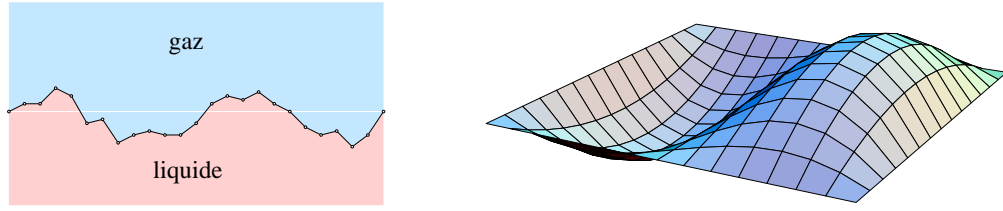


FIGURE 6.1 – Interprétation des modèles gradients comme modèles de surfaces aléatoires.

6.4 Modèles effectifs de surfaces aléatoires

Le théorème précédent montre que l'absence de mesures de Gibbs en volume infini en dimensions 1 et 2 n'est pas limitée au modèle Gaussien, mais est un phénomène commun.

Une interprétation de ces modèles gradients est comme modèles de surfaces aléatoires (discrétisées) d -dimensionnelles dans un espace $(d + 1)$ -dimensionnel : on interprète simplement φ_i comme la hauteur de la surface au-dessus du site i ; en d'autres termes, on voit le graphe de la fonction aléatoire $i \mapsto \varphi_i(\omega)$ comme une surface aléatoire, *cf.* Fig. 6.1.

L'interprétation des résultats sur les modèles non-massifs est qu'une surface aléatoire a des fluctuations non-bornées dans la limite thermodynamique en basse dimension ($d = 1, 2$), alors que celles-ci restent bornées uniformément dans la taille du système en haute dimension ($d \geq 3$).

Positivité sous réflexion

Dans ce chapitre, nous allons étudier une classe intéressante de champs markoviens : ceux qui possèdent la propriété de positivité sous réflexion. Cette dernière donne accès à certains outils particulièrement puissants, en particulier la borne infrarouge et l'estimation de l'échiquier. Nous verrons ensuite comment appliquer ces outils pour démontrer divers résultats. Cette technique est apparue tout d'abord en théorie des champs, puis a été introduite dans le contexte présent à la fin des années 1970. Pour plus de détails sur ce sujet, nous référons le lecteur aux excellents articles de revue [25] et [3], ainsi qu'aux livres [8] et [23].

Le cadre le plus naturel pour ce chapitre est celui des champs aléatoires sur le tore $\mathbb{T}_L^d = (\mathbb{Z}/L\mathbb{Z})^d$ (ce qui revient à considérer la boîte $\{0, \dots, L-1\}^d$ avec condition au bord périodique) à valeur dans un espace topologique S ; on supposera toujours que L est pair. Ce tore est invariant sous réflexion par rapport à tout plan Π orthogonal aux axes de coordonnées et (*cf.* Fig. 7.1)

- soit passant par des sommets de \mathbb{T}_L^d (on parlera de *réflexion à travers les sommets*),
- soit passant par le milieu des arêtes joignant deux sommets voisins (on parlera de *réflexion à travers les arêtes*).

Chaque tel plan Π partage le tore en deux moitiés $\mathbb{T}_{L,+}^d = \mathbb{T}_{L,+}^d(\Pi)$ et $\mathbb{T}_{L,-}^d = \mathbb{T}_{L,-}^d(\Pi)$, images l'une de l'autre sous la réflexion. Dans le cas d'une réflexion à travers les arêtes, $\mathbb{T}_{L,+}^d \cap \mathbb{T}_{L,-}^d = \emptyset$, tandis que dans le cas d'une réflexion à travers les sommets, $\mathbb{T}_{L,+}^d \cap \mathbb{T}_{L,-}^d$ contient les sommets du tore situés sur le plan. On note $\mathfrak{A}_+ = \mathfrak{A}_+(\Pi)$ l'algèbre des fonctions $f : S^{\mathbb{T}_L^d} \rightarrow \mathbb{R}$ ne dépendant que des spins situés dans $\mathbb{T}_{L,+}^d$, et $\mathfrak{A}_- = \mathfrak{A}_-(\Pi)$ l'algèbre de celles ne dépendant que des spins situés dans $\mathbb{T}_{L,-}^d$.

Soit $\Theta : \mathbb{T}_L^d \rightarrow \mathbb{T}_L^d$ l'application associant à un sommet son image par la réflexion de plan Π . On note également $\Theta : \mathfrak{A}_+ \rightarrow \mathfrak{A}_-$ l'opérateur de réflexion défini par $\Theta(f)(\omega) = f(\Theta(\omega))$, où $(\Theta(\omega))_i = \omega_{\Theta(i)}$.

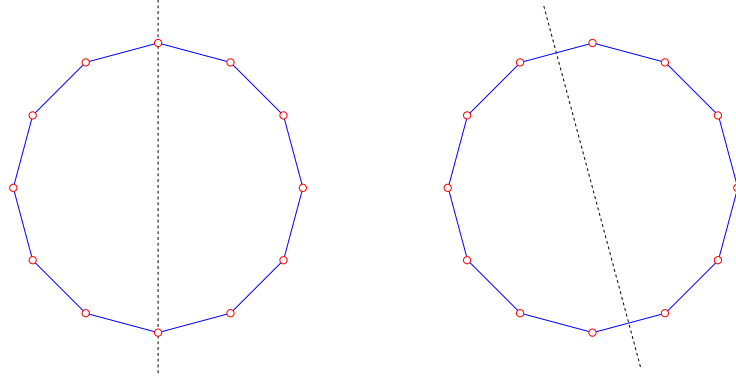


FIGURE 7.1 – Le tore \mathbb{T}_{12}^1 . Gauche : réflexion à travers les sommets. Droite : réflexion à travers les arêtes.

7.1 Mesures positives sous réflexion

7.1.1 Définition et propriétés élémentaires

Soit $\Omega_L \stackrel{\text{déf}}{=} \mathcal{S}^{\mathbb{T}_L^d}$, et \mathcal{F}_L la tribu-produit associée.

Une mesure $\mu \in \mathcal{M}(\Omega_L, \mathcal{F}_L)$ est **positive sous réflexion** par rapport à Θ si

1. $\langle f \Theta(g) \rangle_\mu = \langle g \Theta(f) \rangle_\mu$, pour tout $f, g \in \mathfrak{A}_+$;
2. $\langle f \Theta(f) \rangle_\mu \geq 0$, pour tout $f \in \mathfrak{A}_+$.

L'ensemble des mesures sur le tore positives sous réflexion par rapport à Θ est noté $\mathcal{RP}(\Theta)$. Cette définition garantit que l'application $(f, g) \mapsto \langle f \Theta(g) \rangle_\mu$ est une forme bilinéaire positive, lorsque μ est positive sous réflexion. Ceci implique la validité d'une inégalité de Cauchy-Schwarz.

Lemme 7.1.1. *Soit $\mu \in \mathcal{RP}(\Theta)$. Pour tout $f, g \in \mathfrak{A}_+$,*

$$|\langle f \Theta(g) \rangle_\mu| \leq \sqrt{\langle f \Theta(f) \rangle_\mu \langle g \Theta(g) \rangle_\mu}.$$

Démonstration. Puisque, pour tout $x \in \mathbb{R}$,

$$\langle f \Theta(f) \rangle_\mu x^2 + 2\langle f \Theta(g) \rangle_\mu x + \langle g \Theta(g) \rangle_\mu = \langle (x f + g) \Theta(x f + g) \rangle_\mu \geq 0,$$

le discriminant doit être négatif ou nul, ce qui prouve le résultat. \square

Le résultat suivant peut être vu comme une version améliorée de l'inégalité précédente ; il se révèle très utile dans la pratique.

Lemme 7.1.2. *Soit $\mu \in \mathcal{RP}(\Theta)$, et $A, B, C_\alpha, D_\alpha \in \mathfrak{A}_+$. Alors*

$$\{\langle e^{A+\Theta(B)+\sum_\alpha C_\alpha \Theta(D_\alpha)} \rangle_\mu\}^2 \leq \langle e^{A+\Theta(A)+\sum_\alpha C_\alpha \Theta(C_\alpha)} \rangle_\mu \langle e^{B+\Theta(B)+\sum_\alpha D_\alpha \Theta(D_\alpha)} \rangle_\mu.$$

Démonstration. En développant $e^{\sum_{\alpha} C_{\alpha} \Theta(D_{\alpha})}$ en série, on obtient

$$\langle e^{A+\Theta(B)+\sum_{\alpha} C_{\alpha} \Theta(D_{\alpha})} \rangle_{\mu} = \sum_{n \geq 0} \frac{1}{n!} \sum_{\alpha_1, \dots, \alpha_n} \langle e^A C_{\alpha_1} \dots C_{\alpha_n} \Theta(e^B D_{\alpha_1} \dots D_{\alpha_n}) \rangle_{\mu}$$

Le lemme 7.1.1 implique que

$$\begin{aligned} & \langle e^A C_{\alpha_1} \dots C_{\alpha_n} \Theta(e^B D_{\alpha_1} \dots D_{\alpha_n}) \rangle_{\mu} \\ & \leq [\langle e^A C_{\alpha_1} \dots C_{\alpha_n} \Theta(e^A C_{\alpha_1} \dots C_{\alpha_n}) \rangle_{\mu} \langle e^B D_{\alpha_1} \dots D_{\alpha_n} \Theta(e^B D_{\alpha_1} \dots D_{\alpha_n}) \rangle_{\mu}]^{1/2}. \end{aligned}$$

En insérant cette expression, et en appliquant l'inégalité de Cauchy-Schwarz pour la somme, on voit que

$$\begin{aligned} \langle e^{A+\Theta(B)+\sum_{\alpha} C_{\alpha} \Theta(D_{\alpha})} \rangle_{\mu} & \leq \left[\sum_{n \geq 0} \frac{1}{n!} \sum_{\alpha_1, \dots, \alpha_n} \langle e^A C_{\alpha_1} \dots C_{\alpha_n} \Theta(e^A C_{\alpha_1} \dots C_{\alpha_n}) \rangle_{\mu} \right]^{1/2} \\ & \quad \times \left[\sum_{n \geq 0} \frac{1}{n!} \sum_{\alpha_1, \dots, \alpha_n} \langle e^B D_{\alpha_1} \dots D_{\alpha_n} \Theta(e^B D_{\alpha_1} \dots D_{\alpha_n}) \rangle_{\mu} \right]^{1/2} \\ & = [\langle e^{A+\Theta(A)+\sum_{\alpha} C_{\alpha} \Theta(C_{\alpha})} \rangle_{\mu} \langle e^{B+\Theta(B)+\sum_{\alpha} D_{\alpha} \Theta(D_{\alpha})} \rangle_{\mu}]^{1/2}. \end{aligned}$$

□

7.1.2 Construction de mesures positives sous réflexion

Il convient à présent de construire des mesures possédant la propriété de positivité sous réflexion. Le premier exemple est trivial, mais se révélera utile pour la construction de champs plus intéressants.

Lemme 7.1.3. *Soit ρ une mesure sur S . Alors la mesure produit $\mu_0 = \bigotimes_{i \in \mathbb{T}_L^d} \rho \in \mathcal{RP}(\Theta)$, pour toutes les réflexions Θ du tore.*

Démonstration. Commençons par traiter les réflexions à travers les arêtes. Soient $f, g \in \mathfrak{A}_+$. Dans ce cas, puisque $\mathbb{T}_{L,+}^d \cap \mathbb{T}_{L,-}^d = \emptyset$, les variables aléatoires f et $\Theta(g)$ sont indépendantes sous μ_0 . Par conséquent,

$$\langle f \Theta(g) \rangle_{\mu_0} = \langle f \rangle_{\mu_0} \langle \Theta(g) \rangle_{\mu_0} = \langle f \rangle_{\mu_0} \langle g \rangle_{\mu_0},$$

ce qui implique immédiatement la positivité sous réflexion.

Le cas des réflexions à travers les sommets est similaire, mais le support de f et $\theta(g)$ pouvant à présent s'intersecter, il faut procéder un peu différemment. Soit P l'ensemble des sites de \mathbb{T}_L^d par lesquels le plan passe, et $\omega|_P = (\omega_i)_{i \in P}$ la restriction de la configuration à P . Alors

$$\langle f \Theta(g) | \omega|_P \rangle_{\mu_0} = \langle f | \omega|_P \rangle_{\mu_0} \langle \Theta(g) | \omega|_P \rangle_{\mu_0} = \langle f | \omega|_P \rangle_{\mu_0} \langle g | \omega|_P \rangle_{\mu_0},$$

et la positivité sous réflexion suit à nouveau. □

7.2. L'ESTIMATION DE L'ÉCHIQUIER

Lemme 7.1.4. *On fixe un plan de réflexion Π et l'opérateur Θ correspondant. Soit ρ une mesure sur S . Si μ est une mesure de Gibbs sur \mathbb{T}_L^d ,*

$$\mu(d\omega) = \frac{1}{Z_L} e^{-\mathcal{H}_L(\omega)} \prod_{i \in \mathbb{T}_L^d} \rho(d\omega_i),$$

pour un Hamiltonien de la forme

$$-\mathcal{H}_L = A + \Theta(A) + \sum_{\alpha} C_{\alpha} \Theta(C_{\alpha}),$$

avec $A, C_{\alpha} \in \mathfrak{A}_+$, alors $\mu \in \mathcal{RP}(\Theta)$.

Démonstration. Notons $\mu_0 = \bigotimes_{i \in \mathbb{T}_L^d} \rho$. En développant l'exponentielle,

$$\begin{aligned} \langle f \Theta(g) \rangle_{\mu} &= \frac{1}{Z_L} \langle f \Theta(g) e^{A + \Theta(A) + \sum_{\alpha} C_{\alpha} \Theta(C_{\alpha})} \rangle_{\mu_0} \\ &= \frac{1}{Z_L} \sum_{n \geq 0} \frac{1}{n!} \sum_{\alpha_1, \dots, \alpha_n} \langle f e^A C_{\alpha_1} \cdots C_{\alpha_n} \Theta(g e^A C_{\alpha_1} \cdots C_{\alpha_n}) \rangle_{\mu_0}. \end{aligned}$$

Le résultat suit de la positivité sous réflexion de la mesure produit μ_0 . \square

Remarque 7.1.1. *On vérifie aisément que les mesures de Gibbs sur \mathbb{T}_L^d avec interaction plus-proches-voisins satisfont les hypothèses du lemme précédent, et sont donc positives sous réflexion, pour toute réflexion Θ .*

Remarque 7.1.2. *On n'a considéré que les réflexions orthogonales aux axes de coordonnées, mais on aurait pu également traiter le cas de réflexion à travers les diagonales du tore (à travers les sommets). Celles-ci laissent évidemment le tore invariant (pour cela il est important que le tore considéré soit "carré"). Tout ce qui sera discuté par la suite reste vrai dans le cas de mesures positives sous de telles réflexions. En particulier, c'est le cas des mesures avec interaction plus-proches-voisins.*

7.2 L'estimation de l'échiquier

Dans cette section, nous allons démontrer une première conséquence particulièrement utile de la positivité sous réflexion : l'estimation de l'échiquier (*chessboard estimate*, en anglais). Nous nous contenterons de traiter le cas des réflexions à travers les arêtes ; l'énoncé et la preuve dans le cas de réflexions à travers les sommets est immédiate.

Soient $B < L$ deux entiers tels que $2B$ divise L , et $\Lambda_B = \{0, 1, \dots, B-1\}^d \subset \mathbb{T}_L^d$. Par choix de B et L , on peut décomposer \mathbb{T}_L^d en union disjointe de translatés de Λ_B ¹,

$$\mathbb{T}_L^d = \bigvee_{t \in \mathbb{T}_{L/B}^d} (\Lambda_B + Bt).$$

1. Dans le cas de réflexions à travers les sommets, deux translatés voisins ne sont plus disjoints, mais se partagent les sommets appartenant au plan de réflexion les envoyant l'un sur l'autre.

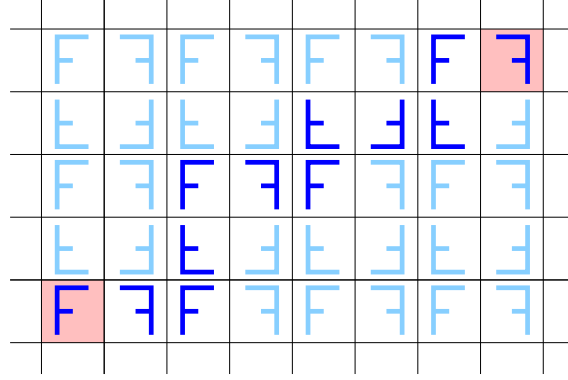


FIGURE 7.2 – La fonction $f^{[t]}$ est définie par une succession de réflexions. La fonction f se trouve ici en bas à droite et la fonction $f^{[t]}$ en haut à gauche (cases marquées). Observez que la définition ne dépend pas du chemin choisi (les F bleus foncés).

On dira qu'une fonction $f : \Omega_L \rightarrow \mathbb{R}$ est une **fonction B -bloc** si elle ne dépend que de la restriction de la configuration au bloc Λ_B . Similairement un événement sera **B -bloc** si son indicatrice est une fonction B -bloc.

Étant donné une fonction B -bloc f et $t \in \mathbb{T}_{L/B}^d$, on définit la fonction $f^{[t]}$ par réflexions successives. Soit $t_0 = 0, t_1, \dots, t_k = t$ un chemin dans $\mathbb{T}_{L/B}^d$ et soit Θ_i l'opérateur correspondant à la réflexion à travers le plan Π_i passant par le milieu des arêtes joignant les blocs $\Lambda_B + t_{i-1}B$ et $\Lambda_B + t_iB$. On pose alors (voir Fig. 7.2)

$$f^{[t]} = \Theta_k \circ \Theta_{k-1} \circ \dots \circ \Theta_1(f).$$

On vérifie facilement que cette définition est indépendante du chemin choisi (ici, il est important que L/B soit pair). Par construction, la fonction $f^{[t]}$ ne dépend que de la restriction de la configuration au bloc $\Lambda_B + tB$.

Théorème 7.2.1 (Estimation de l'échiquier). *Soit $\mu \in \mathcal{RP}(\Theta)$ pour toutes les réflexions Θ entre blocs de la forme $\Lambda_B + tB$, $t \in \mathbb{T}_{L/B}^d$. Alors, pour toute famille f_t , $t \in \mathbb{T}_{L/B}^d$ de fonctions B -blocs,*

$$|\langle \prod_{t \in \mathbb{T}_{L/B}^d} f_t^{[t]} \rangle_\mu| \leq \prod_{t \in \mathbb{T}_{L/B}^d} \left[\langle \prod_{s \in \mathbb{T}_{L/B}^d} f_t^{[s]} \rangle_\mu \right]^{1/|\mathbb{T}_{L/B}^d|}.$$

Remarque 7.2.1. *Le résultat reste vrai si le bloc Λ_B est remplacé par un parallélogramme rectangle dont le double de la longueur de chaque côté divise L .*

Démonstration. On commence par le cas unidimensionnel. On pose $N = L/(2B) \in \mathbb{N}$, et on introduit la fonctionnelle multilinéaire suivante sur l'ensemble des fonctions B -blocs :

$$F(f_1, \dots, f_{2N}) = \langle \prod_{t=1}^{2N} f_t^{[t]} \rangle_\mu.$$

Avec cette notation, l'inégalité à démontrer devient

$$|F(f_1, \dots, f_{2N})| \leq \prod_{t=1}^{2N} F(f_t, \dots, f_t)^{1/2N}. \quad (7.1)$$

Cette fonctionnelle possède les propriétés suivantes :

$$F(f_1, f_2, \dots, f_{2N}) = F(f_2, f_3, \dots, f_{2N}, f_1), \quad (7.2)$$

et, d'après le Lemme 7.1.1,

$$\begin{aligned} & F(f_1, \dots, f_N, f_{N+1}, \dots, f_{2N})^2 \\ & \leq F(f_1, \dots, f_N, f_N, \dots, f_1) F(f_{2N}, \dots, f_{N+1}, f_{N+1}, \dots, f_{2N}). \end{aligned} \quad (7.3)$$

Comme chacune des espérance dans le membre de droite de (7.1) est positive (elle peut être mise sous la forme $\langle g \Theta(g) \rangle$ pour une fonction g et une réflexion Θ appropriée), il n'y a rien à démontrer lorsque $F(f_1, \dots, f_{2N}) = 0$. Supposons donc que $|F(f_1, \dots, f_{2N})| > 0$. Vérifions tout d'abord que cela implique que $F(f_t, \dots, f_t) > 0$ pour tout $t = 1, \dots, 2N$. Par la propriété (7.2), il suffit de montrer que $F(f_N, \dots, f_N) > 0$. En utilisant à plusieurs reprises et alternativement les deux propriétés ci-dessus, on obtient

$$\begin{aligned} |F(f_1, \dots, f_{2N})| & \leq C_1 F(f_1, \dots, f_N, f_N, \dots, f_1)^{1/2} \\ & = C_1 F(f_2, \dots, f_N, f_N, \dots, f_1, f_1)^{1/2} \\ & \leq C_2 F(f_2, \dots, f_N, f_N, f_N, f_N, \dots, f_2)^{1/4} \\ & = C_2 F(f_4, \dots, f_N, f_N, f_N, f_N, \dots, f_2, f_2, f_3, f_3)^{1/4} \\ & \leq C_3 F(f_4, \dots, f_N, f_N, f_N, f_N, f_N, f_N, f_N, f_N, \dots, f_4)^{1/8} \\ & \leq \dots \\ & \leq C_M F(f_N, \dots, f_N)^{2^{-M}}, \end{aligned}$$

pour certaines fonctions positives C_i de f_1, \dots, f_{2N} et M le plus petit entier tel que $2^M \geq 2N$. On a donc bien $F(f_N, \dots, f_N) > 0$. Il est par conséquent possible de définir une autre fonctionnelle

$$G(f_1, \dots, f_{2N}) = \frac{|F(f_1, \dots, f_{2N})|}{\prod_{t=1}^{2N} F(f_t, \dots, f_t)^{1/2N}}.$$

On vérifie immédiatement que G hérite de l'analogie des propriétés (7.2) et (7.3). De plus, $G(f, \dots, f) = 1$ pour toute fonction B -bloc f .

