

ASPECTS RIGoureux DE LA MÉCANIQUE STATISTIQUE À L'ÉQUILIBRE LIMITE THERMODYNAMIQUE ET MESURES DE GIBBS

Guilhem Semerjian

Février-Juin 2013

Notes, préliminaires, non corrigées et non destinées à une diffusion publique, du cours du cursus mixte Maths-Physique 2012-2013, 2013-2014 et 2014-2015 donné par Jérémie Bouttier et Guilhem Semerjian. Ce fichier comprend la partie du cours donnée par GS.

Les deux ouvrages principalement utilisés dans la préparation de ce cours sont :

- Barry Simon, *Statistical mechanics of lattice gases* [1]
- Errico Presutti, *Scaling limits in statistical mechanics and microstructures in continuum mechanics* [2]

Table des matières

1	Rappels de mécanique statistique - motivations du cours	3
1.1	Configurations et énergie	3
1.2	Ensembles statistiques	3
1.3	Transitions de phase	3
1.4	Différentes approches aux transitions de phase	4
2	Limite thermodynamique	5
2.1	Définitions	5
2.2	Lemmes préparatoires	6
2.3	Energie libre pour les interactions à portée finie	8
2.4	Généralisations	9
2.4.1	Limite au sens de van Hove	10
2.4.2	Existence et propriétés de la limite pour des interactions plus générales	11
2.4.3	Conditions aux bords	13
3	Mesures de Gibbs	15
3.1	Mesures de Gibbs-Boltzmann en volume fini	15
3.1.1	Lois marginales et conditionnelles	15
3.1.2	Mesures de l'ensemble canonique	16
3.2	Rappels de topologie et de théorie de la mesure	17
3.3	Définition des mesures de Gibbs	21
3.4	Existence des mesures de Gibbs	24
3.5	Transitions de phase dans l'ensemble des mesures de Gibbs	28
3.5.1	Unicité à haute température	28
3.5.2	Argument de Peierls pour le modèle d'Ising	32
3.6	Extrémalité et trivialité à l'infini	35

4 Principe variationnel	41
4.1 En volume fini	41
4.2 Caractérisation variationnelle de $f(\beta, \Phi)$	41
4.2.1 Entropie d'une mesure en volume infini	41
4.2.2 Energie moyenne d'une mesure en volume infini	43
4.2.3 Principe variationnel pour l'énergie libre	44
4.3 Caractérisation variationnelle de $\mathcal{G}_{\text{inv}}(\beta, \Phi)$	46
5 Equivalence des ensembles	47
5.1 Discussion heuristique	47
5.2 Existence et propriétés de l'entropie microcanonique	47
5.3 Discussion	49
5.4 Autres paires d'ensembles	50
6 Limite de Kac	51
6.1 Motivation - Construction de Maxwell	51
6.2 Interaction de Kac	52
6.3 Limite de Kac pour $f(h)$	52
6.4 Limite de Kac pour $\hat{f}(m)$	55
Bibliographie	57
TD 1 : Modèles unidimensionnels et de champ moyen	58
Corrigé du TD 1	60
TD 2 : Entropie et principe variationnel	63
Corrigé du TD 2	66
TD 3 : Transformées de Legendre	71
Corrigé du TD 3	73
Examen de juin 2013 : Les zéros de Lee et Yang	76
Corrigé de l'examen de juin 2013	79
Examen de juin 2014 : Inégalités de corrélation	82
Corrigé de l'examen de juin 2014	85

1 Rappels de mécanique statistique - motivations du cours

1.1 Configurations et énergie

but de la mécanique statistique : expliquer les propriétés macroscopiques d'un système à partir d'une modélisation microscopique

point de départ de la mécanique statistique : définir un espace de configurations et une fonction énergie (Hamiltonien) sur cet espace

- systèmes de N particules, dans un espace de dimension d
configurations = positions $\vec{r}_i \in \mathcal{V} \subset \mathbb{R}^d$ et impulsions $\vec{p}_i \in \mathbb{R}^d$ de chaque particule
Hamiltonien somme des énergies cinétique et potentielle, $H = \sum_i \frac{\vec{p}_i^2}{2m_i} + \sum_{i < j} V(\vec{r}_i - \vec{r}_j)$ en supposant interactions à deux corps seulement
on n'en parlera plus
- modèles discrets sur réseau
 N variables $\underline{\sigma} = (\sigma_1, \dots, \sigma_N) \in \chi^N$
pour Ising $\chi = \{-1, +1\}$, pour Potts $\chi = \{1, \dots, q\}$.
Modèle d'Ising ferromagnétique, sur un graphe, $H(\underline{\sigma}) = -J \sum_{\langle ij \rangle} \sigma_i \sigma_j - h \sum_{i=1}^N \sigma_i$, somme sur les liens du graphe, interprétation couplage $J > 0$, champ magnétique h : physiquement décrit aimants.
En particulier sur des portions de \mathbb{Z}^d , interactions entre proches voisins. Pour des cubes de longueur L , $N = L^d$.
Mentionner l'équivalence d'un système de spins d'Ising avec un système de particules, $\sigma_i = \pm 1 \leftrightarrow$ particule présente/absente sur le site i , donc ce modèle décrit plus que les aimants, utile aussi pour comprendre les fluides
- on n'abordera pas non plus les aspects quantiques de la mécanique statistique, ni la dynamique

1.2 Ensembles statistiques

i.e. lois de probabilité sur l'espace des configurations. Cette approche se justifie car la description microscopique est (i) ingérable car N est trop grand, on ne peut pas connaître la configuration microscopique du système (ii) inutile car on ne veut/sait mesurer que des observables macroscopiques.

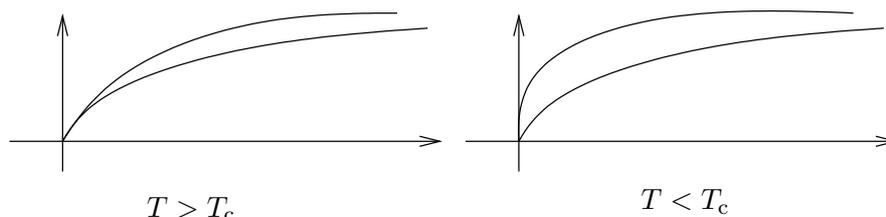
- microcanonique, correspond physiquement à un système isolé
équiprobabilité des configurations avec $H(\underline{\sigma}) = E$ (à ΔE près)
potentiel thermodynamique est (moins) l'entropie
 $S(E) = k_B \ln(\text{nbre configurations autorisées})$, avec souvent $k_B = 1$
- canonique, correspond à un système en contact avec un thermostat de température T
notation $\beta = \frac{1}{k_B T}$
probabilité de Gibbs-Boltzmann $\mu(\underline{\sigma}) = \frac{1}{Z(\beta)} e^{-\beta H(\underline{\sigma})}$
normalisation est la fonction de partition $Z(\beta) = \sum_{\underline{\sigma}} e^{-\beta H(\underline{\sigma})}$
potentiel thermodynamique est l'énergie libre, $F(\beta) = -\frac{1}{\beta} \ln Z(\beta)$
notation $\langle \cdot \rangle$ pour les moyennes d'une fonction de l'espace des configurations par rapport à la mesure de Gibbs-Boltzmann

mentionner équivalence des ensembles (dans la "limite thermodynamique"), structure de transformée de Legendre entre $F(T)$ et $S(E)$, énergie fixée strictement dans l'ensemble microcanonique, seulement en moyenne dans l'ensemble canonique, mais petites fluctuations relatives d'où l'équivalence, sera discutée plus en détails dans le chapitre 5.

1.3 Transitions de phase

- prendre l'exemple d'un modèle d'Ising, par exemple à 2d, tracer l'allure de $m(h)$, où $m = \langle \sigma_0 \rangle$ l'aimantation du spin au centre, justifier impaire (symétrie), croissante (intuitif + inégalités de Griffiths), $\lim_{h \rightarrow \infty} m = 1$

pour différentes températures, différentes tailles du système



pente en 0 (susceptibilité) finie ou d'ordre N dans la limite $N \rightarrow \infty$

- définir l'aimantation spontanée, $\lim_{h \rightarrow 0^+} \lim_{N \rightarrow \infty} \langle \sigma_0 \rangle$, en intervertissant les limites c'est toujours 0, allure en fonction de la température, exposant critique β , à ne pas confondre avec $1/T$
- transitions de phase (singularités des fonctions thermodynamiques) n'apparaissent que pour $N \rightarrow \infty$, une partie du cours s'attachera à discuter l'existence et les propriétés de cette limite, sur les grandeurs thermodynamiques (densité d'énergie libre $f(\beta) = \lim_{N \rightarrow \infty} \frac{1}{N} F(\beta)$, entropie) et sur les mesures de probabilités elles-mêmes (non-trivial, mesures de probabilité sur un espace infini). Pour les systèmes macroscopiques $N \sim 10^{23} \gg 1$: bien sûr ce sont des systèmes finis, mais tellement grand qu'étant donné les résolutions de nos dispositifs expérimentaux (par exemple pente de $m(h)$ finie mais d'ordre N) il est tout-à-fait pertinent de prendre la limite $N \rightarrow \infty$.
- définir et mentionner transitions du premier vs second ordre. Exemple du diagramme de phase Ising en (T, h) .
- mentionner les autres exposants critiques pour les transitions du second ordre, en particulier pour la singularité de $f(\beta)$ en $|T - T_c|^{2-\alpha}$
- mentionner d'autres transitions de phase, corps pur (montrer diagramme des phases, exposant β , le même que pour le ferromagnétisme), cristaux liquides, supraconductivité... et idée d'universalité (mêmes exposants critiques).

1.4 Différentes approches aux transitions de phase

- résolutions exactes en petite dimension : facile à 1d mais pas de transition (cf TD1), très riche en 2d, cf partie du cours faite par Jérémie
- preuves rigoureuses d'existence de transitions de phase pour $d \geq 2$, même si pas de résolution explicite pour $d > 2$. On le verra en partie pour le modèle d'Ising au chapitre 3.5.
- champ moyen (cf TD1), très approché, mais donne qualitativement la phénoménologie des transitions. Correspond à la limite de grande dimension, ou d'interactions à longue portée (on reparlera de ces limites). Pour la supraconductivité, théorie BCS champ moyen quantitativement pertinente.
- groupe de renormalisation, permet de calculer les exposants critiques de manière approchée, et non rigoureuse, on n'en parlera pas. Cette approche permet de comprendre les origines de l'universalité, invariance d'échelle, longueur de corrélation divergente. Les exposants de champ moyen sont exacts en dimension grande mais finie, par exemple pour la classe d'universalité du modèle d'Ising cette dimension critique supérieure est 4.

En résumé l'objectif principal du cours est de caractériser mathématiquement la limite thermodynamique $N \rightarrow \infty$, sur les grandeurs thermodynamiques (densité d'énergie libre, d'entropie) et sur les mesures de probabilités des configurations microscopiques. Et ceci pour deux raisons :

- les systèmes macroscopiques ont $N \sim 10^{23} \gg 1$
- les singularités de type transitions de phase ne peuvent apparaître mathématiquement qu'à la limite

2 Limite thermodynamique

objectif de cette partie : montrer l'existence de la limite thermodynamique pour l'énergie libre, i.e.

$$\lim_{\Lambda \rightarrow \infty} \frac{-1}{\beta |\Lambda|} \ln Z_\Lambda \quad (1)$$

pour une large classe de modèles définis sur \mathbb{Z}^d , il faudra préciser le sens de la limite, et la régularisation de Z dans un volume Λ fini

on ne peut pas travailler directement sur le système infini, si on essaie d'écrire l'Hamiltonien d'Ising

$$H(\underline{\sigma}) = -J \sum_{\langle xy \rangle} \sigma_x \sigma_y - h \sum_x \sigma_x \quad (2)$$

sommes infinies pas contrôlées

2.1 Définitions

Notations élémentaires :

- $x \in \mathbb{Z}^d$ les points du réseau, où vivent les spins :
- $\sigma_x \in \chi$, avec χ fini, exemple d'Ising $\chi = \{-1, +1\}$
- $\underline{\sigma} = \{\sigma_x, x \in \mathbb{Z}^d\}$ configuration globale, $\Sigma = \chi^{\mathbb{Z}^d}$ l'ensemble des configurations
- pour $\Lambda \subset \mathbb{Z}^d$, $\underline{\sigma}_\Lambda = \{\sigma_x, x \in \Lambda\}$, configuration des spins à l'intérieur de Λ
- $\Lambda^c = \mathbb{Z}^d \setminus \Lambda$ le complément de Λ
- on note $X \subset_f \mathbb{Z}^d$ les parties finies X de \mathbb{Z}^d
- on note $|X|$ le cardinal de X , "volume" pour une partie finie
- $\text{dist}(x, y) = \sum_{\alpha \in [1, d]} |x_\alpha - y_\alpha|$ la distance (issue de la norme 1) entre deux points de \mathbb{Z}^d
- $\text{diam}(\Lambda) = \sup\{\text{dist}(x, y), (x, y) \in \Lambda\}$ le diamètre d'une partie de \mathbb{Z}^d
- pour $X \subset_f \mathbb{Z}^d$, \mathcal{C}_X est l'ensemble des fonctions ϕ de Σ vers \mathbb{R} qui ne dépendent que des spins dans X , i.e. $\phi(\underline{\sigma}) = \phi(\underline{\sigma}_X)$ (on utilisera les deux notations par commodité)
- pour $\phi \in \mathcal{C}_X$, $\|\phi\|_\infty = \max_{\underline{\sigma}_X \in \chi^X} \phi(\underline{\sigma}_X)$, c'est bien une norme
- T_a opérateur de translation de $a \in \mathbb{Z}^d$, i.e. $(T_a \underline{\sigma})_x = \sigma_{x-a}$ et pour $X \subset \mathbb{Z}^d$, $T_a X = \{x, x-a \in X\}$

Définition 1. Ensemble \mathcal{I} des interactions : $\Phi \in \mathcal{I}$ est une collection de ϕ_X (interaction élémentaire), où X parcourt les parties finies de \mathbb{Z}^d , avec ϕ_X une fonction réelle de $\underline{\sigma}$, qui ne dépend que de $\underline{\sigma}_X$, énergie due à l'interaction entre les variables dans X , formellement $\Phi = \{\phi_X\}_{X \subset_f \mathbb{Z}^d}$, avec $\phi_X \in \mathcal{C}_X \forall X$.

exemple Ising,

$$\phi_X(\underline{\sigma}) = \begin{cases} -h\sigma_x & \text{si } X = \{x\} \\ -J\sigma_x\sigma_y & \text{si } X = \{x, y\} \text{ avec } \text{dist}(x, y) = 1 \\ 0 & \text{sinon} \end{cases} \quad (3)$$

Restriction dans toute la suite : on ne considèrera que des interactions invariantes par translation, $\phi_{T_a X}(\underline{\sigma}) = \phi_X(T_{-a} \underline{\sigma})$

\mathcal{I} a naturellement une structure d'espace vectoriel, $(\alpha \Phi)_X = \alpha \phi_X$, $(\Phi + \Psi)_X = \phi_X + \psi_X$.

Définition 2. Hamiltonien pour une interaction Φ dans un domaine fini $\Lambda \subset_f \mathbb{Z}^d$ avec conditions aux bords libres :

$$H_\Lambda^\Phi(\underline{\sigma}) = \sum_{X \subset \Lambda} \phi_X(\underline{\sigma}) \quad (4)$$

On notera que $H_\Lambda^\Phi(\underline{\sigma}) \in \mathcal{C}_\Lambda$, ne dépend en fait que de $\underline{\sigma}_\Lambda$.

On enlève complètement l'extérieur, exemple sur Ising.

Définitions des quantités thermodynamiques associées :

$$Z_\Lambda(\beta, \Phi) = \sum_{\underline{\sigma}_\Lambda} e^{-\beta H_\Lambda^\Phi(\underline{\sigma})}, \quad F_\Lambda(\beta, \Phi) = -\frac{1}{\beta} \ln Z_\Lambda(\beta, \Phi), \quad f_\Lambda(\beta, \Phi) = -\frac{1}{\beta|\Lambda|} \ln Z_\Lambda(\beta, \Phi) \quad (5)$$

Classes d'interactions :

H_Λ^Φ toujours fini pour Λ fini, mais pour prendre la limite $\Lambda \rightarrow \mathbb{Z}^d$ il faut imposer des restrictions à l'interaction.

Trois familles d'interaction :

- $\Phi \in \mathcal{I}^{(r)} \subset \mathcal{I}$, à portée r , ssi $\text{diam}(X) > r \Rightarrow \phi_X = 0$. Portée finie arbitraire $\mathcal{I}_f = \bigcup_{r \geq 0} \mathcal{I}^{(r)}$.
- $\|\Phi\|_p = \sum_{X \ni 0} \|\phi_X\|_\infty$, notons $\mathcal{I}_p = \{\Phi \in \mathcal{I}, \|\Phi\|_p < \infty\}$. Par invariance par translation le 0 dans la définition de la norme est arbitraire.
- $\|\Phi\|_g = \sum_{X \ni 0} \frac{\|\phi_X\|_\infty}{|X|}$, notons $\mathcal{I}_g = \{\Phi \in \mathcal{I}, \|\Phi\|_g < \infty\}$. Même remarque.

$\|\cdot\|_p$ et $\|\cdot\|_g$ sont des normes ($\|\Phi\| = 0 \Leftrightarrow \Phi = 0$, $\|\lambda\Phi\| = |\lambda|\|\Phi\|$, $\|\Phi + \Psi\| \leq \|\Phi\| + \|\Psi\|$), \mathcal{I}_p et \mathcal{I}_g sont des espaces de Banach (e.v.n. complet pour les normes associés, i.e. toute suite de Cauchy converge).

$\mathcal{I}_f \subset \mathcal{I}_p \subset \mathcal{I}_g$: si $\Phi \in \mathcal{I}_f$ les sommes dans la définition de $\|\Phi\|_p$ et $\|\Phi\|_g$ sont finies, et par ailleurs $\|\Phi\|_g \leq \|\Phi\|_p$, donc si $\|\Phi\|_p$ est finie alors $\|\Phi\|_g$ l'est aussi.

pertinent de prendre des interactions à portée arbitraire pour des raisons à la fois techniques et physiques (exemple des interactions électrostatiques, gravité, mais Coulomb n'est pas dans \mathcal{I}_g : $|X| = 2$, $\|\phi_{0,x}\|_\infty = O(1/|x|)$, $\|\Phi\|_g \propto \int dr r^{d-1} \frac{1}{r}$ diverge pour $d \geq 2$). Non seulement la portée est arbitraire, mais le nombre de variables impliquées dans une interaction élémentaire ϕ_X est arbitraire.

2.2 Lemmes préparatoires

Lemme 1. Si $\Phi \in \mathcal{I}_g$ et $\Lambda \subset_f \mathbb{Z}^d$, alors $\|H_\Lambda^\Phi\|_\infty \leq |\Lambda| \|\Phi\|_g$.

Démonstration. Permet d'expliquer pourquoi $\|\cdot\|_g$ et \mathcal{I}_g sont "naturels", condition minimale pour avoir un Hamiltonien "extensif".

$$H_\Lambda^\Phi(\underline{\sigma}) = \sum_{X \subset \Lambda} \phi_X(\underline{\sigma}) \quad \Rightarrow \quad \|H_\Lambda^\Phi\|_\infty \leq \sum_{X \subset \Lambda} \|\phi_X\|_\infty \quad (6)$$

(inégalité triangulaire sur la valeur absolue, puis max de la somme \leq somme des max). Or $\sum_{x \in X} \frac{1}{|X|} = 1$, on peut donc écrire

$$\sum_{X \subset \Lambda} \|\phi_X\|_\infty = \sum_{X \subset \Lambda} \sum_{x \in X} \frac{\|\phi_X\|_\infty}{|X|} \leq \sum_{x \in \Lambda} \sum_{X \ni x} \frac{\|\phi_X\|_\infty}{|X|} = |\Lambda| \|\Phi\|_g \quad (7)$$

où l'inégalité vient des termes de bord, on ne comptait que les X entièrement inclus dans Λ , ensuite on compte ceux qui ont une intersection non-nulle avec Λ . Invariance par translation pour retrouver $\|\cdot\|_g$. \square

Lemme 2. Pour $\Lambda \subset_{\mathbb{f}} \mathbb{Z}^d$, α, g_1, g_2 des fonctions de $\underline{\sigma}_\Lambda$ avec $\alpha \geq 0$,

$$\left| \ln \sum_{\underline{\sigma}_\Lambda} \alpha(\underline{\sigma}_\Lambda) e^{g_1(\underline{\sigma}_\Lambda)} - \ln \sum_{\underline{\sigma}_\Lambda} \alpha(\underline{\sigma}_\Lambda) e^{g_2(\underline{\sigma}_\Lambda)} \right| \leq \|g_1 - g_2\|_\infty . \quad (8)$$

Démonstration. Notons que $\|e^{g_1 - g_2}\|_\infty \leq e^{\|g_1 - g_2\|_\infty}$.

$$\ln \sum_{\underline{\sigma}_\Lambda} \alpha(\underline{\sigma}_\Lambda) e^{g_1(\underline{\sigma}_\Lambda)} - \ln \sum_{\underline{\sigma}_\Lambda} \alpha(\underline{\sigma}_\Lambda) e^{g_2(\underline{\sigma}_\Lambda)} = \ln \left(\frac{\sum_{\underline{\sigma}_\Lambda} \alpha(\underline{\sigma}_\Lambda) e^{g_1(\underline{\sigma}_\Lambda) - g_2(\underline{\sigma}_\Lambda)} e^{g_2(\underline{\sigma}_\Lambda)}}{\sum_{\underline{\sigma}_\Lambda} \alpha(\underline{\sigma}_\Lambda) e^{g_2(\underline{\sigma}_\Lambda)}} \right) \quad (9)$$

$$\leq \ln \left(\frac{\sum_{\underline{\sigma}_\Lambda} \alpha(\underline{\sigma}_\Lambda) e^{\|g_1 - g_2\|_\infty} e^{g_2(\underline{\sigma}_\Lambda)}}{\sum_{\underline{\sigma}_\Lambda} \alpha(\underline{\sigma}_\Lambda) e^{g_2(\underline{\sigma}_\Lambda)}} \right) \quad (10)$$

Puis recommencer en intervertissant g_1 et g_2 . □

Lemme 3. Pour $\Lambda \subset_{\mathbb{f}} \mathbb{Z}^d$ et Φ, Ψ deux interactions de \mathcal{I}_g ,

$$|f_\Lambda(\beta, \Phi) - f_\Lambda(\beta, \Psi)| \leq \|\Phi - \Psi\|_g . \quad (11)$$

Démonstration. En appliquant le lemme 2 avec $\alpha = 1$, $g_1 = -\beta H_\Lambda^\Phi$ et $g_2 = -\beta H_\Lambda^\Psi$,

$$|\ln Z_\Lambda(\beta, \Phi) - \ln Z_\Lambda(\beta, \Psi)| \leq \beta \|H_\Lambda^\Phi - H_\Lambda^\Psi\|_\infty = \beta \|H_\Lambda^{\Phi - \Psi}\|_\infty \leq \beta |\Lambda| \|\Phi - \Psi\|_g , \quad (12)$$

où la dernière égalité vient du lemme 1. Il suffit de diviser par $\beta |\Lambda|$ pour conclure. □

Lemme 4. En l'absence d'interaction l'énergie libre est seulement l'entropie divisé par $-\beta$, i.e. $f_\Lambda(\beta, 0) = -\frac{\ln(|\chi|)}{\beta}$.

Démonstration.

$$Z_\Lambda(\beta, 0) = \sum_{\underline{\sigma}_\Lambda} 1 = |\chi|^{|\Lambda|} , \quad (13)$$

puis prendre le logarithme et diviser par $-\beta |\Lambda|$. □

Lemme suivant sera utilisé seulement plus tard, mais la structure de la preuve va revenir très bientôt. Utilisé aussi par Jérémie pour les SAW.

Lemme 5 (Lemme de Fekete). Soit u_n une suite sous-additive, i.e. $u_{n+m} \leq u_n + u_m$. Alors $\frac{u_n}{n}$ admet une limite quand $n \rightarrow \infty$ (éventuellement égale à $-\infty$), et

$$\lim_{n \rightarrow \infty} \frac{u_n}{n} = \inf_{n \geq 1} \frac{u_n}{n} . \quad (14)$$

Résultat similaire pour suite sur-additive.

Démonstration. Notons d'abord qu'en itérant l'inégalité avec $m = 1$, on obtient $u_n \leq nu_1$, soit $\frac{u_n}{n} \leq u_1$, donc $\frac{u_n}{n}$ ne peut partir en $+\infty$. Plus généralement, $\frac{u_{np}}{np} \leq \frac{u_p}{p}$, donc le long de certaines sous-suites la suite $\frac{u_n}{n}$ est en dessous de toutes ses cordes.

Début de la preuve proprement dite. Fixons $p \geq 1$, et faisons la division euclidienne $n = ap + b$ avec $a = \lfloor \frac{n}{p} \rfloor$ et $b \in [0, p - 1]$. Par sous-additivité, $u_n \leq a u_p + u_b$, soit

$$\frac{u_n}{n} \leq \frac{1}{n} \left\lfloor \frac{n}{p} \right\rfloor u_p + \frac{u_b}{n} . \quad (15)$$

En prenant $n \rightarrow \infty$,

$$\limsup_{n \rightarrow \infty} \frac{u_n}{n} \leq \frac{u_p}{p}, \quad (16)$$

car $\frac{1}{n} \lfloor \frac{n}{p} \rfloor \rightarrow \frac{1}{p}$, et $|u_b| \leq \max_{i \in [0, p-1]} |u_i| < \infty$. Comme c'est vrai quelque soit p ,

$$\limsup_{n \rightarrow \infty} \frac{u_n}{n} \leq \inf_{n \geq 1} \frac{u_n}{n}, \quad (17)$$

or en général

$$\inf_{n \geq 1} a_n \leq \liminf_{n \rightarrow \infty} a_n \leq \limsup_{n \rightarrow \infty} a_n \leq \sup_{n \geq 1} a_n, \quad (18)$$

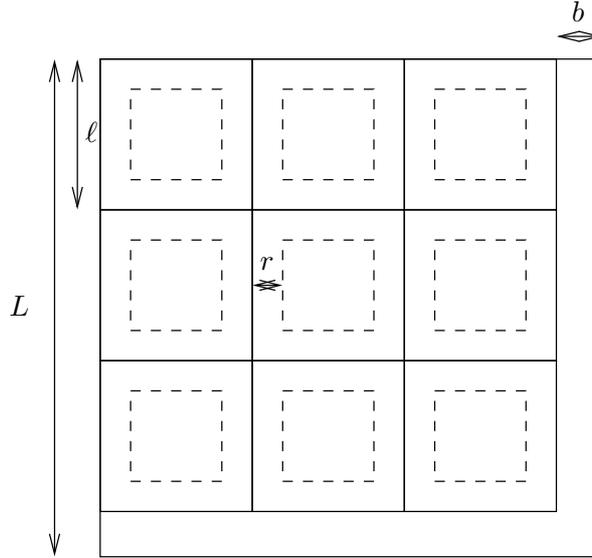
donc lim sup et lim inf coincide, i.e. la suite converge. \square

2.3 Energie libre pour les interactions à portée finie

Théorème 1. *Considérons une interaction à portée finie $\Phi \in \mathcal{I}_f$, et notons C_L un cube de côté L de \mathbb{Z}^d , e.g. $C_L = [0, L-1]^d$. Alors $f_{C_L}(\beta, \Phi)$ admet une limite quand $L \rightarrow \infty$, notée $f(\beta, \Phi)$, qui est finie.*

Démonstration. On va utiliser le fait que F_Λ est ‘‘presque-additive’’ par rapport à Λ . Dans la preuve de l'existence de la limite de Fekete l'inégalité n'est pas le point crucial.

Notons r la portée des interactions de Φ , choisissons $\ell > 2r$, et posons $L = a\ell + b$, où $a = \lfloor \frac{L}{\ell} \rfloor$, $b \in [0, \ell - 1]$. On décompose C_L en a^d cubes de côté ℓ , notés $\{C^{(\alpha)}, \alpha \in [1, a^d]\}$, et un reste $C^{(0)}$.



On peut écrire $H_{C_L}^\Phi = \sum_{\alpha=1}^{a^d} H_{C^{(\alpha)}}^\Phi + R$. Le reste R comprend les interactions incluses dans $C^{(0)}$, et celles à cheval entre deux $C^{(\alpha)}$ (y compris $\alpha = 0$). On veut se débarrasser de ce terme-là, on peut donc borner grossièrement avec

$$\|R\|_\infty \leq [a^d(\ell^d - (\ell - 2r)^d) + (L^d - (a\ell)^d)] \|\Phi\|_g. \quad (19)$$

En effet R est de la forme $\sum_{X \in \mathcal{R}} \phi_X$, et la preuve du lemme 1 montre que l'on a

$$\left\| \sum_{X \in \mathcal{R}} \phi_X \right\|_\infty \leq |\mathcal{R}_s| \|\Phi\|_g, \quad (20)$$

où \mathcal{R}_s désigne les sites qui apparaissent au moins une fois dans l'ensemble \mathcal{R} de parties de \mathbb{Z}^d . Ici les interactions de R ne peuvent faire intervenir que des sites dans $C^{(0)}$ (avec $|C^{(0)}| = (L^d - (a\ell)^d)$) et des sites à distance plus petite que r du bord des petits cubes, il y en a au plus $a^d(\ell^d - (\ell - 2r)^d)$.

En utilisant le lemme 2,

$$\left| \ln \sum_{\underline{\sigma}_{C_L}} e^{-\beta H_{C_L}^{\Phi}(\underline{\sigma}_{C_L})} - \ln \sum_{\underline{\sigma}_{C_L}} e^{-\beta \sum_{\alpha=1}^{a^d} H_{C^{(\alpha)}}^{\Phi}(\underline{\sigma}_{C_L})} \right| \leq \beta \|R\|_{\infty} \leq [a^d(\ell^d - (\ell - 2r)^d) + (L^d - (a\ell)^d)] \beta \|\Phi\|_{\mathfrak{g}}. \quad (21)$$

Les différents cubes de longueur ℓ n'interagissent pas entre eux, et contribuent tous de la même façon à cause de l'invariance par translation, i.e.

$$\sum_{\underline{\sigma}_{C_L}} e^{-\beta \sum_{\alpha=1}^{a^d} H_{C^{(\alpha)}}^{\Phi}(\underline{\sigma}_{C_L})} = \prod_{\alpha=1}^{a^d} \left(\sum_{\underline{\sigma}_{C^{(\alpha)}}} e^{-\beta H_{C^{(\alpha)}}^{\Phi}(\underline{\sigma}_{C^{(\alpha)}})} \right) \sum_{\underline{\sigma}_{C^{(0)}}} 1 = (Z_{C_{\ell}}(\beta, \Phi))^{a^d} |\chi|^{|C^{(0)}|}, \quad (22)$$

où on a la contribution purement entropique des spins de $C^{(0)}$. Le membre de gauche de (21) est donc

$$\left| (-\beta) F_{C_L}(\beta, \Phi) - a^d (-\beta) F_{C_{\ell}}(\beta, \Phi) - (L^d - (a\ell)^d) \ln(|\chi|) \right|. \quad (23)$$

On peut réécrire cette expression :

$$\beta \left| L^d f_{C_L}(\beta, \Phi) - a^d \ell^d f_{C_{\ell}}(\beta, \Phi) + (L^d - (a\ell)^d) \frac{\ln(|\chi|)}{\beta} \right|. \quad (24)$$

En divisant par βL^d , et en n'utilisant qu'une partie de l'inégalité sur la valeur absolue,

$$f_{C_L}(\beta, \Phi) \leq \left(\frac{a\ell}{L} \right)^d f_{C_{\ell}}(\beta, \Phi) + \left[\left(\frac{a\ell}{L} \right)^d \left(1 - \left(1 - \frac{2r}{\ell} \right)^d \right) + \left(1 - \left(\frac{a\ell}{L} \right)^d \right) \right] \|\Phi\|_{\mathfrak{g}} \quad (25)$$

$$- \left(1 - \left(\frac{a\ell}{L} \right)^d \right) \frac{\ln(|\chi|)}{\beta}. \quad (26)$$

En prenant $L \rightarrow \infty$, comme $a\ell/L \rightarrow 1$,

$$\limsup_{L \rightarrow \infty} f_{C_L}(\beta, \Phi) \leq f_{C_{\ell}}(\beta, \Phi) + \left(1 - \left(1 - \frac{2r}{\ell} \right)^d \right) \|\Phi\|_{\mathfrak{g}}. \quad (27)$$

On s'est débarrassé de la partie du reste qui venait de $C^{(0)}$ en prenant $L \rightarrow \infty$, en effet son volume est négligeable devant celui des cubes dans cette limite. Maintenant on se débarrasse des effets de bord sur chaque cube en prenant $\ell \rightarrow \infty$, comme la portée est finie ce sont des effets de surface négligeable devant le volume,

$$\limsup_{L \rightarrow \infty} f_{C_L}(\beta, \Phi) \leq \liminf_{\ell \rightarrow \infty} f_{C_{\ell}}(\beta, \Phi), \quad (28)$$

donc lim sup et lim inf coïncident et la limite existe.

Montrons que la limite est finie. D'après les lemmes 3 et 4 :

$$|f_{\Lambda}(\beta, \Phi) - f_{\Lambda}(\beta, 0)| \leq \|\Phi\|_{\mathfrak{g}} \Rightarrow |f_{\Lambda}(\beta, \Phi)| \leq \|\Phi\|_{\mathfrak{g}} + \frac{\ln(|\chi|)}{\beta}, \quad (29)$$

est vrai pour tout Λ fini, donc aussi pour la limite.

Pour montrer que la limite est finie on peut aussi noter que (27) implique que le limsup est $< +\infty$, en bornant de manière similaire le lim inf $> -\infty$ on aurait aussi la preuve de la finitude de la limite. \square

2.4 Généralisations

Robustesse du résultat obtenue précédemment :

- pour la famille $\mathcal{I}_{\mathfrak{g}}$ plus générale d'interactions
- avec des conditions aux bords arbitraire, et un sens de $\Lambda \rightarrow \infty$ plus général que juste en prenant des cubes.

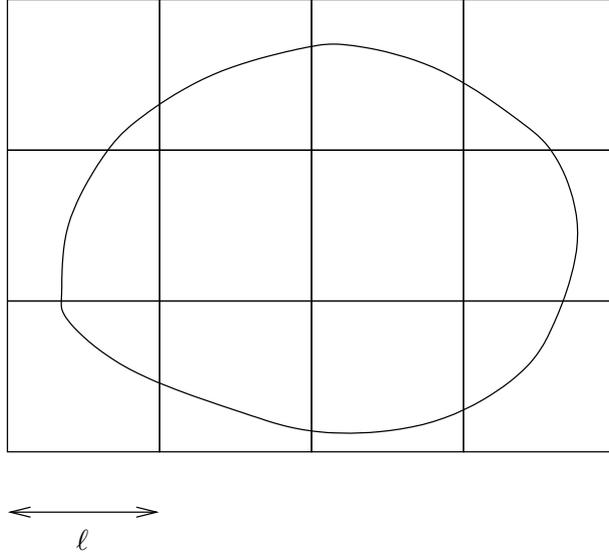
2.4.1 Limite au sens de van Hove

Pour l'instant $\Lambda \rightarrow \infty$ a été pris le long d'une suite particulière, les cubes de taille L avec $L \rightarrow \infty$. Les résultats de convergence restent vraies pour des suites plus générales. Mais il faut quand même imposer des conditions, pour que les effets de surface soient négligeables devant les effets de volume.

Définition 3. Une suite de parties finies Λ_n de \mathbb{Z}^d va à l'infini au sens de Van Hove si pour tout entier $\ell > 0$ et pour tout n on peut trouver un pavage de \mathbb{Z}^d par des cubes de côté ℓ avec $N_\ell^+(\Lambda_n)$ cubes qui ont une intersection non vide avec Λ_n , $N_\ell^-(\Lambda_n)$ cubes inclus dans Λ_n , et telle que

$$\forall \ell, \quad N_\ell^-(\Lambda_n) \xrightarrow{n \rightarrow \infty} +\infty, \quad \frac{N_\ell^+(\Lambda_n)}{N_\ell^-(\Lambda_n)} \xrightarrow{n \rightarrow \infty} 1. \quad (30)$$

Exemple :



$$N_\ell^-(\Lambda) = 2, \quad N_\ell^+(\Lambda) = 12$$

Si Λ_n est un rectangle de côtés L_n^1, \dots, L_n^d , la condition pour que la suite soit de van Hove est que tous les L_n^i tendent vers l'infini quand n va à l'infini, mais ils peuvent le faire avec des scalings différents. En effet $N_\ell^-(\Lambda_n) = \prod_{i=1}^d \lfloor \frac{L_n^i}{\ell} \rfloor$ et $N_\ell^+(\Lambda_n) = \prod_{i=1}^d \lceil \frac{L_n^i}{\ell} \rceil$.

Théorème 2. Considérons une interaction à portée finie $\Phi \in \mathcal{I}_f$, et une suite Λ_n qui va à l'infini au sens de Van Hove. Alors $f_{\Lambda_n}(\beta, \Phi) \rightarrow f(\beta, \Phi)$, la même limite que définie par les cubes.

Démonstration. On suit la même stratégie que pour $\Lambda = C_L$: on fixe $\ell > 2r$, on écrit $H_{\Lambda_n}^\Phi = \sum_{\alpha=1}^{N_\ell^-(\Lambda_n)} H_{C^{(\alpha)}}^\Phi + R$, où la première somme est sur les cubes de taille ℓ inclus dans Λ_n . On peut écrire $\Lambda_n = C^{(0)} \cup \cup_{\alpha=1}^{N_\ell^-(\Lambda_n)} C^{(\alpha)}$, où $C^{(0)}$ est la partie de Λ_n hors des cubes complètement inclus dans Λ_n .

$$\|R\|_\infty \leq [N_\ell^-(\Lambda_n)(\ell^d - (\ell - 2r)^d) + |C^{(0)}|] \|\Phi\|_g. \quad (31)$$

Avec le lemme et les mêmes remarques,

$$|\ln Z_{\Lambda_n}(\beta, \Phi) - N_\ell^-(\Lambda_n) \ln Z_{C_\ell}(\beta, \Phi) - |C^{(0)}| \ln(|\chi|)| \leq \beta \|R\|_\infty \leq \beta [N_\ell^-(\Lambda_n)(\ell^d - (\ell - 2r)^d) + |C^{(0)}|] \|\Phi\|_g. \quad (32)$$

En factorisant $-\beta$ à l'intérieur,

$$\left| F_{\Lambda_n}(\beta, \Phi) - N_\ell^-(\Lambda_n) F_{C_\ell}(\beta, \Phi) + |C^{(0)}| \frac{\ln(|\chi|)}{\beta} \right| \leq [N_\ell^-(\Lambda_n)(\ell^d - (\ell - 2r)^d) + |C^{(0)}|] \|\Phi\|_g. \quad (33)$$

En divisant par $|\Lambda_n|$,

$$f_{\Lambda_n}(\beta, \Phi) \leq \frac{N_\ell^-(\Lambda_n)\ell^d}{|\Lambda_n|} f_{C_\ell}(\beta, \Phi) + \left[\frac{N_\ell^-(\Lambda_n)\ell^d}{|\Lambda_n|} \left(1 - \left(1 - \frac{2r}{\ell} \right)^d \right) + \frac{|C^{(0)}|}{|\Lambda_n|} \right] \|\Phi\|_g \quad (34)$$

$$- \frac{|C^{(0)}| \ln(|\chi|)}{|\Lambda_n| \beta}. \quad (35)$$

Comme $|C^{(0)}| \leq \ell^d(N_\ell^+(\Lambda_n) - N_\ell^-(\Lambda_n))$ et $|\Lambda_n| \geq \ell^d N_\ell^-(\Lambda_n)$, on a

$$\frac{|C^{(0)}|}{|\Lambda_n|} \leq \frac{N_\ell^+(\Lambda_n)}{N_\ell^-(\Lambda_n)} - 1 \xrightarrow{n \rightarrow \infty} 0 \quad (36)$$

par définition de la limite de Van Hove les deux dernier termes disparaissent. De plus,

$$N_\ell^-(\Lambda_n)\ell^d \leq |\Lambda_n| \leq N_\ell^+(\Lambda_n)\ell^d, \quad 1 \leq \frac{|\Lambda_n|}{N_\ell^-(\Lambda_n)\ell^d} \leq \frac{N_\ell^+(\Lambda_n)}{N_\ell^-(\Lambda_n)}, \quad (37)$$

le dernier terme tend vers 1 par définition de la limite de Van Hove, donc $\frac{N_\ell^-(\Lambda_n)\ell^d}{|\Lambda_n|} \rightarrow 1$ on a donc

$$\limsup_{n \rightarrow \infty} f_{\Lambda_n}(\beta, \Phi) \leq f_{C_\ell}(\beta, \Phi) + \left(1 - \left(1 - \frac{2r}{\ell} \right)^d \right) \|\Phi\|_p. \quad (38)$$

C'est vrai pour tout ℓ , on peut prendre la limite $\ell \rightarrow \infty$, on a déjà montré qu'elle valait $f(\beta, \Phi)$. On peut borner inférieurement le lim inf de la même façon, donc la limite au sens de Van Hove existe et est égale à celle obtenue par les cubes. \square

2.4.2 Existence et propriétés de la limite pour des interactions plus générales

pour les autres interactions, en utilisant la densité des interactions à portée finie à l'intérieur de \mathcal{I}_g

Théorème 3. *Considérons une interaction $\Phi \in \mathcal{I}_g$, et notons Λ_n une suite qui va à l'infini au sens de Van Hove. Alors $f_{\Lambda_n}(\beta, \Phi)$ admet une limite, notée $f(\beta, \Phi)$, quand $n \rightarrow \infty$. Cette limite est finie, et $f(\beta, \Phi)$ vérifie par ailleurs*

(a) $|f(\beta, \Phi) - f(\beta, \Psi)| \leq \|\Phi - \Psi\|_g$ (Lipschitz)

(b) concavité dans les interactions, i.e. $f(\beta, \alpha\Phi + (1 - \alpha)\Psi) \geq \alpha f(\beta, \Phi) + (1 - \alpha)f(\beta, \Psi)$ pour tout $\alpha \in [0, 1]$.

Démonstration. On va s'appuyer sur le résultat obtenu pour les interactions à portée finie, qui sont en fait denses dans \mathcal{I}_g .

De $\Phi \in \mathcal{I}_g$ on définit une version régularisée, en coupant les interactions à trop longue portée, $\Phi^{(r)} \in \mathcal{I}^{(r)}$, avec

$$\phi_X^{(r)} = \begin{cases} \phi_X & \text{si } \text{diam}(X) \leq r \\ 0 & \text{sinon} \end{cases} \quad (39)$$

On a $\|\Phi - \Phi^{(r)}\|_g \xrightarrow{r \rightarrow \infty} 0$: écrivons $u_n = \sum_{\substack{X \ni 0 \\ \text{diam}(X)=n}} \frac{\|\phi_X\|_\infty}{|X|}$, les u_n sont positifs,

$$\|\Phi\|_g = \sum_{n=1}^{\infty} u_n < +\infty \text{ puisque } \Phi \in \mathcal{I}_g, \|\Phi - \Phi^{(r)}\|_g = \sum_{n=r+1}^{\infty} u_n = \|\Phi\|_g - \sum_{n=1}^r u_n, \quad (40)$$

donc la dernière quantité va à 0 quand $r \rightarrow \infty$ par la définition des séries pour des suites positives.

Comme $\|\Phi - \Phi^{(r_0)}\|_g \xrightarrow{r \rightarrow \infty} 0$, pour tout $\epsilon > 0$ on peut trouver r_0 tel que $\|\Phi - \Phi^{(r_0)}\|_g \leq \frac{\epsilon}{2}$. Alors en utilisant le lemme 3,

$$f_{\Lambda_n}(\beta, \Phi) \leq f_{\Lambda_n}(\beta, \Phi^{(r_0)}) + \frac{\epsilon}{2} \Rightarrow \limsup_{n \rightarrow \infty} f_{\Lambda_n}(\beta, \Phi) \leq f(\beta, \Phi^{(r_0)}) + \frac{\epsilon}{2} \quad (41)$$

en prenant $n \rightarrow \infty$ et en utilisant la convergence précédemment démontrée de $f_{\Lambda_n}(\beta, \Phi^{(r_0)})$. De la même façon on obtient

$$\liminf_{n \rightarrow \infty} f_{\Lambda_n}(\beta, \Phi) \geq f(\beta, \Phi^{(r_0)}) - \frac{\epsilon}{2}, \quad (42)$$

soit en combinant les deux

$$|\liminf_{n \rightarrow \infty} f_{\Lambda_n}(\beta, \Phi) - \limsup_{n \rightarrow \infty} f_{\Lambda_n}(\beta, \Phi)| \leq \epsilon. \quad (43)$$

Comme c'est vrai pour tout $\epsilon > 0$ les lim sup et lim inf sont égaux, et donc la limite existe.

La preuve que la limite est finie est exactement la même que dans le cas des interactions à portée finie.

La propriété (a) découle du lemme 3, puisque la limite existe l'inégalité valable pour tout Λ fini l'est aussi à la limite.

Pour prouver la concavité, il suffit de montrer l'inégalité à volume fini

$$f_{\Lambda}(\beta, \alpha\Phi + (1 - \alpha)\Psi) \geq \alpha f_{\Lambda}(\beta, \Phi) + (1 - \alpha)f_{\Lambda}(\beta, \Psi) \quad \forall \alpha \in [0, 1], \quad (44)$$

la limite (simple) quand $\Lambda \rightarrow \infty$ vérifiera encore l'inégalité.