Le théorème sera démontré si l'on parvient à établir que $G(f_1, \dots, f_{2N}) \leq 1$, ce qui revient à dire que G est maximisée pour les $2N$ -tuples composés d'une même fonction.

Étant donné un $2N$ -tuple (f_1, \dots, f_{2N}) , soit (g_1, \dots, g_{2N}) tel que

- (i) $g_i \in \{f_1, \dots, f_{2N}\}$ pour chaque $i \in \{1, \dots, 2N\}$;
- (ii) $G(g_1, \dots, g_{2N})$ maximise G pour de tels choix de g_1, \dots, g_{2N} ;

- (iii) (g_1, \dots, g_{2N}) est minimal dans le sens qu'il contient la plus grande chaîne (en tenant compte de la périodicité) de la forme f_i, \dots, f_i pour un $i \in \{1, \dots, 2N\}$.

Soit k la longueur de cette chaîne. en utilisant l'invariance sous permutation cyclique, on peut supposer qu'elle a lieu au début de la suite (g_1, \dots, g_{2N}) , c'est-à-dire que $g_1 = g_2 = \dots = g_k = f_i$. Nous allons à présent vérifier que $k = 2N$, ce qui terminera la démonstration du cas unidimensionnel.

Supposons que $k < 2N$. Dans ce cas, $g_{2N} \neq f$. On obtient donc

$$\begin{aligned} G(g_1, \dots, g_{2N})^2 &\leq G(g_1, \dots, g_N, g_N, \dots, g_1) G(g_{2N}, \dots, g_{N+1}, g_{N+1}, \dots, g_{2N}) \\ &\leq G(g_1, \dots, g_N, g_N, \dots, g_1) G(g_1, \dots, g_{2N}), \end{aligned}$$

puisque (g_1, \dots, g_{2N}) maximise G . On a donc

$$G(g_1, \dots, g_{2N}) \leq G(g_1, \dots, g_N, g_N, \dots, g_1),$$

ce qui signifie que $(g_1, \dots, g_N, g_N, \dots, g_1)$ maximise également G . Mais ceci entraîne une contradiction puisque $(g_1, \dots, g_N, g_N, \dots, g_1)$ possède une chaîne de f_i de longueur au moins $\min\{2N, 2k\} > k$.

Pour traiter le cas multidimensionnel, on écrit

$$\prod_{t \in \mathbb{T}_{L/B}^d} f_t^{[t]} = \prod_{j=1}^{L/B} \left(\prod_{\substack{t \in \mathbb{T}_{L/B}^d \\ t_1=j}} f_t^{[t]} \right),$$

et on applique le résultat unidimensionnel au produit sur j . Ceci homogénéise le produit des f_t le long de la première coordonnée. En procédant ainsi dans chaque direction, le théorème suit. \square

Un cas particulier de l'estimation de l'échiquier qui est souvent utilisé est lorsque les fonctions B -blocs sont des indicatrices d'événements B -blocs \mathcal{A}_t , $t \in \mathbb{T}_{L/B}^d$. Dans ce cas, l'estimation prend la forme

$$\mu \left(\bigcap_{t \in \mathbb{T}_{L/B}^d} \mathcal{A}_t^{[t]} \right) \leq \prod_{t \in \mathbb{T}_{L/B}^d} \left[\mu \left(\bigcap_{s \in \mathbb{T}_{L/B}^d} \mathcal{A}_t^{[s]} \right) \right]^{1/|\mathbb{T}_{L/B}^d|},$$

où $\mathcal{A}^{[s]}$ est l'événement dont la fonction indicatrice est $(\mathbf{1}_{\mathcal{A}})^{[s]}$. Afin de pouvoir l'utiliser en pratique, il convient donc de pouvoir estimer des expressions de la forme

$$\mathfrak{z}(\mathcal{A}) = \mu \left(\bigcap_{t \in \mathbb{T}_{L/B}^d} \mathcal{A}^{[t]} \right)^{1/|\mathbb{T}_{L/B}^d|}.$$

Le lemme suivant permet de simplifier cette tâche, puisqu'il permet de décomposer l'événement \mathcal{A} en sous-événements pour lesquels le calcul est plus simple.

Lemme 7.2.1. Soient \mathcal{A} et $\mathcal{A}_1, \mathcal{A}_2, \dots$ des événements B -blocs tels que

$$\mathcal{A} \subseteq \bigcup_k \mathcal{A}_k.$$

Alors,

$$\mathfrak{z}(\mathcal{A}) \leq \sum_k \mathfrak{z}(\mathcal{A}_k).$$

Démonstration. Puisque

$$\bigcap_{t \in \mathbb{T}_{L/B}^d} \mathcal{A}^{[t]} \subseteq \bigcap_{t \in \mathbb{T}_{L/B}^d} \bigcup_k \mathcal{A}_k^{[t]} = \bigcup_{(k_t)_{t \in \mathbb{T}_{L/B}^d}} \bigcap_{t \in \mathbb{T}_{L/B}^d} \mathcal{A}_{k_t}^{[t]},$$

il suit de la σ -additivité de μ que

$$\mathfrak{z}(\mathcal{A})^{|\mathbb{T}_{L/B}^d|} = \mu\left(\bigcap_{t \in \mathbb{T}_{L/B}^d} \mathcal{A}^{[t]}\right) \leq \sum_{(k_t)_{t \in \mathbb{T}_{L/B}^d}} \mu\left(\bigcap_{t \in \mathbb{T}_{L/B}^d} \mathcal{A}_{k_t}^{[t]}\right).$$

Une application de l'estimation de l'échiquier donne alors

$$\mu\left(\bigcap_{t \in \mathbb{T}_{L/B}^d} \mathcal{A}_{k_t}^{[t]}\right) \leq \prod_{t \in \mathbb{T}_{L/B}^d} \mathfrak{z}(\mathcal{A}_{k_t}).$$

En insérant cette dernière expression, on obtient finalement

$$\mathfrak{z}(\mathcal{A})^{|\mathbb{T}_{L/B}^d|} \leq \sum_{(k_t)_{t \in \mathbb{T}_{L/B}^d}} \prod_{t \in \mathbb{T}_{L/B}^d} \mathfrak{z}(\mathcal{A}_{k_t}) = \prod_{t \in \mathbb{T}_{L/B}^d} \sum_k \mathfrak{z}(\mathcal{A}_k) = \left(\sum_k \mathfrak{z}(\mathcal{A}_k)\right)^{|\mathbb{T}_{L/B}^d|}.$$

□

7.3 Domination gaussienne et borne infrarouge

Dans cette section, nous allons dériver une autre conséquence essentielle de la positivité sous réflexion : la borne infrarouge (*infrared bound* en anglais). Contrairement à l'estimation de l'échiquier, nous allons ici nous restreindre à une classe spécifique de mesures positives sous réflexion.

On choisit comme espace de spin $S = \mathbb{R}^\nu$, et on suppose que les spins sont distribués *a priori* selon une mesure ρ supportée sur un sous-ensemble compact $\Omega \subset \mathbb{R}^\nu$. On considère un Hamiltonien de la forme

$$\mathcal{H}_{L,\beta}(\omega) = \frac{1}{2}\beta \sum_{i,j \in \mathbb{T}_L^d, i \sim j} |\omega_i - \omega_j|^2.$$

On note $\mu_{L,\beta}$ la mesure de Gibbs définie par

$$\frac{d\mu_{L,\beta}}{d\mu_0}(\omega) = \frac{1}{Z_{L,\beta}} e^{-\mathcal{H}_{L,\beta}(\omega)}. \quad (7.4)$$

Avant de poursuivre, donnons quelques exemples montrant que cette classe de mesures contient de nombreux modèles d'intérêts.

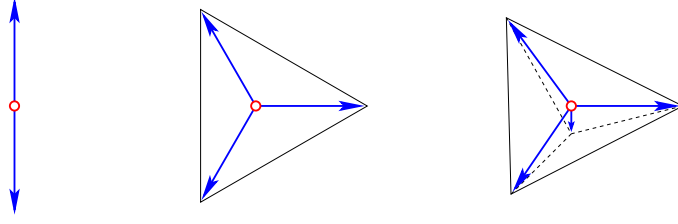


FIGURE 7.3 – Les vecteurs pour le modèle de Potts à 2, 3 et 4 états.

- Exemple 7.3.1.** 1. **Modèle d'Ising.** Dans ce cas, $\nu = 1$, et $\rho = \delta_{-1} + \delta_1$.
2. **Modèle $O(N)$.** Dans ce cas, $\nu = N$, et ρ est la mesure uniforme sur la sphère $\mathbb{S}^{N-1} \subset \mathbb{R}^N$. Le cas $N = 1$ correspond à nouveau au modèle d'Ising, le cas $N = 2$ au modèle XY et le cas $N = 3$ au modèle de Heisenberg classique.
3. **Modèle de Potts à q états.** Dans ce cas, $\nu = q - 1$, et ρ est la mesure uniforme sur les sommets du polyèdre régulier à q sommets inscrit dans la sphère $\mathbb{S}^{\nu-1}$ (cf. Figure 7.3). Observez que ce modèle peut être décrit de façon équivalente de la manière suivante : les spins prennent valeurs dans $\{1, \dots, q\}$, et l'Hamiltonien prend la forme

$$\mathcal{H} = -\frac{1}{2}\tilde{\beta} \sum_{i \sim j} \delta_{\omega_i, \omega_j}.$$

Le cas $q = 2$ correspond évidemment au modèle d'Ising. Il peut également être lié au modèle de percolation via la FK percolation discutée dans la série 8.

Théorème 7.3.1 (Domination gaussienne). Soit $h = (h_i)_{i \in \mathbb{T}_L^d} \in (\mathbb{R}^\nu)^{\mathbb{T}_L^d}$, et

$$\mathbf{Z}_{L,\beta}(h) = \langle \exp(-\frac{1}{2}\beta \sum_{i,j \in \mathbb{T}_L^d, i \sim j} |\omega_i - \omega_j + h_i - h_j|^2) \rangle_{\mu_0}.$$

Alors,

$$\mathbf{Z}_{L,\beta}(h) \leq \mathbf{Z}_{L,\beta}(0).$$

Remarque 7.3.1. Observez que $\mathbf{Z}_{L,\beta}(0)$ est précisément la fonction de partition apparaissant dans (7.4).

Démonstration. Manifestement, lorsque $h_i \rightarrow \pm\infty$, pour n'importe quel $i \in \mathbb{T}_L^d$, $\mathbf{Z}_{L,\beta}(h) \rightarrow 0$. Le maximum de $\mathbf{Z}_{L,\beta}(h)$ est donc réalisé pour un h fini. Soit h^* un tel maximiseur, pour lequel

$$N(h) = \#\{\{i, j\} : i \sim j, h_i \neq h_j\}$$

est minimal parmi tous les maximiseurs. Nous allons voir que $N(h^*) = 0$. En effet, si $N(h^*) > 0$, alors il existe un plan de réflexion Π passant par les arêtes et intersectant au moins une arête dont les extrémités i, j satisfont $h_i^* \neq h_j^*$. Soit Θ la réflexion correspondante. On vérifie aisément que, la somme dans l'exponentielle peut être mise sous la forme

$$A + \Theta(B) + \sum_{\alpha} C_{\alpha} \Theta(D_{\alpha}).$$

Par conséquent, on peut appliquer le Lemme 7.1.2 pour obtenir

$$\mathbf{Z}_{L,\beta}(h^*)^2 \leq \mathbf{Z}_{L,\beta}(h_+^*)\mathbf{Z}_{L,\beta}(h_-^*),$$

où $h_+^* = h^*$ sur $\mathbb{T}_{L,+}^d$ et $h_+^* = \Theta(h^*)$ sur $\mathbb{T}_{L,-}^d$, et similairement pour h_-^* . À présent, on observe que

$$\min\{N(h_+^*), N(h_-^*)\} < N(h^*).$$

Pour être spécifique, supposons que $N(h_+^*) < N(h^*)$. Alors, puisque h^* est un maximiseur,

$$\mathbf{Z}_{L,\beta}(h^*)^2 \leq \mathbf{Z}_{L,\beta}(h_+^*)\mathbf{Z}_{L,\beta}(h_-^*) \leq \mathbf{Z}_{L,\beta}(h_+^*)\mathbf{Z}_{L,\beta}(h^*),$$

c'est-à-dire $\mathbf{Z}_{L,\beta}(h_+^*) \geq \mathbf{Z}_{L,\beta}(h^*)$, et donc h_+^* est également un maximiseur, ce qui contredit notre choix de h^* .

On a donc bien $N(h^*) = 0$, ce qui signifie que h^* est constante. Comme $\mathbf{Z}_{L,\beta}(h+c) = \mathbf{Z}_{L,\beta}(h)$, pour tout $c \in \mathbb{R}$, la conclusion suit. \square

À l'aide du résultat précédent, on peut facilement démontrer la borne suivante, qui joue un rôle crucial dans de nombreuses applications. Comme sa formulation naturelle fait intervenir la transformation de Fourier, nous aurons besoin du tore réciproque,

$$\mathbb{T}_L^{d,*} \stackrel{\text{déf}}{=} \left\{ \frac{2\pi}{L}(n_1, \dots, n_d) : 0 \leq n_i < L \right\}.$$

Théorème 7.3.2 (Borne infrarouge). *Soit $\mu_{L,\beta}$ une mesure du type (7.4). Alors, pour tout $p \in \mathbb{T}_L^{d,*} \setminus \{0\}$,*

$$\sum_{j \in \mathbb{T}_L^d} e^{ip \cdot j} \langle \omega_0 \cdot \omega_j \rangle_{\mu_{L,\beta}} \leq \frac{\nu}{4\beta d} \left(1 - \frac{1}{2d} \sum_{j \sim 0} \cos(p \cdot j) \right)^{-1}.$$

Démonstration. D'après le Théorème 7.3.1, $\mathbf{Z}_{L,\beta}(h)$ est maximum en $h = 0$. Par conséquent, à h fixé,

$$\begin{aligned} \frac{\partial}{\partial \lambda} \mathbf{Z}_{L,\beta}(\lambda h) \Big|_{\lambda=0} &= 0, \\ \frac{\partial^2}{\partial \lambda^2} \mathbf{Z}_{L,\beta}(\lambda h) \Big|_{\lambda=0} &\leq 0. \end{aligned}$$

Après un peu de calcul, on obtient que

$$\begin{aligned} \frac{\partial^2}{\partial \lambda^2} \mathbf{Z}_{L,\beta}(\lambda h) \Big|_{\lambda=0} &= \beta^2 \left\langle \left| \sum_{\substack{i,j \in \mathbb{T}_L^d \\ i \sim j}} (\omega_i - \omega_j) \cdot (h_i - h_j) \right|^2 \exp \left\{ -\frac{\beta}{2} \sum_{\substack{i,j \in \mathbb{T}_L^d \\ i \sim j}} |\omega_i - \omega_j|^2 \right\} \right\rangle_{\mu_0} \\ &\quad - \beta \sum_{\substack{i,j \in \mathbb{T}_L^d \\ i \sim j}} |h_i - h_j|^2 \left\langle \exp \left\{ -\frac{\beta}{2} \sum_{\substack{i,j \in \mathbb{T}_L^d \\ i \sim j}} |\omega_i - \omega_j|^2 \right\} \right\rangle_{\mu_0}. \end{aligned}$$

L'inégalité précédente devient donc

$$\langle |\sum_{\substack{i,j \in \mathbb{T}_L^d \\ i \sim j}} (\omega_i - \omega_j) \cdot (h_i - h_j)|^2 \rangle_\mu \leq \frac{1}{\beta} \sum_{\substack{i,j \in \mathbb{T}_L^d \\ i \sim j}} |h_i - h_j|^2. \quad (7.5)$$

Cette inégalité est vérifiée pour toute collection h de nombres réels, mais on voit, en séparant les parties réelle et imaginaire, qu'elle l'est également pour des h complexes. On choisit $h_j = e^{ip \cdot j} \mathbf{e}_\ell$, avec $p \in \mathbb{T}_L^{d,*} \setminus \{0\}$ et $\ell \in \{1, \dots, \nu\}$.

Le membre de droite de (7.5) devient

$$\frac{1}{\beta} \sum_{\substack{i,j \in \mathbb{T}_L^d \\ i \sim j}} |h_i - h_j|^2 = \frac{1}{\beta} \sum_{\substack{i,j \in \mathbb{T}_L^d \\ i \sim j}} (2 - 2 \cos(p \cdot (i - j))) = \frac{4d|\mathbb{T}_L^d|}{\beta} (1 - \frac{1}{2d} \sum_{j \sim 0} \cos(p \cdot j)).$$

Similairement,

$$\sum_{\substack{i,j \in \mathbb{T}_L^d \\ i \sim j}} (\omega_i - \omega_j) \cdot (h_i - h_j) = 4d(1 - \frac{1}{2d} \sum_{j \sim 0} \cos(p \cdot j)) \sum_{j \in \mathbb{T}_L^d} (\omega_j(\ell) e^{ip \cdot j}),$$

et donc

$$\langle |\sum_{\substack{i,j \in \mathbb{T}_L^d \\ i \sim j}} (\omega_i - \omega_j) \cdot (h_i - h_j)|^2 \rangle_{\mu_{L,\beta}} = 16d^2 |1 - \frac{1}{2d} \sum_{j \sim 0} \cos(p \cdot j)|^2 \langle |\sum_{j \in \mathbb{T}_L^d} \omega_j(\ell) e^{ip \cdot j}|^2 \rangle_{\mu_{L,\beta}}.$$

La conclusion suit en observant que l'invariance sous translation de $\mu_{L,\beta}$ implique que

$$\langle |\sum_{j \in \mathbb{T}_L^d} \omega_j(\ell) e^{ip \cdot j}|^2 \rangle_{\mu_{L,\beta}} = \sum_{i,j \in \mathbb{T}_L^d} e^{ip \cdot (j-i)} \langle \omega_i(\ell) \omega_j(\ell) \rangle_{\mu_{L,\beta}} = |\mathbb{T}_L^d| \sum_{j \in \mathbb{T}_L^d} e^{ip \cdot j} \langle \omega_0(\ell) \omega_j(\ell) \rangle_{\mu_{L,\beta}},$$

et en sommant sur ℓ de 1 à ν pour récupérer le produit scalaire. \square

7.4 Applications

7.4.1 Condensation d'ondes de spin et transitions de phase

Comme première application de la positivité sous réflexion, et plus particulièrement de la borne infrarouge, nous allons présenter une preuve de l'existence d'une transition de phase dans toute la classe des modèles $O(N)$ et des modèles de Potts, lorsque β est grand et la dimension au moins égale à 3.

Le premier résultat montre que, dans ces conditions, il y a bien un ordre à grande distance (les spins restant alignés quelle que soit la distance les séparant). Nous verrons ensuite comment démontrer l'existence de plusieurs mesures de Gibbs.

7.4. APPLICATIONS

Théorème 7.4.1. Soit $\mu_{L,\beta}$ la mesure de Gibbs associée à un modèle $O(N)$ ou à un modèle de Potts. On suppose que $d \geq 3$. Soit²

$$\beta_0 = \frac{\nu}{4d} \int_{[-\pi,\pi]^d} \left(1 - \frac{1}{2d} \sum_{j \sim 0} \cos(p \cdot j)\right)^{-1} dp.$$

Alors, pour tout $\beta > \beta_0$,

$$\lim_{L \rightarrow \infty} \left\langle \left| \frac{1}{|\mathbb{T}_L^d|} \sum_{i \in \mathbb{T}_L^d} \omega_i \right|^2 \right\rangle_{\mu_{\beta,L}} \geq 1 - \frac{\beta_0}{\beta}.$$

Démonstration. On note

$$\hat{\omega}_p = \frac{1}{|\mathbb{T}_L^d|^{1/2}} \sum_{j \in \mathbb{T}_L^d} \omega_j e^{ip \cdot j}.$$

On a bien sûr, par invariance sous translations,

$$\langle |\hat{\omega}_p|^2 \rangle_{\mu_{\beta,L}} = \frac{1}{|\mathbb{T}_L^d|} \sum_{i,j \in \mathbb{T}_L^d} e^{ip \cdot (j-i)} \langle \omega_i \cdot \omega_j \rangle_{\mu_{\beta,L}} = \sum_{j \in \mathbb{T}_L^d} e^{ip \cdot j} \langle \omega_0 \cdot \omega_j \rangle_{\mu_{\beta,L}}.$$

Avec ces notations,

$$\left\langle \left| \frac{1}{|\mathbb{T}_L^d|} \sum_{i \in \mathbb{T}_L^d} \omega_i \right|^2 \right\rangle_{\mu_{\beta,L}} = \frac{1}{|\mathbb{T}_L^d|^2} \sum_{i,j \in \mathbb{T}_L^d} \langle \omega_i \cdot \omega_j \rangle_{\mu_{\beta,L}} = \frac{1}{|\mathbb{T}_L^d|} \langle |\hat{\omega}_0|^2 \rangle_{\mu_{\beta,L}}.$$

Afin d'estimer le membre de droite, on observe que, par la formule de Plancherel,

$$\sum_{p \in \mathbb{T}_L^{d,*}} |\hat{\omega}_p|^2 = \sum_{j \in \mathbb{T}_L^d} |\omega_j|^2 = |\mathbb{T}_L^d|,$$

puisque les spins satisfont $|\omega_i| = 1$. Par conséquent,

$$\frac{1}{|\mathbb{T}_L^d|} \langle |\hat{\omega}_0|^2 \rangle_{\mu_{\beta,L}} = 1 - \frac{1}{|\mathbb{T}_L^d|} \sum_{\substack{p \in \mathbb{T}_L^{d,*} \\ p \neq 0}} \langle |\hat{\omega}_p|^2 \rangle_{\mu_{\beta,L}}.$$

À présent, la borne infrarouge implique que, pour $p \neq 0$,

$$\langle |\hat{\omega}_p|^2 \rangle_{\mu_{\beta,L}} \leq \frac{\nu}{4\beta d} \left(1 - \frac{1}{2d} \sum_{j \sim 0} \cos(p \cdot j)\right)^{-1}.$$

Comme

$$\lim_{L \rightarrow \infty} \frac{1}{|\mathbb{T}_L^d|} \sum_{\substack{p \in \mathbb{T}_L^{d,*} \\ p \neq 0}} \left(1 - \frac{1}{2d} \sum_{j \sim 0} \cos(p \cdot j)\right)^{-1} = \int_{[-\pi,\pi]^d} \left(1 - \frac{1}{2d} \sum_{j \sim 0} \cos(p \cdot j)\right)^{-1} dp,$$

la conclusion suit. □

2. β_0 est fini lorsque $d \geq 3$, cf. exercice 2, série 9.