En volume fini la fonction $t \rightarrow f_{\Lambda}(\beta, \Phi + t\Psi)$ est analytique en t car somme d'un nombre fini d'exponentielles,

$$f_{\Lambda}(\beta, \Phi + t\Psi) = -\frac{1}{\beta|\Lambda|} \ln \left(\sum_{\underline{\sigma}_{\Lambda}} e^{-\beta H_{\Lambda}^{\Phi+t\Psi}(\underline{\sigma})} \right) = -\frac{1}{\beta|\Lambda|} \ln \left(\sum_{\underline{\sigma}_{\Lambda}} e^{-\beta H_{\Lambda}^{\Phi}(\underline{\sigma}) - \beta t H_{\Lambda}^{\Psi}(\underline{\sigma})} \right), \quad (45)$$

on peut notamment calculer explicitement sa dérivée seconde

$$\frac{d^2}{dt^2} f_{\Lambda}(\beta, \Phi + t\Psi) = -\frac{\beta}{|\Lambda|} \langle (H_{\Lambda}^{\Psi} - \langle H_{\Lambda}^{\Psi} \rangle)^2 \rangle \leq 0, \quad \text{où } \langle \cdot \rangle = \frac{\sum_{\underline{\sigma}_{\Lambda}} \cdot e^{-\beta H_{\Lambda}^{\Phi+t\Psi}(\underline{\sigma})}}{\sum_{\underline{\sigma}_{\Lambda}} e^{-\beta H_{\Lambda}^{\Phi+t\Psi}(\underline{\sigma})}}. \quad (46)$$

Cela implique la concavité par des arguments standards :

$$\frac{d^2}{d\alpha^2} f_{\Lambda}(\beta, \alpha\Phi + (1 - \alpha)\Psi) = \frac{d^2}{d\alpha^2} f_{\Lambda}(\beta, \Psi + \alpha(\Phi - \Psi)) \leq 0, \quad (47)$$

or si $h(\alpha)$ est \mathcal{C}^2 sur $[0, 1]$ avec $h''(\alpha) \leq 0$ alors $h(\alpha) \geq \alpha h(1) + (1 - \alpha)h(0)$. En effet par le théorème de Rolle $h(1) - h(0) = h'(\alpha_0)$ pour un $\alpha_0 \in [0, 1]$. Pour $\alpha \in (0, \alpha_0]$ on peut écrire $\frac{h(\alpha) - h(0)}{\alpha} = h'(\alpha'_0)$ avec $\alpha'_0 \in [0, \alpha_0]$. Comme $h'' < 0$ $h'(\alpha'_0) \geq h'(\alpha_0)$, ce qui donne l'inégalité recherchée. Pour $\alpha \in [\alpha_0, 1]$ on choisit $\alpha'_0 \in [\alpha_0, 1]$ tel que $\frac{h(1) - h(\alpha)}{1 - \alpha} = h'(\alpha_0)$.

Remarque : on peut aussi obtenir la concavité en utilisant l'inégalité de Hölder, qu'on peut écrire

$$\sum_{\underline{\sigma}_{\Lambda}} e^{g_1(\underline{\sigma}) + g_2(\underline{\sigma})} \leq \left(\sum_{\underline{\sigma}_{\Lambda}} e^{p g_1(\underline{\sigma})} \right)^{\frac{1}{p}} \left(\sum_{\underline{\sigma}_{\Lambda}} e^{q g_2(\underline{\sigma})} \right)^{\frac{1}{q}}, \quad (48)$$

avec $g_1, g_2 \in \mathcal{C}_{\Lambda}$ et $\frac{1}{p} + \frac{1}{q} = 1$. En utilisant $g_1 = -\alpha\beta H_{\Lambda}^{\Phi}$, $g_2 = -(1 - \alpha)\beta H_{\Lambda}^{\Psi}$, $p = \frac{1}{\alpha}$ et $q = \frac{1}{1 - \alpha}$, l'inégalité devient

$$Z_{\Lambda}(\beta, \alpha\Phi + (1 - \alpha)\Psi) \leq (Z_{\Lambda}(\beta, \Phi))^{\alpha} (Z_{\Lambda}(\beta, \Psi))^{1 - \alpha}, \quad (49)$$

d'où l'inégalité recherchée en prenant le logarithme et en divisant par $-\beta|\Lambda|$. \square

2.4.3 Conditions aux bords

Dans la définition du Hamiltonien avec conditions aux bords libres H_Λ^Φ on a complètement éliminé l'extérieur Λ^c . On peut aussi considérer que le système de volume fini Λ baigne dans un environnement fixé avec lequel il interagit, et définir un Hamiltonien avec conditions aux bords fixées.

On définit la concaténation de deux configurations, $(\underline{\sigma}_\Lambda, \underline{\tau}_{\Lambda^c})$, par

$$(\underline{\sigma}_\Lambda, \underline{\tau}_{\Lambda^c})_x = \begin{cases} \sigma_x & \text{si } x \in \Lambda \\ \tau_x & \text{si } x \in \Lambda^c \end{cases}, \quad (50)$$

i.e. on prend la configuration $\underline{\sigma}$ à l'intérieur de Λ et $\underline{\tau}$ à l'extérieur.

Définition 4. *Hamiltonien pour une interaction $\Phi \in \mathcal{I}_p$ dans un domaine fini $\Lambda \subset_f \mathbb{Z}^d$ avec conditions aux bords $\underline{\tau}$:*

$$H_\Lambda^{\Phi, \underline{\tau}}(\underline{\sigma}) = \sum_{X, X \cap \Lambda \neq \emptyset} \phi_X((\underline{\sigma}_\Lambda, \underline{\tau}_{\Lambda^c})) \quad (51)$$

vu comme fonction de $\underline{\sigma}$ c'est dans \mathcal{C}_Λ , ne dépend que de $\underline{\sigma}_\Lambda$, vu comme fonction de $\underline{\tau}$ c'est dans \mathcal{C}_{Λ^c} .

Exemple pour Ising,

$$H_\Lambda^{\Phi, \underline{\tau}}(\underline{\sigma}) = H_\Lambda^\Phi(\underline{\sigma}) - J \sum_{\langle x, y \rangle, x \in \Lambda, y \notin \Lambda} \sigma_x \tau_y, \quad (52)$$

on rajoute donc un champ magnétique $-h\tau_y$ sur les sites x à la frontière intérieure de Λ .

On s'est limitée aux interactions dans \mathcal{I}_p , le petit espace des interactions, car même quand Λ est fini la somme dans (51) est infinie; $\Phi \in \mathcal{I}_p$ est une condition suffisante pour la faire converger, uniformément, avec

$$\|H_\Lambda^{\Phi, \underline{\tau}}\|_\infty \leq \sum_{X, X \cap \Lambda \neq \emptyset} \|\phi_X\|_\infty \leq \sum_{x \in \Lambda} \sum_{X \ni x} \|\phi_X\|_\infty = |\Lambda| \|\Phi\|_p. \quad (53)$$

Définitions des quantités thermodynamiques associées :

$$Z_\Lambda^\tau(\beta, \Phi) = \sum_{\underline{\sigma}_\Lambda} e^{-\beta H_\Lambda^{\Phi, \underline{\tau}}(\underline{\sigma})}, \quad F_\Lambda^\tau(\beta, \Phi) = -\frac{1}{\beta} \ln Z_\Lambda^\tau(\beta, \Phi), \quad f_\Lambda^\tau(\beta, \Phi) = -\frac{1}{\beta|\Lambda|} \ln Z_\Lambda^\tau(\beta, \Phi) \quad (54)$$

Comme c'est un effet de surface on s'attend à ce que ça continue à converger, et vers la même limite que dans le cas des conditions aux limites libre. En effet on a le théorème suivant :

Théorème 4. *Considérons une interaction $\Phi \in \mathcal{I}_p$, Λ_n une suite qui va à l'infini au sens de Van Hove, et $\underline{\tau}_n$ une suite arbitraire de conditions aux bords. Alors $f_{\Lambda_n}^{\tau_n}(\beta, \Phi)$ admet $f(\beta, \Phi)$ comme limite quand $n \rightarrow \infty$.*

Démonstration. Considérons pour commencer le cas où $\Phi \in \mathcal{I}^{(r)}$, interaction à portée maximale r , et où Λ_n est une suite de cubes de côté L_n . On a, pour $L_n > 2r$

$$\|H_{\Lambda_n}^{\Phi, \underline{\tau}_n} - H_{\Lambda_n}^\Phi\|_\infty \leq [L_n^d - (L_n - 2r)^d] \|\Phi\|_p, \quad (55)$$

car les interactions de bord non comptées dans $H_{\Lambda_n}^\Phi$ font intervenir les sites de Λ_n à distance au plus r du bord. Donc

$$|f_{\Lambda_n}^{\tau_n}(\beta, \Phi) - f_{\Lambda_n}(\beta, \Phi)| \leq \left[1 - \left(1 - \frac{L_n}{r}\right)^d\right] \|\Phi\|_p. \quad (56)$$

Comme le membre de droite va à 0 et que dans le membre de gauche une des deux suites converge vers $f(\beta, \Phi)$, l'autre aussi.

Si Λ_n est une suite de Van Hove générique le même type de raisonnement s'applique, le bord à distance r de Λ_n est toujours négligeable devant Λ_n quand $n \rightarrow \infty$.

Finalement, si $\Phi \in \mathcal{I}_p$, on raisonne de nouveau par densité, en utilisant les troncations $\Phi^{(r)}$ définies en (39).

□

Conditions aux limites périodiques

Si Λ est un rectangle de côtés L_1, \dots, L_d , on peut définir à partir de $\underline{\sigma}_\Lambda$ une configuration $\underline{\tau} = \text{per}_\Lambda(\underline{\sigma})$ par

$$\tau_{x_1, \dots, x_d} = \sigma_{x_1 + n_1 L_1, \dots, x_d + n_d L_d} , \quad (57)$$

où les $(n_1, \dots, n_d) \in \mathbb{Z}^d$ sont choisis de telle façon que $(x_1 + n_1 L_1, \dots, x_d + n_d L_d) \in \Lambda$. On peut donc construire un Hamiltonien avec conditions aux limites périodiques selon

$$H_\Lambda^{\Phi, \text{per}}(\underline{\sigma}) = H_\Lambda^{\Phi, \text{per}_\Lambda(\underline{\sigma})}(\underline{\sigma}) , \quad (58)$$

et les grandeurs thermodynamiques associées. Le résultat de convergence précédent s'applique à ce cas particulier de conditions aux limites.

3 Mesures de Gibbs

Dans l'introduction on avait présenté les deux aspects de la mécanique statistique, i.e. les propriétés thermodynamiques (énergie libre) et les lois de probabilité dont on voulait décrire rigoureusement le passage à la limite thermodynamique. On attaque ici le deuxième aspect, à savoir la généralisation de la formule

$$\mu(\underline{\sigma}) = \frac{1}{Z} e^{-\beta H(\underline{\sigma})} \quad (59)$$

au cas d'un système en volume infini. Telle quelle cette formule n'a plus de sens, les configurations ont des énergies infinies, on ne peut définir une mesure de probabilité sur l'espace des configurations en volume infini par les probabilités de configurations individuelles, qui sont a priori toutes nulles. On va donc avoir besoin de s'appuyer sur la théorie de la mesure au sens mathématique du terme. Avant cela il sera utile de considérer le cas élémentaire des probabilités de Gibbs-Boltzmann en volume fini, où ces difficultés ne se présentent pas mais qui nous serviront de guide pour le passage ultérieur en volume infini.

3.1 Mesures de Gibbs-Boltzmann en volume fini

Dans cette section on se fixe $\Delta \subset_f \mathbb{Z}^d$ une fois pour toute, et on note $\underline{\sigma}$ au lieu de $\underline{\sigma}_\Delta$ une configuration, il y en a un nombre fini puisque Δ est supposé fini. On peut donc définir une probabilité η de manière élémentaire en se donnant les nombres $\eta(\underline{\sigma})$, probabilité pour chacune des configurations (pour simplifier on supposera $\eta(\underline{\sigma}) > 0$ pour toutes les configurations).

3.1.1 Lois marginales et conditionnelles

La loi marginale de η pour un sous-ensemble de variables dans $\Lambda \subset \Delta$ est définie naturellement par

$$\eta_\Lambda(\underline{\sigma}_\Lambda) = \sum_{\underline{\sigma}_{\Lambda^c}} \eta((\underline{\sigma}_\Lambda, \underline{\sigma}_{\Lambda^c})) , \quad (60)$$

où Λ^c est ici $\Delta \setminus \Lambda$.

Introduisons aussi la loi conditionnelle

$$\eta_\Lambda(\underline{\sigma}_\Lambda | \underline{\sigma}_{\Lambda^c}) = \frac{\eta((\underline{\sigma}_\Lambda, \underline{\sigma}_{\Lambda^c}))}{\eta_{\Lambda^c}(\underline{\sigma}_{\Lambda^c})} , \quad (61)$$

qui est bien normalisée par rapport à son premier argument,

$$\sum_{\underline{\sigma}_\Lambda} \eta_\Lambda(\underline{\sigma}_\Lambda | \underline{\sigma}_{\Lambda^c}) = \frac{\sum_{\underline{\sigma}_\Lambda} \eta((\underline{\sigma}_\Lambda, \underline{\sigma}_{\Lambda^c}))}{\eta_{\Lambda^c}(\underline{\sigma}_{\Lambda^c})} = 1 , \quad (62)$$

par définition de la loi marginale η_{Λ^c} . On peut utiliser cette loi conditionnelle pour décomposer les probabilités sous la forme

$$\eta_\Lambda(\underline{\sigma}_\Lambda) = \sum_{\underline{\sigma}_{\Lambda^c}} \eta_{\Lambda^c}(\underline{\sigma}_{\Lambda^c}) \eta_\Lambda(\underline{\sigma}_\Lambda | \underline{\sigma}_{\Lambda^c}) , \quad (63)$$

autrement dit tirer une configuration dans un sous-volume Λ est équivalent à tirer une condition aux bords $\underline{\sigma}_{\Lambda^c}$ à l'extérieur, puis tirer l'intérieur de Λ conditionné à son extérieur. On peut aussi écrire :

$$\eta(\underline{\sigma}) = \sum_{\underline{\tau}} \eta(\underline{\tau}) \sum_{\underline{\tau}'_\Lambda} \eta_\Lambda(\underline{\tau}'_\Lambda | \underline{\tau}_{\Lambda^c}) \mathbb{I}(\underline{\sigma} = (\underline{\tau}'_\Lambda, \underline{\tau}_{\Lambda^c})) , \quad (64)$$

qui s'interprète de la façon suivante : pour tirer une configuration $\underline{\sigma}$ du système on peut tirer une autre configuration $\underline{\tau}$, garder une partie $\underline{\tau}_{\Lambda^c}$ de cette configuration et l'utiliser comme condition aux bords pour tirer $\underline{\tau}'_\Lambda$ à l'intérieur, et prendre pour $\underline{\sigma}$ la concaténation $(\underline{\tau}'_\Lambda, \underline{\tau}_{\Lambda^c})$. C'est bien sûr une tautologie ici, mais on va suivre cette intuition pour la généralisation en volume infini.

3.1.2 Mesures de l'ensemble canonique

La discussion ci-dessus était valable pour n'importe quelle loi de probabilité η . Considérons le cas particulier

$$\eta(\underline{\sigma}) = \frac{1}{Z} e^{-\beta H(\underline{\sigma})}, \quad H(\underline{\sigma}) = \sum_{X \subset \Delta} \phi_X(\underline{\sigma}), \quad (65)$$

i.e. la mesure de Gibbs-Boltzmann associée à l'Hamiltonien avec conditions aux bords libres dans Δ . En général le calcul de la fonction de partition Z est difficile (il nécessite la sommation sur un nombre de configurations exponentiellement grand dans le volume $|\Delta|$), et donc même si on connaît $H(\underline{\sigma})$ pour toutes les configurations on ne sait pas calculer facilement $\eta(\underline{\sigma})$. Le même problème se pose pour les marginales, par exemple pour calculer la marginale sur une seule variable $\eta_{\{i\}}(\sigma_i)$ il faut quand même sommer sur toutes les configurations. Par contre les lois conditionnelles sont, de ce point de vue, beaucoup plus simples. En effet,

$$\eta_\Lambda(\underline{\sigma}_\Lambda | \underline{\sigma}_{\Lambda^c}) = \frac{\eta((\underline{\sigma}_\Lambda, \underline{\sigma}_{\Lambda^c}))}{\eta_{\Lambda^c}(\underline{\sigma}_{\Lambda^c})} = \frac{e^{-\beta H((\underline{\sigma}_\Lambda, \underline{\sigma}_{\Lambda^c}))}}{\sum_{\underline{\sigma}'_\Lambda} e^{-\beta H((\underline{\sigma}'_\Lambda, \underline{\sigma}_{\Lambda^c}))}}, \quad (66)$$

en simplifiant les Z au numérateur et au dénominateur. Dans ce dernier la sommation n'est que sur les configurations de $\underline{\sigma}'_\Lambda$, donc si Λ est petit, quelque soit la taille du système total Δ le nombre de configurations sur lesquelles on doit sommer ne croît pas. Par ailleurs, on voit que si l'on ajoute à H n'importe quelle fonction g de $\underline{\sigma}_{\Lambda^c}$, indépendante de $\underline{\sigma}_\Lambda$, la loi conditionnelle ne change pas car les facteurs $e^{-\beta g}$ se simplifient au numérateur et au dénominateur. On peut donc remplacer $H((\underline{\sigma}_\Lambda, \underline{\sigma}_{\Lambda^c}))$ par sa contribution qui dépend explicitement de $\underline{\sigma}_\Lambda$, i.e.

$$\begin{aligned} \eta_\Lambda(\underline{\sigma}_\Lambda | \underline{\sigma}_{\Lambda^c}) &= \frac{1}{Z_\Lambda(\underline{\sigma}_{\Lambda^c})} e^{-\beta H_\Lambda^{\underline{\sigma}_{\Lambda^c}}(\underline{\sigma}_\Lambda)}, \quad \text{avec} \\ H_\Lambda^{\underline{\sigma}_{\Lambda^c}}(\underline{\sigma}_\Lambda) &= \sum_{\substack{X \cap \Lambda \neq \emptyset \\ X \subset \Delta}} \phi_X((\underline{\sigma}_\Lambda, \underline{\sigma}_{\Lambda^c})), \quad Z_\Lambda(\underline{\sigma}_{\Lambda^c}) = \sum_{\underline{\sigma}'_\Lambda} e^{-\beta H_\Lambda^{\underline{\sigma}_{\Lambda^c}}(\underline{\sigma}'_\Lambda)}. \end{aligned} \quad (67)$$

On retrouve l'Hamiltonien avec conditions aux bords fixées, à la seule différence qu'on travaillait ici dans un volume total fixé Δ . Par exemple dans le cas du modèle d'Ising $\eta_\Lambda(\underline{\sigma}_\Lambda | \underline{\sigma}_{\Lambda^c})$ dépend de la condition aux bords $\underline{\sigma}_{\Lambda^c}$ par l'intermédiaire des spins à la frontière extérieure de Λ , qui créent des champs magnétiques effectifs sur les spins à la frontière intérieure.

Dans la suite on va reprendre $\Delta \rightarrow \mathbb{Z}^d$, mais la définition des lois conditionnelles dans ce cas-là est plus subtile (on ne peut pas diviser par $\eta_{\Lambda^c}(\underline{\sigma}_{\Lambda^c})$ qui est en général nulle si $\Lambda^c = \mathbb{Z}^d \setminus \Lambda$ est infinie). On va d'abord donner quelques définitions plus formelles pour manipuler des lois de probabilité sur un espace de configurations infini.

3.2 Rappels de topologie et de théorie de la mesure

On veut maintenant définir des mesures de probabilité pour $\Delta \rightarrow \mathbb{Z}^d$, en fait il est plus satisfaisant de travailler directement avec $\Delta = \mathbb{Z}^d$, mais du coup il faut être plus précis sur la notion de loi de probabilité, et utiliser des notions de théorie de la mesure. On va mettre deux structures sur l'espace de configurations $\Sigma = \chi^{\mathbb{Z}^d}$, en en faisant un espace topologique et mesurable.

On notera les cylindres élémentaires

$$C_\Lambda(\underline{\sigma}_\Lambda) = \{\underline{\tau} \in \Sigma : \tau_\Lambda = \underline{\sigma}_\Lambda\}, \quad (68)$$

i.e. les parties de Σ où l'on a contraint la valeur des spins dans une région Λ (pas forcément finie) de \mathbb{Z}^d . On dira que Λ est le support du cylindre C_Λ . Quand Λ est réduit à un seul point on écrira avec un léger abus de notation :

$$C_x(\sigma) = \{\underline{\tau} \in \Sigma : \tau_x = \sigma\}. \quad (69)$$

Les projections canoniques p_x de Σ dans χ sont définies comme :

$$p_x : \begin{cases} \Sigma & \rightarrow \chi \\ \underline{\sigma} & \rightarrow \sigma_x \end{cases}, \quad (70)$$

de telle sorte que les images réciproques de p_x sont reliées aux cylindres élémentaires C_x , par exemple :

$$p_x^{-1}(\{\sigma\}) = C_x(\sigma). \quad (71)$$

Σ comme espace topologique

- Rappelons qu'une topologie sur un ensemble X est une collection de parties de X (ses ouverts), contenant \emptyset et X , stable par intersection finie et union arbitraire. Un ensemble muni d'une topologie est dit espace topologique ; une fonction d'un espace topologique dans un autre est dite continue si l'image réciproque de tout ouvert est ouvert. L'intersection de topologies est une topologie, on peut donc définir la topologie la moins fine contenant une partie des parties de X comme l'intersection de toutes les topologies ayant cette propriété.
- Sur χ fini il est naturel d'introduire la topologie discrète (toutes les parties de χ sont ouvertes). Toute fonction de χ dans \mathbb{R} est alors continue au sens topologique du terme, même si cette notion de "continuité" n'est pas celle intuitive puisque l'espace de départ χ est discret.
- Sur Σ qui est un produit cartésien de copies de χ on utilisera la topologie produit, i.e. la topologie la moins fine telle que les projections p_x soient continues. Explicitons les ouverts de cette topologie. Comme vu plus haut les cylindres élémentaires C_x sont des images réciproques des projections, ils sont donc ouverts. Une topologie étant stable par intersection finie, les cylindres élémentaires C_Λ sont aussi ouverts quand leur support Λ est une partie *finie* de \mathbb{Z}^d . Les unions (arbitraires a priori) de cylindres élémentaires à support fini sont donc aussi des ouverts de Σ . Comme ce dernier ensemble forme une topologie (il est stable par union et intersection finie, comme toute union d'intersections finies de parties d'un ensemble), c'est bien la topologie la moins fine rendant les projections continues. En résumé, les ouverts de la topologie produit sur Σ sont les unions de cylindres élémentaires à support fini.
- Les cylindres élémentaires $C_\Lambda(\underline{\sigma}_\Lambda)$ à support Λ fini sont dénombrables, ils forment donc une base dénombrable d'ouverts pour Σ , qui est en conséquence un espace séparable. En général une base d'ouverts d'un espace topologique est un ensemble de parties de l'espace tel que tout ouvert est une union d'éléments de la base, et un espace séparable est un espace topologique qui possède une partie dénombrable dense (i.e. dont l'adhérence, le plus petit fermé le contenant, est tout l'espace).

- Les voisinages d'une configuration $\underline{\sigma}$ sont les (unions de) cylindres élémentaires $C_\Lambda(\underline{\sigma}_\Lambda)$ à support Λ fini. Ces derniers forment un système fondamental de voisinages de $\underline{\sigma}$. En général les voisinages d'un point dans un espace topologique sont les parties qui contiennent un ouvert qui contient le point, et les systèmes fondamentaux de voisinage sont tels que tout voisinage contienne un élément du système fondamental.
- Dans cette topologie une suite de configurations $\underline{\sigma}^{(n)}$ converge vers $\underline{\sigma}$ ssi

$$\forall x \in \mathbb{Z}^d \quad \exists N \quad \forall n \geq N \quad \sigma_x^{(n)} = \sigma_x . \quad (72)$$

Le rang N au delà duquel la suite $\sigma_x^{(n)}$ devient constante peut bien sûr dépendre du site x . En général dans un espace topologique une suite x_n converge vers x si pour tout voisinage de x (ou tout élément d'un système fondamental de voisinage de x) il existe un rang N tel que la suite est dans ce voisinage pour tous les $n \geq N$.

- Une propriété importante et utile pour la suite est que Σ est compact. En effet le théorème de Tychonov assure qu'un produit (avec la topologie produit) d'espaces compacts est compact, et clairement χ est compact (tout recouvrement ouvert de χ est par lui-même fini, l'existence d'un sous-recouvrement fini est triviale).
- Notons \mathcal{C} l'ensemble des fonctions continues de Σ dans \mathbb{R} . C'est un espace vectoriel, que l'on munira de la norme $\|f\|_\infty = \sup_{\underline{\sigma}} |f(\underline{\sigma})|$ (dans cette section on notera souvent f des fonctions de Σ dans \mathbb{R} , il ne devrait pas y avoir de confusion possible avec l'énergie libre $f(\beta, \Phi)$). Puisque Σ est compact toute fonction continue est bornée, $\|\cdot\|_\infty$ est bien une norme, et \mathcal{C} est un Banach pour la norme $\|\cdot\|_\infty$. On avait défini l'ensemble \mathcal{C}_Λ des fonctions qui ne dépendent que de $\underline{\sigma}_\Lambda$; clairement toutes ces fonctions sont continues quand Λ est fini (leurs images réciproques sont des unions de cylindres élémentaires C_Λ qui sont bien des ouverts). Par ailleurs $\mathcal{C}_f = \cup_{\Lambda \subset_f \mathbb{Z}^d} \mathcal{C}_\Lambda$ est dense dans \mathcal{C} , pour la topologie induite par la norme $\|\cdot\|_\infty$, en d'autres termes toute fonction continue est approximable par des fonctions qui ne dépendent que d'un nombre fini de variables (c'est le théorème de Stone-Weierstrass qui l'assure).
- La topologie sur Σ est métrisable, avec la distance

$$d(\underline{\sigma}, \underline{\tau}) = \sum_{x \in \mathbb{Z}^d} e^{-\text{dist}(0,x)} \mathbb{I}(\sigma_x \neq \tau_x) . \quad (73)$$

Vérifions que c'est en effet une distance. Le nombre de sommets de \mathbb{Z}^d à distance ℓ de l'origine croît comme ℓ^{d-1} , donc grâce au terme $e^{-\ell}$ la somme est convergente quelque soit les configurations $\underline{\sigma}$ et $\underline{\tau}$. C'est une somme convergente de termes non-négatifs, elle est donc non-négative, et nulle si et seulement si tous les termes sont nuls, ce qui implique $\sigma_x = \tau_x$ pour tout les x , i.e. $\underline{\sigma} = \underline{\tau}$. La symétrie $d(\underline{\sigma}, \underline{\tau}) = d(\underline{\tau}, \underline{\sigma})$ est évidente, et finalement $\mathbb{I}(\sigma_x \neq \sigma'_x) \leq \mathbb{I}(\sigma_x \neq \sigma''_x) + \mathbb{I}(\sigma'_x \neq \sigma''_x)$ (en vérifiant les différents cas possibles), ce qui donne l'inégalité triangulaire $d(\underline{\sigma}, \underline{\sigma}'') \leq d(\underline{\sigma}, \underline{\sigma}') + d(\underline{\sigma}', \underline{\sigma}'')$.

Vérifions maintenant que la topologie introduite ci-dessus correspond à celle induite par cette distance. En général dans la topologie induite par une distance les ouverts sont les parties U telles que pour tout point x de U il existe une boule ouverte (pour la distance considérée) centrée en x , de rayon strictement positif, et incluse dans U .

Considérons un voisinage de la configuration $\underline{\tau}$, i.e. $\{\underline{\sigma} : \underline{\sigma}_\Lambda = \underline{\tau}_\Lambda\}$ pour un Λ fini. Notons $\epsilon = \min_{x \in \Lambda} e^{-x}$. Alors la boule ouverte de rayon ϵ (qui est bien > 0) autour de $\underline{\tau}$ est incluse dans ce voisinage : s'il y avait $x \in \Lambda$ avec $\sigma_x \neq \tau_x$, nécessairement $d(\underline{\sigma}, \underline{\tau}) \geq \epsilon$. Réciproquement toute boule ouverte autour de $\underline{\tau}$ contient un voisinage de $\underline{\tau}$: pour $\epsilon > 0$ on peut trouver Λ fini tel que $\sum_{x \notin \Lambda} e^{-x} < \epsilon$, et si $\underline{\sigma}_\Lambda = \underline{\tau}_\Lambda$ alors $d(\underline{\sigma}, \underline{\tau}) < \epsilon$.

- Σ est complet car métrisable et compact.
- Les fonctions continues de Σ dans \mathbb{R} sont uniformément continues, d'après le théorème de Heine, puisque Σ est compact.

Σ comme espace mesurable

- Rappelons qu'une tribu (σ -algèbre) sur un ensemble X est une collection \mathcal{B} de parties de X (les parties mesurables) contenant \emptyset et X , stable par passage au complémentaire et par intersection et union dénombrable. Un ensemble (X, \mathcal{B}) muni d'une tribu est appelé espace mesurable; une fonction d'un espace mesurable dans un autre est dite mesurable si l'image réciproque d'une partie mesurable est une partie mesurable. L'intersection de tribus est une tribu, on peut donc définir la tribu engendrée par une famille \mathcal{A} de parties de X comme l'intersection de toutes les tribus contenant cette famille de parties, que l'on note $\sigma(\mathcal{A})$. Si X est aussi un espace topologique on appelle tribu Borélienne celle engendrée par les ouverts de X . Si X et Y sont deux espaces topologiques, rendus mesurables par leurs tribus Boréliennes respectives, alors toute fonction continue de X vers Y est aussi mesurable (il suffit de vérifier que $f^{-1}(B)$ est mesurable pour les ensembles B appartenant à une classe engendrant la tribu de Y , les ouverts de Y conviennent).
- Considérons deux espaces mesurables (X_1, \mathcal{B}_1) et (X_2, \mathcal{B}_2) . Sur le produit cartésien $X_1 \times X_2$ on définit la tribu produit $\mathcal{B}_1 \otimes \mathcal{B}_2$ comme la plus petite tribu rendant mesurable les projections canoniques de $X_1 \times X_2$ vers X_1 et X_2 . Elle est donc engendrée par les pavés mesurables, $\mathcal{B}_1 \otimes \mathcal{B}_2 = \sigma(\{B_1 \times B_2\}_{B_1 \in \mathcal{B}_1, B_2 \in \mathcal{B}_2})$. Rappelons deux propriétés de cette construction :
 - Si $C \in \mathcal{B}_1 \otimes \mathcal{B}_2$ est une partie mesurable pour la tribu produit, x_1 un point de X_1 , alors la section $C_{x_1} = \{x_2 \in X_2 : (x_1, x_2) \in C\}$ est mesurable pour \mathcal{B}_2 , i.e. $C_{x_1} \in \mathcal{B}_2$.
 - Si f est une fonction mesurable de $(X_1 \times X_2, \mathcal{B}_1 \otimes \mathcal{B}_2)$, x_1 un point de X_1 , alors la fonction $f_{x_1}(x_2) = f(x_1, x_2)$ est \mathcal{B}_2 -mesurable.
- Sur χ fini on prend naturellement la tribu contenant toutes les parties de χ , sur Σ la tribu produit dénotée \mathcal{F} , i.e. la plus petite tribu telle que les projections p_x soient mesurables pour tout $x \in \mathbb{Z}^d$ (c'est une généralisation du cas vu ci-dessus à un produit de plus de deux espaces mesurables). C'est donc la tribu engendrée par les cylindres élémentaires supportés par un seul point, $\mathcal{F} = \sigma(\{C_x(\tau)\}_{x \in \mathbb{Z}^d, \tau \in \chi})$. Par stabilité des tribus sous les intersections dénombrables les cylindres C_Λ sont mesurables, que Λ soit fini ou pas. En particulier les singletons $\{\underline{x}\}$ sont mesurables (puisque \mathbb{Z}^d est dénombrable). Contrairement au cas de la topologie on ne peut pas expliciter complètement la tribu produit \mathcal{F} . Cette dernière est plus petite que l'ensemble des parties de Σ , on peut en effet construire des parties non-mesurables (mais ces constructions reposent habituellement sur l'axiome du choix). Si Σ était dénombrable la tribu \mathcal{F} , qui contient les unions dénombrables des singletons, contiendrait donc toutes les parties de Σ ; mais si E est fini, F dénombrable, E^F n'est pas dénombrable (de cardinal aleph 1 et pas aleph 0), par exemple $\{0, 1\}^{\mathbb{N}}$ est isomorphe à $[0, 1]$, avec l'écriture binaire des réels.
- Les deux structures introduites ci-dessus sur Σ (espace topologique et espace mesurable) sont compatibles, dans le sens que \mathcal{F} est la tribu Borélienne pour la topologie produit (la plus petite tribu contenant tous les ouverts). Plus généralement la tribu produit de tribus Boréliennes est la tribu Borélienne de la topologie produit si celle-ci admet une base dénombrable d'ouverts (i.e. est séparable). Les fonctions continues de Σ dans \mathbb{R} seront donc toujours mesurables (d'après la propriété générale des espaces topologiques munis de leur tribu Borélienne rappelée plus haut).
- Pour une partie Λ (finie ou pas) de \mathbb{Z}^d on notera \mathcal{F}_Λ la sous-tribu de \mathcal{F} engendrée par les cylindres élémentaires supportés par les $x \in \Lambda$, $\mathcal{F}_\Lambda = \sigma(\{C_x(\tau)\}_{x \in \Lambda, \tau \in \chi})$. Les fonctions \mathcal{F}_Λ mesurables ne dépendront donc que des variables dans Λ . Lorsque Λ est finie les parties mesurables de \mathcal{F}_Λ sont les produits cartésiens de parties arbitraires de χ^Λ avec tout χ^{Λ^c} .
- Que Λ soit finie ou pas on peut voir \mathcal{F}_Λ soit comme une tribu de Σ , soit comme la tribu produit de χ^Λ . Avec un léger abus de notation on a donc par associativité du produit $\mathcal{F} = \mathcal{F}_\Lambda \otimes \mathcal{F}_{\Lambda^c}$.

- Une mesure (finie) sur un espace mesurable X quelconque (plus précisément sur (X, \mathcal{B})) est une application μ de \mathcal{B} dans $[0, \infty[$, avec $\mu(\emptyset) = 0$ et μ σ -additive (si les B_i sont une famille dénombrable d'éléments de \mathcal{B} , tous disjoints, $\mu(\cup B_i) = \sum \mu(B_i)$). Une mesure est dite de probabilité si $\mu(X) = 1$. On note \mathcal{M} l'ensemble de ces mesures de probabilité. On appelle souvent évènements les parties mesurables.
- On peut maintenant définir la notion d'intégrale d'une fonction par rapport à une mesure, que l'on notera

$$\mu(f) = \int d\mu(x)f(x) = \int \mu(dx)f(x) , \quad (74)$$

On note dx même si x est discret, ce n'est évidemment pas un élément différentiel infinitésimal. Si f est une fonction mesurable étagée (i.e. qui prennent un nombre fini de valeurs distinctes), $f = \sum_{i=1}^n f_i \mathbb{I}_{B_i}$ avec des parties B_i mesurables disjointes, l'intégrale de f est définie comme $\mu(f) = \sum_i f_i \mu(B_i)$. On étend cette définition aux fonctions $f(x)$ mesurables positives quelconques comme le sup des intégrales de fonctions étagées inférieures à la fonction f . Comme ici les fonctions continues sont bornées (puisque Σ est compact) on peut définir $\mu(f)$ pour toute fonction $f \in \mathcal{C}$ pas forcément positive, comme $\mu(f^+) - \mu(f^-)$, avec $f^+(\underline{\sigma}) = \max(f(\underline{\sigma}), 0)$ et $f^-(\underline{\sigma}) = \max(-f(\underline{\sigma}), 0)$.

- On utilisera la même notation $\mu(B)$ et $\mu(f)$ pour la probabilité d'un évènement mesurable et la moyenne d'une fonction ; la première peut être vue comme un cas particulier de la deuxième, en prenant pour f la fonction indicatrice de la partie B .
- Rappelons deux constructions de mesure :
 - Si μ est une mesure sur (X, \mathcal{B}) et g une fonction mesurable positive, on peut définir une mesure notée $(\mu \cdot g)$ selon

$$(\mu \cdot g)(B) = \int \mu(dx) g(x) \mathbb{I}(x \in B) \quad \forall B \in \mathcal{B} . \quad (75)$$

- Si μ_1 est une mesure sur (X_1, \mathcal{B}_1) et μ_2 une mesure sur (X_2, \mathcal{B}_2) , on peut construire une (unique) mesure notée $(\mu_1 \otimes \mu_2)$ sur $(X_1 \times X_2, \mathcal{B}_1 \otimes \mathcal{B}_2)$, telle que

$$(\mu_1 \otimes \mu_2)(B_1 \times B_2) = \mu_1(B_1) \mu_2(B_2) \quad \forall B_1 \in \mathcal{B}_1 , B_2 \in \mathcal{B}_2 . \quad (76)$$

Alors pour $C \in \mathcal{B}_1 \times \mathcal{B}_2$,

$$(\mu_1 \otimes \mu_2)(C) = \int_{X_1} \mu_1(dx_1) \mu_2(C_{x_1}) , \quad (77)$$

où la notation C_{x_1} pour la section de C a été définie précédemment.

- Dans la suite il sera souvent plus facile de caractériser une mesure $\mu \in \mathcal{M}$ non par sa valeur sur les parties mesurables de la tribu \mathcal{F} , mais par la valeur des intégrales $\mu(f)$ sur une classe bien choisie de fonctions. Ici on a en effet pour deux mesures μ, ν l'équivalence entre $\mu = \nu$ et $\mu(f) = \nu(f)$ pour toute fonction $f \in \mathcal{C}_f$ qui dépend d'un nombre fini de variables. En effet si on revient sur la définition de la tribu produit, $\mathcal{F} = \sigma(\{C_x(\tau)\}_{x \in \mathbb{Z}^d, \tau \in \chi})$, on se rend compte que $\mathcal{F} = \sigma(\{C_\Lambda(\underline{\tau}_\Lambda)\}_{\Lambda \subset \mathbb{Z}^d, \underline{\tau}_\Lambda})$: on a seulement ajouté aux parties engendrant \mathcal{F} des parties qui appartiennent à \mathcal{F} par stabilité de la tribu sous les intersections finies. Or $\{C_\Lambda(\underline{\tau}_\Lambda)\}_{\Lambda \subset \mathbb{Z}^d, \underline{\tau}_\Lambda}$ est une famille stable par intersection finie, engendrant \mathcal{F} , une conséquence du lemme de classe monotone est que $\mu = \nu$ si et seulement si elles donnent la même mesure à tous ces cylindres de support fini. Cette condition est clairement équivalente à l'égalité $\mu(f) = \nu(f) \forall f \in \mathcal{C}_f$.

Une manière moins directe s'appuyant sur des résultats généraux d'aboutir à cette conclusion est la suivante. Sur un espace métrique deux mesures de probabilité μ, ν sont égales si et seulement si $\mu(f) = \nu(f)$ pour toutes les fonctions réelles f bornées et uniformément continues (Th 1.2 dans

Billingsley, Convergence of probability measures). Dans notre cas Σ est métrisable est compact, toutes les fonctions continues satisfont ces deux propriétés, donc $\mu = \nu$ ssi $\mu(f) = \nu(f)$ pour toutes les fonctions continues $f \in \mathcal{C}$. Comme les fonctions cylindriques à support fini, notées \mathcal{C}_f , sont denses dans \mathcal{C} on a $\mu = \nu \Leftrightarrow \mu(f) = \nu(f) \forall f \in \mathcal{C}_f$ (il suffit d'approximer $f \in \mathcal{C}$ par $g \in \mathcal{C}_f$ avec $\|f - g\|_\infty \leq \epsilon$, et d'utiliser $|\mu(f) - \mu(g)| \leq \|f - g\|_\infty$, i.e. l'application $f \rightarrow \mu(f)$ est continue de \mathcal{C} avec $\|\cdot\|_\infty$ vers \mathbb{R}).

- La réciproque de cette caractérisation d'une mesure par son action sur les fonctions continues est fournie par le théorème de représentation de Riesz : si L est une forme linéaire positive sur \mathcal{C} (la positivité signifiant $L(f) \geq 0$ quand $f \geq 0$), alors il existe une unique mesure μ telle que $\mu(f) = L(f)$ pour toutes les fonctions $f \in \mathcal{C}$. Ici l'espace étant métrique et compact il n'y a pas d'hypothèses supplémentaires au théorème.
- De $\mu \in \mathcal{M}$ on peut définir la mesure de probabilité marginale μ_Λ , où $\Lambda \subset \mathbb{Z}^d$ n'est pas forcément fini, comme la mesure sur $(\chi^\Lambda, \mathcal{F}_\Lambda)$ telle que $\mu_\Lambda(f) = \mu(f)$ pour toutes les fonctions $f \in \mathcal{C}_f \cap \mathcal{C}_\Lambda$, i.e. qui ne dépendent que d'un nombre fini de variables dans Λ : on a vu ci-dessus que c'était suffisant pour caractériser uniquement une mesure de probabilité sur $(\chi^\Lambda, \mathcal{F}_\Lambda)$. On peut aussi définir les lois conditionnelles de μ mais c'est nettement plus difficile techniquement, dans la suite on va les utiliser de manière implicite.

Topologie sur \mathcal{M}

- On munit finalement l'ensemble \mathcal{M} des mesures de probabilité sur (Σ, \mathcal{F}) d'une structure d'espace topologique. On utilise pour cela la topologie dite de la convergence faible, où une suite μ_n de mesures ($\mu_n \in \mathcal{M}$) converge vers $\mu \in \mathcal{M}$ si et seulement si $\mu_n(f) \rightarrow \mu(f)$ pour tout f dans \mathcal{C} . Par la densité de \mathcal{C}_f dans \mathcal{C} cette condition est équivalente à $\mu_n(f) \rightarrow \mu(f) \forall f \in \mathcal{C}_f$.
- C'est une topologie métrisable, avec la distance de Lévy-Prokhorov

$$d(\mu, \nu) = \inf\{\epsilon > 0 : \mu(A) \leq \nu(A^\epsilon) + \epsilon, \nu(A) \leq \mu(A^\epsilon) + \epsilon \forall A \in \mathcal{F}\}, \quad (78)$$

où A^ϵ est le voisinage à distance ϵ de A . La caractérisation par les suites est donc suffisante, car dans un espace métrique les fermés sont caractérisés de manière univoque par le fait que leurs points d'adhérence leur appartiennent (et donc les ouverts sont bien définis comme complémentaires des fermés).

- Un résultat qui sera crucialement utile dans la suite est que \mathcal{M} est compact pour cette topologie. C'est un cas particulier du théorème de Prokhorov, comme Σ est compact toute famille de mesures de probabilité est étroite (tendue, tight). Cf Billingsley, Convergence of probability measures, page 59.

3.3 Définition des mesures de Gibbs

Lemme 6. Rappelons que pour $\Phi \in \mathcal{I}_p$ et β la température inverse le Hamiltonien dans $\Lambda \subset_f \mathbb{Z}^d$ avec conditions aux bords $\underline{\tau}$ et la fonction de partition associée sont définis par

$$H_\Lambda^{\Phi, \underline{\tau}}(\underline{\sigma}) = \sum_{X, X \cap \Lambda \neq \emptyset} \phi_X((\underline{\sigma}_\Lambda, \underline{\tau}_{\Lambda^c})), \quad Z_\Lambda^{\underline{\tau}}(\beta, \Phi) = \sum_{\underline{\sigma}_\Lambda} e^{-\beta H_\Lambda^{\Phi, \underline{\tau}}(\underline{\sigma})}. \quad (79)$$

On définit maintenant, en gardant Φ et β sous-entendus pour alléger les notations,

$$\eta_\Lambda(\underline{\sigma} | \underline{\tau}) = \frac{1}{Z_\Lambda^{\underline{\tau}}(\beta, \Phi)} e^{-\beta H_\Lambda^{\Phi, \underline{\tau}}(\underline{\sigma})}, \quad (80)$$

qui dépend en fait seulement de $\underline{\sigma}_\Lambda$ et de $\underline{\tau}_{\Lambda^c}$. Alors :

- pour tout $\underline{\tau}$, $\sum_{\underline{\sigma}_\Lambda} \eta_\Lambda(\underline{\sigma} | \underline{\tau}) = 1$.

– pour tout $\underline{\sigma}, \underline{\tau} \rightarrow \eta_\Lambda(\underline{\sigma}|\underline{\tau})$ est une fonction continue, et donc mesurable.

Pour les interactions dans \mathcal{I}_f la dépendance en $\underline{\tau}$ ne porte que sur un nombre fini de variables (celles à l'extérieur de Λ et à distance de Λ inférieure à la portée de l'interaction), la continuité est donc évidente.

Démonstration. On avait vu précédemment que la condition $\Phi \in \mathcal{I}_p$ assurait la convergence de la somme dans $H_\Lambda^{\Phi, \underline{\tau}}(\underline{\sigma})$, uniformément en $\underline{\sigma}$ et $\underline{\tau}$, η_Λ est donc bien définie.

Le premier point est trivial par la définition.