Le résultat précédent indique clairement qu'une transition de phase a lieu dans ces systèmes à basse température, puisque les spins sont globalement orientés. Malheureusement, il n'est pas si élémentaire d'en déduire la multiplicité des mesures de Gibbs. La raison est que la positivité sous réflexion nous force à travailler avec la condition au bord périodique. Afin de remédier à ce problème, il est nécessaire de faire appel à des résultats plus abstraits concernant les mesures de Gibbs. Le théorème fondamental dans ce contexte est le suivant. Nous ne le démontrerons pas entièrement, car il fait appel à des résultats sur la théorie abstraite des mesures de Gibbs que nous n'avons pas vus, mais en esquisserons l'argument. Il permet de démontrer l'existence d'une transition de phase du premier ordre en montrant que l'énergie libre associée au modèle a une dérivée discontinue par rapport à un "champ magnétique" approprié, similairement à ce que l'on a vu dans le cas du modèle d'Ising.

Théorème 7.4.2. *On considère une famille de spins bornés (ce qui signifie que la mesure a priori sur S est supportée sur une partie bornée de \mathbb{R}^ν). Soit \mathfrak{G}_β l'ensemble des mesures de Gibbs en volume infini pour une famille donnée de potentiels plus-proches-voisins invariants sous translation³, et soit $\mu_{L,\beta}$ une mesure de Gibbs sur le tore \mathbb{T}_L^d pour cette même famille de potentiels. Soit τ_i , $i \in \mathbb{Z}^d$, la translation par i , $(\tau_i(\omega))_j = \omega_{j-i}$. Soit $g : \Omega \rightarrow \mathbb{R}$ une fonction locale bornée. Alors*

1. *La limite*

$$f(h) = \lim_{L \rightarrow \infty} \frac{1}{|\mathbb{T}_L^d|} \log \langle \exp(h \sum_{i \in \mathbb{T}_L^d} g \circ \tau_i) \rangle_{\mu_{L,\beta}}$$

existe pour tout $h \in \mathbb{R}$ et est convexe en h .

2. *Si $\mu \in \mathfrak{G}_\beta$ est invariante sous les translations, alors*

$$\frac{\partial f}{\partial h^-} \Big|_{h=0} \leq \langle g \rangle_\mu \leq \frac{\partial f}{\partial h^+} \Big|_{h=0}.$$

3. *Il existe des mesures $\mu^+, \mu^- \in \mathfrak{G}_\beta$, invariantes sous translations et ergodiques⁴, telles que*

$$\langle g \rangle_{\mu^\pm} = \frac{\partial f}{\partial h^\pm} \Big|_{h=0}.$$

Esquisse de preuve. Soit $\Lambda_L = \{0, \dots, L-1\}^d$.

1. Sous les hypothèses du Théorème, l'existence et la convexité de $f(h)$ se démontre exactement de la même manière qu'on l'a fait pour l'énergie libre du modèle d'Ising, dans le Chapitre 2. De plus, comme dans ce chapitre, on montre en fait que la limite est inchangée si on remplace la mesure $\mu_{\beta,L}$ avec condition au bord périodique, par une suite de mesures de Gibbs dans Λ_L avec conditions au bord quelconque.

2. Soit $\mu \in \mathfrak{G}_\beta$ une mesure invariante sous les translations. En introduisant la notation

$$\mathbf{Z}_L(h) = \langle \exp(h \sum_{i \in \Lambda_L} g \circ \tau_i) \rangle_\mu,$$

3. Un potentiel Φ est invariant sous translation si $\Phi_A = \Phi_{\tau_i A}$, pour tout $i \in \mathbb{Z}^d$ et $A \in \mathbb{Z}^d$.

4. Rappel : une mesure est dite **ergodique** sous l'action d'un groupe G (ici, le groupe des translations) si tous les événements G -invariants ont probabilité 0 ou 1.

7.4. APPLICATIONS

et en observant que $\log \mathbf{Z}_L$ est convexe en h , on peut écrire, pour tout $h > 0$,

$$\log \mathbf{Z}_L(h) - \log \mathbf{Z}_L(0) \geq h \frac{\partial}{\partial h} \log \mathbf{Z}_L(h) \Big|_{h=0} = h \left\langle \sum_{i \in \Lambda_L} g \circ \tau_i \right\rangle_\mu = h |\Lambda_L| \langle g \rangle_\mu.$$

Puisque f est indépendante de la condition au bord, on en déduit que

$$f(h) - f(0) \geq h \langle g \rangle_\mu,$$

et donc

$$\frac{\partial f}{\partial h^+} \Big|_{h=0} = \lim_{h \downarrow 0} \frac{1}{h} (f(h) - f(0)) \geq \langle g \rangle_\mu.$$

L'autre borne est démontrée de la même façon.

3. Soit \mathcal{H} l'Hamiltonien associé aux potentiels de l'énoncé. Soit $\mathfrak{G}_{\beta,h}$ l'ensemble des mesures de Gibbs en volume infini correspondant à l'Hamiltonien $\mathcal{H} - h \sum_i g \circ \tau_i$. Alors, une variante de l'argument du point (2) permet de montrer que si $\mu_h \in \mathfrak{G}_{\beta,h}$ est invariante sous les translations, alors

$$\frac{\partial f}{\partial h^-} \leq \langle g \rangle_{\mu_h} \leq \frac{\partial f}{\partial h^+}.$$

En particulier, si $h > 0$, on en déduit que

$$\langle g \rangle_{\mu_h} \geq \frac{\partial f}{\partial h^-} \geq \frac{\partial f}{\partial h^+} \Big|_{h=0},$$

par monotonie des dérivées d'une fonction convexe (*cf.* Théorème 2.11.1). On considère une suite $h_n \downarrow 0$. Par un argument de compacité, on peut extraire une sous-suite h_{n_k} telle que $\mu_{h_{n_k}}$ converge vers une mesure $\mu^+ \in \mathfrak{G}_\beta$ (ces résultats font appel à la théorie abstraite des mesures de Gibbs que l'on verra plus tard). Par conséquent, puisque g est locale,

$$\langle g \rangle_{\mu^+} = \lim_{k \rightarrow \infty} \langle g \rangle_{\mu_{h_{n_k}}} \geq \frac{\partial f}{\partial h^+} \Big|_{h=0}.$$

Il suit alors du point 2. que

$$\langle g \rangle_{\mu^+} = \frac{\partial f}{\partial h^+} \Big|_{h=0}.$$

La mesure μ^+ est invariante sous les translations. La preuve qu'on peut en fait la choisir ergodique est assez abstraite, et repose à nouveau sur la théorie générale des mesures de Gibbs. \square

Armé de ce théorème, nous sommes en mesure de montrer que le Théorème 7.4.1 implique l'existence d'une transition de phase de premier ordre. Nous ne traitons que les modèles $O(N)$, mais un résultat analogue (avec une preuve analogue) est vraie dans les autres cas également.

Théorème 7.4.3. *On considère le modèle $O(N)$ en dimension 3 ou plus. Soit β_0 comme dans le Théorème 7.4.1. Il existe une famille de mesures de Gibbs $(\mu_\psi)_{\psi \in \mathbb{S}^\nu}$ telles que, pour tout $\beta > \beta_0$,*

$$\lim_{L \rightarrow \infty} \frac{1}{|\Lambda_L|} \sum_{i \in \Lambda_L} \omega_i = m^* \psi, \quad \mu_\psi\text{-presque sûrement,}$$

pour un $m^* = m^*(\beta) > 0$.

Démonstration. Fixons un vecteur $\psi \in \mathbb{S}^\nu$, et considérons la fonction locale $g(\omega) = \omega_0 \cdot \psi$. Par le Théorème 7.4.2, la limite

$$f(h) = \lim_{L \rightarrow \infty} \frac{1}{|\mathbb{T}_L^d|} \log \langle \exp(h \sum_{i \in \mathbb{T}_L^d} g \circ \tau_i) \rangle_{\mu_{L,\beta}}$$

existe. Nous allons montrer que $\frac{\partial f}{\partial h^+}|_{h=0} > 0$ (et donc $\frac{\partial f}{\partial h^-}|_{h=0} < 0$ par symétrie). Bien entendu,

$$\sum_{i \in \mathbb{T}_L^d} g \circ \tau_i = \psi \cdot \sum_{i \in \mathbb{T}_L^d} \omega_i \equiv \psi \cdot M_L.$$

Nous savons du Théorème 7.4.1 que

$$\frac{1}{|\mathbb{T}_L^d|^2} \langle |M_L|^2 \rangle_{\mu_{\beta,L}} \geq 1 - \frac{\beta_0}{\beta} - o(1),$$

le dernier terme disparaissant lorsque $L \rightarrow \infty$. Puisque $|M_L| \leq |\mathbb{T}_L^d|$, il suit que, pour tout $0 < \epsilon < 1$,

$$\frac{1}{|\mathbb{T}_L^d|^2} \langle |M_L|^2 \rangle_{\mu_{\beta,L}} \leq \epsilon^2 + \mu_{\beta,L}(|M_L| \geq \epsilon |\mathbb{T}_L^d|).$$

Par conséquent,

$$\mu_{\beta,L}(|M_L| \geq \frac{1}{2}(1 - \frac{\beta_0}{\beta})|\mathbb{T}_L^d|) \geq \frac{1}{2}(1 - \frac{\beta_0}{\beta}) - o(1).$$

À présent, par invariance sous les rotations de $\mu_{\beta,L}$, il existe $c = c(\nu) > 0$ tel que

$$\mu_{\beta,L}(\psi \cdot M_L \geq \frac{1}{4}(1 - \frac{\beta_0}{\beta})|\mathbb{T}_L^d| \mid |M_L| \geq \frac{1}{2}(1 - \frac{\beta_0}{\beta})|\mathbb{T}_L^d|) \geq c,$$

et donc

$$\mu_{\beta,L}(\psi \cdot M_L \geq \frac{1}{4}(1 - \frac{\beta_0}{\beta})|\mathbb{T}_L^d|) \geq \frac{c}{2}(1 - \frac{\beta_0}{\beta}) - o(1).$$

On en déduit que, pour $h > 0$,

$$\begin{aligned} \log \langle \exp(h \psi \cdot M_L) \rangle_{\mu_{L,\beta}} &\geq \log \langle \exp(h \psi \cdot M_L) \mathbf{1}_{\{\psi \cdot M_L \geq \frac{1}{4}(1 - \frac{\beta_0}{\beta})|\mathbb{T}_L^d|\}} \rangle_{\mu_{L,\beta}} \\ &\geq \frac{1}{4}(1 - \frac{\beta_0}{\beta})|\mathbb{T}_L^d| + \log(\frac{c}{2}(1 - \frac{\beta_0}{\beta}) - o(1)), \end{aligned}$$

et on a bien

$$\frac{\partial f}{\partial h^+}|_{h=0} \geq \frac{h}{4} \left(1 - \frac{\beta_0}{\beta}\right).$$

Une application du Théorème 7.4.2 implique donc, lorsque $\beta > \beta_0$, l'existence d'une mesure de Gibbs invariante sous les translations, ergodique et telle que

$$\langle \omega_0 \cdot \psi \rangle_{\mu_\psi} = \frac{\partial f}{\partial h^+}|_{h=0} > 0. \quad (7.6)$$

Le théorème ergodique⁵ implique que

$$\lim_{L \rightarrow \infty} \frac{1}{|\Lambda_L|} \sum_{i \in \Lambda_L} \omega_i = m^* \tilde{\psi}, \quad \mu_\psi\text{-presque sûrement}, \quad (7.7)$$

pour un $\tilde{\psi} \in \mathcal{S}'$ et m^* est un nombre strictement positif tel que, au vu de (7.6),

$$m^* \psi \cdot \tilde{\psi} = \frac{\partial f}{\partial h^+}|_{h=0}.$$

Nous allons montrer que $\tilde{\psi} = \psi$, ce qui conclura l'argument. Supposons que ce ne soit pas le cas. Soit $A \in O(N)$ une rotation telle que $A\psi = \psi$, et soit $\tilde{\mu} = \mu_\psi \circ A^{-1}$. Par invariance du modèle sous $O(N)$, on doit avoir $\tilde{\mu} \in \mathfrak{G}_\beta$. Il suit de (7.7) que $\langle \omega_0 \rangle_{\mu_\psi} = \tilde{\psi}$. Par conséquent,

$$\langle \omega_0 \cdot \psi \rangle_{\tilde{\mu}} = \langle A\omega_0 \cdot \psi \rangle_{\mu_\psi} = m^* |\psi|^2 > m^* \psi \cdot \tilde{\psi} = \frac{\partial f}{\partial h^+}|_{h=0}.$$

Mais ceci est impossible, car $\tilde{\mu}$ est une mesure de Gibbs et doit donc satisfaire aux bornes données au point 2. du Théorème 7.4.2. \square

7.4.2 Transition de phase dans les modèles $O(N)$ anisotropes

Dans cette sous-section et la suivante, nous allons voir deux applications de l'estimation de l'échiquier. Une autre application célèbre de cet outil est pour prouver que l'aimantation du modèle de Potts (pour q grand) est discontinue à la température critique, ce qui correspond à la coexistence des q phases ordonnées de basse température avec la phase désordonnée de haute température. Malheureusement, nous ne discuterons pas de cette application par manque de temps.

5. Dans le présent contexte, celui affirme que, pour toute mesure ergodique μ et toute fonction $g \in \mathcal{L}_1(\mu)$, on a, pour μ -presque tout ω ,

$$\lim_{L \rightarrow \infty} \frac{1}{|\mathbb{T}_L^d|} \sum_{i \in \mathbb{T}_L^d} g \circ \tau_i(\omega) = \langle g \rangle_{\mu_{L,\beta}}.$$

Nous allons commencer par démontrer l'existence d'une transition de phase dans une variante du modèle $O(N)$ bidimensionnel appelée modèle $O(N)$ anisotrope. Dans ce modèle, les spins prennent valeurs dans la sphère \mathbb{S}^{N-1} . L'hamiltonien est alors de la forme

$$\mathcal{H}_{L;\beta,\alpha}(\omega) = -\frac{1}{2}\beta \sum_{\substack{i,j \in \mathbb{T}_L^2 \\ i \sim j}} (\omega_i(1)\omega_j(1) + \alpha \sum_{k=2}^N \omega_i(k)\omega_j(k)).$$

$0 \leq \alpha \leq 1$ est le coefficient d'anisotropie. On voit que lorsque $\alpha = 1$ on retrouve le produit scalaire, et, par conséquent, il suit de Théorème de Mermin-Wagner qu'il n'y a pas d'aimantation spontanée à basse température. Le théorème suivant montre que ceci n'est plus vrai dès que $\alpha < 1$ (évidemment, dans ce cas, le Théorème de Mermin-Wagner ne s'applique plus).

Théorème 7.4.4. *Pour tout $0 \leq \alpha < 1$ et tout $\epsilon > 0$, il existe $\beta_0 = \beta_0(\alpha, \epsilon)$ tel que, pour tout $\beta > \beta_0$,*

$$\lim_{L \rightarrow \infty} \langle |M_L(1)|^2 \rangle_{L;\beta,\alpha} > 1 - \epsilon,$$

où $M_L(1) = |\mathbb{T}_L^2|^{-1} \sum_{i \in \mathbb{T}_L^2} S_i(1)$ est la projection de l'aimantation moyenne le long de la première coordonnée.

Remarque 7.4.1. *On peut à nouveau reformuler, à l'aide du Théorème 7.4.2, le résultat précédent en terme de multiplicité des mesures de Gibbs. Comme cela se fait de manière similaire à ce que l'on a déjà fait dans la sous-section précédente, nous nous contenterons ici du résultat sur l'aimantation.*

Démonstration. Il suffit de montrer que

$$\langle \omega_0(1)\omega_j(1) \rangle_{L;\beta,\alpha} \geq c(\beta),$$

uniformément en L et en $j \in \mathbb{T}_L^2$, avec $c(\beta) \rightarrow 1$ lorsque $\beta \rightarrow \infty$. En effet, cela implique que

$$\begin{aligned} \lim_{L \rightarrow \infty} \langle |M_L(1)|^2 \rangle_{L;\beta,\alpha} &= \lim_{L \rightarrow \infty} |\mathbb{T}_L^2|^{-2} \langle \sum_{i,j \in \mathbb{T}_L^2} \omega_i(1)\omega_j(1) \rangle_{L;\beta,\alpha} \\ &= \lim_{L \rightarrow \infty} |\mathbb{T}_L^2|^{-1} \langle \sum_{j \in \mathbb{T}_L^2} \omega_0(1)\omega_j(1) \rangle_{L;\beta,\alpha} \geq c(\beta). \end{aligned}$$

En introduisant les indicatrices suivantes ($0 \leq \delta \leq 1$ et $i \in \mathbb{T}_L^2$)

$$P_i^{\pm\delta}(\omega) = \mathbf{1}_{\{\pm\omega_i(1) \geq \delta\}}, \quad P_i^{<\delta}(\omega) = 1 - P_i^{+\delta}(\omega) - P_i^{-\delta}(\omega) = \mathbf{1}_{\{|\omega_i(1)| < \delta\}},$$

on peut écrire

$$\begin{aligned} \langle \omega_0(1)\omega_j(1) \rangle_{L;\beta,\alpha} &= \langle \omega_0(1)\omega_j(1) (P_0^{+\delta} + P_0^{-\delta} + P_0^{<\delta})(P_j^{+\delta} + P_j^{-\delta} + P_j^{<\delta}) \rangle_{L;\beta,\alpha} \\ &\geq 2\delta^2 \langle P_0^{+\delta} P_j^{+\delta} \rangle_{L;\beta,\alpha} - 2 \langle P_0^{+\delta} P_j^{-\delta} \rangle_{L;\beta,\alpha} - \delta^2 \langle P_0^{<\delta} P_j^{<\delta} \rangle_{L;\beta,\alpha} \\ &\quad - 4\delta \langle P_0^{+\delta} P_j^{<\delta} \rangle_{L;\beta,\alpha}, \end{aligned}$$

7.4. APPLICATIONS

où l'on a utilisé la symétrie sous l'échange $0 \leftrightarrow j$, la symétrie sous $\omega(1) \rightarrow -\omega(1)$, et le fait que $|\omega_i(1)| \leq 1$ pour tout $i \in \mathbb{T}_L^2$.

On observe à présent que

$$\begin{aligned} \langle P_0^{+\delta} P_j^{<\delta} \rangle_{L;\beta,\alpha} &\leq \langle P_0^{<\delta} \rangle_{L;\beta,\alpha}, \\ \langle P_0^{<\delta} P_j^{<\delta} \rangle_{L;\beta,\alpha} &\leq \sqrt{\langle P_0^{<\delta} \rangle_{L;\beta,\alpha}} \sqrt{\langle P_j^{<\delta} \rangle_{L;\beta,\alpha}} = \langle P_0^{<\delta} \rangle_{L;\beta,\alpha}, \end{aligned}$$

la deuxième ligne suivant de Cauchy-Schwarz. Comme on a également

$$\begin{aligned} \langle P_0^{+\delta} P_j^{+\delta} \rangle_{L;\beta,\alpha} &= \langle P_0^{+\delta} (1 - P_j^{-\delta} - P_j^{<\delta}) \rangle_{L;\beta,\alpha} \\ &= \langle P_0^{+\delta} \rangle_{L;\beta,\alpha} - \langle P_0^{+\delta} P_j^{-\delta} \rangle_{L;\beta,\alpha} - \langle P_0^{+\delta} P_j^{<\delta} \rangle_{L;\beta,\alpha}, \end{aligned}$$

et que

$$\langle P_0^{+\delta} \rangle_{L;\beta,\alpha} = \frac{1}{2} - \frac{1}{2} \langle P_j^{<\delta} \rangle_{L;\beta,\alpha},$$

on voit que le théorème suit des deux lemmes suivants.

Lemme 7.4.1. *Pour tout $0 \leq \alpha < 1$, $0 < \delta < 1$ et tout $\epsilon > 0$, il existe $\beta_0 = \beta_0(\epsilon, \alpha, \delta)$ tel que, pour tout $\beta > \beta_0$,*

$$\langle P_0^{<\delta} \rangle_{L;\beta,\alpha} < \epsilon.$$

Lemme 7.4.2. *Pour tout $0 \leq \alpha < 1$, $0 < \delta < 1$, tout $\epsilon > 0$ et tout $j \in \mathbb{T}_L^2$, il existe $\beta_0 = \beta_0(\epsilon, \alpha)$ tel que, pour tout $\beta > \beta_0$,*

$$\langle P_0^{+\delta} P_j^{-\delta} \rangle_{L;\beta,\alpha} < \epsilon.$$

Démonstration du Lemme 7.4.1. On considère l'événement

$$P_L^{<\delta} = \prod_{i \in \mathbb{T}_L^2} P_i^{<\delta}.$$

L'estimation de l'échiquier implique que

$$\langle P_0^{<\delta} \rangle_{L;\beta,\alpha} \leq (\langle P_L^{<\delta} \rangle_{L;\beta,\alpha})^{1/|\mathbb{T}_L^2|}.$$

On est donc ramené à l'estimation d'un événement "global", ce qui peut se faire beaucoup plus simplement. On écrit

$$\langle P_L^{<\delta} \rangle_{L;\beta,\alpha} = \frac{\mathbf{Z}_{L;\beta,\alpha}^{<\delta}}{\mathbf{Z}_{L;\beta,\alpha}},$$

où

$$\mathbf{Z}_{L;\beta,\alpha}^{<\delta} = \langle e^{-\mathcal{H}_{L;\beta,\alpha}} P_L^{<\delta} \rangle_{\mu_0}, \quad \mathbf{Z}_{L;\beta,\alpha} = \langle e^{-\mathcal{H}_{L;\beta,\alpha}} \rangle_{\mu_0},$$

avec μ_0 le produit des mesures uniformes sur \mathbb{S}^{N-1} , supposées normalisées, $\langle 1 \rangle_{\mu_0} = 1$.