Pour le deuxième, on note que $\underline{\tau} \rightarrow \phi_X((\underline{\sigma}_\Lambda, \underline{\tau}_{\Lambda^c}))$ est une fonction continue car elle ne dépend que d'un nombre fini de variables dans $X \cap \Lambda^c$, une fonction d'un nombre fini de variables est continue, et l'on utilise le fait que pour $\Phi \in \mathcal{I}_p$ la somme converge normalement en $\underline{\tau}$ pour montrer que $\underline{\tau} \rightarrow H_\Lambda^{\Phi, \underline{\tau}}(\underline{\sigma})$ est continue. Comme la somme sur $\underline{\sigma}_\Lambda$ contient un nombre fini de termes on a aussi $\underline{\tau} \rightarrow Z_\Lambda^{\underline{\tau}}(\beta, \Phi)$ continue, et donc $\underline{\tau} \rightarrow \eta_\Lambda(\underline{\sigma}|\underline{\tau})$ est continue. \square

Lemme 7. *Si f est une fonction réelle continue (resp. mesurable), $\Lambda \subset_f \mathbb{Z}^d$ et $\underline{\sigma}_\Lambda \in \chi^\Lambda$ fixée, alors $\underline{\tau} \rightarrow f((\underline{\sigma}_\Lambda, \underline{\tau}_{\Lambda^c}))$ est continue (resp. mesurable).*

Démonstration. Notons, en ayant fixé $\Lambda \subset_f \mathbb{Z}^d$ et $\underline{\sigma}_\Lambda \in \chi^\Lambda$, g la fonction $\underline{\tau} \rightarrow f((\underline{\sigma}_\Lambda, \underline{\tau}_{\Lambda^c}))$.

Pour la continuité, il suffit de remarquer que si $\underline{\tau}^{(n)} \rightarrow \underline{\tau}$, alors $(\underline{\sigma}_\Lambda, \underline{\tau}_{\Lambda^c}^{(n)}) \rightarrow (\underline{\sigma}_\Lambda, \underline{\tau}_{\Lambda^c})$ (c'est une convergence simple), comme f est continue $g(\underline{\tau}^{(n)}) = f((\underline{\sigma}_\Lambda, \underline{\tau}_{\Lambda^c}^{(n)})) \rightarrow f((\underline{\sigma}_\Lambda, \underline{\tau}_{\Lambda^c})) = g(\underline{\tau})$. Comme Σ est métrisable c'est suffisant pour prouver la continuité.

Pour la mesurabilité, comme on l'a vu plus haut la tribu \mathcal{F} peut être vue comme le produit $\mathcal{F}_\Lambda \otimes \mathcal{F}_{\Lambda^c}$, et de manière générale la restriction d'une fonction mesurable d'un produit de deux espaces mesurables est mesurable en chacun de ses arguments, l'autre étant fixé (cf. les rappels de théorie de la mesure ci-dessus). \square

Définition 5. *Une mesure de probabilité $\mu \in \mathcal{M}$ sur (Σ, \mathcal{F}) est dite à l'équilibre dans $\Lambda \subset_f \mathbb{Z}^d$ pour une interaction $\Phi \in \mathcal{I}_p$ à la température inverse β si elle vérifie la condition dite de Dobrushin-Lanford-Ruelle (DLR) :*

$$\mu(B) = \int \mu(d\underline{\tau}) \sum_{\underline{\sigma}_\Lambda} \eta_\Lambda(\underline{\sigma}|\underline{\tau}) \mathbb{I}((\underline{\sigma}_\Lambda, \underline{\tau}_{\Lambda^c}) \in B) \quad \forall B \in \mathcal{F}. \quad (81)$$

On note alors $\mu \in \mathcal{G}_\Lambda(\beta, \Phi)$.

Cette condition traduit l'idée qu'à l'intérieur de Λ le système est à l'équilibre, avec des conditions aux bords elles-mêmes tirées selon la mesure d'équilibre. C'est une façon d'écrire que les probabilités conditionnelles de μ sont données par les lois de Gibbs-Boltzmann. C'est le pendant de (64), qui était une tautologie dans le cas d'un espace de configuration fini, mais qui est non-triviale ici.

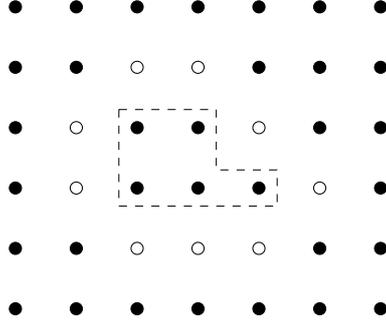
Le membre de droite est bien défini car c'est l'intégrale d'une fonction de $\underline{\tau}$ qui est positive et mesurable. C'est une somme finie, sur $\underline{\sigma}_\Lambda$, de fonctions, il suffit de vérifier que chacune de ces fonctions de $\underline{\tau}_{\Lambda^c}$ est mesurable, ce qui est assuré par les deux lemmes précédents (B étant un évènement mesurable, sa fonction indicatrice est mesurable).

On peut aussi vérifier que $\mu(\Sigma) = 1$ est bien respectée par l'équation, car alors la fonction indicatrice vaut 1, $\sum_{\underline{\sigma}_\Lambda} \eta_\Lambda(\underline{\sigma}|\underline{\tau}) = 1$, et l'intégrale de 1 vaut 1.

Définition 6. *Une mesure de probabilité $\mu \in \mathcal{M}$ sur (Σ, \mathcal{F}) est dite mesure de Gibbs pour une interaction $\Phi \in \mathcal{I}_p$ à la température inverse β si elle est à l'équilibre dans tout volume fini, on note alors $\mu \in \mathcal{G}(\beta, \Phi)$. Autrement dit $\mathcal{G}(\beta, \Phi) = \bigcap_{\Lambda \subset_f \mathbb{Z}^d} \mathcal{G}_\Lambda(\beta, \Phi)$.*

On va voir que c'est le "bon niveau de contraintes" : il existe des mesures de Gibbs, mais elles ne sont pas forcément uniques, au contraire les transitions de phase vont se traduire par l'existence de plusieurs mesures de Gibbs associées à la même interaction Φ et à la même température inverse β .

Explicitons le sens des conditions DLR, en supposant pour simplifier l'interaction à portée finie. Notons $\partial_e\Lambda$ les sites hors d'une partie finie Λ de \mathbb{Z}^d , mais à distance de Λ plus petite que la portée de l'interaction, de sorte que $\eta_\Lambda(\underline{\sigma}|\underline{\tau})$ ne dépend de $\underline{\tau}$ que par l'intermédiaire de $\underline{\tau}_{\partial_e\Lambda}$. Par exemple dans le cas de l'interaction d'Ising $\partial_e\Lambda$ est l'ensemble des sites hors de Λ à distance 1 de Λ , sur la figure :



Λ correspond aux sites entourées, les sites de $\partial_e\Lambda$ sont en blanc.

Considérons donc μ une mesure de Gibbs pour une interaction à portée finie, $\underline{x}_{\Lambda\cup\partial_e\Lambda}$ une configuration de spins sur les variables dans Λ et sa frontière extérieure $\partial_e\Lambda$, et notons $B = C_{\Lambda\cup\partial_e\Lambda}(\underline{x}_{\Lambda\cup\partial_e\Lambda})$ le cylindre élémentaire. En insérant dans l'équation (81), vérifiée puisque μ est à l'équilibre dans Λ , il vient

$$\mu(B) = \int \mu(d\underline{\tau}) \sum_{\underline{\sigma}_\Lambda} \eta_\Lambda(\underline{\sigma}|\underline{\tau}) \mathbb{I}(\underline{\sigma}_\Lambda = \underline{x}_\Lambda) \mathbb{I}(\underline{\tau}_{\partial_e\Lambda} = \underline{x}_{\partial_e\Lambda}) \quad (82)$$

$$= \eta_\Lambda(\underline{x}_\Lambda|\underline{x}_{\partial_e\Lambda}) \int \mu(d\underline{\tau}) \mathbb{I}(\underline{\tau}_{\partial_e\Lambda} = \underline{x}_{\partial_e\Lambda}) \quad (83)$$

$$= \eta_\Lambda(\underline{x}_\Lambda|\underline{x}_{\partial_e\Lambda}) \mu(C_{\partial_e\Lambda}(\underline{x}_{\partial_e\Lambda})) . \quad (84)$$

On a donc, si μ est une mesure de Gibbs pour l'interaction d'Ising, avec un léger abus de notation,

$$\mu(\underline{\sigma}_{\Lambda\cup\partial_e\Lambda}) = \mu(\underline{\sigma}_{\partial_e\Lambda}) \eta_\Lambda(\underline{\sigma}_\Lambda|\underline{\sigma}_{\partial_e\Lambda}) , \quad (85)$$

en utilisant le fait que μ est à l'équilibre dans Λ . On a noté $\mu(\underline{\sigma}_{\Lambda\cup\partial_e\Lambda})$ ce qui formellement serait $\mu(C_{\Lambda\cup\partial_e\Lambda}(\underline{\sigma}_{\Lambda\cup\partial_e\Lambda}))$ la probabilité du cylindre élémentaire correspondant, et de même pour $\mu(\underline{\sigma}_{\partial_e\Lambda})$.

Insistons sur l'interprétation des conditions DLR : on a renoncé à écrire $\mu(\underline{\sigma}) = \frac{1}{Z} e^{-\beta H(\underline{\sigma})}$ qui n'a pas de sens en volume infini, on impose la forme de Gibbs-Boltzmann seulement sur les probabilités conditionnelles : en supposant $\mu(\underline{\sigma}_{\partial_e\Lambda}) > 0$ on peut écrire (85) comme

$$\frac{\mu(\underline{\sigma}_{\Lambda\cup\partial_e\Lambda})}{\mu(\underline{\sigma}_{\partial_e\Lambda})} = \mu(\underline{\sigma}_\Lambda|\underline{\sigma}_{\partial_e\Lambda}) = \eta_\Lambda(\underline{\sigma}_\Lambda|\underline{\sigma}_{\partial_e\Lambda}) . \quad (86)$$

Les équations DLR (81) sont des équations d'auto-cohérence, dans le sens que μ est présent dans les deux membres de l'équation. On les impose dans tout Λ fini, ce qui laisse donc la possibilité que deux mesures de Gibbs "diffèrent à l'infini" (on donnera un sens précis à cette affirmation un peu plus tard.)

Explicitons pour la suite la loi conditionnelle dans le cas particulier $\Lambda = \{x\}$:

$$\eta_{\{x\}}(\sigma_x|\underline{\tau}) = \frac{e^{\beta(h+J \sum_{y \in \partial x} \tau_y) \sigma_x}}{2 \cosh(\beta(h+J \sum_{y \in \partial x} \tau_y))} = \frac{1 + \sigma_x \tanh(\beta(h+J \sum_{y \in \partial x} \tau_y))}{2} , \quad (87)$$

les voisins de x (que l'on note ici et dans la suite ∂x) créent un champ magnétique effectif sur σ_x .

3.4 Existence des mesures de Gibbs

Lemme 8. Pour $\Lambda \subset_f \mathbb{Z}^d$, f une fonction de Σ dans \mathbb{R} , on note

$$f_\Lambda(\underline{\tau}) = \sum_{\underline{\sigma}_\Lambda} \eta_\Lambda(\underline{\sigma}|\underline{\tau}) f((\underline{\sigma}_\Lambda, \underline{\tau}_{\Lambda^c})) . \quad (88)$$

Si f est continue alors f_Λ l'est aussi (et donc mesurable), et f_Λ ne dépend que de $\underline{\tau}_{\Lambda^c}$, i.e. elle est \mathcal{F}_{Λ^c} -mesurable. f_Λ est donc la moyenne conditionnée de f , en re-tirant la configuration au sein de Λ avec la loi conditionnelle étant donnée la condition aux bords $\underline{\tau}_{\Lambda^c}$.

Démonstration. La continuité est une conséquence des lemmes 6 et 7, puisqu'il n'y a qu'un nombre fini de termes dans la somme sur $\underline{\sigma}_\Lambda$, la \mathcal{F}_{Λ^c} -mesurabilité est évidente par la définition. \square

Si f est indépendante des variables dans Λ (i.e. si elle est \mathcal{F}_{Λ^c} mesurable) alors $f_\Lambda = f$.

Lemme 9. Pour Λ fini

$$\mu \in \mathcal{G}_\Lambda(\beta, \Phi) \Leftrightarrow \mu(f) = \mu(f_\Lambda) \quad \forall f \in \mathcal{C} \Leftrightarrow \mu(f) = \mu(f_\Lambda) \quad \forall f \in \mathcal{C}_f , \quad (89)$$

i.e. on peut tester la condition DLR sur toutes les fonctions continues (voire continues et dépendant d'un nombre fini de variables) plutôt que sur tous les événements mesurables.

Démonstration. On a vu dans les rappels que $\mu = \nu$ ssi $\mu(f) = \nu(f)$ pour toutes les fonctions f continues, et qu'il était même suffisant de se restreindre aux fonctions dépendant d'un nombre fini de variables. Notons $\nu(B)$ le membre de droite de (81). On a

$$\nu(B) = \int \mu(d\underline{\tau}) \sum_{\underline{\sigma}_\Lambda} \eta_\Lambda(\underline{\sigma}|\underline{\tau}) \mathbb{I}((\underline{\sigma}_\Lambda, \underline{\tau}_{\Lambda^c}) \in B) \quad (90)$$

$$= \sum_{\underline{\sigma}_\Lambda} \nu_{\underline{\sigma}_\Lambda}(B) , \quad \text{avec } \nu_{\underline{\sigma}_\Lambda}(B) = \int \mu(d\underline{\tau}) \eta_\Lambda(\underline{\sigma}|\underline{\tau}) \mathbb{I}((\underline{\sigma}_\Lambda, \underline{\tau}_{\Lambda^c}) \in B) , \quad (91)$$

puisque l'on peut intervertir intégration et somme d'un nombre fini de fonctions mesurables. Etudions un des termes de la somme, en fixant donc $\underline{\sigma}_\Lambda$. On peut écrire

$$\nu_{\underline{\sigma}_\Lambda}(B) = \int_{\chi^{\Lambda^c}} \mu_{\Lambda^c}(d\underline{\tau}_{\Lambda^c}) \eta_\Lambda(\underline{\sigma}_\Lambda|\underline{\tau}_{\Lambda^c}) \mathbb{I}(\underline{\tau}_{\Lambda^c} \in B_{\underline{\sigma}_\Lambda}) , \quad (92)$$

en utilisant la loi marginale μ_{Λ^c} de μ et la section $B_{\underline{\sigma}_\Lambda}$ de B . En se souvenant des deux constructions de mesure rappelées en (75-77) on voit que

$$\nu_{\underline{\sigma}_\Lambda} = \mathbb{I}_{\underline{\sigma}_\Lambda} \otimes (\mu_{\Lambda^c} \cdot \eta_\Lambda(\underline{\sigma}_\Lambda|\bullet)) , \quad (93)$$

le produit étant par rapport à la décomposition $\mathcal{F} = \mathcal{F}_\Lambda \otimes \mathcal{F}_{\Lambda^c}$. On a donc montré que $\nu_{\underline{\sigma}_\Lambda}$ est une mesure sur (Σ, \mathcal{F}) , et par linéarité c'est aussi le cas de ν . Cette dernière est de plus une mesure de probabilité (on avait précédemment vérifié que $\nu(\Sigma) = 1$). On conclut finalement grâce au théorème de Fubini :

$$\nu(f) = \sum_{\underline{\sigma}_\Lambda} \nu_{\underline{\sigma}_\Lambda}(f) \quad (94)$$

$$= \sum_{\underline{\sigma}_\Lambda} \int_{\chi^{\Lambda^c}} \mu_{\Lambda^c}(d\underline{\tau}_{\Lambda^c}) \eta_\Lambda(\underline{\sigma}_\Lambda|\underline{\tau}_{\Lambda^c}) f((\underline{\sigma}_\Lambda, \underline{\tau}_{\Lambda^c})) \quad (95)$$

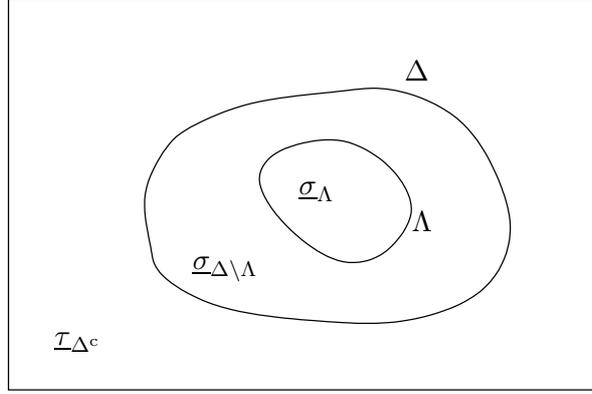
$$= \int \mu(d\underline{\tau}) \sum_{\underline{\sigma}_\Lambda} \eta_\Lambda(\underline{\sigma}|\underline{\tau}) f((\underline{\sigma}_\Lambda, \underline{\tau}_{\Lambda^c})) = \mu(f_\Lambda) . \quad (96)$$

\square

Lemme 10. Les ensembles $\mathcal{G}_\Lambda(\beta, \Phi)$ des mesures à l'équilibre dans un volume fini Λ vérifient les deux propriétés suivantes :

- Si $\Lambda \subset \Delta$ alors $\mathcal{G}_\Delta(\beta, \Phi) \subset \mathcal{G}_\Lambda(\beta, \Phi)$, i.e. si la condition DLR est vérifiée dans un volume fini elle l'est automatiquement dans tout volume à l'intérieur.
- $\mathcal{G}_\Lambda(\beta, \Phi)$ est une partie fermée (et donc compacte) de \mathcal{M} .

Démonstration. Pour le premier point, notons que pour $\Lambda \subset \Delta$, comme représenté sur ce schéma,



$$\eta_\Delta(\underline{\sigma}_\Delta | \mathcal{I}_{\Delta^c}) = \eta_\Lambda(\underline{\sigma}_\Lambda | \underline{\sigma}_{\Delta \setminus \Lambda}, \mathcal{I}_{\Delta^c}) \sum_{\underline{\sigma}'_\Lambda} \eta_\Delta(\underline{\sigma}'_\Lambda, \underline{\sigma}_{\Delta \setminus \Lambda} | \mathcal{I}_{\Delta^c}). \quad (97)$$

En effet, en gardant implicite l'interaction Φ pour alléger les notations mais en écrivant explicitement de quelles variables les fonctions dépendent, on a

$$\begin{aligned} H_\Delta^{\mathcal{I}_{\Delta^c}}(\underline{\sigma}_\Delta) &= \sum_{X, X \cap \Delta \neq \emptyset} \phi_X((\underline{\sigma}_\Delta, \mathcal{I}_{\Delta^c})) = \sum_{X, X \cap \Lambda \neq \emptyset} \phi_X((\underline{\sigma}_\Delta, \mathcal{I}_{\Delta^c})) + \sum_{\substack{X, X \cap \Delta \neq \emptyset \\ X \cap \Lambda = \emptyset}} \phi_X((\underline{\sigma}_\Delta, \mathcal{I}_{\Delta^c})) \\ &= H_\Lambda^{(\underline{\sigma}_{\Delta \setminus \Lambda}, \mathcal{I}_{\Delta^c})}(\underline{\sigma}_\Lambda) + \sum_{\substack{X, X \cap \Delta \neq \emptyset \\ X \cap \Lambda = \emptyset}} \phi_X((\underline{\sigma}_{\Delta \setminus \Lambda}, \mathcal{I}_{\Delta^c})), \end{aligned} \quad (98)$$

où par construction les termes de la dernière somme sont indépendants de $\underline{\sigma}_\Lambda$. Il vient donc

$$\eta_\Delta(\underline{\sigma}_\Delta | \mathcal{I}_{\Delta^c}) = \eta_\Lambda(\underline{\sigma}_\Lambda | \underline{\sigma}_{\Delta \setminus \Lambda}, \mathcal{I}_{\Delta^c}) g(\underline{\sigma}_{\Delta \setminus \Lambda}, \mathcal{I}_{\Delta^c}), \quad (99)$$

où g ne dépend pas de $\underline{\sigma}_\Lambda$. Il suffit de sommer sur $\underline{\sigma}_\Lambda$ dans cette dernière équation et d'utiliser la condition de normalisation de η_Λ par rapport à son premier argument pour obtenir l'expression de g , et donc (97). De cette équation on peut obtenir :

$$\forall f \in \mathcal{C}, \quad f_\Delta = (f_\Lambda)_\Delta. \quad (100)$$

En effet

$$\begin{aligned} f_\Delta(\underline{\tau}) &= \sum_{\underline{\sigma}_\Delta} f((\underline{\sigma}_\Delta, \mathcal{I}_{\Delta^c})) \eta_\Delta(\underline{\sigma} | \underline{\tau}) = \sum_{\underline{\sigma}_\Lambda, \underline{\sigma}_{\Delta \setminus \Lambda}, \underline{\sigma}'_\Lambda} \eta_\Lambda(\underline{\sigma}_\Lambda | \underline{\sigma}_{\Delta \setminus \Lambda}, \mathcal{I}_{\Delta^c}) f((\underline{\sigma}_\Lambda, \underline{\sigma}_{\Delta \setminus \Lambda}, \mathcal{I}_{\Delta^c})) \eta_\Delta(\underline{\sigma}'_\Lambda, \underline{\sigma}_{\Delta \setminus \Lambda} | \mathcal{I}_{\Delta^c}) \\ &= \sum_{\underline{\sigma}'_\Lambda, \underline{\sigma}_{\Delta \setminus \Lambda}} f_\Lambda((\underline{\sigma}_{\Delta \setminus \Lambda}, \mathcal{I}_{\Delta^c})) \eta_\Delta(\underline{\sigma}'_\Lambda, \underline{\sigma}_{\Delta \setminus \Lambda} | \mathcal{I}_{\Delta^c}) = (f_\Lambda)_\Delta(\underline{\tau}). \end{aligned} \quad (101)$$

Considérons maintenant $\mu \in \mathcal{G}_\Delta(\beta, \Phi)$ et $f \in \mathcal{C}$. Par la caractérisation (89) de \mathcal{G}_Δ appliquée à f on a $\mu(f) = \mu(f_\Delta)$, et en l'appliquant à f_Λ il vient $\mu(f_\Lambda) = \mu((f_\Lambda)_\Delta)$. En combinant ces deux égalités et l'équation (100), il vient $\mu(f) = \mu(f_\Lambda)$. Comme c'est vrai pour toute fonction f continue, on a montré $\mu \in \mathcal{G}_\Lambda(\beta, \Phi)$, et donc l'inclusion $\mathcal{G}_\Delta(\beta, \Phi) \subset \mathcal{G}_\Lambda(\beta, \Phi)$.

Pour le deuxième point, comme \mathcal{M} est métrisable, \mathcal{G}_Λ est fermée si et seulement si toute suite d'éléments de \mathcal{G}_Λ qui converge dans \mathcal{M} a pour limite un élément de \mathcal{G}_Λ . Considérons donc une suite μ_n d'éléments de \mathcal{G}_Λ , qui converge dans \mathcal{M} vers μ . Par définition de la topologie de \mathcal{M} , pour toute fonction continue f on a $\mu_n(f) \rightarrow \mu(f)$. Or pour tout n on a $\mu_n(f_\Lambda) = \mu_n(f)$ puisque $\mu_n \in \mathcal{G}_\Lambda$, et donc en passant à la limite $\mu(f_\Lambda) = \mu(f)$, ce qui démontre $\mu \in \mathcal{G}_\Lambda$.