Commençons par chercher une borne supérieure sur le numérateur. Pour une configuration ω telle que $P_L^{<\delta}(\omega) = 1$, on a

$$\mathcal{H}_{L;\beta,\alpha}(\omega) \geq -\frac{1}{2}\beta \sum_{\substack{i,j \in \mathbb{T}_L^2 \\ i \sim j}} (\delta^2 \vee \alpha) = -2\beta(\delta^2 \vee \alpha) |\mathbb{T}_L^2|.$$

Par conséquent,

$$\mathbf{Z}_{L;\beta,\alpha}^{<\delta} \leq e^{2\beta(\delta^2 \vee \alpha) |\mathbb{T}_L^2|}.$$

Cherchons à présent une borne inférieure sur $\mathbf{Z}_{L;\beta,\alpha}$. Soit $0 < \tilde{\delta} < 1$, et

$$P_L^{+\tilde{\delta}} = \prod_{i \in \mathbb{T}_L^2} P_i^{+\tilde{\delta}}.$$

On peut ainsi écrire

$$\mathbf{Z}_{L;\beta,\alpha} \geq \langle e^{-\mathcal{H}_{L;\beta,\alpha}} P_L^{+\tilde{\delta}} \rangle_{\mu_0} = \langle e^{-\mathcal{H}_{L;\beta,\alpha}} \rangle_{\tilde{\mu}_0} \langle P_L^{+\tilde{\delta}} \rangle_{\mu_0},$$

où l'on a introduit la mesure produit $\tilde{\mu}_0$ définie par

$$\frac{d\tilde{\mu}_0}{d\mu_0} = P_L^{+\tilde{\delta}}.$$

Par l'inégalité de Jensen, on a alors

$$\langle e^{-\mathcal{H}_{L;\beta,\alpha}} \rangle_{\tilde{\mu}_0} \geq \exp(-\langle \mathcal{H}_{L;\beta,\alpha} \rangle_{\tilde{\mu}_0}).$$

Or,

$$\langle \mathcal{H}_{L;\beta,\alpha} \rangle_{\tilde{\mu}_0} = -\frac{1}{2}\beta \sum_{\substack{i,j \in \mathbb{T}_L^2 \\ i \sim j}} \langle \omega_i(1) \rangle_{\tilde{\mu}_0} \langle \omega_j(1) \rangle_{\tilde{\mu}_0} \leq -2\beta\tilde{\delta}^2 |\mathbb{T}_L^2|,$$

puisque $\langle \omega_i(k) \rangle_{\tilde{\mu}_0} = 0$ par symétrie, pour $k = 2, \dots, N$.

Finalement, pour tout $0 < \tilde{\delta} < 1$, il existe une constante $0 < c(\tilde{\delta}) < 1$, ne dépendant que de la dimension N des spins, telle que

$$\langle P_L^{+\tilde{\delta}} \rangle_{\mu_0} = (1 - c(\tilde{\delta}))^{|\mathbb{T}_L^2|}.$$

En mettant tout cela ensemble et en choisissant $\tilde{\delta} = \frac{1}{2}(1 + (\delta^2 \vee \alpha))$, on obtient finalement la borne suivante

$$\langle P_L^{<\delta} \rangle_{L;\beta,\alpha} \leq \exp(-2\beta(\tilde{\delta}^2 - (\delta^2 \vee \alpha)) - \frac{1}{2\beta} \log(1 - c(\tilde{\delta})) |\mathbb{T}_L^2|) \leq \exp(-\frac{2}{3}\beta(1 - (\delta^2 \vee \alpha)) |\mathbb{T}_L^2|).$$

Ceci implique donc que

$$\langle P_0^{<\delta} \rangle_{L;\beta,\alpha} \leq \exp(-\frac{2}{3}\beta(1 - (\delta^2 \vee \alpha))),$$

ce qui tend bien vers 0, lorsque $\beta \rightarrow \infty$, pour tout $\alpha, \delta < 1$. □

Démonstration du Lemme 7.4.2. La preuve de ce lemme repose sur un argument de Peierls couplé à l'estimation de l'échiquier.

Soit $P_j^\pm = \mathbf{1}_{\{\pm\omega_j(1) \geq 0\}}$. On a évidemment

$$\langle P_0^{+\delta} P_j^{-\delta} \rangle_{L;\beta,\alpha} \leq \langle P_0^+ P_j^- \rangle_{L;\beta,\alpha}.$$

Puisque $P_i^+ + P_i^- = 1$, pour tout i , on peut écrire

$$\begin{aligned} \langle P_0^+ P_j^- \rangle_{L;\beta,\alpha} &= \langle P_0^+ P_j^- \prod_{i \in \mathbb{T}_L^2 \setminus \{0,j\}} (P_i^+ + P_i^-) \rangle_{L;\beta,\alpha} \\ &= \sum_{\substack{\eta \in \{-1,1\}^{\mathbb{T}_L^2} \\ \eta_0=1, \eta_j=-1}} \langle \prod_{i \in \mathbb{T}_L^2} P_i^{\eta_i} \rangle_{L;\beta,\alpha}. \end{aligned}$$

À chacune des configurations η apparaissant dans la dernière somme, on associe ses contours $\Gamma(\eta)$, comme dans l'argument de Peierls. Manifestement, puisque $\eta_0 \neq \eta_j$, il existe (au moins) un contour γ séparant 0 et j . On peut alors écrire

$$\langle P_0^+ P_j^- \rangle_{L;\beta,\alpha} \leq \sum_{\gamma} \sum_{\eta: \Gamma(\eta) \ni \gamma} \langle \prod_{i \in \mathbb{T}_L^2} P_i^{\eta_i} \rangle_{L;\beta,\alpha},$$

où la somme est prise sur tous les contours séparant 0 et j . On note γ_{\perp} l'ensemble des arêtes (i, j) intersectant un contour γ ; on suppose que i est toujours du côté “+”, et j du côté “−” de γ . En procédant de la même façon que ci-dessus, on obtient

$$\begin{aligned} \langle \prod_{(i,j) \in \gamma_{\perp}} P_i^+ P_j^- \rangle_{L;\beta,\alpha} &= \langle \prod_{(i,j) \in \gamma_{\perp}} P_i^+ P_j^- \prod_{\substack{k \in \mathbb{T}_L^2 \\ k \cap \gamma_{\perp} = \emptyset}} (P_k^+ + P_k^-) \rangle_{L;\beta,\alpha} \\ &= \sum_{\eta: \Gamma(\eta) \ni \gamma} \langle \prod_{i \in \mathbb{T}_L^2} P_i^{\eta_i} \rangle_{L;\beta,\alpha}. \end{aligned}$$

Par conséquent, on a

$$\langle P_0^{+\delta} P_j^{-\delta} \rangle_{L;\beta,\alpha} \leq \sum_{\gamma} \langle \prod_{(i,j) \in \gamma_{\perp}} P_i^+ P_j^- \rangle_{L;\beta,\alpha}.$$

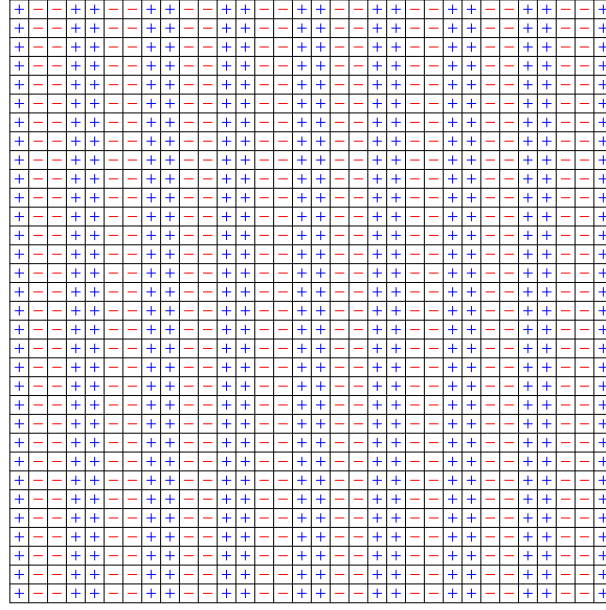
On va à présent montrer que

$$\langle \prod_{(i,j) \in \gamma_{\perp}} P_i^+ P_j^- \rangle_{L;\beta,\alpha} \leq e^{-c\beta|\gamma|}, \quad (7.8)$$

pour une constante $c > 0$. La preuve se termine alors comme dans l'argument de Peierls pour le modèle d'Ising⁶

Pour montrer (7.8), on utilise l'estimation de l'échiquier. Étant donné une arête horizontale (i, j) , on note $i \wedge j = \min\{i(1), j(1)\}$. On fait une définition analogue pour les arêtes verticales. On décompose γ_{\perp} en quatre parties disjointes : $\gamma_{\perp}^{\text{h,e}}$, $\gamma_{\perp}^{\text{h,o}}$, $\gamma_{\perp}^{\text{v,e}}$ et $\gamma_{\perp}^{\text{v,o}}$. Les

6. La seule différence mineure avec la cas du modèle d'Ising traité dans le chapitre 2 est qu'on se trouve sur un tore. Par conséquent, il convient de considérer trois situations : γ entoure 0, γ entoure j , et γ s'enroule autour du tore. Les deux premières sont traitées comme on l'a fait pour le modèle d'Ising. Pour la dernière, il suffit de réaliser qu'un tel contour passe nécessairement par un sommet de la forme $(\frac{1}{2}, k)$, et par conséquent le nombre de tels contours de longueur ℓ est borné par $L4^{\ell}$. Comme le plus court contour s'enroulant autour du tore est de longueur $\ell = L$, on voit que la contribution de ces contours est exponentiellement petite en L , lorsque β est grand.


 FIGURE 7.4 – La configuration $\eta^{h,o}$.

ensembles $\gamma_{\perp}^{h,e/o}$ contiennent les arêtes horizontales de γ_{\perp} : le premier contient les arêtes (i, j) telles que $i \wedge j$ est pair, le second celles pour lesquelles $i \wedge j$ est impair. On procède similairement pour les arêtes verticales.

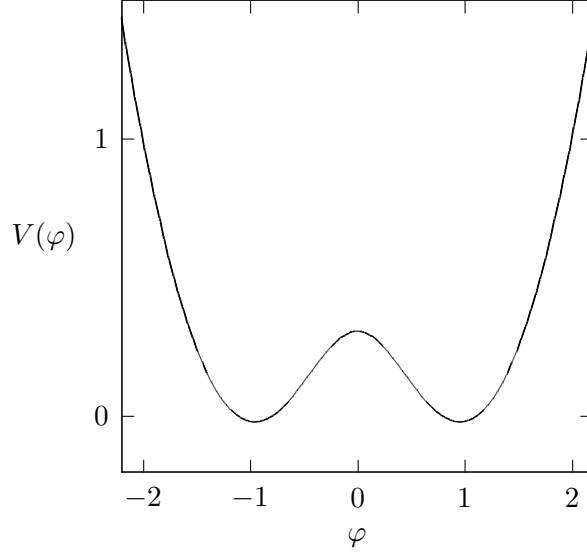
Par Cauchy-Schwarz, on a alors

$$\left\langle \prod_{(i,j) \in \gamma_{\perp}} P_i^+ P_j^- \right\rangle_{L;\beta,\alpha} \leq \prod_{\substack{a \in \{h,v\} \\ b \in \{e,o\}}} \left(\left\langle \prod_{(i,j) \in \gamma_{\perp}^{a,b}} P_i^+ P_j^- \right\rangle_{L;\beta,\alpha} \right)^{1/4}.$$

Les termes $\langle \prod_{(i,j) \in \gamma_{\perp}^{a,b}} P_i^+ P_j^- \rangle_{L;\beta,\alpha}$ se prêtent bien à une application de l'estimation de l'échiquier. Considérons par exemple $\langle \prod_{(i,j) \in \gamma_{\perp}^{h,e}} P_i^+ P_j^- \rangle_{L;\beta,\alpha}$. Chacun des produits $P_i^+ P_j^-$ est une fonction B -bloc pour un bloc 2×1 ou la réflexion d'une telle fonction à travers les arêtes, comme cela a été introduit dans la Section 7.2. Une application de l'estimation de l'échiquier donne donc

$$\left\langle \prod_{(i,j) \in \gamma_{\perp}^{h,e}} P_i^+ P_j^- \right\rangle_{L;\beta,\alpha} \leq \left(\left\langle \prod_{i \in \mathbb{T}_L^2} P_i^{\eta_i^{h,e}} \right\rangle_{L;\beta,\alpha} \right)^{|\gamma_{\perp}^{h,e}|/|\mathbb{T}_L^2|},$$

où la configuration $\eta^{h,e}$ est représentée sur la Figure 7.4. Pour évaluer cette dernière espérance, on procède de façon similaire à la preuve du lemme précédent. On estime inférieurement l'énergie des configurations contribuant à l'événement : chaque arête entre deux sommets auxquels $\eta^{h,e}$ prend la même valeur contribue au minimum $-\beta$ à l'énergie, et elles représentent les $\frac{3}{4}$ des arêtes ; pour minimiser l'énergie entre spins situés aux extrémités des arêtes restantes, il convient d'annuler leur composante 1, ce qui donne


 FIGURE 7.5 – Le potentiel $V_2(\varphi)$.

une contribution minimale de $-\alpha\beta$ par arête. On obtient donc que l'énergie minimale est toujours supérieure à $-2(\frac{3}{4} + \frac{1}{4}\alpha)\beta|\mathbb{T}_L^2|$. En utilisant la borne inférieure sur la fonction de partition de ce modèle établie dans la preuve du lemme précédent, on conclut que la probabilité est bornée par

$$\exp\left\{-2\beta(\tilde{\delta}^2 - \frac{1}{2\beta}\log(1 - c(\tilde{\delta})) - (\frac{3}{4} + \frac{1}{4}\alpha))|\mathbb{T}_L^2|\right\} \leq \exp\left\{-c\beta(1 - \alpha)|\mathbb{T}_L^2|\right\},$$

pour une constante $c > 0$ et tout β suffisamment grand, dès que $\tilde{\delta}$ est choisi assez proche de 1. □

□

7.4.3 Transition de phase dans un modèle gaussien avec double-puits

Dans cette section, nous allons étudier une variante du modèle gaussien introduit dans le chapitre 6. Le modèle est défini sur le tore \mathbb{T}_L^2 , les spins sont à valeurs dans \mathbb{R} , et l'Hamiltonien prend la forme

$$\mathcal{H}_L(\omega) \stackrel{\text{déf}}{=} \frac{\beta}{4d} \sum_{\substack{\{i,j\} \subset \mathbb{T}_L^2 \\ i \sim j}} (\varphi_i(\omega) - \varphi_j(\omega))^2 + \sum_{i \in \mathbb{T}_L^2} V_\kappa(\varphi_i), \quad (7.9)$$

où le potentiel V_κ est défini par

$$e^{-V_\kappa(\varphi)} = e^{-\frac{\kappa}{2}(\varphi-1)^2} + e^{-\frac{\kappa}{2}(\varphi+1)^2}$$

La raison pour laquelle on a pris la fonction V_κ de cette forme particulière et parce qu'elle permet ainsi très facilement de passer à une représentation en termes de spins d'Ising. Plus précisément, on observe que

$$\begin{aligned} \prod_i e^{-V_\kappa(\varphi_i)} &= \prod_i \sum_{\sigma_i \in \{-1,1\}} e^{-\frac{\kappa}{2}(\varphi_i - \sigma_i)^2} = \sum_{\sigma \in \{-1,1\}^{\mathbb{T}_L^2}} \prod_i e^{-\frac{\kappa}{2}(\varphi_i - \sigma_i)^2} \\ &= e^{-\kappa|\mathbb{T}_L^2|} \sum_{\sigma \in \{-1,1\}^{\mathbb{T}_L^2}} \prod_i e^{-\frac{\kappa}{2}\varphi_i^2 - \kappa\sigma_i\varphi_i}. \end{aligned}$$

Par conséquent, on peut considérer un système dont les spins sont à valeurs dans $\mathbb{R} \times \{-1, 1\}$, les variables aléatoires correspondantes étant notées (φ_i, σ_i) , et dont l'Hamiltonien est donné par

$$\frac{\beta}{4d} \sum_{\substack{\{i,j\} \subset \mathbb{T}_L^2 \\ i \sim j}} (\varphi_i - \varphi_j)^2 + \sum_{i \in \mathbb{T}_L^2} \frac{\kappa}{2} \varphi_i^2 - \kappa \sigma_i \varphi_i.$$

On notera $\mu_{L;\beta,\kappa}$ la mesure de Gibbs associée. Clairement, la φ -marginale de cette mesure n'est autre que la mesure de Gibbs associée à l'Hamiltonien (7.9). Il est également immédiat de vérifier que $\mu_{L;\beta,\kappa}$ est positive sous toutes les réflexions (horizontales et verticales (à travers les sites ou les arêtes), ou diagonales (à travers les sites)).

Le résultat principal de cette section est le suivant. Il montre qu'à basse température et à grand κ , les spins tendent à se regrouper dans un des deux puits. Comme précédemment, on se contente d'un résultat sur l'aimantation, mais il n'est pas difficile d'en déduire un résultat sur les mesures de Gibbs en volume infini.

Théorème 7.4.5. *Pour tout $\epsilon > 0$, il existe $\beta_0, \kappa_0 > 0$ tels que, pour tout $\beta > \beta_0$ et tout $\kappa > \kappa_0$,*

$$1 \geq \lim_{L \rightarrow \infty} \langle \left| \frac{1}{|\mathbb{T}_L^2|} \sum_{i \in \mathbb{T}_L^2} \varphi_i \right| \rangle_{L;\beta,\kappa} \geq 1 - \epsilon.$$

Démonstration. Appelons plaquette en $i \in \mathbb{T}_L^2$ l'ensemble de 4 sommets voisins $\{i, i + (1, 0), i + (1, 1), i + (0, 1)\}$. Le motif d'une plaquette dans une configuration donnée est la restriction de la configuration σ à la plaquette. Nous allons montrer que lorsque β et κ tendent vers l'infini, seuls les motifs $\begin{smallmatrix} + & + \\ + & + \end{smallmatrix}$ et $\begin{smallmatrix} - & - \\ - & - \end{smallmatrix}$ survivent. Nous dirons que ces deux derniers motifs sont bons, et les autres mauvais. Nous dirons également qu'une plaquette est bonne ou mauvaise selon que son motif est bon ou mauvais.

Lemme 7.4.3. *Pour tout $\beta, \kappa > 0$,*

$$\begin{aligned} \mathfrak{z}\left(\begin{smallmatrix} + & - \\ - & + \end{smallmatrix}\right) &\leq \exp\left(-\frac{\beta\kappa}{2\beta + \kappa}\right), \\ \mathfrak{z}\left(\begin{smallmatrix} + & - \\ + & - \end{smallmatrix}\right) &\leq \exp\left(-\frac{\beta\kappa}{2\beta + 2\kappa}\right), \\ \mathfrak{z}\left(\begin{smallmatrix} + & + \\ + & - \end{smallmatrix}\right) &\leq \exp\left(-\frac{\beta\kappa}{4\beta + 2\kappa}\right). \end{aligned}$$

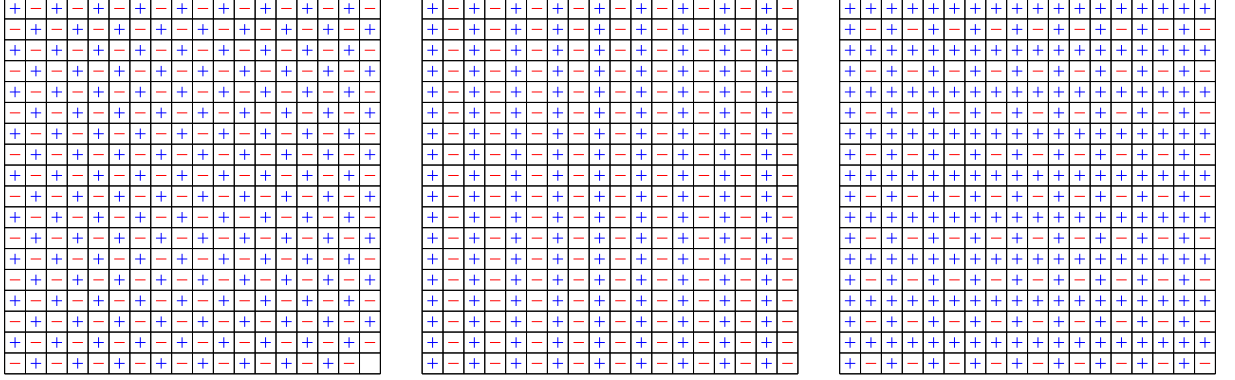


FIGURE 7.6 – Les configurations induites par les trois motifs du Lemme 7.4.3 par réflexions horizontales et verticales à travers les sommets.

Remarque 7.4.2. Bien entendu, le modèle étant invariant sous l'échange $+ \leftrightarrow -$, ainsi que sous les réflexions, ce lemme permet de traiter tous les mauvais motifs.

Démonstration du Lemme 7.4.3. Soit $\mathbf{Z}_{L;\beta,\kappa} = \sum_{\sigma} \int e^{-\mathcal{H}(\varphi,\sigma)} \prod_{i \in \mathbb{T}_L^2} d\varphi_i$ la fonction de partition associée à $\mu_{L;\beta,\kappa}$. On note $\mathbf{Z}_{L;\beta,\kappa}(\frac{+}{-} \frac{-}{+})$ la fonction partition dans laquelle la somme sur σ est restreinte à la seule configuration obtenue en étendant par réflexion à travers les sommets le motif $\frac{+}{-} \frac{-}{+}$ à tout le tore ; par exemple, dans le cas du motif $\frac{+}{-} \frac{-}{+}$, la fonction de partition est restreinte à la configuration σ ayant une structure en damier avec des $-$ sur les sommets pairs et des $+$ sur les sommets impairs (cf. Fig. 7.6). On peut ainsi écrire, par définition de \mathfrak{z} ,

$$\mathfrak{z}(\frac{+}{-} \frac{-}{+})^{|\mathbb{T}_L^2|} = \frac{\mathbf{Z}_{L;\beta,\kappa}(\frac{+}{-} \frac{-}{+})}{\mathbf{Z}_{L;\beta,\kappa}} \leq \frac{\mathbf{Z}_{L;\beta,\kappa}(\frac{+}{-} \frac{-}{+})}{\mathbf{Z}_{L;\beta,\kappa}(\frac{+}{+} \frac{+}{+})} = \frac{\langle e^{\kappa \sum_i (-1)^{\|i\|_1} \varphi_i} \rangle_{\text{GM}}}{\langle e^{\kappa \sum_i \varphi_i} \rangle_{\text{GM}}},$$

les espérances dans la dernière expression étant prises par rapport à la mesure du modèle gaussien massif introduite dans la section 6.1, avec masse $\sqrt{\kappa}$. Afin d'estimer ces espérances, il convient de se rappeler que la fonction génératrice des moments d'un vecteur gaussien X quelconque satisfait

$$\langle e^{t \cdot X} \rangle = e^{t \cdot \langle X \rangle + \frac{1}{2} \text{Var}(t \cdot X)}, \quad \forall t.$$

Puisque dans notre cas $\langle \varphi_i \rangle_{\text{GM}} = 0$, il ne reste plus qu'à estimer $\langle (\sum_i t_i \varphi_i)^2 \rangle_{\text{GM}}$ pour les vecteurs $t \equiv 1$ et $t_i = (-1)^{\|i\|_1}$. Bien entendu,

$$\langle (\sum_i t_i \varphi_i)^2 \rangle_{\text{GM}} = \sum_{i,j} t_i t_j \langle \varphi_i \varphi_j \rangle_{\text{GM}} = \sum_{i,j} t_i t_j G(i,j),$$

où $\mathcal{G}_L = (G(i,j))_{i,j \in \mathbb{T}_L^2}$ est l'inverse de la matrice $(\beta + \kappa)(\mathcal{I}_L - \mathcal{P}_L)$, où $\mathcal{I}_L = (\delta_{ij})_{i,j \in \mathbb{T}_L^2}$

et $\mathcal{P}_L = (P(i, j))_{i, j \in \mathbb{T}_L^2}$ avec

$$P(i, j) = \begin{cases} \frac{\beta}{\beta + \kappa} \frac{1}{2d} & \text{si } i, j \in \mathbb{T}_L^2, i \sim j, \\ 1 - \frac{\beta}{\beta + \kappa} & \text{si } i \in \mathbb{T}_L^2 \text{ et } j = \dagger, \\ 1 & \text{si } i = j = \dagger, \\ 0 & \text{sinon.} \end{cases}$$

(Voir la Section 6.2 pour le cas d'une boîte dans \mathbb{Z}^2 ; ici, on a gardé explicitement la dépendance en β et utilisé le fait que $m^2 = \kappa$.) La façon la plus simple d'obtenir les estimations désirées et d'observer que, sur le tore, il est très facile de déterminer les vecteurs propres de $(\beta + \kappa)(\mathcal{I}_L - \mathcal{P}_L)$, et donc de \mathcal{G}_L . En effet, ceux-ci sont donnés par les fonctions $e^{ip \cdot j}$, pour $p \in \mathbb{T}_L^{2, \star}$, puisque

$$\sum_j (\beta + \kappa)(\mathcal{I}_L - \mathcal{P}_L)_{ij} e^{ip \cdot j} = e^{ip \cdot i} (\beta + \kappa - \frac{\beta}{2d} \sum_{j \sim 0} e^{ip \cdot j}) = (\beta \hat{D}_p + \kappa) e^{ip \cdot i},$$

où l'on a introduit la transformée de Fourier du Laplacien discret,

$$\hat{D}_p \stackrel{\text{déf}}{=} 1 - \frac{1}{2d} \sum_{j \sim 0} e^{ip \cdot j}.$$

On en déduit que les fonctions propres de \mathcal{G}_L sont données par $e^{ip \cdot j}$, $p \in \mathbb{T}_L^{2, \star}$, avec valeurs propres $1/(\beta \hat{D}_p + \kappa)$.

On obtient à présent, pour $t \equiv 1$,

$$\sum_{i, j} t_i t_j G(i, j) = \sum_{i, j} e^{i0 \cdot i} G(i, j) e^{i0 \cdot j} = \frac{|\mathbb{T}_L^2|}{\kappa},$$

puisque $\hat{D}_0 = 0$. Similairement, pour $t_i = (-1)^{\|i\|_1}$,

$$\sum_{i, j} t_i t_j G(i, j) = \sum_{i, j} e^{i(\pi, \pi) \cdot i} G(i, j) e^{i(\pi, \pi) \cdot j} = \frac{|\mathbb{T}_L^2|}{2\beta + \kappa},$$

étant donné que $\hat{D}_{(\pi, \pi)} = 2$. On obtient donc finalement que

$$\mathfrak{z}(\begin{smallmatrix} + & - \\ - & + \end{smallmatrix})^{|\mathbb{T}_L^2|} \leq \exp\left\{\frac{1}{2}\kappa^2 |\mathbb{T}_L^2| \left(\frac{1}{2\beta + \kappa} - \frac{1}{\kappa}\right)\right\} = \exp\left\{-\frac{\beta\kappa}{2\beta + \kappa} |\mathbb{T}_L^2|\right\},$$

et le premier motif est sous contrôle.

On traite les deux autres de la même manière. Pour $\begin{smallmatrix} + & - \\ + & - \end{smallmatrix}$, on se retrouve avec $t_i = (-1)^{|i_1|}$, et donc

$$\sum_{i, j} t_i t_j G(i, j) = \sum_{i, j} e^{i(\pi, 0) \cdot i} G(i, j) e^{i(\pi, 0) \cdot j} = \frac{|\mathbb{T}_L^2|}{\beta + \kappa},$$

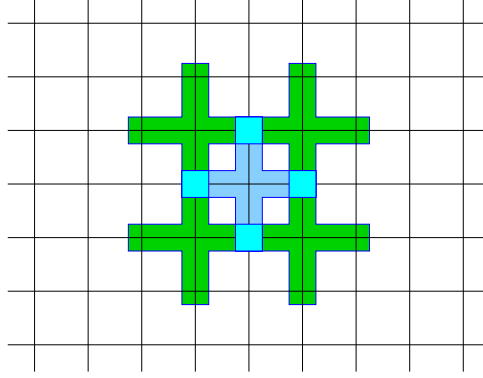


FIGURE 7.7 – Un bloc adapté aux réflexions diagonales (en bleu) et les quatre blocs voisins (en vert). Observez que ces blocs se chevauchent, deux blocs voisins se partageant les sommets (en cyan) appartenant à la diagonale les interchangeant par réflexion.

car $\hat{D}_{\pi,0} = 1$. On en déduit que

$$\mathfrak{z}\left(\begin{smallmatrix} + & - \\ + & - \end{smallmatrix}\right)^{|\mathbb{T}_L^2|} \leq \exp\left\{\frac{1}{2}\kappa^2|\mathbb{T}_L^2|\left(\frac{1}{\beta + \kappa} - \frac{1}{\kappa}\right)\right\} = \exp\left\{-\frac{\beta\kappa}{2\beta + 2\kappa}|\mathbb{T}_L^2|\right\}.$$

Le dernier motif, $\begin{smallmatrix} + & + \\ + & - \end{smallmatrix}$, est plus ennuyeux à analyser, car il ne génère pas une fonction propre de \mathcal{G}_L . On peut cependant s'en sortir à moindre coût en utilisant le fait que la mesure de Gibbs est positive sous réflexion à travers les diagonales passant par les sommets du tore. On décompose donc le tore en translatés d'un bloc de base (pas disjoints, car on considère ici des réflexions par des sommets) comme indiqué sur la Figure 7.7. On observe alors que dans la configuration correspondant au motif $\begin{smallmatrix} + & + \\ + & - \end{smallmatrix}$, la moitié de ces blocs contient le motif $\begin{smallmatrix} + & + \\ + & + \end{smallmatrix}$, et l'autre moitié le motif $\begin{smallmatrix} + & + \\ + & + \end{smallmatrix}$ (cf. Fig. 7.8). Ces motifs étant invariants sous les réflexions diagonales, une application de l'estimation de l'échiquier donne

$$\mathfrak{z}\left(\begin{smallmatrix} + & + \\ + & - \end{smallmatrix}\right) \leq \mathfrak{z}\left(\begin{smallmatrix} + & + \\ + & + \end{smallmatrix}\right)^{1/2} \mathfrak{z}\left(\begin{smallmatrix} - & + \\ + & - \end{smallmatrix}\right)^{1/2},$$

puisque les motifs $\begin{smallmatrix} + & + \\ + & + \end{smallmatrix}$ et $\begin{smallmatrix} + & + \\ + & + \end{smallmatrix}$ donnent respectivement lieu, par réflexions diagonales, aux configurations correspondant aux motifs $\begin{smallmatrix} + & + \\ + & + \end{smallmatrix}$ et $\begin{smallmatrix} - & + \\ + & - \end{smallmatrix}$. La preuve se ramène donc au premier motif analysé. \square

Corollaire 7.4.1. *Soit \mathfrak{B} l'événement que la plaquette en 0 soit mauvaise. Alors, pour tout $\epsilon > 0$, il existe $\beta_0, \kappa_0 > 0$ tels que $\mathfrak{z}(\mathfrak{B}) \leq \epsilon$, pour tout $\beta > \beta_0, \kappa > \kappa_0$.*

Démonstration du Corollaire 7.4.1. L'événement \mathfrak{B} pouvant être décomposé en un nombre fini de mauvais motifs, le résultat suit du lemme précédent et du Lemme 7.2.1. \square

Afin d'exploiter les estimations du corollaire précédent, on va utiliser un argument de Peierls. Soit $i \in \mathbb{T}_L^2 \setminus \{0\}$; lorsque $\sigma_0 = 1$ et $\sigma_i = -1$, on est assuré de l'existence d'un circuit γ de mauvaises plaquettes séparant 0 et i . En effet, soit toutes les plaquettes

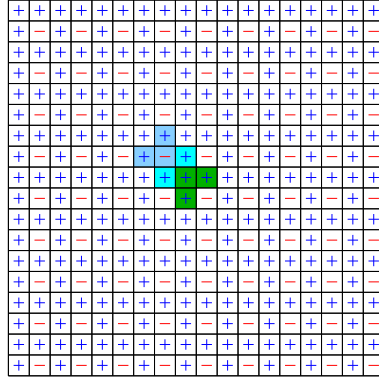


FIGURE 7.8 – Les deux motifs apparaissant dans les translatés du bloc de base.

contenant 0 sont mauvaises, soit il existe une composante connexe de bonnes plaquettes contenant 0 et pas i (puisque $\sigma_0 \neq \sigma_i$), et dans ce cas le bord de cette composante est formé de mauvaises plaquettes. On a donc

$$\mu_{L;\beta,\kappa}(\sigma_0 = 1, \sigma_i = -1) \leq \sum_{\gamma} \mu_{L;\beta,\kappa}(\bigcap_{t \in \gamma} \mathfrak{B}^{[t]}) \leq \sum_{\gamma} \mathfrak{z}(\mathfrak{B})^{|\gamma|},$$

la dernière inégalité résultant d'une application de l'estimation de l'échiquier. En raisonnant comme d'habitude, on en conclut que pour tout $\epsilon > 0$, il est possible de choisir β et κ suffisamment grands pour garantir que

$$\mu_{L;\beta,\kappa}(\sigma_0 = 1, \sigma_i = -1) \leq \epsilon.$$

D'un autre côté, par symétrie, $\mu_{L;\beta,\kappa}(\sigma_0 = 1) = 1/2$. Par conséquent,

$$\mu_{L;\beta,\kappa}(\sigma_i = -1 \mid \sigma_0 = 1) \leq 2\epsilon,$$

d'où l'on déduit que

$$\langle \sigma_i \mid \sigma_0 = 1 \rangle = 1 - 2\mu_{L;\beta,\kappa}(\sigma_i = -1 \mid \sigma_0 = 1) \geq 1 - 2\epsilon. \quad (7.10)$$

On a donc établi que la grande majorité des spins σ prend une des valeurs $+1$ ou -1 . Il reste à en déduire que les spins φ se concentrent dans le puits correspondant. La façon la plus simple de faire cela est d'étendre la représentation en marche aléatoire au cas présent. La procédure est tout à fait similaire (compléter le carré dans l'exposant de la gaussienne). Pour alléger les notations, on notera $\langle\langle a, b \rangle\rangle_L \stackrel{\text{déf}}{=} \sum_{i \in \mathbb{T}_L^2} a_i b_i$ le produit scalaire de deux fonctions a et b définies sur le tore. Fixons une configuration de spins σ . La mesure de Gibbs pour les spins φ , étant donnée cette configuration σ , correspond à l'Hamiltonien

$$-\frac{\beta + \kappa}{2} \langle\langle \varphi, (\mathcal{I}_L - \mathcal{P}_L) \varphi \rangle\rangle_L + \kappa \langle\langle \sigma, \varphi \rangle\rangle_L.$$

L'inverse \mathcal{G}_L de $\mathcal{I}_L - \mathcal{P}_L$ étant symétrique, on peut écrire

$$\langle\langle \sigma, \varphi \rangle\rangle_L = \langle\langle \sigma, \mathcal{G}_L(\mathcal{I}_L - \mathcal{P}_L) \varphi \rangle\rangle_L = \langle\langle \mathcal{G}_L \sigma, (\mathcal{I}_L - \mathcal{P}_L) \varphi \rangle\rangle_L,$$

et l'Hamiltonien peut donc être réécrit

$$-\frac{\beta + \kappa}{2} \langle \langle \varphi - \frac{\kappa}{\beta + \kappa} \mathcal{G}_L \sigma, (\mathcal{J}_L - \mathcal{P}_L) (\varphi - \frac{\kappa}{\beta + \kappa} \mathcal{G}_L \sigma) \rangle \rangle_L.$$

On voit donc que la moyenne de φ_i , conditionnellement à la configuration σ , est donnée par

$$\langle \varphi_i | \sigma \rangle_{L;\beta,\kappa} = \frac{\kappa}{\beta + \kappa} (\mathcal{G}_L \sigma)_i = \frac{\kappa}{\beta + \kappa} \sum_j G(i, j) \sigma_j.$$

On en déduit immédiatement une première borne, valable pour tout β, κ , et toute configuration σ :

$$\langle |\sum_i \varphi_i| | \sigma \rangle_{L;\beta,\kappa} = \frac{\kappa}{\beta + \kappa} \langle |\sum_{i,j} G(i, j) \sigma_j| | \sigma \rangle_{L;\beta,\kappa} \leq \frac{\kappa}{\beta + \kappa} \sum_{i,j} G(i, j).$$

Or,

$$\begin{aligned} \sum_{i,j} G(i, j) &= |\mathbb{T}_L^2| \sum_j G(0, j) = |\mathbb{T}_L^2| \sum_{n \geq 0} \sum_j \mathbb{P}_0(Z_n = j) \\ &= |\mathbb{T}_L^2| \sum_{n \geq 0} \mathbb{P}_0(\Upsilon_{\dagger} > n) = |\mathbb{T}_L^2| \mathbb{E}_0(\Upsilon_{\dagger}) = |\mathbb{T}_L^2| \frac{\beta + \kappa}{\kappa}. \end{aligned} \quad (7.11)$$

(Rappelons que Υ_{\dagger} est le moment où le processus Z_n meurt, cf. Section 6.3.1.) Par conséquent, pour tout β, κ , et toute configuration σ , on a

$$\langle |\sum_i \varphi_i| | \sigma \rangle_{L;\beta,\kappa} \leq |\mathbb{T}_L^2|.$$

Pour obtenir la borne inférieure, on utilise une fois de plus la symétrie :

$$\begin{aligned} \langle |\sum_i \varphi_i| \rangle_{L;\beta,\kappa} &= \sum_{\sigma: \sigma_0=1} \langle |\sum_i \varphi_i| | \sigma \rangle_{L;\beta,\kappa} \mu_{L;\beta,\kappa}(\sigma | \sigma_0 = 1) \\ &= \frac{\kappa}{\beta + \kappa} \sum_{\sigma: \sigma_0=1} |\sum_{i,j} G(i, j) \sigma_j| \mu_{L;\beta,\kappa}(\sigma | \sigma_0 = 1) \\ &\geq \frac{\kappa}{\beta + \kappa} \sum_{\sigma: \sigma_0=1} \sum_{i,j} G(i, j) \sigma_j \mu_{L;\beta,\kappa}(\sigma | \sigma_0 = 1) \\ &= \frac{\kappa}{\beta + \kappa} \sum_{i,j} G(i, j) \langle \sigma_j | \sigma_0 = 1 \rangle_{L;\beta,\kappa} \end{aligned}$$

Il suit donc de (7.10) et (7.11) que

$$\langle |\sum_i \varphi_i| \rangle_{L;\beta,\kappa} \geq (1 - 2\epsilon) \frac{\kappa}{\beta + \kappa} \sum_{i,j} G(i, j) = (1 - 2\epsilon) |\mathbb{T}_L^2|.$$

□

Les champs gibbsiens

Nous avons vu dans les chapitres précédents divers exemples de mesures de Gibbs, essentiellement associées à des champs markoviens en volume fini. Jusqu'à présent, la construction de mesures de Gibbs en volume infini nécessitait de considérer l'ensemble des points d'accumulation de suites de mesures en volume fini, pour des suites arbitraires de sous-ensembles du réseau infini et de conditions au bord. Cette approche, quoique intuitivement et physiquement satisfaisante, n'est pas très efficace pour en développer une théorie abstraite, et il serait préférable d'avoir un critère permettant de caractériser les mesures de Gibbs directement en volume infini.

Nous avons vu dans le chapitre 3 que les mesures de Gibbs (en volume fini) associées à des potentiels plus-proches-voisins possèdent automatiquement la propriété de Markov (et réciproquement, qu'à tout champ markovien en volume fini correspond une mesure de Gibbs avec potentiels plus-proches-voisins). On vérifie cependant aisément que les mesures en volume fini, associées à des potentiels quelconques, satisfont une sorte de généralisation de la propriété de Markov, dans le sens suivant. Soit $V \subseteq \mathbb{Z}^d$, et \mathbb{P}_V est une mesure de Gibbs sur V , avec espace de spin S (supposé discret pour simplifier) et potentiels $(\Phi_A)_{A \subseteq V}$. Alors, pour tout $D \subset V$ et toute configuration $\bar{\omega} \in S^{V \setminus D}$,

$$\mathbb{P}_V(\omega_i = \eta_i, \forall i \in D \mid \omega_i = \bar{\omega}_i, \forall i \in V \setminus D) = \frac{e^{-\mathcal{H}_D(\eta|\bar{\omega})}}{\mathbf{Z}_D(\bar{\omega})}, \quad \forall \eta \in S^D,$$

où on a posé

$$\mathcal{H}_D(\eta|\bar{\omega}) = \sum_{\substack{A \subseteq V \\ A \cap D \neq \emptyset}} \Phi_A(\eta\bar{\omega}),$$

et

$$\mathbf{Z}_D(\bar{\omega}) = \sum_{\eta \in S^D} e^{-\mathcal{H}_D(\eta|\bar{\omega})},$$

et noté $\eta\bar{\omega}$ la configuration coïncidant avec η sur D et avec $\bar{\omega}$ sur $V \setminus D$.

L'idée est de définir une mesure de Gibbs \mathbb{P} en volume infini, associée aux mêmes potentiels, par la généralisation naturelle de cette propriété : pour tout $D \subseteq \mathbb{Z}^d$ et toute

configuration $\bar{\omega}$ sur $\mathbb{Z}^d \setminus D$,

$$\mathbb{P}(\omega_i = \eta_i, \forall i \in D \mid \omega_i = \bar{\omega}_i, \forall i \notin D) = \frac{e^{-\mathcal{H}(\eta \mid \bar{\omega})}}{\mathbf{Z}_D(\bar{\omega})}, \quad \forall \eta \in S^D.$$

En d'autres termes, on cherche à définir un champ aléatoire sur \mathbb{Z}^d à partir de ses probabilités conditionnelles sur des sous-ensembles finis. Dans le reste de ce chapitre, nous allons donner un sens précis à cette définition, et investiguer les propriétés principales des champs ainsi définis. Nous ferons bien sûr aussi le lien avec la définition utilisée dans les chapitres précédents.

Afin de maintenir la discussion à un niveau aussi élémentaire que possible, nous nous restreindrons au cas où les spins prennent valeur dans un espace fini. Le cas général nécessite des outils plus élaborés d'analyse fonctionnelle, mais les idées et résultats principaux demeurent essentiellement inchangés (au moins tant que S est compact). Nous supposons également que le réseau sur lequel le champ aléatoire est défini est toujours \mathbb{Z}^d .

8.1 Cadre mathématique

Nous commençons par décrire la structure mathématique nécessaire, et introduire les notations correspondantes. Pour des raisons pratiques, certaines de ces dernières diffèrent quelque peu des notations correspondantes utilisées précédemment.