Une partie fermée d'un compact étant compacte, comme \mathcal{M} est compact et \mathcal{G}_Λ fermée dans \mathcal{M} on a bien \mathcal{G}_Λ compacte. □

Théorème 5. $\mathcal{G}(\beta, \Phi)$ est une partie compacte non-vide de \mathcal{M} .

Démonstration. Notons μ_Λ^τ la mesure avec la condition aux bord strictement fixée à τ à l'extérieur de Λ , i.e.

$$\mu_\Lambda^\tau(B) = \sum_{\underline{\sigma}_\Lambda} \eta_\Lambda(\underline{\sigma}|\tau) \mathbb{I}((\underline{\sigma}_\Lambda, \tau_{\Lambda^c}) \in B) \quad \forall B \in \mathcal{F}. \quad (102)$$

On a $\mu_\Lambda^\tau(f) = f_\Lambda(\tau) = \mu_\Lambda^\tau(f_\Lambda)$ pour toute fonction $f \in \mathcal{C}$, et donc $\mu_\Lambda^\tau \in \mathcal{G}_\Lambda(\beta, \Phi)$ (quelque soit τ), ce qui montre $\mathcal{G}_\Lambda(\beta, \Phi) \neq \emptyset$.

Ici et dans toute la suite on notera $\Lambda_n \nearrow \mathbb{Z}^d$ si Λ_n est une suite de parties finies de \mathbb{Z}^d , croissante pour l'inclusion, qui envahit \mathbb{Z}^d (i.e. $\forall x \in \mathbb{Z}^d \exists N \forall n \geq N$ on a $x \in \Lambda_n$), par exemple $\Lambda_n = \{-n, -n+1, \dots, n\}^d$. C'est une notion strictement plus forte que la convergence à l'infini au sens de Van Hove.

Considérons donc une telle suite, et la suite de mesures associées $\mu_{\Lambda_n}^\tau$. Comme \mathcal{M} est compact il existe une sous-suite $n(p)$ telle que $\mu_{\Lambda_{n(p)}}^\tau$ converge quand $p \rightarrow \infty$, vers une limite notée μ . Montrons que $\mu \in \mathcal{G}(\beta, \Phi)$, i.e. que $\mu \in \mathcal{G}_\Lambda(\beta, \Phi)$ pour toute partie finie Λ de \mathbb{Z}^d . Comme $\Lambda_n \nearrow \mathbb{Z}^d$, pour tout Λ il existe p_0 tel que $\Lambda \subset \Lambda_{n(p)}$ pour tout $p \geq p_0$, donc $\mu_{n(p)} \in \mathcal{G}_\Lambda(\beta, \Phi)$ pour tout $p \geq p_0$. Les \mathcal{G}_Λ étant fermés d'après le lemme précédent, la limite μ appartient à \mathcal{G}_Λ . On a donc montré que $\mathcal{G}(\beta, \Phi)$ est non-vide. Comme c'est l'intersection des $\mathcal{G}_\Lambda(\beta, \Phi)$, qui sont des parties fermées de \mathcal{M} , elle est elle aussi fermée, et donc compacte puisque \mathcal{M} l'est.

On a en fait reprobé ici un lemme plus général : une suite décroissante de compacts non-vides a une intersection non-vide. □

Remarquons que dans cette preuve τ peut changer pour chaque Λ_n . Si on choisit de garder τ constante, on peut se demander à quelles conditions $\mu_{\Lambda_n}^\tau$ converge, sans devoir passer par une sous-suite. Pour le modèle d'Ising, il est naturel de considérer la configuration $\tau = \pm$ (resp. $\tau = \underline{\quad}$) avec $\tau_x = +1 \forall x$ (resp. $\tau_x = -1 \forall x$). On peut démontrer à l'aide des inégalités de corrélations FKGG (pour Fortuyn-Kastelein-Ginibre), similaires à celle de Harris discutée par Jérémie dans le cadre de la percolation, que les suites $\mu_{\Lambda_n}^+$ et $\mu_{\Lambda_n}^-$ convergent pour toute suite $\Lambda_n \nearrow \mathbb{Z}^d$ (cf. l'examen de juin 2014).

Existence des mesures de Gibbs invariantes par translation

On s'est restreint à des interactions invariantes par translation, on peut se demander ce qui devient de cette invariance pour les mesures de Gibbs. En fait toutes les mesures de Gibbs ne sont pas nécessairement invariantes, car les conditions aux bords peuvent briser cette invariance.

On a défini $T_a \underline{\sigma}$ la translation par $a \in \mathbb{Z}^d$ d'une configuration, selon $(T_a \underline{\sigma})_x = \sigma_{x-a}$. Pour un évènement mesurable $B \in \mathcal{F}$ on définit $T_a B = \{\underline{\sigma} : T_{-a} \underline{\sigma} \in B\}$, qui est lui aussi mesurable (\mathcal{F} est engendrée par une famille de parties de Σ stable sous les translations) et la translation d'une mesure $\mu \in \mathcal{M}$ selon $(T_a \mu)(B) = \mu(T_{-a} B)$. De manière équivalente on peut définir l'action des translations sur les fonctions continues, $(T_a f)(\underline{\sigma}) = f(T_{-a} \underline{\sigma})$, et sur les mesures de probabilité par $(T_a \mu)(f) = \mu(T_{-a} f)$. On notera \mathcal{M}_{inv} les mesures de probabilité invariantes par translation, i.e. $\mathcal{M}_{\text{inv}} = \{\mu \in \mathcal{M} : T_a \mu = \mu \forall a \in \mathbb{Z}^d\}$. Notons que \mathcal{M}_{inv} est une partie fermée de \mathcal{M} , car intersection de $\{\mu \in \mathcal{M} : T_a \mu = \mu\}$ sur $a \in \mathbb{Z}^d$, et ces parties sont fermées. En effet, si μ_n est une suite convergente vers μ de mesures invariantes par translation de a ,

$$\forall f \in \mathcal{C} \quad (T_a \mu)(f) = \mu(T_{-a} f) = \lim_{n \rightarrow \infty} \mu_n(T_{-a} f) = \lim_{n \rightarrow \infty} (T_a \mu_n)(f) = \lim_{n \rightarrow \infty} \mu_n(f) = \mu(f). \quad (103)$$

On définira les mesures de Gibbs invariantes par translation, notées $\mathcal{G}_{\text{inv}}(\beta, \Phi)$ comme $\mathcal{G}(\beta, \Phi) \cap \mathcal{M}_{\text{inv}}$.

Théorème 6. $\mathcal{G}_{\text{inv}}(\beta, \Phi)$ est une partie compacte non-vide de \mathcal{M} .

Démonstration. Comme \mathcal{G}_{inv} est l'intersection des fermés \mathcal{G} et \mathcal{M}_{inv} elle est aussi fermée dans \mathcal{M} et donc compacte, il reste à montrer qu'elle est non-vide.

Considérons $\nu \in \mathcal{G}(\beta, \Phi)$ une mesure de Gibbs (qui existe d'après le précédent théorème). Les translatées $T_a\nu$ sont aussi des mesures de Gibbs pour tout $a \in \mathbb{Z}^d$: soit $\Lambda \subset \mathbb{Z}^d$, alors $\nu \in \mathcal{G}_{T_{-a}\Lambda}(\beta, \Phi)$ puisque ν est dans l'intersection de tous les \mathcal{G}_Λ . Donc $T_a\nu \in \mathcal{G}_\Lambda$, quelque soit Λ , et donc $T_a\nu \in \mathcal{G}$: l'ensemble des mesures de Gibbs est stable par translation.

Considérons maintenant une suite $\Lambda_n \nearrow \mathbb{Z}^d$, et la suite de mesures

$$\mu_n = \frac{1}{|\Lambda_n|} \sum_{b \in \Lambda_n} T_b\nu . \quad (104)$$

Par l'observation précédente les μ_n sont pour tout n dans $\mathcal{G}(\beta, \Phi)$ compact, on peut supposer, quitte à prendre une sous-suite convergente, que les μ_n convergent, vers une limite que l'on notera μ . Montrons que μ est invariante par translation. Soit $a \in \mathbb{Z}^d$, montrons que $T_a\mu = \mu$. On prend $f \in \mathcal{C}$,

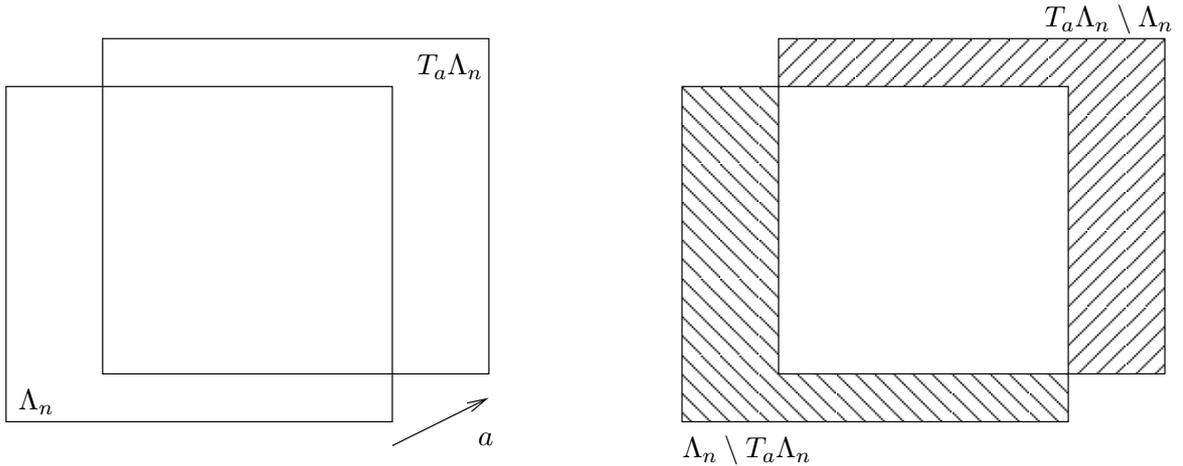
$$T_a(\mu)(f) - \mu(f) = \mu(T_{-a}f) - \mu(f) = \lim_{n \rightarrow \infty} \mu_n(T_{-a}f) - \mu_n(f) \quad (105)$$

$$= \lim_{n \rightarrow \infty} \frac{1}{|\Lambda_n|} \sum_{b \in \Lambda_n} [(T_b\nu)(T_{-a}f) - (T_b\nu)(f)] \quad (106)$$

$$= \lim_{n \rightarrow \infty} \frac{1}{|\Lambda_n|} \sum_{b \in \Lambda_n} [(T_{a+b}\nu)(f) - (T_b\nu)(f)] \quad (107)$$

$$= \lim_{n \rightarrow \infty} \frac{1}{|\Lambda_n|} \left[\sum_{b \in T_a\Lambda_n \setminus \Lambda_n} (T_b\nu)(f) - \sum_{b \in \Lambda_n \setminus T_a\Lambda_n} (T_b\nu)(f) \right] . \quad (108)$$

Or comme $\Lambda_n \nearrow \mathbb{Z}^d$, $|T_a\Lambda_n \setminus \Lambda_n|/|\Lambda_n|$ et $|\Lambda_n \setminus T_a\Lambda_n|/|\Lambda_n|$ vont à 0 quand n va à l'infini, ce dont on peut se convaincre graphiquement :



Chacun des termes de la somme étant borné, on a bien $(T_a\mu)(f) = \mu(f)$ pour toute $f \in \mathcal{C}$, donc $T_a\mu = \mu$, et comme c'est vrai pour tout a , μ est invariante par translation. \square

3.5 Transitions de phase dans l'ensemble des mesures de Gibbs

3.5.1 Unicité à haute température

Quand $\beta\Phi = 0$ (i.e. à température infinie) la seule mesure de Gibbs est la mesure produit des mesures uniformes sur χ pour chacun des spins. En effet on a alors $\eta_\Lambda(\underline{\sigma}|\underline{\tau}) = \frac{1}{|\chi|^{|\Lambda|}}$. Pour une fonction $f \in \mathcal{C}_\Lambda$, $f_\Lambda(\underline{\tau}) = \frac{1}{|\chi|^{|\Lambda|}} \sum_{\underline{\sigma}_\Lambda} f(\underline{\sigma}_\Lambda)$ est indépendante de τ , donc si μ est une mesure de Gibbs $\mu(f) = \mu(f_\Lambda) = f_\Lambda$ il vient

$$\forall \Lambda \subset_f \mathbb{Z}^d, \forall f \in \mathcal{C}_\Lambda, \quad \mu(f) = \frac{1}{|\chi|^{|\Lambda|}} \sum_{\underline{\sigma}_\Lambda} f(\underline{\sigma}_\Lambda) = \nu(f), \quad (109)$$

où ν est la mesure produit des mesures uniformes. Alors $\mu = \nu$ d'après la caractérisation donnée dans les rappels des mesures par leur comportement sur les fonctions continues dépendant d'un nombre fini de variables.

Pour β suffisamment petit, i.e. à suffisamment haute température, et pour des interactions suffisamment régulières, il n'y a aussi qu'une seule mesure de Gibbs. Une façon de prouver ce résultat se base sur un argument de Dobrushin. On va dans la suite en donner les idées principales, et le mettre en œuvre dans le cas du modèle d'Ising.

Pour une fonction $f \in \mathcal{C}$ on définit sa variation en un point x de \mathbb{Z}^d comme

$$\delta_x(f) = \sup_{\underline{\sigma}_{\mathbb{Z}^d \setminus x}, \sigma_x, \sigma'_x} |f((\underline{\sigma}_{\mathbb{Z}^d \setminus x}, \sigma_x)) - f((\underline{\sigma}_{\mathbb{Z}^d \setminus x}, \sigma'_x))|, \quad (110)$$

et sa variation totale comme

$$\Delta(f) = \sum_{x \in \mathbb{Z}^d} \delta_x(f), \quad (111)$$

qui est une semi-norme (sur la partie de \mathcal{C} où $\Delta(f) < \infty$).

Rappelons que $\mathcal{C}_f = \bigcup_{\Lambda \subset_f \mathbb{Z}^d} \mathcal{C}_\Lambda$ désigne l'ensemble des fonctions qui dépendent d'un nombre fini (arbitraire) de variables.

Lemme 11. Si $f \in \mathcal{C}_f$:

- $\Delta(f) < \infty$
- $\sup(f) - \inf(f) \leq \Delta(f)$
- $\Delta(f) = 0$ implique que f est constante

Démonstration. En effet, avec $f \in \mathcal{C}_\Lambda$,

- $\delta_x(f) = 0$ si $x \notin \Lambda$, la somme est donc seulement sur un nombre fini de termes, eux-mêmes bornés.
- Notons $\underline{\sigma}_\Lambda$ et $\underline{\tau}_\Lambda$ deux configurations qui atteignent le sup et l'inf respectivement, $\Lambda = \{x_1, \dots, x_n\}$ une énumération de Λ , et définissons la suite de configurations $\underline{\sigma}_\Lambda^{(i)}$ pour $i = 0, 1, \dots, n$ qui passent de $\underline{\sigma}_\Lambda$ à $\underline{\tau}_\Lambda$ en changeant une variable à la fois :
 $\underline{\sigma}_\Lambda^{(i)} = (\tau_{x_1}, \dots, \tau_{x_i}, \sigma_{x_{i+1}}, \dots, \sigma_{x_n})$. Alors

$$\sup(f) - \inf(f) = f(\underline{\sigma}_\Lambda^{(0)}) - f(\underline{\sigma}_\Lambda^{(n)}) = \sum_{i=0}^{n-1} f(\underline{\sigma}_\Lambda^{(i)}) - f(\underline{\sigma}_\Lambda^{(i+1)}) \leq \sum_{i=1}^n \delta_{x_i}(f) = \Delta(f), \quad (112)$$

car de $\underline{\sigma}_\Lambda^{(i)}$ à $\underline{\sigma}_\Lambda^{(i+1)}$ on ne change l'argument de f qu'en x_{i+1} (voire pas du tout si $\sigma_{x_{i+1}} = \tau_{x_{i+1}}$).

- Si $\Delta(f) = 0$ alors $\sup(f) = \inf(f)$, donc f est constante. □

Lemme 12. Supposons qu'il existe une application \mathbb{T} de \mathcal{C}_f dans \mathcal{C}_f , vérifiant :

- \mathbb{T} est une contraction pour Δ : $\forall f \in \mathcal{C}_f, \Delta(\mathbb{T}(f)) \leq \alpha \Delta(f)$ avec $\alpha < 1$
- si $\mu \in \mathcal{G}(\beta, \Phi)$, $\mu(\mathbb{T}(f)) = \mu(f) \forall f \in \mathcal{C}_f$

Alors $|\mathcal{G}(\beta, \Phi)| = 1$.

Démonstration. Notons d'abord que si μ_1 et μ_2 sont deux mesures de probabilités et $f \in \mathcal{C}_f$,

$$|\mu_1(f) - \mu_2(f)| \leq \Delta(f). \quad (113)$$

En effet comme $\mu_1(f) \geq \inf(f)$ et $\mu_2(f) \leq \sup(f)$, on a $\mu_2(f) - \mu_1(f) \leq \sup(f) - \inf(f) \leq \Delta(f)$. On obtient l'inégalité sur la valeur absolue en recommençant avec les rôles de μ_1 et μ_2 intervertis.

Considérons maintenant μ_1 et μ_2 deux mesures de Gibbs, $f \in \mathcal{C}_f$. On a

$$|\mu_1(f) - \mu_2(f)| = |\mu_1(\mathbb{T}(f)) - \mu_2(\mathbb{T}(f))| = \dots = |\mu_1(\mathbb{T}^n(f)) - \mu_2(\mathbb{T}^n(f))| \leq \Delta(\mathbb{T}^n(f)), \quad (114)$$

où dans la dernière étape on a utilisé l'inégalité (113). Il vient donc

$$|\mu_1(f) - \mu_2(f)| \leq \alpha^n \Delta(f), \quad (115)$$

et en faisant tendre n vers l'infini on a $\mu_1(f) = \mu_2(f)$. Comme c'est vrai pour tout $f \in \mathcal{C}_f$, $\mu_1 = \mu_2$, il n'existe donc qu'une mesure de Gibbs. \square

Théorème 7. *Pour le modèle d'Ising en dimension d , si $2d \tanh(\beta J) < 1$, il existe une unique mesure de Gibbs (quelque soit le champ magnétique h).*

Démonstration. On va construire une application \mathbb{T} qui vérifie les propriétés du lemme précédent.

Pour $f \in \mathcal{C}$ et x_1, \dots, x_n des points de \mathbb{Z}^d on notera

$$f_{x_1, x_2, \dots, x_n} = (\dots (f_{\{x_1\}})_{\{x_2\}} \dots)_{\{x_n\}} \quad (116)$$

la fonction obtenue par les transformations $f \rightarrow f_\Lambda$ définie en (88), avec successivement $\Lambda = \{x_1\}$, puis $\Lambda = \{x_2\}$, et ainsi de suite jusqu'à $\Lambda = \{x_n\}$. Notons que $f_{x_1, x_2, \dots, x_n} \neq f_{\{x_1, x_2, \dots, x_n\}}$, dans le deuxième cas on retire simultanément tous les n spins, dans le premier cas on le fait séquentiellement. On peut se convaincre de la différence avec $n = 2$.

\mathbb{T} est définie de la manière suivante : si $f \in \mathcal{C}_\Lambda$, on énumère les points de Λ comme x_1, \dots, x_n , et on pose $\mathbb{T}(f) = f_{x_1, x_2, \dots, x_n}$. Alors :

- $\mathbb{T}(f) \in \mathcal{C}_f$. En effet, pour le modèle d'Ising, si $f \in \mathcal{C}_\Lambda$, $f_x \in \mathcal{C}_{(\Lambda \setminus x) \cup \partial x}$, donc si on part d'une fonction qui dépend d'un nombre fini de variables et qu'on applique un nombre fini de telles opérations la dépendance reste sur un nombre fini de variables.
- si $\mu \in \mathcal{G}(\beta, \Phi)$,

$$\mu(\mathbb{T}(f)) = \mu((\dots (f_{\{x_1\}})_{\{x_2\}} \dots)_{\{x_n\}}) = \mu((\dots (f_{\{x_1\}})_{\{x_2\}} \dots)_{\{x_{n-1}\}}) = \dots = \mu(f_{\{x_1\}}) = \mu(f), \quad (117)$$

en utilisant à chaque pas $\mu \in \mathcal{G}_{\{x_i\}}(\beta, \Phi)$.

- on va montrer avec une série de lemmes dans la suite que c'est une contraction avec $\alpha = 2d \tanh(\beta J)$. On peut juste remarquer ici que si $J = 0$, $\mathbb{T}(f)$ est constante (c'est la moyenne de f avec la mesure produit biaisée par h), et alors $\Delta(\mathbb{T}(f)) = 0$, ce qui est cohérent.

En admettant le caractère contractant on peut donc utiliser le lemme précédent pour conclure sur l'unicité de la mesure d'Ising. \square

Pour comprendre d'où vient la contraction considérons le cas le plus simple d'une fonction qui ne dépend que d'un spin, $f(\underline{\sigma}) = f(\sigma_x)$. Pour des spins d'Ising une telle fonction peut toujours se décomposer comme $f(\sigma_x) = a + b\sigma_x$ avec a et b deux constantes, on supposera donc pour simplifier et sans perte de généralité que $f(\underline{\sigma}) = \sigma_x$. Alors $\mathbb{T}(f) = f_x$ est une fonction de $\underline{\sigma}_{\partial x}$, que l'on peut expliciter :

$$f_x(\underline{\sigma}_{\partial x}) = f(+1)\eta_x(+1|\underline{\sigma}_{\partial x}) + f(-1)\eta_x(-1|\underline{\sigma}_{\partial x}) \quad (118)$$

$$= \frac{1 + \tanh\left(\beta\left(h + J \sum_{y \in \partial x} \sigma_y\right)\right)}{2} - \frac{1 - \tanh\left(\beta\left(h + J \sum_{y \in \partial x} \sigma_y\right)\right)}{2} \quad (119)$$

$$= \tanh\left(\beta\left(h + J \sum_{y \in \partial x} \sigma_y\right)\right) \quad (120)$$

On peut calculer maintenant les variations $\delta_z(f_x)$: elles sont évidemment nulles si $z \notin \partial x$. Dans le cas contraire,

$$\begin{aligned} \delta_z(f_x) &= \sup_{\underline{\sigma}_{\partial x \setminus z}} \left| \tanh \left(\beta \left(h + J \sum_{y \in \partial x \setminus z} \sigma_y \right) + \beta J \right) - \tanh \left(\beta \left(h + J \sum_{y \in \partial x \setminus z} \sigma_y \right) - \beta J \right) \right| \\ &= 2 \sup_{p=-2d+1, -2d+3, \dots, 2d-1} \left| \frac{\tanh(\gamma + \beta J) - \tanh(\gamma - \beta J)}{2} \right| \quad \text{où } \gamma = \beta(h + Jp) \end{aligned} \quad (121)$$

$$\leq 2 \tanh(\beta J) \quad (122)$$

où l'on a utilisé l'identité $|\tanh(a+b) - \tanh(a-b)|/2 \leq \tanh(b)$ pour $b \geq 0$ qui se prouve en considérant la fonction $a \rightarrow \frac{\tanh(a+b) - \tanh(a-b)}{2} - \tanh(b)$, qui est nulle en 0 et dont la dérivée est négative pour $a \geq 0$ (car $\tanh'(x)$ est décroissante pour $x \geq 0$).