8.1.1 Espace de configurations.

Soit S un ensemble fini, que l'on appellera **espace de spin**. On appelle **espace de configurations** l'ensemble $\Omega \stackrel{\text{def}}{=} S^{\mathbb{Z}^d}$. Étant donné une configuration $\omega \in \Omega$ et un sous-ensemble $\Lambda \subset \mathbb{Z}^d$, on notera $\omega|_{\Lambda} \in S^{\Lambda} \equiv \Omega_{\Lambda}$ la restriction de ω à Λ . Étant donné $\Lambda, \Delta \subset \mathbb{Z}^d$ avec $\Lambda \cap \Delta = \emptyset$, la configuration $\omega|_{\Lambda} \omega'|_{\Delta} \in \Omega_{\Lambda \cup \Delta}$ est celle coïncidant avec $\omega|_{\Lambda}$ sur Λ et avec $\omega'|_{\Delta}$ sur Δ . On utilisera aussi la convention que $\omega|_{\Lambda} \omega' \equiv \omega|_{\Lambda} \omega'|_{\Lambda^c}$, où $\Lambda^c \stackrel{\text{def}}{=} \mathbb{Z}^d \setminus \Lambda$.

8.1.2 Structure d'espace de probabilité.

On note \mathcal{S} la tribu associée à S , c'est-à-dire l'ensemble des parties de S . La tribu associée à Ω est le produit $\mathcal{F} \stackrel{\text{def}}{=} \mathcal{S}^{\mathbb{Z}^d}$; elle est engendrée par l'ensemble $\text{Cyl} \stackrel{\text{def}}{=} \{\mathcal{C}_{\omega, \Lambda} : \Lambda \Subset \mathbb{Z}^d, \omega \in \Omega_{\Lambda}\}$ des cylindres $\mathcal{C}_{\omega, \Lambda} \stackrel{\text{def}}{=} \{\omega' \in \Omega : \omega'|_{\Lambda} = \omega\}$.

Un **champ aléatoire** sur \mathbb{Z}^d est une mesure de probabilité sur (Ω, \mathcal{F}) ; on note $\mathcal{M}_1(\Omega, \mathcal{F})$ l'ensemble de ces champs aléatoires.

Deux sous-tribus de \mathcal{F} jouent un rôle important dans la suite :

- La tribu $\mathcal{F}_{\Lambda} = \mathcal{S}^{\Lambda}$, $\Lambda \subset \mathbb{Z}^d$, générée par l'ensemble $\text{Cyl}_{\Lambda} \stackrel{\text{def}}{=} \{\mathcal{C}_{\omega, \Delta} : \Delta \Subset \Lambda, \omega \in \Omega_{\Delta}\}$. Cette tribu contient tous les événements ne dépendant que de la restriction de la configuration à la région Λ . En particulier, une fonction $f : \Omega \rightarrow \mathbb{R}$ \mathcal{F} -mesurable est \mathcal{F}_{Λ} -mesurable si et seulement si $f(\omega) = f(\omega')$ dès que $\omega|_{\Lambda} = \omega'|_{\Lambda}$. On emploiera

souvent dans ce cas l'abus de notation $f(\omega) = f(\omega_\Lambda)$. Dans le cas où $\Lambda \Subset \mathbb{Z}^d$, ces fonctions correspondent aux **observables microscopiques**; on les appelle fonctions **locales**.

- La tribu asymptotique $\mathcal{F}_\infty \stackrel{\text{déf}}{=} \bigcap_{\Lambda \in \mathbb{Z}^d} \mathcal{F}_{\Lambda^c}$. Elle contient les événements dont la réalisation est indépendante de l'état des spins de n'importe quelle région finie. En particulier, une fonction $f : \Omega \rightarrow \mathbb{R}$ \mathcal{F} -mesurable sera \mathcal{F}_∞ -mesurable si et seulement si $f(\omega) = f(\omega')$ dès que $\omega|_{\Lambda^c} = \omega'|_{\Lambda^c}$ pour un $\Lambda \Subset \mathbb{Z}^d$. Un exemple d'une telle fonction est $f(\omega) = \lim_{n \rightarrow \infty} |\Lambda_n|^{-1} \sum_{i \in \Lambda_n} \omega_i$ si la limite existe, et 0 sinon. Ces fonctions correspondent aux **observables macroscopiques**.

Dans la suite, on utilisera aussi les notations suivantes : si $\mu \in \mathcal{M}_1(\Omega, \mathcal{F})$, et $f : \Omega \rightarrow \mathbb{R}$ est une fonction \mathcal{F} -mesurable μ -intégrable positive, on écrit

$$\mu(f) \stackrel{\text{déf}}{=} \int \mu(d\omega) f(\omega).$$

Si de plus $\mu(f) = 1$, alors on définit $f\mu \in \mathcal{M}_1(\Omega, \mathcal{F})$ par

$$f\mu(A) \stackrel{\text{déf}}{=} \int \mu(d\omega) f(\omega) \mathbf{1}_A(\omega),$$

pour tout $A \in \mathcal{F}$.

8.1.3 Structure topologique et métrique.

On munit à présent l'espace Ω d'une topologie \mathcal{T} , compatible avec la tribu \mathcal{F} , dans le sens que \mathcal{F} est la tribu borélienne associée à la topologie \mathcal{T} (c'est-à-dire \mathcal{F} est la plus petite tribu telle que les ouverts de \mathcal{T} soient mesurables). Concrètement, ceci est assuré en définissant \mathcal{T} comme la topologie engendrée par l'ensemble des cylindres, Cyl , puisque celui-ci est dénombrable. Ceci correspond en fait à prendre comme topologie sur S la topologie discrète (toute partie de S est déclarée ouverte), et sur Ω le produit correspondant.

Lemme 8.1.1. *La topologie \mathcal{T} est induite par la métrique*

$$d(\omega, \omega') = \sum_{i \in \mathbb{Z}^d} 2^{-\|i\|} \mathbf{1}_{\{\omega_i \neq \omega'_i\}}. \quad (8.1)$$

Démonstration. Soit $B(\omega, r)$ la boule ouverte de rayon r centrée en $\omega \in \Omega$. Il suffit de montrer que l'ensemble de ces boules forme une base pour \mathcal{T} . Soit $\Lambda \Subset \mathbb{Z}^d$ et $C \in \text{Cyl}_\Lambda$ un cylindre. Posons $r_\Lambda = \inf_{i \in \Lambda} 2^{-\|i\|}$. Alors, pour tout $\omega \in C$,

$$\{\omega' \in \Omega : d(\omega, \omega') < r_\Lambda\} \subset \{\omega' \in \Omega : \omega'|_\Lambda = \omega|_\Lambda\} \subset C.$$

Par conséquent, on a bien

$$C = \bigcup_{\omega \in C} B(\omega, r_\Lambda).$$

□

Il est important de comprendre intuitivement ce que signifie “converger” dans cette topologie : une suite de configurations $(\omega^n)_n$ converge vers une configuration ω si et seulement si pour toute région finie $\Lambda \Subset \mathbb{Z}^d$, on peut trouver n_0 tel que

$$\omega^n|_{\Lambda} = \omega|_{\Lambda}, \quad \forall n \geq n_0. \quad (8.2)$$

Cette topologie munit (Ω, \mathcal{T}) d’une riche structure.

Théorème 8.1.1. (Ω, \mathcal{T}) est un espace polonais¹ compact.

Démonstration. Soit $(\omega^n)_n$ une suite de Cauchy pour la métrique (8.1). Soit $\Lambda \Subset \mathbb{Z}^d$ et $C \in \text{Cyl}_{\Lambda}$; on pose comme précédemment $r_{\Lambda} = \inf_{i \in \Lambda} 2^{-\|i\|}$. On a alors, $\forall m, n > N(\Lambda)$,

$$d(\omega^n, \omega^m) < r_{\Lambda} \implies \omega^n|_{\Lambda} = \omega^m|_{\Lambda},$$

et donc la suite $(\omega^n)_n$ converge vers la configuration ω telle que $\omega_i = \lim_{n \rightarrow \infty} \omega_i^n$. Ceci montre que l’espace métrique (Ω, d) est complet.

Soient $\omega \in \Omega$ une configuration et $\epsilon > 0$. On peut alors trouver $\Lambda \Subset \mathbb{Z}^d$ tel que la configuration ω' définie par $\omega'|_{\Lambda} = \omega|_{\Lambda}$ et $\omega'|_{\Lambda^c} \equiv s$, pour $s \in S$, satisfasse $d(\omega, \omega') < \epsilon$. Par conséquent, l’ensemble dénombrable des configurations asymptotiquement constantes, $\{\omega \in \Omega : \omega|_{\Lambda^c} \equiv s, s \in S, \Lambda \Subset \mathbb{Z}^d\}$, est dense, et (Ω, \mathcal{T}) est séparable.

La compacité suit évidemment du Théorème de Tychonoff, cependant, dans le cadre considéré ici, on peut en donner des preuves plus directes. Ici, nous allons utiliser un argument standard de diagonalisation. Puisque (Ω, \mathcal{T}) est un espace métrisable, il est suffisant de montrer que toute suite $(\omega^n)_n$ admet une sous-suite convergente. On numérote les sommets de \mathbb{Z}^d de façon arbitraire : t_1, t_2, \dots . S étant fini, on peut extraire une sous-suite indicée par $n_k^1, k \geq 1$, et $s_1 \in S$, tels que $\omega_{t_1}^{n_k^1} = s_1$, pour tout $k \geq 1$. On peut ensuite extraire de cette sous-suite une nouvelle sous-suite d’indices $n_k^2, k \geq 1$, et $s_2 \in S$, tels que $\omega_{t_2}^{n_k^2} = s_2$, pour tout $k \geq 1$. En poursuivant cette procédure on obtient, pour tout $j \geq 1$, une suite d’indices $n_k^j, k \geq 1$, et $s_j \in S$, tels que $\omega_{t_i}^{n_k^j} = s_i$ pour tout $k \geq 1$ et tout $i \leq j$. On vérifie alors immédiatement que la sous-suite $(\omega^{n_k^j})_k$ converge vers la configuration donnée par s_i au sommet t_i . \square

Une fonction $f : \Omega \rightarrow \mathbb{R}$ est donc continue en ω si et seulement si, pour tout $\epsilon > 0$, on peut trouver $\Lambda \Subset \mathbb{Z}^d$ tel que $\sup_{\omega' \in \Omega} |f(\omega|_{\Lambda} \omega') - f(\omega)| < \epsilon$. L’espace (Ω, \mathcal{T}) étant compact, une fonction continue partout est nécessairement uniformément continue et bornée. On note $\mathcal{C}(\Omega)$ l’ensemble des fonctions continues ; $\mathcal{C}(\Omega)$ est un espace de Banach pour la norme $\|f\|_{\infty} \stackrel{\text{déf}}{=} \sup_{\omega} |f(\omega)|$. De plus, \mathcal{T} étant la tribu Borélienne associée à la topologie \mathcal{T} , on en déduit que toute fonction $f \in \mathcal{C}(\Omega)$ est mesurable.

Manifestement, les fonctions locales sont toutes continues. En fait, on vérifie immédiatement qu’elles forment un sous-ensemble dense de $(\mathcal{C}(\Omega), \|\cdot\|_{\infty})$, puisque la continuité de f signifie que pour tout $\epsilon > 0$ et toute configuration $\omega' \in \Omega$, il existe $\Lambda \Subset \mathbb{Z}^d$ tel que

$$\sup_{\omega} |f(\omega|_{\Lambda} \omega') - f(\omega)| < \epsilon,$$

1. Un espace polonais est un espace séparable, homéomorphe à un espace métrique complet.

et que la fonction $\omega \mapsto f(\omega_\Lambda \omega')$ est locale.

À l'opposé, les observables macroscopiques non constantes sont toutes des fonctions discontinues.

Remarque 8.1.1. *Dans un cadre général, la fermeture de l'ensemble des fonctions locales, par rapport à la norme $\|\cdot\|_\infty$, est appelé ensemble des fonctions **quasilocales**. Dans le cas présent (S fini), les fonctions quasilocales coïncident avec les fonctions continues.*

8.1.4 Convergence des champs aléatoires

On va à présent munir l'espace $\mathcal{M}_1(\Omega, \mathcal{F})$ des champs aléatoires sur Ω d'une structure topologique appropriée. Cette topologie est celle générée par la notion de convergence suivante : une suite $(\mu_n)_{n \geq 1}$ de mesures dans $\mathcal{M}_1(\Omega, \mathcal{F})$ converge vers la mesure $\mu \in \mathcal{M}_1(\Omega, \mathcal{F})$ si et seulement si $\mu_n(f) \rightarrow \mu(f)$ pour toute fonction quasilocale. Cette topologie est appelée **topologie de la convergence locale**. Puisque les fonctions quasilocales coïncident ici avec les fonctions continues bornées, cette topologie n'est autre que la topologie faible.

Évidemment, les fonctions locales étant denses dans $\mathcal{C}(\Omega)$, il suffit d'avoir convergence de l'espérance des fonctions locales. En fait, puisque celles-ci s'écrivent comme combinaisons linéaires finies d'indicatrices de cylindre, il est même suffisant d'avoir $\mu_n(\mathcal{C}_{\omega, \Lambda}) \rightarrow \mu(\mathcal{C}_{\omega, \Lambda})$ pour tout cylindre $\mathcal{C}_{\omega, \Lambda}$.

Théorème 8.1.2. *Muni de cette topologie, l'espace $\mathcal{M}_1(\Omega, \mathcal{F})$ est compact et métrisable.*

Démonstration. Soit $(\Lambda_k)_{k \geq 1}$ une suite arbitraire de sous-ensembles finis de \mathbb{Z}^d telle que $\Lambda_k \nearrow \mathbb{Z}^d$. Nous allons vérifier que la métrique

$$d(\mu, \nu) \stackrel{\text{déf}}{=} \sup_{k \geq 1} \frac{1}{k} \max_{\omega \in \Omega_{\Lambda_k}} |\mu(\mathcal{C}_{\omega, \Lambda_k}) - \nu(\mathcal{C}_{\omega, \Lambda_k})|$$

convient. Supposons que $\mu_n \rightarrow \mu$. Fixons $\epsilon > 0$, et $K > 2/\epsilon$. Soit $N = N(\epsilon)$ tel que

$$\max_{\omega \in \Omega_{\Lambda_k}} |\mu_n(\mathcal{C}_{\omega, \Lambda_k}) - \mu(\mathcal{C}_{\omega, \Lambda_k})| \leq \epsilon,$$

pour tout $k \leq K$. On a bien, pour tout $n > N$,

$$d(\mu_n, \mu) = \sup_{k \geq 1} \frac{1}{k} \max_{\omega \in \Omega_{\Lambda_k}} |\mu_n(\mathcal{C}_{\omega, \Lambda_k}) - \mu(\mathcal{C}_{\omega, \Lambda_k})| \leq \epsilon,$$

en utilisant la borne triviale $|\mu_n(\mathcal{C}_{\omega, \Lambda_k}) - \mu(\mathcal{C}_{\omega, \Lambda_k})| \leq 2$ (et la définition de K) pour traiter les $k > K$.

Réciproquement, supposons que $d(\mu_n, \mu) \rightarrow 0$. Soit $\Lambda \Subset \mathbb{Z}^d$ et $C \in \text{Cyl}_\Lambda$. Soit k suffisamment grand pour que $\Lambda \subseteq \Lambda_k$. On a alors

$$|\mu_n(C) - \mu(C)| \leq \sum_{\omega \in \Omega_{\Lambda_k} \cap C} |\mu_n(\mathcal{C}_{\omega, \Lambda_k}) - \mu(\mathcal{C}_{\omega, \Lambda_k})|,$$

et chacun des termes de la somme tend vers 0 lorsque $n \rightarrow \infty$.

$\mathcal{M}_1(\Omega, \mathcal{F})$ étant métrisable, il suffit de vérifier que cet espace est séquentiellement compact. Soit $(\mu_n)_{n \geq 1}$ une suite dans $\mathcal{M}_1(\Omega, \mathcal{F})$. L'ensemble des événements locaux étant dénombrable, on peut en extraire, en utilisant à nouveau un procédé de diagonalisation, une sous-suite $(\mu_{n_k})_{k \geq 1}$ telle que $\mu_{n_k}(A) \rightarrow \mu^*(A)$, pour tout événement A local. La conclusion suit en répétant l'argument donné dans la Remarque 2.4.1. \square

8.2 Probabilités conditionnelles et spécifications

8.2.1 Noyaux de probabilité

Dans cette sous-section, nous allons introduire une classe d'objets importants en théorie des probabilités, les noyaux de probabilité, ainsi que quelques notations très pratiques qui leurs sont associées.

Un **noyau de probabilité** d'un espace probabilisable (Ω, \mathcal{F}) vers un espace probabilisable (Ω', \mathcal{F}') est une fonction

$$\pi(\cdot | \cdot) : \mathcal{F}' \times \Omega \rightarrow [0, 1]$$

telle que

- $\pi(\cdot | \omega)$ est une mesure de probabilité sur (Ω', \mathcal{F}') pour chaque $\omega \in \Omega$;
- $\pi(A' | \cdot)$ est \mathcal{F} -mesurable pour chaque $A' \in \mathcal{F}'$.

Étant donné deux noyaux de probabilité, π de (Ω, \mathcal{F}) vers (Ω', \mathcal{F}') et π' de (Ω', \mathcal{F}') vers $(\Omega'', \mathcal{F}'')$, on définit un nouveau noyau de probabilité $\pi\pi'$ de (Ω, \mathcal{F}) vers $(\Omega'', \mathcal{F}'')$ par

$$\pi\pi'(A'' | \omega) \stackrel{\text{déf}}{=} \pi(\pi'(A'' | \cdot) | \omega) = \int \pi(d\omega' | \omega) \pi'(A'' | \omega').$$

En particulier, lorsque le premier de ces noyaux de probabilité est simplement une mesure de probabilité, on obtient une application entre mesures de probabilité : si $\mu \in \mathcal{M}_1(\Omega, \mathcal{F})$ et π est un noyau de probabilité de (Ω, \mathcal{F}) vers (Ω', \mathcal{F}') , alors $\mu\pi \in \mathcal{M}_1(\Omega', \mathcal{F}')$ est définie par

$$\mu\pi(A') \stackrel{\text{déf}}{=} \int \mu(d\omega) \pi(A' | \omega).$$

Similairement, on définit l'action d'un noyau de probabilité π sur une fonction $f : \Omega \rightarrow \mathbb{R}$ par $\pi f(\cdot) \stackrel{\text{déf}}{=} \pi(f | \cdot) = \int \pi(d\omega' | \cdot) f(\omega')$.

8.2.2 Probabilités conditionnelles

Commençons par rappeler la notion cruciale d'espérance conditionnelle. Soit $(\Omega, \mathcal{F}, \mu)$ un espace de probabilité, \mathcal{B} une sous-tribu de \mathcal{F} et $f : \Omega \rightarrow \mathbb{R}$ une fonction \mathcal{F} -mesurable, μ -intégrable. Une **espérance conditionnelle** de f étant donné \mathcal{B} est une fonction

$$\mu(f | \mathcal{B})(\cdot) : \Omega \rightarrow \mathbb{R}$$

telle que

1. $\mu(f | \mathcal{B})$ est \mathcal{B} -mesurable.

2. Pour toute fonction g \mathcal{B} -mesurable et bornée,

$$\mu(g\mu(f|\mathcal{B})) = \mu(gf). \quad (8.3)$$

L'existence de telles espérances conditionnelles suit du Théorème de Radon-Nikodým². Bien sûr, il n'y a pas unicité de l'espérance conditionnelle, mais on vérifie aisément que deux versions de l'espérance conditionnelle diffèrent au plus sur un ensemble de mesure 0. Il est facile de voir que la condition (8.3) implique que

$$\mu(fg|\mathcal{B}) = g\mu(f|\mathcal{B}), \quad \mu\text{-p.s.}, \quad (8.4)$$

pour toute fonction bornée g \mathcal{B} -mesurable. Une autre propriété importante (et immédiate à vérifier) est que

$$\mu(\mu(f|\mathcal{B})|\mathcal{B}') = \mu(f|\mathcal{B}'), \quad \mu\text{-p.s.}, \quad (8.5)$$

pour toute sous-tribu $\mathcal{B}' \subset \mathcal{B}$.

On pourrait être tenté de passer des espérances conditionnelles ainsi construites à des probabilités conditionnelles $\mu_{|\mathcal{B}}(\cdot|\omega)$, en posant pour $A \in \mathcal{F}$,

$$\mu_{|\mathcal{B}}(A|\omega) = \mu(\mathbf{1}_A|\mathcal{B})(\omega).$$

On vérifie en particulier qu'on a les propriétés suivantes :

- $\mu_{|\mathcal{B}}(\Omega|\cdot) = 1$ and $\mu_{|\mathcal{B}}(\emptyset|\cdot) = 0$, $\mu\text{-p.s.}$;
- pour tout $A \in \mathcal{F}$, $0 \leq \mu_{|\mathcal{B}}(A|\cdot) \leq 1$, $\mu\text{-p.s.}$;
- pour toute collection $\{A_i\}$ d'ensembles disjoints de \mathcal{F} ,

$$\mu_{|\mathcal{B}}\left(\bigcup_i A_i \mid \cdot\right) = \sum_i \mu_{|\mathcal{B}}(A_i|\cdot), \quad \mu\text{-p.s.};$$

- Si $A \in \mathcal{B}$, $\mu_{|\mathcal{B}}(A|\cdot) = \mathbf{1}_A(\cdot)$, $\mu\text{-p.s.}$.

Le problème est que ces propriétés ne sont valables que μ -presque partout, et que les ensembles de mesure nulle apparaissant ci-dessus *dépendent* des ensembles considérés. Comme il y en a une quantité non-dénombrable, on ne peut plus rien dire du tout.

Ce que l'on désire, c'est avoir un unique ensemble de mesure nulle à l'extérieur duquel toutes ces propriétés sont vérifiées simultanément. En d'autres termes, on voudrait avoir une bonne version de la probabilité conditionnelle :

Une probabilité conditionnelle régulière de μ étant donné une sous-tribu \mathcal{B} de \mathcal{F} est un noyau de probabilité $\mu_{|\mathcal{B}}$ de (Ω, \mathcal{F}) vers lui-même tel que, pour toute fonction f \mathcal{F} -mesurable μ -intégrable

$$\mu_{|\mathcal{B}}(f|\cdot) = \mu(f|\mathcal{B})(\cdot), \quad \mu\text{-p.s.}$$

Il n'est pas évident que de telles probabilités conditionnelles régulières existent. Dans le cadre qui nous intéresse, ceci suit du théorème suivant, que l'on ne démontrera pas.

Théorème 8.2.1. *Toute mesure sur un espace polonais possède des probabilités conditionnelles régulières par rapport à toute sous-tribu dénombrablement engendrée.*

2. Si $f \geq 0$, on introduit la mesure ν sur (Ω, \mathcal{B}) par $\nu(B) \stackrel{\text{déf}}{=} \mu(f\mathbf{1}_B)$, $\forall B \in \mathcal{B}$. Manifestement, ν est finie, positive et $\nu \ll \mu$. Par conséquent, le théorème de Radon-Nikodým implique l'existence d'une fonction g \mathcal{B} -mesurable et positive telle que $\nu(B) = \mu(g\mathbf{1}_B)$, $\forall B \in \mathcal{B}$. On peut alors prendre $\mu(f|\mathcal{B}) = g$. Pour une fonction f générale, on prend $\mu(f|\mathcal{B}) = \mu(f^+|\mathcal{B}) - \mu(f^-|\mathcal{B})$.

8.2.3 Spécifications

Une notion centrale dans la théorie des champs gibbsiens est celle de spécification : une **spécification** sur (Ω, \mathcal{F}) est une famille $\Pi = \{\pi_\Lambda : \Lambda \Subset \mathbb{Z}^d\}$ de noyaux de probabilité de (Ω, \mathcal{F}) vers lui-même telle que

- $\pi_\Lambda(f | \cdot)$ est \mathcal{F}_{Λ^c} -mesurable pour chaque $\Lambda \Subset \mathbb{Z}^d$ et chaque fonction \mathcal{F} -mesurable bornée f ;
- chaque π_Λ est **propre** :

$$\pi_\Lambda(fg | \omega) = g(\omega)\pi_\Lambda(f | \omega), \quad (8.6)$$

pour tout $\omega \in \Omega$, g \mathcal{F}_{Λ^c} -mesurable, et f mesurable bornée ;

- la famille Π est **cohérente** :

$$\pi_\Lambda \pi_{\Lambda'} = \pi_\Lambda, \quad (8.7)$$

si $\Lambda' \subseteq \Lambda$.

On dira qu'un champ aléatoire $\mu \in \mathcal{M}_1(\Omega, \mathcal{F})$ est **compatible** avec la spécification Π si

$$\mu \pi_\Lambda = \mu, \quad (8.8)$$

pour tout $\Lambda \Subset \mathbb{Z}^d$. On note $\mathcal{G}(\Pi)$ l'ensemble des champs aléatoires compatibles avec Π . On appelle **équations DLR** (pour Dobrushin, Lanford et Ruelle) le système d'équations (8.8).

L'intérêt des spécifications est rendu manifeste par le théorème suivant.

Théorème 8.2.2. *Soit $\Pi = \{\pi_\Lambda : \Lambda \Subset \mathbb{Z}^d\}$ une spécification et μ un champ aléatoire sur (Ω, \mathcal{F}) . Les propriétés suivantes sont équivalentes :*

1. μ est compatible avec Π .
2. $\pi_\Lambda(\cdot | \omega) = \mu|_{\mathcal{F}_{\Lambda^c}}(\cdot | \omega)$ pour μ -presque tout $\omega \in \Omega$, et tout $\Lambda \Subset \mathbb{Z}^d$.
3. $\pi_\Lambda(f | \cdot) = \mu(f | \mathcal{F}_\Lambda^c)(\cdot)$, μ -presque sûrement pour chaque $\Lambda \Subset \mathbb{Z}^d$ et chaque f μ -intégrable.

Démonstration. L'équivalence entre 2 et 3 est claire.

Montrons $1 \implies 3$. Puisque π_Λ est propre et $\mu \pi_\Lambda = \mu$, on voit que

$$\int \mu(d\omega) g(\omega) \pi_\Lambda(f | \omega) = \int \mu(d\omega) \pi_\Lambda(gf | \omega) = \mu \pi_\Lambda(fg) = \mu(fg),$$

pour tout f μ -intégrable et g bornée et \mathcal{F}_{Λ^c} -mesurable. π_Λ fournit donc bien une version de l'espérance conditionnelle.

Montrons que $3 \implies 1$. Par définition de l'espérance conditionnelle,

$$\mu \pi_\Lambda(f) = \int \mu(d\omega) \pi_\Lambda(f | \omega) = \int \mu(d\omega) \mu(f | \mathcal{F}_\Lambda^c)(\omega) = \mu(f),$$

pour tout f \mathcal{F} -mesurable μ -intégrable, ce qui montre que μ est compatible avec Π . \square

Ce théorème montre l'intérêt des spécifications : fournir une famille cohérente de probabilités conditionnelles régulières *sans faire référence à un champ aléatoire*. Ceci est particulièrement important puisque notre but est de construire un champ aléatoire à partir de ses probabilités conditionnelles, et qu'on ne connaît donc pas ce champ *a priori*. Si l'on se donne une spécification Π , le problème auquel nous sommes confrontés est la détermination de l'ensemble $\mathcal{G}(\Pi)$ des champs aléatoires compatibles avec Π . Le théorème nous assure que les probabilités conditionnelles de ces champs seront bien données par la spécification de départ.

Dans le cadre qui nous intéresse pour ce chapitre, S est un ensemble fini. Par conséquent, chaque noyau propre de probabilité π_Λ est absolument continu par rapport au produit de la mesure de comptage sur Ω_Λ et de la masse de Dirac sur Ω_{Λ^c} . Il est donc particulièrement agréable de travailler en termes des densités correspondantes : les **densités de spécifications** d'une spécification $\Pi = \{\pi_\Lambda : \Lambda \in \mathbb{Z}^d\}$ sont les fonctions $\gamma_\Lambda(\cdot | \cdot) : \Omega_\Lambda \times \Omega_{\Lambda^c} \rightarrow [0, 1]$ définies par

$$\gamma_\Lambda(\omega|_\Lambda | \omega|_{\Lambda^c}) \stackrel{\text{déf}}{=} \pi_\Lambda(\mathcal{C}_{\omega, \Lambda} | \omega),$$

c'est-à-dire les fonctions telles que

$$\pi_\Lambda(f | \bar{\omega}) = \sum_{\omega|_\Lambda \in \Omega_\Lambda} f(\omega|_\Lambda \bar{\omega}) \gamma_\Lambda(\omega|_\Lambda | \bar{\omega}|_{\Lambda^c}),$$

pour toute fonction f bornée mesurable.

8.3 Potentiels et mesures de Gibbs

Les spécifications gibbsiennes auxquelles nous allons nous intéresser sont construites à partir d'une famille de fonctions (Hamiltoniens) \mathcal{H}_Λ , $\Lambda \in \mathbb{Z}^d$, par la formule $\gamma_\Lambda \propto e^{-\mathcal{H}_\Lambda}$, comme nous l'avons fait dans les précédents chapitres du cours. En fait, les Hamiltoniens qui vont nous intéresser auront une structure particulière : ce seront, comme dans le chapitre 3, des sommes de termes locaux, appelés potentiels, qui sont les véritables quantités fondamentales dans cette approche.

Un **potentiel d'interaction** (ou simplement potentiel, ou interaction) est une famille $\Phi = \{\phi_A : A \in \mathbb{Z}^d\}$ de fonctions $\phi_A : \Omega \rightarrow \mathbb{R}$ telles que ϕ_A est \mathcal{F}_A -mesurable pour chaque $A \in \mathbb{Z}^d$.

On définit ensuite l'**Hamiltonien** dans $\Lambda \in \mathbb{Z}^d$ avec condition au bord $\bar{\omega} \in \Omega$ par

$$\mathcal{H}_\Lambda^\Phi(\omega|_\Lambda | \bar{\omega}|_{\Lambda^c}) \stackrel{\text{déf}}{=} \sum_{\substack{A \in \mathbb{Z}^d \\ A \cap \Lambda \neq \emptyset}} \phi_A(\omega|_\Lambda \bar{\omega}),$$

pour $\omega, \bar{\omega}$ tels que cette somme existe.

Nous dirons que Φ est **sommable** en $\bar{\omega}$ si $\mathcal{H}_\Lambda^\Phi(\omega|_\Lambda | \bar{\omega}|_{\Lambda^c})$ existe pour tout $\Lambda \in \mathbb{Z}^d$ et $\omega|_\Lambda \in \Omega_\Lambda$. On notera Ω^Φ l'ensemble des configurations en lesquelles Φ est sommable.

8.3. POTENTIELS ET MESURES DE GIBBS

Nous dirons que Φ est uniformément absolument sommable si

$$\sum_{A \ni i} \|\phi_A\|_\infty < \infty,$$

pour tout $i \in \mathbb{Z}^d$. On notera \mathfrak{B}_1 l'ensemble des potentiels uniformément absolument sommables.

Les poids de Boltzmann associés à l'interaction $\Phi \in \mathfrak{B}_1$ sont définis par

$$\gamma_\Lambda^\Phi(\omega|_\Lambda | \bar{\omega}|_{\Lambda^c}) \stackrel{\text{déf}}{=} \frac{e^{-\mathcal{H}_\Lambda^\Phi(\omega|_\Lambda | \bar{\omega}|_{\Lambda^c})}}{Z_\Lambda^\Phi(\bar{\omega})},$$

pour $\Lambda \Subset \mathbb{Z}^d$ et $\omega, \bar{\omega} \in \Omega$. La fonction de partition $Z_\Lambda^\Phi(\bar{\omega})$ est donnée par

$$\sum_{\omega|_\Lambda \in \Omega_\Lambda} e^{-\mathcal{H}_\Lambda^\Phi(\omega|_\Lambda | \bar{\omega}|_{\Lambda^c})}.$$

On vérifie facilement que les poids de Boltzmann forment les densités d'une spécification, notée Π^Φ et appelée la **spécification gibbsienne** associée à l'interaction Φ . Il est également important de noter que les poids de Boltzmann correspondent exactement aux mesures de Gibbs en volume fini des chapitres précédents.

Les **mesures de Gibbs** associées à une interaction $\Phi \in \mathfrak{B}_1$ sont les mesures compatibles avec Π^Φ . Afin d'alléger la notation, on écrit dans ce cas $\mathcal{G}(\Phi)$ au lieu de $\mathcal{G}(\Pi^\Phi)$.

Un champ aléatoire est une **mesure de Gibbs** (ou champ aléatoire gibbsien) s'il existe une interaction $\Phi \in \mathfrak{B}_1$ telle que $\mu \in \mathcal{G}(\Pi^\Phi)$.

Pour $\Lambda \subset \mathbb{Z}^d$, notons $\Delta_r(\Lambda) \stackrel{\text{déf}}{=} \{i \in \mathbb{Z}^d \setminus \Lambda : \inf_{j \in \Lambda} \|j - i\|_1 \leq r\}$. On dit qu'une spécification $\Pi = \{\pi_\Lambda : \Lambda \Subset \mathbb{Z}^d\}$ est r -markovienne, $r \geq 0$, si $\pi_\Lambda(A|\cdot)$ est $\mathcal{F}_{\Delta_r(\Lambda)}$ -mesurable, pour tout $\Lambda \Subset \mathbb{Z}^d$ et $A \in \mathcal{F}_\Lambda$. Ceci est équivalent à demander que, pour tout $\Lambda \Subset \mathbb{Z}^d$, tout $A \in \mathcal{F}_\Lambda$, et toute paire de configurations $\omega, \omega' \in \Omega$ telles que $\omega|_{\Delta_r(\Lambda)} = \omega'|_{\Delta_r(\Lambda)}$,

$$\pi_\Lambda(A|\omega) = \pi_\Lambda(A|\omega').$$

On dit que Π est markovienne si elle est r -markovienne pour un $r \geq 0$.

On dit qu'une spécification $\Pi = \{\pi_\Lambda : \Lambda \Subset \mathbb{Z}^d\}$ est **quasilocale** si $f \in \mathcal{C}(\Omega) \implies \pi_\Lambda(f|\cdot) \in \mathcal{C}(\Omega)$, pour tout $\Lambda \Subset \mathbb{Z}^d$. Ceci est équivalent à demander que, pour tout A cylindrique et tout $\Lambda \Subset \mathbb{Z}^d$,

$$\lim_{\Delta \nearrow \mathbb{Z}^d} \sup_{\substack{\omega, \omega': \\ \omega|_\Delta = \omega'|_\Delta}} |\pi_\Lambda(A|\omega) - \pi_\Lambda(A|\omega')| = 0.$$

On dit qu'une spécification $\Pi = \{\pi_\Lambda : \Lambda \Subset \mathbb{Z}^d\}$ est **uniformément non-nulle** si et seulement si

$$\inf_{\substack{\omega|_\Lambda \in \Omega_\Lambda \\ \bar{\omega}|_{\Lambda^c} \in \Omega_{\Lambda^c}}} \gamma_\Lambda(\omega|_\Lambda | \bar{\omega}|_{\Lambda^c}) > 0,$$

pour tout $\Lambda \in \mathbb{Z}^d$.

Le théorème suivant, qui généralise le Théorème 3.2.1 au présent contexte, montre que les mesures de Gibbs forment une classe très naturelle de spécifications. La partie “seulement si” est élémentaire, mais l’autre demande davantage de travail et on ne la démontrera pas.

Théorème 8.3.1 (Théorème de Kozlov–Sullivan). *Une spécification est gibbsienne si et seulement si elle est quasilocale et uniformément non-nulle.*

Remarque 8.3.1. *Comme on l’a vu dans le Chapitre 3, l’application $\Phi \mapsto \Pi^\Phi$ est loin d’être injective : on peut recombinaison les termes locaux définissant l’interaction d’une infinité de façons sans modifier les poids de Boltzmann. Deux interactions donnant lieu aux mêmes poids de Boltzmann sont dites **physiquement équivalentes**. D’un autre côté, on peut montrer qu’un champ aléatoire μ tel que $\mu(A) > 0$ pour tout ouvert $A \neq \emptyset$ est compatible avec au plus une spécification.*

8.4 Propriétés des mesures de Gibbs

8.4.1 Structure de $\mathcal{G}(\Pi)$

Nous allons à présent étudier quelques propriétés de l’ensemble $\mathcal{G}(\Pi)$ des mesures associées à une spécification Π .

Théorème 8.4.1. *Soit Π une spécification quasilocale sur (Ω, \mathcal{F}) . Alors, $\mathcal{G}(\Pi)$ est non-vide, compact et convexe.*

Démonstration. On fixe $\omega \in \Omega$ et $(\Lambda_n)_{n \geq 1}$ telle que $\Lambda_n \nearrow \mathbb{Z}^d$, et on considère la suite de mesures définies par $\nu_n^\omega(\cdot) = \pi_{\Lambda_n}(\cdot | \omega)$. $\mathcal{M}_1(\Omega, \mathcal{F})$ étant compact (cf. Théorème 8.1.2), on peut extraire une sous-suite, que l’on continue de noter $(\nu_n^\omega)_{n \geq 1}$, convergeant vers un champ aléatoire $\nu \in \mathcal{M}_1(\Omega, \mathcal{F})$. Vérifions que $\nu \in \mathcal{G}(\Pi)$. Soit $\Lambda \in \mathbb{Z}^d$ et $f \in \mathcal{C}(\Omega)$. La continuité de $\pi_\Lambda f$ implique que

$$\nu \pi_\Lambda(f) = \nu(\pi_\Lambda f) = \lim_{n \rightarrow \infty} \pi_{\Lambda_n}(\pi_\Lambda(f) | \omega) = \lim_{n \rightarrow \infty} \pi_{\Lambda_n} \pi_\Lambda(f | \omega) = \lim_{n \rightarrow \infty} \pi_{\Lambda_n}(f | \omega) = \nu(f).$$

(La quatrième identité suit de la propriété de cohérence de la spécification.) On a donc bien $\mathcal{G}(\Pi) \neq \emptyset$.

$\mathcal{M}_1(\Omega, \mathcal{F})$ étant compact, la compacité de $\mathcal{G}(\Pi)$ sera établie si on montre que cet ensemble est fermé. Soit $(\nu_n)_{n \geq 1}$ une suite d’éléments de $\mathcal{G}(\Pi)$ convergeant vers $\nu \in \mathcal{M}_1(\Omega, \mathcal{F})$. Alors, pour tout $\Lambda \in \mathbb{Z}^d$ et toute fonction $f \in \mathcal{C}(\Omega)$,

$$\nu \pi_\Lambda(f) = \nu(\pi_\Lambda f) = \lim_{n \rightarrow \infty} \nu_n(\pi_\Lambda f) = \lim_{n \rightarrow \infty} \nu_n \pi_\Lambda(f) = \lim_{n \rightarrow \infty} \nu_n(f) = \nu(f),$$

et donc $\nu \in \mathcal{G}(\Pi)$ et $\mathcal{G}(\Pi)$ est fermé.

La convexité de \mathcal{G} est aussi élémentaire : soient $\nu_1, \nu_2 \in \mathcal{G}(\Pi)$ et $\alpha \in [0, 1]$; alors

$$(\alpha \nu_1 + (1 - \alpha) \nu_2) \pi_\Lambda = \alpha \nu_1 \pi_\Lambda + (1 - \alpha) \nu_2 \pi_\Lambda = \alpha \nu_1 + (1 - \alpha) \nu_2,$$

et donc $\alpha \nu_1 + (1 - \alpha) \nu_2 \in \mathcal{G}(\Pi)$. □

L'ensemble $\mathcal{G}(\Pi)$ étant convexe, il est clair qu'il convient à présent de porter notre attention sur ses éléments extrémaux, c'est-à-dire les champs aléatoires $\mu \in \mathcal{G}(\Pi)$ tels que la seule façon d'avoir $\mu = \alpha\nu_1 + (1 - \alpha)\nu_2$, avec $\alpha \in (0, 1)$ et $\nu_1, \nu_2 \in \mathcal{G}(\Pi)$, est lorsque $\nu_1 = \nu_2$. L'ensemble des éléments extrémaux de $\mathcal{G}(\Pi)$, appelés **mesures extrémales**, est noté $\text{ex } \mathcal{G}(\Pi)$.

Avant d'étudier les propriétés des mesures extrémales, il est important de considérer la relation entre les propriétés globales des éléments de $\mathcal{G}(\Pi)$ et la tribu asymptotique.

8.4.2 Propriétés globales et tribu asymptotique

Il est utile de reformuler les équations de DLR en termes de deux opérations sur les champs aléatoires : la restriction de \mathcal{F} à une sous-tribu \mathcal{F}_{Λ^c} et l'extension de la sous-tribu \mathcal{F}_{Λ^c} vers la tribu \mathcal{F} .

Soit $\Lambda \in \mathbb{Z}^d$; la restriction $r_\Lambda : \mathcal{M}_1(\Omega, \mathcal{F}) \rightarrow \mathcal{M}_1(\Omega, \mathcal{F}_{\Lambda^c})$ est définie par

$$r_\Lambda(\mu)(A) \stackrel{\text{déf}}{=} \mu(A), \quad \forall A \in \mathcal{F}_{\Lambda^c}.$$

Soit $\Lambda \in \mathbb{Z}^d$; l'extension (ou prédiction) $t_\Lambda^\Pi : \mathcal{M}_1(\Omega, \mathcal{F}_{\Lambda^c}) \rightarrow \mathcal{M}_1(\Omega, \mathcal{F})$ associée à une spécification Π est définie par

$$t_\Lambda^\Pi(\mu)(A) \stackrel{\text{déf}}{=} \mu\pi_\Lambda(A), \quad \forall A \in \mathcal{F}.$$

Ces deux applications possèdent les propriétés suivantes.

Lemme 8.4.1. 1. *La restriction est conditionnellement linéaire : pour toute mesure $\mu \in \mathcal{M}_1(\Omega, \mathcal{F})$, pour toute fonction $g \geq 0$ \mathcal{F} -mesurable telle que $\mu(g) = 1$,*

$$r_\Lambda(g\mu) = \mu(g | \mathcal{F}_{\Lambda^c})r_\Lambda(\mu).$$

2. *L'extension est linéaire : pour tout $\mu \in \mathcal{M}_1(\Omega, \mathcal{F}_{\Lambda^c})$, pour toute fonction $g \geq 0$ \mathcal{F}_{Λ^c} -mesurable telle que $\mu(g) = 1$,*

$$t_\Lambda^\Pi(g\mu) = g \cdot t_\Lambda^\Pi(\mu).$$

Démonstration. Pour tout $A \in \mathcal{F}_{\Lambda^c}$, on a

$$\begin{aligned} \mu(g | \mathcal{F}_{\Lambda^c})r_\Lambda(\mu)(A) &= \int_A r_\Lambda(\mu)(d\omega) \mu(g | \mathcal{F}_{\Lambda^c})(\omega) = \int_A \mu(d\omega) \mu(g | \mathcal{F}_{\Lambda^c})(\omega) \\ &= \int_A \mu(d\omega) g(\omega) = g\mu(A) = r_\Lambda(g\mu)(A), \end{aligned}$$

ce qui prouve la première affirmation.

Soit f une fonction simple, \mathcal{F}_{Λ^c} -mesurable : $f = \sum_i f_i \mathbf{1}_{B_i}$, $B_i \in \mathcal{F}_{\Lambda^c}$, $\forall i$. Alors, pour tout $A \in \mathcal{F}$,

$$\begin{aligned} f \cdot t_\Lambda^\Pi(\mu)(A) &= \sum_i f_i \int \mu(d\omega) \int \pi_\Lambda(d\omega' | \omega) \mathbf{1}_{B_i}(\omega') \mathbf{1}_A(\omega') \\ &= \sum_i f_i \int \mu(d\omega) \pi_\Lambda(A \cap B_i | \omega) \\ &= \sum_i f_i \int \mu(d\omega) \pi_\Lambda(A | \omega) \mathbf{1}_{B_i}(\omega) = \int \mu(d\omega) \pi_\Lambda(A | \omega) f(\omega) = t_\Lambda^\Pi(f\mu)(A). \end{aligned}$$

(La troisième identité utilise le fait que le noyau de probabilité π_Λ est propre.) Le cas des fonctions f non-simples est obtenu par passage à la limite. \square

On peut à présent énoncer un résultat donnant une caractérisation équivalente des champs compatibles avec une spécification, montrant qu'un tel champ "se prédit lui-même".