En sommant sur les $2d$ voisins de x il vient bien

$$\Delta(f_x) \leq 2d \tanh(\beta J) \Delta(f), \quad (123)$$

ce qui démontre la contraction dans ce cas simple (l'inégalité est valable pour une fonction $f(\underline{\sigma}) = a + b\sigma_x$ avec a et b générique). Intuitivement, on passe de f qui ne dépend que de σ_x à f_x qui dépend de plus de variables, mais qui est plus proche d'une constante (dans le sens mesuré par Δ).

On va maintenant prouver de manière générale que l'application \mathbb{T} définie ci-dessus est une contraction quand $\alpha = 2d \tanh(\beta J) < 1$.

Lemme 13. *Pour toute fonction $f \in \mathcal{C}$, pour tous points x, z de \mathbb{Z}^d , on a*

$$\delta_z(f_x) \leq \mathbb{I}(z \neq x) \delta_z(f) + \mathbb{I}(z \in \partial x) \tanh(\beta J) \delta_x(f). \quad (124)$$

Démonstration. Notons en particulier que $\delta_x(f_x) = 0$, comme il se doit, et que si f ne dépend que de σ_x on retrouve le résultat de l'observation précédente (cf. Eq.(122))

Rappelons la définition de f_x , en écrivant \mathbb{Z}^d comme l'union disjointe de $x, \partial x$ et d'un domaine noté Λ :

$$f_x(\underline{\sigma}_{\partial x}, \underline{\sigma}_{\Lambda}) = \sum_{\sigma_x} \eta_x(\sigma_x | \underline{\sigma}_{\partial x}) f(\sigma_x, \underline{\sigma}_{\partial x}, \underline{\sigma}_{\Lambda}). \quad (125)$$

Il est clair que $\delta_x(f_x) = 0$, on va donc calculer $\delta_z(f_x)$ pour $z \neq x$. La dépendance de f_x en σ_z prend des formes différentes selon que $z \in \partial x$ ou pas, on va donc distinguer ces deux cas.

– Supposons $z \in \Lambda$ (i.e. $z \neq x$ et $z \notin \partial x$). Alors notons $\Lambda' = \Lambda \setminus z$,

$$\delta_z(f_x) = \sup_{\underline{\sigma}_{\partial x}, \underline{\sigma}_{\Lambda'}} \left| \sum_{\sigma_x} \eta_x(\sigma_x | \underline{\sigma}_{\partial x}) (f(\sigma_x, \underline{\sigma}_{\partial x}, \underline{\sigma}_{\Lambda'}, +1) - f(\sigma_x, \underline{\sigma}_{\partial x}, \underline{\sigma}_{\Lambda'}, -1)) \right| \quad (126)$$

$$\leq \delta_z(f), \quad (127)$$

car $\delta_z(f)$ majore la valeur absolue de la différence des deux f , quelque soient les arguments $\sigma_x, \underline{\sigma}_{\partial x}, \underline{\sigma}_{\Lambda'}$.

– Supposons $z \in \partial x$. On notera $\partial' x = \partial x \setminus z$.

$$\delta_z(f_x) = \sup_{\underline{\sigma}_{\partial' x}, \underline{\sigma}_{\Lambda}} \left| \sum_{\sigma_x} [\eta_x(\sigma_x | \underline{\sigma}_{\partial' x}, +1) f(\sigma_x, \underline{\sigma}_{\partial' x}, +1, \underline{\sigma}_{\Lambda}) - \eta_x(\sigma_x | \underline{\sigma}_{\partial' x}, -1) f(\sigma_x, \underline{\sigma}_{\partial' x}, -1, \underline{\sigma}_{\Lambda})] \right|$$

$$= \sup_{\underline{\sigma}_{\partial' x}, \underline{\sigma}_{\Lambda}} \left| \sum_{\sigma_x} \eta_x(\sigma_x | \underline{\sigma}_{\partial' x}, +1) (f(\sigma_x, \underline{\sigma}_{\partial' x}, +1, \underline{\sigma}_{\Lambda}) - f(\sigma_x, \underline{\sigma}_{\partial' x}, -1, \underline{\sigma}_{\Lambda})) \right| \quad (128)$$

$$+ \left| \sum_{\sigma_x} (\eta_x(\sigma_x | \underline{\sigma}_{\partial' x}, +1) - \eta_x(\sigma_x | \underline{\sigma}_{\partial' x}, -1)) f(\sigma_x, \underline{\sigma}_{\partial' x}, -1, \underline{\sigma}_{\Lambda}) \right| \quad (129)$$

$$\leq \sup_{\underline{\sigma}_{\partial' x}, \underline{\sigma}_{\Lambda}} \left| \sum_{\sigma_x} \eta_x(\sigma_x | \underline{\sigma}_{\partial' x}, +1) (f(\sigma_x, \underline{\sigma}_{\partial' x}, +1, \underline{\sigma}_{\Lambda}) - f(\sigma_x, \underline{\sigma}_{\partial' x}, -1, \underline{\sigma}_{\Lambda})) \right| \quad (130)$$

$$+ \sup_{\underline{\sigma}_{\partial' x}, \underline{\sigma}_{\Lambda}} \left| \sum_{\sigma_x} (\eta_x(\sigma_x | \underline{\sigma}_{\partial' x}, +1) - \eta_x(\sigma_x | \underline{\sigma}_{\partial' x}, -1)) f(\sigma_x, \underline{\sigma}_{\partial' x}, -1, \underline{\sigma}_{\Lambda}) \right|. \quad (131)$$

On peut majorer la première ligne par $\delta_z(f)$, comme dans le cas précédent. Pour la deuxième, on remarque que l'on peut ajouter une quelconque fonction indépendante de σ_x à une somme pondérée par la différence de deux mesures de probabilité pour σ_x . On peut donc réécrire la deuxième ligne

$$\begin{aligned} & \sup_{\underline{\sigma}'_x, \underline{\sigma}_\Lambda} \left| \sum_{\sigma_x} (\eta_x(\sigma_x | \underline{\sigma}'_x, +1) - \eta_x(\sigma_x | \underline{\sigma}'_x, -1)) (f(\sigma_x, \underline{\sigma}'_x, -1, \underline{\sigma}_\Lambda) - f(-1, \underline{\sigma}'_x, -1, \underline{\sigma}_\Lambda)) \right| \\ &= \sup_{\underline{\sigma}'_x, \underline{\sigma}_\Lambda} |\eta_x(+1 | \underline{\sigma}'_x, +1) - \eta_x(+1 | \underline{\sigma}'_x, -1)| |f(+1, \underline{\sigma}'_x, -1, \underline{\sigma}_\Lambda) - f(-1, \underline{\sigma}'_x, -1, \underline{\sigma}_\Lambda)| \\ &\leq \tanh(\beta J) \delta_x(f), \end{aligned} \tag{132}$$

où le $\tanh(\beta J)$ provient de la borne sur la différence des deux η_x , comme expliqué à l'équation (122).

En regroupant les différents cas on obtient la formulation dans l'énoncé du lemme. \square

Lemme 14. *Si $\alpha = 2d \tanh(\beta J) \leq 1$, pour toute fonction $f \in \mathcal{C}$ et pour des points x_1, \dots, x_n distincts de \mathbb{Z}^d , on a*

$$\Delta(f_{x_1, \dots, x_n}) \leq \sum_{z \neq x_1, \dots, x_n} \delta_z(f) + \alpha \sum_{i=1}^n \delta_{x_i}(f). \tag{133}$$

Démonstration. En sommant sur $z \in \mathbb{Z}^d$ le résultat du lemme précédent il vient

$$\Delta(f_x) \leq \sum_{z \neq x} \delta_z(f) + \alpha \delta_x(f), \tag{134}$$

ce qui est bien ce que l'on veut prouver pour $n = 1$. Raisonnons maintenant par récurrence. En appliquant l'hypothèse de récurrence au rang $n - 1$ à la fonction f_{x_1} , il vient

$$\Delta(f_{x_1, \dots, x_n}) \leq \sum_{z \neq x_2, \dots, x_n} \delta_z(f_{x_1}) + \alpha \sum_{i=2}^n \delta_{x_i}(f_{x_1}). \tag{135}$$

On applique maintenant le lemme précédent aux différents $\delta_{\cdot}(f_{x_1})$:

$$\Delta(f_{x_1, \dots, x_n}) \leq \sum_{z \neq x_1, \dots, x_n} \delta_z(f) + \tanh(\beta J) \sum_{z \neq x_2, \dots, x_n} \mathbb{I}(z \in \partial x_1) \delta_{x_1}(f) \tag{136}$$

$$+ \alpha \sum_{i=2}^n [\delta_{x_i}(f) + \mathbb{I}(x_i \in \partial x_1) \tanh(\beta J) \delta_{x_1}(f)] \tag{137}$$

$$\leq \sum_{z \neq x_1, \dots, x_n} \delta_z(f) + \alpha \sum_{i=2}^n \delta_{x_i}(f) \tag{138}$$

$$+ \delta_{x_1}(f) \tanh(\beta J) \left[\sum_{z \neq x_2, \dots, x_n} \mathbb{I}(z \in \partial x_1) + \alpha \sum_{i=2}^n \mathbb{I}(x_i \in \partial x_1) \right]. \tag{139}$$

Comme $\alpha \leq 1$, l'expression entre crochets de la dernière ligne est plus petite que

$$\sum_{z \neq x_2, \dots, x_n} \mathbb{I}(z \in \partial x_1) + \sum_{i=2}^n \mathbb{I}(x_i \in \partial x_1) = \sum_z \mathbb{I}(z \in \partial x_1) = 2d, \tag{140}$$

ce qui conclut la preuve du lemme. \square

On peut maintenant finir la preuve du théorème : si $f \in \mathcal{C}_\Lambda$, avec $\Lambda = \{x_1, \dots, x_n\}$, en appliquant le dernier lemme et en notant que $\delta_z(f) = 0$ si $z \notin \Lambda$, il vient

$$\Delta(\mathbb{T}(f)) \leq \alpha \sum_{x \in \Lambda} \delta_x(f) = \alpha \Delta(f). \tag{141}$$

Cette preuve, due à Dobrushin, se généralise bien au-delà du modèle d'Ising. Citons sans démonstration le résultat suivant :

Théorème 8. *Si l'interaction $\Phi \in \mathcal{I}_p$ et la température inverse β vérifient*

$$\sum_{X \ni 0} (|X| - 1) \beta \|\phi_X\|_\infty < 1, \quad (142)$$

alors $|\mathcal{G}(\beta, \Phi)| = 1$, il n'existe qu'une mesure de Gibbs dans ce cas.

Appliqué au modèle d'Ising cette condition devient $2d\beta J < 1$, le résultat général est donc strictement moins fort dans ce cas que celui obtenu ci-dessus (avec la la condition $2d \tanh(\beta J) < 1$ suffisante pour établir l'unicité) dans le cas particulier du modèle d'Ising.

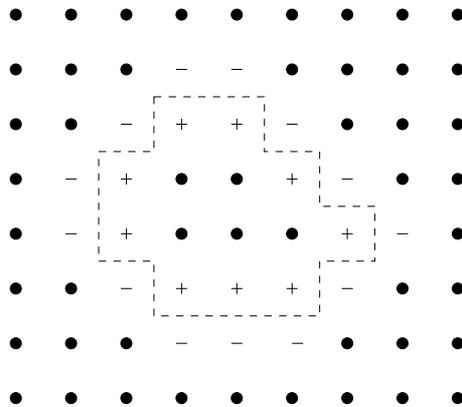
3.5.2 Argument de Peierls pour le modèle d'Ising

On va montrer que pour le modèle d'Ising, en l'absence de champ magnétique, il existe au moins deux mesures de Gibbs à basse température. Pour simplifier la discussion on se place à $d = 2$, l'argument se généralise sans trop de difficultés pour $d \geq 2$.

Dans cette section l'interaction Φ est donc définie par

$$\phi_X(\underline{\sigma}) = \begin{cases} -J\sigma_x\sigma_y & \text{si } X = \{x, y\} \text{ avec } \text{dist}(x, y) = 1 \\ 0 & \text{sinon} \end{cases} \quad (143)$$

A toute configuration $\underline{\sigma} \in \{-1, +1\}^{\mathbb{Z}^2}$ on associe l'ensemble $\Gamma(\underline{\sigma})$ des arêtes du réseau dual qui intersectent les arêtes $\langle xy \rangle$ du réseau initial dont les spins prennent des valeurs opposés, autrement dit les parois de domaine séparant les deux valeurs des spins. Parmi les quatre arêtes autour d'un sommet du réseau dual seulement 0, 2 ou 4 arêtes peuvent appartenir à Γ ((-1) à la puissance le nombre d'arêtes présentes est le produit des $\sigma_x\sigma_y$ le long de la plaquette, toujours égal à un). Un tel ensemble $\Gamma(\underline{\sigma})$ d'arêtes du réseau dual peut se décomposer comme l'union de boucles autoévitantes (qui peuvent se toucher aux angles) de longueur finie et éventuellement de chemins autoévitants qui partent à l'infini. Réciproquement pour tout Γ vérifiant cette contrainte sur le nombre d'arêtes autour de chaque sommet du réseau dual il y a deux configurations telles que $\Gamma(\underline{\sigma}) = \Gamma$, reliées par la symétrie $\underline{\sigma} \leftrightarrow -\underline{\sigma}$. Pour γ une boucle autoévitante de longueur finie (que l'on appellera aussi contour, en pointillés sur la figure) on note $|\gamma|$ son nombre d'arêtes, $\Lambda(\gamma)$ la portion du réseau initial à l'intérieur de γ , $\partial_i\Lambda(\gamma)$ la frontière intérieure de $\Lambda(\gamma)$ (les sites notés + sur la figure), et $\partial_e\Lambda(\gamma)$ sa frontière extérieure (les sites notés - sur la figure).



A basse température on s'attend à ce que les configurations typiques des mesures de Gibbs soient proches de celles qui minimisent l'Hamiltonien, i.e. $\underline{\sigma} = \underline{+}$ et $\underline{\sigma} = \underline{-}$, autrement dit qu'il y ait peu de contours dans $\Gamma(\underline{\sigma})$. Pour alléger les notations on notera parfois dans la suite + plutôt que $\underline{+}$ (resp. - plutôt que $\underline{-}$) les configurations de Σ avec tous les spins égaux à +1 (resp. -1).

Lemme 15. Si γ est un contour et $\mu \in \mathcal{G}_{\Lambda(\gamma)}(\beta, \Phi)$

$$\mu(\{\gamma \subset \Gamma(\underline{\sigma})\}) \leq 2 e^{-2\beta J|\gamma|} . \quad (144)$$

Démonstration. Notons tout d'abord que l'évènement est bien mesurable, il ne dépend que des spins dans le volume fini $\Lambda(\gamma) \cup \partial_e \Lambda(\gamma)$. Plus précisément,

$$\begin{aligned} \mu(\{\gamma \subset \Gamma(\underline{\sigma})\}) &= \mu(\{\underline{\sigma} \in \Sigma : \underline{\sigma}_{\partial_e \Lambda(\gamma)} = +\partial_e \Lambda(\gamma), \underline{\sigma}_{\partial_i \Lambda(\gamma)} = -\partial_i \Lambda(\gamma)\}) \\ &+ \mu(\{\underline{\sigma} \in \Sigma : \underline{\sigma}_{\partial_e \Lambda(\gamma)} = -\partial_e \Lambda(\gamma), \underline{\sigma}_{\partial_i \Lambda(\gamma)} = +\partial_i \Lambda(\gamma)\}) . \end{aligned} \quad (145)$$

En utilisant la condition DLR dans $\Lambda(\gamma)$, comme on l'avait fait à l'équation (85), la première ligne vaut

$$\begin{aligned} &\mu(\{\underline{\sigma} \in \Sigma : \underline{\sigma}_{\partial_e \Lambda(\gamma)} = +\partial_e \Lambda(\gamma)\}) \sum_{\underline{\sigma}_{\Lambda(\gamma) \setminus \partial_i \Lambda(\gamma)}} \eta_{\Lambda(\gamma)}((\underline{\sigma}_{\Lambda(\gamma) \setminus \partial_i \Lambda(\gamma)}, -\partial_i \Lambda(\gamma) | +) \\ &= \mu(\{\underline{\sigma} \in \Sigma : \underline{\sigma}_{\partial_e \Lambda(\gamma)} = +\partial_e \Lambda(\gamma)\}) \sum_{\underline{\sigma}_{\Lambda'}} \eta_{\Lambda}((\underline{\sigma}_{\Lambda'}, -\partial_i \Lambda | \pm) , \end{aligned} \quad (146)$$

où pour alléger les notations on a noté Λ au lieu de $\Lambda(\gamma)$ et $\Lambda' = \Lambda \setminus \partial_i \Lambda$. Explicitons $\eta_{\Lambda}(\cdot | +)$:

$$\eta_{\Lambda}((\underline{\sigma}'_{\Lambda}, \underline{\sigma}_{\partial_i \Lambda} | +) = \frac{e^{-\beta H((\underline{\sigma}'_{\Lambda}, \underline{\sigma}_{\partial_i \Lambda})) + \beta J \sum_{x \in \partial_i \Lambda} \sigma_x}}{Z} , \quad (147)$$

où H contient les interactions incluses dans Λ , et Z est défini par normalisation.

La probabilité que tous les spins de la frontière intérieure de Λ soit égaux à -1 , sachant que la condition aux bords sur la frontière extérieure est $+$, vaut donc

$$\sum_{\underline{\sigma}_{\Lambda'}} \eta_{\Lambda}((\underline{\sigma}_{\Lambda'}, -\partial_i \Lambda | +) = \frac{\sum_{\underline{\sigma}_{\Lambda'}} e^{-\beta H((\underline{\sigma}'_{\Lambda}, -\partial_i \Lambda)) - \beta J |\gamma|}}{\sum_{\underline{\sigma}_{\Lambda'}, \underline{\sigma}_{\partial_i \Lambda}} e^{-\beta H((\underline{\sigma}'_{\Lambda}, \underline{\sigma}_{\partial_i \Lambda})) + \beta J \sum_{x \in \partial_i \Lambda} \sigma_x}} \quad (148)$$

$$\leq \frac{e^{-\beta J |\gamma|} \sum_{\underline{\sigma}_{\Lambda'}} e^{-\beta H((\underline{\sigma}'_{\Lambda}, -\partial_i \Lambda))}}{e^{+\beta J |\gamma|} \sum_{\underline{\sigma}_{\Lambda'}} e^{-\beta H((\underline{\sigma}'_{\Lambda}, +\partial_i \Lambda))}} = e^{-2\beta J |\gamma|} , \quad (149)$$

en notant que $|\partial_i \Lambda(\gamma)| = |\gamma|$, puis en minorant la somme au dénominateur par sa valeur pour $\underline{\sigma}_{\partial_i \Lambda} = +\partial_i \Lambda$, et en remarquant finalement que par invariance de l'interaction sous la symétrie $\underline{\sigma} \leftrightarrow -\underline{\sigma}$ les deux fonctions de partition à l'intérieur de Λ avec $\underline{\sigma}_{\partial_i \Lambda} = +\partial_i \Lambda$ ou $-\partial_i \Lambda$ sont égales. On conclut en utilisant le fait que la probabilité dans (146) d'avoir la condition aux bords égale à $+$ est évidemment plus petite que 1, et en notant que par symétrie les deux lignes de (145) contribuent de la même façon. \square

Lemme 16. Il existe β_* tel que $\forall \beta \geq \beta_*$, $\forall \Lambda \subset_f \mathbb{Z}^d$, $\forall \mu \in \mathcal{G}_{\Lambda}(\beta, \Phi)$,

$$\mu(\{\exists \gamma : 0 \in \Lambda(\gamma), \gamma \subset \Gamma(\underline{\sigma}), \Lambda(\gamma) \subset \Lambda\}) \leq \frac{1}{4} . \quad (150)$$

Démonstration. C'est un évènement mesurable car union dénombrable d'évènements mesurables. En utilisant l'union bound (inégalité de Boole), et le fait que μ est à l'équilibre dans Λ ,

$$\mu(\{\exists \gamma : 0 \in \Lambda(\gamma), \gamma \subset \Gamma(\underline{\sigma}), \Lambda(\gamma) \subset \Lambda\}) \leq \sum_{\gamma, 0 \in \Lambda(\gamma), \Lambda(\gamma) \subset \Lambda} \mu(\{\gamma \subset \Gamma(\underline{\sigma})\}) \leq \sum_{\gamma, 0 \in \Lambda(\gamma), \Lambda(\gamma) \subset \Lambda} 2 e^{-2\beta J |\gamma|} . \quad (151)$$

Or le nombre de boucles auto-évitantes incluses dans Λ , de longueur L et entourant l'origine est certainement plus petit que $4 \times 3^{L-1} \times (2L)^2$, car il y en a moins que le nombre de marches auto-évitantes de longueur L (elle-mêmes sont moins nombreuses que $4 \times 3^{L-1}$), avec un point de départ

se situant nécessairement dans un carré de côté $2L$ centré sur l'origine. En oubliant la contrainte $\Lambda(\gamma) \subset \Lambda$ on ne fait qu'augmenter le majorant, on peut donc écrire

$$\mu(\{\exists \gamma : 0 \in \Lambda(\gamma), \gamma \subset \Gamma(\underline{\sigma}), \Lambda(\gamma) \subset \Lambda\}) \leq \text{cste} \sum_{L=4}^{\infty} L^2 e^{L(\ln 3 - 2\beta J)} \quad (152)$$

En prenant β suffisamment grand on peut rendre la somme arbitrairement petite : $L^2 = e^{2 \ln L} \leq e^{2L}$, la somme est donc majorée par une constante fois

$$\sum_{L=4}^{\infty} e^{aL} = \frac{e^{4a}}{1 - e^a}, \quad (153)$$

qui tend vers 0 quand $a \rightarrow -\infty$, i.e. quand $\beta \rightarrow +\infty$. Le lemme reste vrai quand $1/4$ est remplacée par n'importe quelle constante positive. Notons aussi que le choix de β_* se fait indépendamment de Λ . \square

Théorème 9. $\forall \beta \geq \beta_*$ il n'y a pas unicité des mesures de Gibbs, i.e. $|\mathcal{G}(\beta, \Phi)| > 1$.

Démonstration. Supposons donc $\beta \geq \beta_*$, et considérons une suite $\Lambda_n \nearrow \mathbb{Z}^2$ telle que pour tout n $0 \in \Lambda_n$ et telle que $\mu_{\Lambda_n}^+$ converge (par compacité et en extrayant une sous-suite on peut toujours trouver une telle suite, en fait avec les inégalités FKG on peut montrer que $\mu_{\Lambda_n}^+$ converge quelque soit la suite $\Lambda_n \nearrow \mathbb{Z}^2$, cf l'examen de juin 2014). On notera μ^+ la limite (faible) des $\mu_{\Lambda_n}^+$, qui appartient à $\mathcal{G}(\beta, \Phi)$ comme on l'a vu précédemment.