Proposition 8.4.1. *Soit Π une spécification et $\mu \in \mathcal{M}_1(\Omega, \mathcal{F})$. Alors,*

$$\mu \in \mathcal{G}(\Pi) \iff \forall \Lambda \in \mathbb{Z}^d : \mu = t_\Lambda^\Pi(r_\Lambda(\mu)).$$

Démonstration. Pour tout $A \in \mathcal{F}$,

$$t_\Lambda^\Pi(r_\Lambda(\mu))(A) = \int \pi_\Lambda(A | \omega) r_\Lambda(\mu)(d\omega) = \int \pi_\Lambda(A | \omega) \mu(d\omega) = \mu\pi_\Lambda(A).$$

Par conséquent, $\mu = t_\Lambda^\Pi(r_\Lambda(\mu)) \iff \mu = \mu\pi_\Lambda$ et le résultat est démontré. \square

Le théorème suivant montre la relation entre les propriétés globales des champs aléatoires appartenant à $\mathcal{G}(\Pi)$ et la tribu asymptotique \mathcal{F}_∞ .

Théorème 8.4.2. *Soit Π une spécification.*

1. *Soit $\mu \in \mathcal{G}(\Pi)$, et soit g une fonction \mathcal{F} -mesurable telle que $g \geq 0$ et $\mu(g) = 1$. Alors $g\mu \in \mathcal{G}(\Pi)$ si et seulement si g est \mathcal{F}_∞ -mesurable.*
2. *Les éléments de $\mathcal{G}(\Pi)$ sont entièrement déterminés par leur comportement macroscopique : si $\mu, \nu \in \mathcal{G}(\Pi)$ coïncident sur \mathcal{F}_∞ , alors $\mu = \nu$.*

Démonstration. Soit $\mu \in \mathcal{G}(\Pi)$, et $g \geq 0$ \mathcal{F} -mesurable telle que $\mu(g) = 1$.

Supposons tout d'abord que $g\mu \in \mathcal{G}(\Pi)$. La proposition 8.4.1 et le Lemme 8.4.1 impliquent que

$$g\mu = t_\Lambda^\Pi(r_\Lambda(g\mu)) = t_\Lambda^\Pi(\mu(g | \mathcal{F}_{\Lambda^c}) r_\Lambda(\mu)) = \mu(g | \mathcal{F}_{\Lambda^c}) t_\Lambda^\Pi(r_\Lambda(\mu)) = \mu(g | \mathcal{F}_{\Lambda^c}) \mu,$$

et donc $g = \mu(g | \mathcal{F}_{\Lambda^c})$, μ -presque sûrement, pour tout $\Lambda \in \mathbb{Z}^d$. Ceci n'est possible que si $g = \mu(g | \mathcal{F}_\infty)$, μ -presque sûrement.

Réciproquement, supposons que g soit \mathcal{F}_∞ -mesurable. Alors,

$$(g\mu)\pi_\Lambda(f) = \int \pi_\Lambda(f | \omega) g(\omega) \mu(d\omega) = \int \pi_\Lambda(fg | \omega) \mu(d\omega) = \mu\pi_\Lambda(fg) = \mu(fg) = g\mu(f),$$

où l'on a utilisé le fait que g est \mathcal{F}_{Λ^c} -mesurable et que π_Λ est propre pour la seconde identité. Ceci montre que $g\mu \in \mathcal{G}(\Pi)$.

Pour démontrer la seconde partie du théorème, soient $\mu, \nu \in \mathcal{G}(\Pi)$ telles que $\mu(A) = \nu(A)$, pour tout $A \in \mathcal{F}_\infty$. La mesure $\mu^* = \frac{1}{2}(\mu + \nu)$ appartient également à $\mathcal{G}(\Pi)$, et μ et ν sont toutes deux absolument continues par rapport à μ^* . Par la première partie du théorème, on sait donc qu'il existe deux fonctions f et g \mathcal{F}_∞ -mesurables telles que $\mu = f\mu^*$ et $\nu = g\mu^*$. Par conséquent, pour tout $A \in \mathcal{F}_\infty$,

$$\int_A (f - g) d\mu^* = \mu(A) - \nu(A) = 0.$$

En choisissant successivement $A = \{f - g \geq 0\}$ et $A = \{f - g < 0\}$, on en déduit que $f = g$, μ^* -presque sûrement. On conclut alors aisément : pour tout $B \in \mathcal{F}$,

$$\mu(B) = \mu^*(f\mathbf{1}_B) = \mu^*(g\mathbf{1}_B) = \nu(B).$$

□

8.4.3 Propriétés des mesures extrémales

Nous allons à présent caractériser les mesures extrémales. Nous avons tout d'abord besoin d'un peu de terminologie.

Une mesure $\mu \in \mathcal{M}_1(\Omega, \mathcal{F})$ est triviale sur la tribu $\mathcal{F}' \subseteq \mathcal{F}$ si et seulement si $\mu(A) \in \{0, 1\}$, pour tout $A \in \mathcal{F}'$.

Une mesure $\mu \in \mathcal{M}_1(\Omega, \mathcal{F})$ est mélangeante si et seulement si, pour tout $A \in \mathcal{F}$,

$$\lim_{\Lambda \nearrow \mathbb{Z}^d} \sup_{B \in \mathcal{F}_{\Lambda^c}} |\mu(A \cap B) - \mu(A)\mu(B)| = 0.$$

Remarque 8.4.1. On peut montrer qu'il suffit en fait de vérifier cette dernière propriété pour les cylindres.

Théorème 8.4.3. Soit Π une spécification et $\mu \in \mathcal{G}(\Pi)$. Les propriétés suivantes sont équivalentes :

1. μ est extrémale.
2. μ est triviale sur \mathcal{F}_∞ .
3. μ est mélangeante.

Démonstration. 1 \implies 2. Si μ n'est pas triviale sur \mathcal{F}_∞ , il existe $A \in \mathcal{F}_\infty$ tel que $0 < \alpha \equiv \mu(A) < 1$. Par le théorème 8.4.2, les deux mesures $\mu_1 = \frac{1}{\alpha}\mathbf{1}_A\mu$ et $\mu_2 = \frac{1}{1-\alpha}\mathbf{1}_{A^c}\mu$ appartiennent à $\mathcal{G}(\Pi)$; Comme $\mu = \alpha\mu_1 + (1-\alpha)\mu_2$, μ n'est donc pas extrémale.

2 \implies 1. Supposons μ triviale sur \mathcal{F}_∞ , et supposons que $\mu = \alpha\mu_1 + (1-\alpha)\mu_2$, avec $\alpha \in (0, 1)$ et $\mu_1, \mu_2 \in \mathcal{G}(\Pi)$. Manifestement μ_1 et μ_2 sont toutes deux absolument continues par rapport à μ . Par conséquent, pour tout $A \in \mathcal{F}_\infty$, soit $\mu(A) = 0$ et donc $\mu_1(A) = \mu_2(A) = 0$, soit $\mu(A) = 1$ et donc $\mu_1(A) = \mu_2(A) = 1$. On en conclut que μ_1 et μ_2 coïncident sur \mathcal{F}_∞ , ce qui implique, par le théorème 8.4.2, que $\mu_1 = \mu_2$ et donc que μ est extrémale.

3 \implies 2. En prenant $A = B \in \mathcal{F}_\infty$, la propriété de mélange implique que

$$|\mu(A) - \mu(A)^2| = 0,$$

ce qui n'est possible que si $\mu(A) \in \{0, 1\}$.

2 \implies 3. Soit $(\Lambda_n)_{n \geq 1}$ telle que $\Lambda_n \nearrow \mathbb{Z}^d$, μ triviale sur \mathcal{F}_∞ , et $A \in \mathcal{F}$. On vérifie immédiatement que la suite de variables aléatoires $(X_n)_{n \geq 1}$, $X_n = \mu(A | \mathcal{F}_{\Lambda_n^c})$, forme une martingale renversée³ pour la filtration $(\mathcal{F}_n)_{n \geq 1}$, $\mathcal{F}_n = \mathcal{F}_{\Lambda_n^c}$. Par le théorème de

3. C'est-à-dire qu'on a, pour tout $n \geq 1$, $\mathcal{F}_{n+1} \subset \mathcal{F}_n$, $\mu(|X_n|) < \infty$ et $\mu(X_n | \mathcal{F}_{n+1}) = X_{n+1}$.

convergence des martingales renversées, $X_n \rightarrow X_\infty = \mu(A | \mathcal{F}_\infty)$, μ -presque sûrement et dans $\mathbb{L}^1(\mu)$. En particulier, pour tout $\epsilon > 0$, il existe $\Delta \Subset \mathbb{Z}^d$ tel que

$$\mu(|\mu(A | \mathcal{F}_{\Delta^c}) - \mu(A | \mathcal{F}_\infty)|) < \epsilon.$$

En outre, μ étant triviale sur \mathcal{F}_∞ , on doit avoir $\mu(A | \mathcal{F}_\infty) = \mu(A)$, μ -presque sûrement. Ainsi, pour tout $\epsilon > 0$, il existe $\Delta \Subset \mathbb{Z}^d$ tel que

$$\mu(|\mu(A | \mathcal{F}_{\Delta^c}) - \mu(A)|) < \epsilon.$$

Pout tout $\Delta \subseteq \Lambda \Subset \mathbb{Z}^d$, on a

$$\begin{aligned} \sup_{B \in \mathcal{F}_{\Lambda^c}} |\mu(A \cap B) - \mu(A)\mu(B)| &= \sup_{B \in \mathcal{F}_{\Lambda^c}} \left| \int_B d\mu(\mu(A | \mathcal{F}_{\Delta^c}) - \mu(A)) \right| \\ &\leq \int d\mu |\mu(A | \mathcal{F}_{\Delta^c}) - \mu(A)| < \epsilon. \end{aligned}$$

□

Corollaire 8.4.1. *Soit Π une spécification et $\mu, \nu \in \text{ex } \mathcal{G}(\Pi)$. Si $\mu \neq \nu$, il existe $A \in \mathcal{F}_\infty$ tel que $\mu(A) = 1$ et $\nu(A) = 0$.*

Démonstration. On sait par le théorème 8.4.2 qu'il existe $A \in \mathcal{F}_\infty$ tel que $\mu(A) \neq \nu(A)$. La trivialité de μ et ν sur \mathcal{F}_∞ implique alors que soit $\mu(A) = 1$ et $\nu(A) = 0$, soit $\mu(A^c) = 1$ et $\nu(A^c) = 0$. □

On va voir à présent que l'ensemble $\mathcal{G}(\Pi)$ est bien plus qu'un simple ensemble convexe. On appelle **simplexe de Choquet** un ensemble métrisable, compact, convexe G tel que tout élément de G puisse s'écrire *de façon unique* comme intégrale sur l'ensemble $\text{ex } G$: pour tout $u_0 \in G$, il existe une unique mesure ρ_0 sur $\text{ex } G$ telle que

$$u_0 = \int_{\text{ex } G} u \rho_0(du).$$

Théorème 8.4.4. *Soit Π une spécification quasilocale. Alors $\mathcal{G}(\Pi)$ est un simplexe de Choquet.*

Démonstration. La preuve repose sur la propriété suivante, que nous ne démontrerons pas [?] : *une partie métrisable compacte convexe G d'un espace vectoriel E est un simplexe de Choquet si et seulement si le cône engendré par G dans E est un treillis.*

Nous allons à présent expliquer les termes de cône et de treillis, puis montrer que, dans le cas qui nous intéresse, cette propriété est vérifiée.

Le cône \mathcal{K} engendré par $\mathcal{G}(\Pi)$ dans l'ensemble $\mathcal{M}_+(\Omega, \mathcal{F})$ des mesures σ -finies positives sur (Ω, \mathcal{F}) est défini par

$$\mathcal{K} \stackrel{\text{def}}{=} \left\{ \alpha \mu : \alpha \in \mathbb{R}^+, \mu \in \mathcal{G}(\Pi) \right\}.$$

Ce cône définit un ordre sur $\mathcal{M}_+(\Omega, \mathcal{F})$: $\nu \leq \nu' \iff \nu' - \nu \in \mathcal{K}$.

On dit que \mathcal{K} est un treillis si pour tout couple (ν_1, ν_2) de mesures dans \mathcal{K} , il existe un plus petit majorant commun et un plus grand minorant commun.

Montrons que \mathcal{K} est bien un treillis. Soient $\nu_1, \nu_2 \in \mathcal{K}$. Posons $\nu = \nu_1 + \nu_2$, $\nu_1 = g_1\nu$, $\nu_2 = g_2\nu$. Le théorème 8.4.2 implique (une fois les mesures normalisées) que g_1 et g_2 sont \mathcal{F}_∞ -mesurables. Donc c'est également le cas de $\sup(g_1, g_2)$ et de $\inf(g_1, g_2)$. Par conséquent $\sup(g_1, g_2)\nu$ et $\inf(g_1, g_2)\nu$ appartiennent tous deux à \mathcal{K} et \mathcal{K} est un treillis. \square

Remarque 8.4.2. *L'interprétation usuelle de ce résultat est la suivante. Les véritables états macroscopiques d'un système physique sont décrits par les mesures de Gibbs extrémales ; ces dernières sont en effet caractérisées par la propriété de donner une valeur déterministe aux observables macroscopiques (théorème 8.4.3). On interprète alors les mesures non-extrémales comme décrivant une incertitude sur la préparation du système. Prenons un exemple simple : le modèle d'Ising bidimensionnel en-dessous de la température critique possède exactement deux mesures de Gibbs extrémales, μ^+ et μ^- . Une mesure de Gibbs générale est donc de la forme $\alpha\mu^+ + (1 - \alpha)\mu^-$, un facteur $\alpha \in [0, 1]$ donné étant associé à une unique mesure de Gibbs, par le théorème 8.4.4. Cette mesure décrit une situation où le système a été préparé dans l'état $+$ avec probabilité α et dans l'état $-$ avec probabilité $1 - \alpha$. Les mesures non-extrémales représentent ainsi une incertitude de l'expérimentateur.*

8.4.4 Mesures de Gibbs et limite thermodynamique

Dans cette sous-section, nous allons étudier la relation entre la notion de mesures de Gibbs en volume infini utilisée dans les chapitres précédents, c'est-à-dire l'ensemble des points d'accumulation de mesures de Gibbs en volume fini (ou poids de Boltzmann) pour toute suite de conditions au bord et de volumes $\Lambda_n \nearrow \mathbb{Z}^d$, et celle introduite dans ce chapitre.

Soit Π une spécification quasilocale, $(\omega_n)_{n \geq 1}$, $\omega_n \in \Omega$, et $(\Lambda_n)_{n \geq 1}$, $\Lambda_n \nearrow \mathbb{Z}^d$. On considère la suite de mesures de Gibbs en volume fini $\nu_n = \pi_{\Lambda_n}(\cdot | \omega_n)$. On a déjà vu dans la preuve du théorème 8.4.1 que les points d'accumulation de cette suite appartiennent tous à $\mathcal{G}(\Pi)$. Donc les mesures de Gibbs en volume infini définies via une limite thermodynamique font bien partie de $\mathcal{G}(\Pi)$. Il reste à déterminer quels sont les éléments de $\mathcal{G}(\Pi)$ qui peuvent être atteints de cette manière. Le théorème suivant montre que c'est le cas pour toutes les mesures extrémales. De plus, il suffit pour cela de prendre comme condition au bord une configuration typique de la mesure que l'on désire atteindre.

Théorème 8.4.5. *Soit Π une spécification quasilocale et $\mu \in \text{ex } \mathcal{G}(\Pi)$. Alors,*

$$\lim_{\Lambda \nearrow \mathbb{Z}^d} \pi_\Lambda(\cdot | \omega) \rightarrow \mu,$$

pour μ -presque tout ω .

Démonstration. Soit f une fonction locale. On a déjà observé que les variables aléatoires $\pi_{\Lambda_n}(f | \cdot) = \mu(f | \mathcal{F}_{\Lambda_n^c})(\cdot)$ forment une martingale renversée pour la filtration $(\mathcal{F}_{\Lambda_n^c})_{n \geq 1}$. Le théorème de convergence pour les martingales renversées implique donc que

$$\pi_{\Lambda_n}(f | \cdot) \rightarrow \mu(f | \mathcal{F}_\infty)(\cdot) = \mu(f),$$

μ -presque sûrement (la dernière identité suit de la trivialité de μ sur \mathcal{F}_∞). Comme l'ensemble des fonctions locales est dénombrable, il suit que cette convergence a lieu μ -presque sûrement pour toutes ces fonctions simultanément. \square

Bibliographie

- [1] M. Aizenman. Translation invariance and instability of phase coexistence in the two-dimensional Ising system. *Comm. Math. Phys.*, 73(1) :83–94, 1980.
- [2] M. Aizenman and R. Fernández. On the critical behavior of the magnetization in high-dimensional Ising models. *J. Statist. Phys.*, 44(3-4) :393–454, 1986.
- [3] M. Biskup. Reflection positivity and phase transitions in lattice spin models. Notes of the lectures given at Prague Summer School on Mathematical Statistical Mechanics, September 2006. Arxiv :math-ph/0610025.
- [4] J. Bricmont, J. R. Fontaine, and L. J. Landau. On the uniqueness of the equilibrium state for plane rotators. *Comm. Math. Phys.*, 56(3) :281–296, 1977.
- [5] R. L. Dobrushin. Gibbs states describing a coexistence of phases for the three-dimensional Ising model. *Th.Prob. and its Appl.*, 17(3) :582–600, 1972.
- [6] R. L. Dobrushin and S. B. Shlosman. Absence of breakdown of continuous symmetry in two-dimensional models of statistical physics. *Comm. Math. Phys.*, 42 :31–40, 1975.
- [7] C. M. Fortuin, P. W. Kasteleyn, and J. Ginibre. Correlation inequalities on some partially ordered sets. *Comm. Math. Phys.*, 22 :89–103, 1971.
- [8] H.-O. Georgii. *Gibbs measures and phase transitions*, volume 9 of *de Gruyter Studies in Mathematics*. Walter de Gruyter & Co., Berlin, 1988.
- [9] R. B. Griffiths. Correlation in Ising ferromagnets I, II. *J. Math. Phys.*, 8 :478–489, 1967.
- [10] Y. Higuchi. On the absence of non-translation invariant Gibbs states for the two-dimensional Ising model. In *Random fields, Vol. I, II (Esztergom, 1979)*, volume 27 of *Colloq. Math. Soc. János Bolyai*, pages 517–534. North-Holland, Amsterdam, 1981.
- [11] D. Ioffe, S. Shlosman, and Y. Velenik. 2d models of statistical physics with continuous symmetry : The case of singular interactions. *Comm. Math. Phys.*, 226(2) :433–454, 2002.
- [12] E. Ising. Beitrag zur Theorie des Ferromagnetismus. *Zeit. für Physik*, 31 :253–258, 1925.

- [13] R. B. Israel. *Convexity in the theory of lattice gases*. Princeton University Press, Princeton, N.J., 1979. Princeton Series in Physics, With an introduction by Arthur S. Wightman.
- [14] D. G. Kelly and S. Sherman. General Griffiths's inequality on correlation in Ising ferromagnets. *J. Math. Phys.*, 9 :466–484, 1968.
- [15] R. Kindermann and J. L. Snell. *Markov random fields and their applications*, volume 1 of *Contemporary Mathematics*. American Mathematical Society, Providence, R.I., 1980.
- [16] R. Kotecký and S. B. Shlosman. First-order phase transitions in large entropy lattice models. *Comm. Math. Phys.*, 83(4) :493–515, 1982.
- [17] G. F. Lawler. *Intersections of random walks*. Probability and its Applications. Birkhäuser Boston Inc., Boston, MA, 1991.
- [18] T. D. Lee and C. N. Yang. Statistical theory of equations of state and phase transitions. ii. lattice gas and ising model. *Phys. Rev.*, 87(3) :410–419, 1952.
- [19] L. Onsager. Crystal statistics, i. a two-dimensional model with an order-disorder transition. *Phys. Rev.*, 65 :117–149, 1944.
- [20] R. E. Peierls. On Ising's ferromagnet model. *Proc. Camb. Phil. Soc.*, 32 :477–481, 1936.
- [21] C. É. Pfister. On the symmetry of the Gibbs states in two-dimensional lattice systems. *Comm. Math. Phys.*, 79(2) :181–188, 1981.
- [22] J. G. Propp and D. B. Wilson. Exact sampling with coupled Markov chains and applications to statistical mechanics. In *Proceedings of the Seventh International Conference on Random Structures and Algorithms (Atlanta, GA, 1995)*, volume 9, pages 223–252, 1996.
- [23] B. Prum. *Processus sur un réseau et mesures de Gibbs*. Techniques Stochastiques. [Stochastic Techniques]. Masson, Paris, 1986. Applications.
- [24] D. Ruelle. *Statistical mechanics*. World Scientific Publishing Co. Inc., River Edge, NJ, 1999. Rigorous results, Reprint of the 1989 edition.
- [25] S. B. Shlosman. The method of reflection positivity in the mathematical theory of first-order phase transitions. *Russ. Math. Surv.*, 41(3) :83–134, 1986.
- [26] B. Simon. *The statistical mechanics of lattice gases. Vol. I*. Princeton Series in Physics. Princeton University Press, Princeton, NJ, 1993.
- [27] Ya. G. Sinaï. *Theory of phase transitions : rigorous results*, volume 108 of *International Series in Natural Philosophy*. Pergamon Press, Oxford, 1982. Translated from the Russian by J. Fritz, A. Kramli, P. Major and D. Szász.
- [28] F. Spitzer. *Principles of random walks*. Springer-Verlag, New York, second edition, 1976. Graduate Texts in Mathematics, Vol. 34.
- [29] C.N. Yang. The spontaneous magnetization of the two-dimensional Ising model. *Phys. Rev.*, 85 :808–816, 1952.