$\mu_{\Lambda_n}^+$ est à l'équilibre dans Λ_n , d'après le lemme précédent on a donc

$$\mu_{\Lambda_n}^+(\{\exists \gamma : 0 \in \Lambda(\gamma), \gamma \subset \Gamma(\underline{\sigma}), \Lambda(\gamma) \subset \Lambda_n\}) \leq \frac{1}{4} \quad (154)$$

On a par ailleurs

$$\mu_{\Lambda_n}^+(\{\underline{\sigma} \in \Sigma : \sigma_0 = -1\}) \leq \mu_{\Lambda_n}^+(\{\exists \gamma : 0 \in \Lambda(\gamma), \gamma \subset \Gamma(\underline{\sigma}), \Lambda(\gamma) \subset \Lambda_n\}) : \quad (155)$$

sous cette mesure la condition aux bords est fixée à $+1$, un spin égal à -1 est nécessairement entouré par un contour. Ce n'est qu'une inégalité car si par exemple il y a exactement deux contours concentriques autour de l'origine alors $\sigma_0 = +1$. On en déduit

$$\mu_{\Lambda_n}^+(\sigma_0) = \mu_{\Lambda_n}^+(\{\sigma_0 = +1\}) - \mu_{\Lambda_n}^+(\{\sigma_0 = -1\}) = 1 - 2\mu_{\Lambda_n}^+(\{\sigma_0 = -1\}) \geq \frac{1}{2} \quad (156)$$

Comme $\underline{\sigma} \rightarrow \sigma_0$ est continue, par définition de la convergence faible,

$$\mu^+(\sigma_0) = \lim_{n \rightarrow \infty} \mu_{\Lambda_n}^+(\sigma_0) \geq \frac{1}{2}. \quad (157)$$

De nouveau on aurait pu obtenir une borne inférieure sur l'aimantation arbitrairement proche de 1 en changeant la définition de β_* .

Pour conclure sur la coexistence de plus d'une mesure de Gibbs, il suffit de noter que si l'on construit de même μ^- comme une limite de $\mu_{\Lambda_n}^-$, avec aussi $\mu^- \in \mathcal{G}(\beta, \Phi)$, on obtiendrait de la même façon $\mu^-(\sigma_0) \leq -\frac{1}{2}$. Donc nécessairement $\mu^+ \neq \mu^-$ ($\underline{\sigma} \rightarrow \sigma_0$ est continue donc les aimantations moyennes à l'origine coïncideraient si $\mu^+ = \mu^-$). \square

Cette méthode des contours pour la preuve de la non-unicité des mesures de Gibbs à basse température se généralise pour des modèles différents de celui d'Ising. La généralisation porte le nom de théorie de Pirogov-Sinai.

3.6 Extrémalité et trivialité à l'infini

Rappelons qu'une partie X d'un espace vectoriel est dite convexe si pour toute paire d'éléments $x_1, x_2 \in X$ et tout $\alpha \in [0, 1]$ la combinaison linéaire $\alpha x_1 + (1 - \alpha)x_2$ appartient à X .

Lemme 17. $\mathcal{G}(\beta, \Phi)$ est un ensemble convexe.

Démonstration. C'est trivial, en effet si μ_1 et μ_2 sont deux mesures de probabilité $\alpha\mu_1 + (1-\alpha)\mu_2$ en est une autre, et si les conditions DLR sont vérifiées par μ_1 et μ_2 elles le sont aussi par $\alpha\mu_1 + (1-\alpha)\mu_2$. \square

Les éléments extrémaux d'un ensemble convexe X , notés $\text{ext}(X)$, sont par définition ceux qui n'admettent pas de décomposition non-triviale de la forme $\alpha x_1 + (1 - \alpha)x_2$ (avec $x_1 \neq x_2$ et $\alpha \in]0, 1[$). Par exemple un segment admet deux éléments extrémaux, ses extrémités. On va voir que les éléments extrémaux de $\mathcal{G}(\beta, \Phi)$ jouent un rôle très important dans la théorie des mesures de Gibbs.

Avant cela introduisons un autre concept a priori très différent. La tribu asymptotique (à l'infini) est définie comme $\mathcal{F}_\infty = \bigcap_{\Lambda \subset_f \mathbb{Z}^d} \mathcal{F}_{\Lambda^c}$, c'est une sous-tribu de \mathcal{F} en tant qu'intersection de sous-tribus. Les fonctions mesurables par rapport à \mathcal{F}_∞ sont celles qui ne sont pas modifiées si on change la valeur des spins sur un nombre fini de sites. Considérons par exemple dans le cas du modèle d'Ising la fonction :

$$m(\underline{\sigma}) = \limsup_{n \rightarrow \infty} \frac{1}{|\Lambda_n|} \sum_{x \in \Lambda_n} \sigma_x, \quad \text{où } \Lambda_n \nearrow \mathbb{Z}^d, \quad (158)$$

qui correspond à la densité d'aimantation. Comme pour tout Λ fini on peut écrire

$$m(\underline{\sigma}) = \limsup_{n \rightarrow \infty} \frac{1}{|\Lambda_n|} \sum_{x \in \Lambda_n \setminus \Lambda} \sigma_x, \quad \text{puisque } \frac{|\Lambda|}{|\Lambda_n|} \rightarrow 0, \quad (159)$$

m est bien \mathcal{F}_∞ -mesurable. Plus généralement les fonctions mesurables par rapport à \mathcal{F}_∞ correspondent aux observables macroscopiques d'un système physique, qui ne doivent pas changer si on modifie un nombre fini (arbitrairement grand) de variables d'un système infini.

Définition 7. On dit qu'une mesure $\mu \in \mathcal{M}$ est triviale (ou grossière) à l'infini ssi pour tout événement $A \in \mathcal{F}_\infty$ on a $\mu(A) = 0$ ou $\mu(A) = 1$.

La loi du 0-1 de Kolmogorov énonce que les mesures produits, i.e. pour des variables σ_x indépendantes d'un site à l'autre, sont triviales à l'infini. Mais en général ces mesures produits ne sont pas des mesures de Gibbs (sauf si $\beta\Phi = 0$), on va donc dans la suite chercher à caractériser les mesures de Gibbs triviales à l'infini.

Lemme 18. Si $\mu \in \mathcal{G}(\beta, \Phi)$, $B \in \mathcal{F}_\infty$ avec $\mu(B) \neq 0$, alors la mesure de probabilité $\mu(\cdot|B)$ définie par

$$\mu(A|B) = \frac{\mu(A \cap B)}{\mu(B)} \quad \forall A \in \mathcal{F} \quad (160)$$

est elle aussi une mesure de Gibbs, i.e. $\mu(\cdot|B) \in \mathcal{G}(\beta, \Phi)$.

Démonstration. Intuitivement, B est indépendant de ce qui se passe dans tout volume fini, alors que les conditions DLR expriment la notion d'équilibre dans les volumes finis.

Plus formellement, on peut d'abord voir facilement que $\mu(\cdot|B)$ est bien une mesure de probabilité. Soit $\Lambda \subset_f \mathbb{Z}^d$, montrons que $\mu(\cdot|B) \in \mathcal{G}_\Lambda(\beta, \Phi)$. Comme $B \in \mathcal{F}_\infty \subset \mathcal{F}_{\Lambda^c}$, la fonction caractéristique de B est indépendante de $\underline{\sigma}_\Lambda$, d'où $\forall A \in \mathcal{F}$, en utilisant la condition DLR sur μ ,

$$\mu(A|B) = \frac{\mu(A \cap B)}{\mu(B)} = \frac{1}{\mu(B)} \int \mu(d\underline{\tau}) \sum_{\underline{\sigma}_\Lambda} \eta_\Lambda(\underline{\sigma}|\underline{\tau}) \mathbb{I}((\underline{\sigma}_\Lambda, \underline{\tau}_{\Lambda^c}) \in A) \mathbb{I}((\underline{\sigma}_\Lambda, \underline{\tau}_{\Lambda^c}) \in B) \quad (161)$$

$$= \int \frac{\mu(d\underline{\tau})}{\mu(B)} \mathbb{I}((\cdot, \underline{\tau}_{\Lambda^c}) \in B) \sum_{\underline{\sigma}_\Lambda} \eta_\Lambda(\underline{\sigma}|\underline{\tau}) \mathbb{I}((\underline{\sigma}_\Lambda, \underline{\tau}_{\Lambda^c}) \in A) \quad (162)$$

$$= \int \mu(d\underline{\tau}|B) \sum_{\underline{\sigma}_\Lambda} \eta_\Lambda(\underline{\sigma}|\underline{\tau}) \mathbb{I}((\underline{\sigma}_\Lambda, \underline{\tau}_{\Lambda^c}) \in A), \quad (163)$$

ce qui est bien la condition DLR pour $\mu(\cdot|B)$ dans Λ . Comme c'est vrai pour toutes les parties finies Λ , on a bien $\mu(\cdot|B) \in \mathcal{G}(\beta, \Phi)$. \square

Lemme 19. *Deux mesures de Gibbs sont égales si et seulement si elles coïncident sur \mathcal{F}_∞ .*

Démonstration. L'implication directe est triviale. Supposons que μ_1, μ_2 sont deux éléments de $\mathcal{G}(\beta, \Phi)$ avec $\mu_1(A) = \mu_2(A)$ pour tout $A \in \mathcal{F}_\infty$, montrons que $\mu_1 = \mu_2$, en montrant que pour tout $f \in \mathcal{C}$ on a $\mu_1(f) = \mu_2(f)$. Comme μ_1 et μ_2 sont deux mesures de Gibbs, $\mu_1(f_\Lambda) = \mu_1(f)$ et $\mu_2(f_\Lambda) = \mu_2(f)$ pour tout Λ fini. Considérons maintenant une suite $\Lambda_n \nearrow \mathbb{Z}^d$. On a évidemment $\mu_1(f) = \mu_1(f_{\Lambda_n})$ pour tout n , et donc $\mu_1(f) = \lim_{n \rightarrow \infty} \mu_1(f_{\Lambda_n})$. Il suffit donc pour conclure de montrer que

$$\lim_{n \rightarrow \infty} \mu_1(f_{\Lambda_n}) = \lim_{n \rightarrow \infty} \mu_2(f_{\Lambda_n}). \quad (164)$$

C'est naturel intuitivement, car f_{Λ_n} est $\mathcal{F}_{\Lambda_n^c}$ -mesurable, comme $\Lambda_n \nearrow \mathbb{Z}^d$ à la limite les f_{Λ_n} "devraient devenir" \mathcal{F}_∞ -mesurables, tribu sur laquelle par hypothèse μ_1 et μ_2 coïncide.

La justification rigoureuse de l'égalité (164) repose sur un théorème de convergence pour les martingales rétrogrades. Rappelons donc brièvement quelques notions générales à ce sujet (en gardant sous silence quelques hypothèses techniques qui sont vérifiées dans le cas qui nous intéresse) :

- Si g est une variable aléatoire (i.e. une fonction mesurable réelle) sur un espace probabiliste (X, \mathcal{F}, μ) et \mathcal{F}' une sous-tribu de \mathcal{F} , on définit la moyenne conditionnelle $\mu(g|\mathcal{F}')$ comme une variable aléatoire \mathcal{F}' -mesurable telle que $\mu(h\mu(g|\mathcal{F}')) = \mu(hg)$ pour toute variable aléatoire h qui est \mathcal{F}' -mesurable. Cette variable aléatoire est unique à un ensemble de probabilité nulle près.
- Une filtration rétrograde de l'espace mesurable (X, \mathcal{F}) est une suite décroissante de sous-tribus de \mathcal{F} , i.e. $\mathcal{F}_1 \supset \mathcal{F}_2 \supset \dots$.
- Une martingale rétrograde (par rapport à une filtration donnée) est une suite de variables aléatoires g_1, g_2, \dots , adaptée à la filtration (i.e. g_n est \mathcal{F}_n -mesurable) et telle que si $n \geq m$ alors $g_n = \mu(g_m|\mathcal{F}_n)$.
- On a le résultat de convergence suivant : une martingale rétrograde converge presque sûrement et dans L^1 vers une variable aléatoire g_∞ qui est \mathcal{F}_∞ -mesurable, où $\mathcal{F}_\infty = \bigcap_n \mathcal{F}_n$, et $g_n = \mu(g_\infty|\mathcal{F}_n)$.

On peut appliquer ces résultats à la preuve du lemme. En effet si $\Lambda_n \nearrow \mathbb{Z}^d$ on a $\mathcal{F}_{\Lambda_1^c} \supset \mathcal{F}_{\Lambda_2^c} \supset \dots$ qui forme une filtration rétrograde, et comme Λ_n envahit \mathbb{Z}^d l'intersection des $\mathcal{F}_{\Lambda_n^c}$ est bien la tribu asymptotique définie précédemment. De plus si on note $g_n = f_{\Lambda_n}$ alors la suite des g_n est une martingale rétrograde pour toute mesure de Gibbs :

- f_{Λ_n} est indépendant des variables dans Λ_n , i.e. $\mathcal{F}_{\Lambda_n^c}$ -mesurable, donc g_n est adaptée.
- Notons que si μ est une mesure de Gibbs, $f \in \mathcal{C}$ une fonction continue arbitraire et $\Lambda \subset_f \mathbb{Z}^d$ alors $\mu(f|\mathcal{F}_{\Lambda^c}) = f_\Lambda$, en utilisant la condition DLR dans Λ . Pour montrer le caractère de martingale rétrograde de la suite g_n considérons donc $n \geq m$, et montrons que $f_{\Lambda_n} = \mu(f_{\Lambda_m}|\mathcal{F}_{\Lambda_n^c}) = (f_{\Lambda_m})_{\Lambda_n}$. Or $\Lambda_m \subset \Lambda_n$, l'égalité a donc été prouvée à l'équation (100).

Comme f_{Λ_n} est une martingale rétrograde pour μ_1 et μ_2 , le résultat rappelé ci-dessus assure que cette suite converge dans L^1 (et presque sûrement) vers f_∞ , une fonction \mathcal{F}_∞ -mesurable. Comme μ_1 et μ_2 sont égales sur \mathcal{F}_∞ elles donnent la même moyenne aux fonctions \mathcal{F}_∞ -mesurables, et donc :

$$\lim_{n \rightarrow \infty} \mu_1(f_{\Lambda_n}) = \mu_1(f_\infty) = \mu_2(f_\infty) = \lim_{n \rightarrow \infty} \mu_2(f_{\Lambda_n}). \quad (165)$$

\square

Ces deux lemmes permettent d'obtenir finalement :

Théorème 10. (a) *Une mesure de Gibbs $\mu \in \mathcal{G}(\beta, \Phi)$ est extrémale (dans $\mathcal{G}(\beta, \Phi)$) si et seulement si elle est triviale à l'infini.*

(b) *Deux mesures de Gibbs μ_1, μ_2 extrémales distinctes sont mutuellement singulières (i.e. $\exists A \in \mathcal{F}$ avec $\mu_1(A) = 1, \mu_2(A) = 0$).*

Démonstration. (a) Supposons que μ soit une mesure de Gibbs non triviale à l'infini, il existe donc $B \in \mathcal{F}_\infty$ avec $\mu(B) = \alpha \in]0, 1[$. On peut décomposer tout évènement $A \in \mathcal{F}$ comme $(A \cap B) \cup (A \cap B^c)$, et donc $\mu = \alpha\mu(\cdot|B) + (1 - \alpha)\mu(\cdot|B^c)$. Comme \mathcal{F}_∞ est une tribu elle est stable par passage au complémentaire et donc $B^c \in \mathcal{F}_\infty$. Les deux lemmes précédents montrent respectivement que :

- $\mu(\cdot|B)$ et $\mu(\cdot|B^c)$ sont deux mesures de Gibbs.
- comme $\mu(B|B) = 1 \neq \mu(B|B^c) = 0$ elles ne coïncident pas sur \mathcal{F}_∞ et sont donc distinctes.

Comme $\alpha \notin \{0, 1\}$ la décomposition de μ est non-triviale, donc μ n'est pas extrémale.

Supposons maintenant que μ n'est pas extrémale, i.e. $\mu = \alpha\mu_1 + (1 - \alpha)\mu_2$ avec $\alpha \in]0, 1[$, μ_1, μ_2 deux éléments distincts de $\mathcal{G}(\beta, \Phi)$. Comme elles ne peuvent coïncider sur \mathcal{F}_∞ d'après le lemme (19), il existe $B \in \mathcal{F}_\infty$ avec $\mu_1(B) \neq \mu_2(B)$, et alors $\mu(B)$ ne peut être dans $\{0, 1\}$.

(b) μ_1 et μ_2 étant extrémales elles sont triviales à l'infini, i.e. $\forall A \in \mathcal{F}_\infty$ on a $\mu_1(A) \in \{0, 1\}$ et $\mu_2(A) \in \{0, 1\}$. Mais comme elles sont distinctes elles ne peuvent coïncider sur tous les A de \mathcal{F}_∞ , il existe donc $A \in \mathcal{F}_\infty$ avec $\mu_1(A) = 1, \mu_2(A) = 0$. □

Ce théorème relie deux notions a priori très différentes : l'extrémalité de μ concerne la "position" de μ au sein de la structure d'espace convexe de \mathcal{G} , alors que la trivialité à l'infini dit quelque chose sur μ elle-même (plus précisément sur la structure de ses corrélations spatiales, comme on va le voir).

On qualifie d'"états purs" les éléments extrémaux de $\mathcal{G}(\beta, \Phi)$: comme elles sont triviales à l'infini les observables macroscopiques (dans le sens défini ci-dessus) sont presque sûrement constantes. Or si l'on prend un échantillon macroscopique d'un matériau dans un état bien défini, on utilise une description probabiliste au niveau microscopique, mais toutes les grandeurs mesurables macroscopiquement sont déterministes. En cas de coexistence de phase, on peut notamment penser au cas de l'eau à 0 degré, les états purs correspondent aux états solide (glace) et liquide de l'eau, pour lesquels la densité volumique prend deux valeurs déterministes distinctes. La densité volumique est alors un "paramètre d'ordre", comme l'aimantation pour le modèle d'Ising.

Trivialité à l'infini et corrélations spatiales

Revenons maintenant sur le lien entre trivialité à l'infini et corrélations spatiales d'une mesure. L'exemple le plus simple de mesure triviale à l'infini est une mesure produit (cf loi du 0/1 de Kolmogorov), pour laquelle il n'y a en effet aucune corrélation. En fait les mesures triviales à l'infini sont celles pour lesquelles les corrélations décroissent à grande distance, dans le sens suivant :

Lemme 20. $\mu \in \mathcal{M}$ est triviale à l'infini si et seulement si

$$\forall A \in \mathcal{F}, \quad \lim_{\Lambda \nearrow \mathbb{Z}^d} \sup_{B \in \mathcal{F}_{\Lambda^c}} |\mu(A \cap B) - \mu(A)\mu(B)| = 0 \quad (166)$$

cf Georgii proposition 7.9, Simon théorème III.1.6.

On ne prouvera pas ce lemme, on se contentera d'en donner une conséquence et l'idée de la preuve de ce cas particulier :

Lemme 21. Si μ est une mesure sur $\{-1, 1\}^{\mathbb{Z}^d}$ triviale à l'infini, $x \in \mathbb{Z}^d$, alors

$$\mu(\sigma_x \sigma_y) - \mu(\sigma_x)\mu(\sigma_y) \xrightarrow{y \rightarrow \infty} 0, \quad (167)$$

i.e. les spins décorrelent à grande distance.

Démonstration. Notons $f(\underline{\sigma}) = \sigma_x - \mu(\sigma_x)$, de telle manière que $\mu(f) = 0$. $\mu(f|\mathcal{F}_\infty)$ est presque sûrement constante puisque μ est triviale à l'infini, donc presque sûrement nulle. Pour tout $\epsilon > 0$, en utilisant le théorème de convergence pour les martingales rétrogrades, on peut trouver $\Lambda \subset_f \mathbb{Z}^d$ telle

que la norme L^1 de $\mu(f|\mathcal{F}_{\Lambda^c})$ soit $\leq \epsilon$. Alors il suffit de prendre y suffisamment éloigné de x pour que l'on ait $y \notin \Lambda$, de telle sorte que la fonction notée $g(\underline{x}) = \sigma_y$ soit \mathcal{F}_{Λ^c} mesurable. Alors

$$|\mu(fg)| = |\mu(\mu(fg|\mathcal{F}_{\Lambda^c}))| = |\mu(g\mu(f|\mathcal{F}_{\Lambda^c}))| \leq \epsilon, \quad (168)$$

puisque $|g| \leq 1$ et d'après la borne sur la norme de $\mu(f|\mathcal{F}_{\Lambda^c})$. Par construction $\mu(fg)$ est égale à la fonction de corrélation dans le membre de gauche de (167). \square

En général deux variables aléatoires f et g indépendantes sont décorréelées ($\mu(fg) = \mu(f)\mu(g)$), mais la réciproque est la plupart du temps fautive. Toutefois pour des variables d'Ising ces deux notions sont équivalentes, comme le montre le petit calcul suivant où l'on note \mathbb{P} la loi de probabilité pour deux spins d'Ising σ_1 et σ_2 , et $\langle \bullet \rangle$ les moyennes correspondantes :

$$\mathbb{P}(\sigma_1 = s_1) = \frac{1 + s_1 \langle \sigma_1 \rangle}{2} \quad (169)$$

$$\mathbb{P}((\sigma_1, \sigma_2) = (s_1, s_2)) = \frac{1 + s_1 \langle \sigma_1 \rangle + s_2 \langle \sigma_2 \rangle + s_1 s_2 \langle \sigma_1 \sigma_2 \rangle}{4} \quad (170)$$

$$\mathbb{P}((\sigma_1, \sigma_2) = (s_1, s_2)) - \mathbb{P}(\sigma_1 = s_1)\mathbb{P}(\sigma_2 = s_2) = \frac{1}{4}s_1 s_2 (\langle \sigma_1 \sigma_2 \rangle - \langle \sigma_1 \rangle \langle \sigma_2 \rangle) \quad (171)$$

Pour des interactions à courte portée on s'attend à ce que la fonction de corrélation au sein d'une mesure triviale à l'infini s'annule à grande distance de manière exponentielle, avec une longueur de corrélation finie (mais ce n'est prouvé que dans le régime d'unicité, à haute température). Il est alors assez naturel de s'attendre à ce que la longueur de corrélation diverge au point critique : on doit passer d'une situation où la mesure de Gibbs est unique à une situation où il y a au moins deux mesures extrémales distinctes. Si la longueur de corrélation était toujours finie cette bifurcation serait difficile à interpréter.

Insistons sur le fait que la décorrélation à l'infini ne survit pas à une combinaison linéaire (qui en effet n'est plus extrémale et donc plus triviale à l'infini). Prenons par exemple le modèle d'Ising à basse température et les deux mesures extrémales μ^+ et μ^- , avec l'aimantation spontanée $m_{\text{sp}} = \mu^+(\sigma_x) = -\mu^-(\sigma_x) > 0$. Ces mesures étant extrémales on a $\mu^+(\sigma_x \sigma_y) \rightarrow \mu^+(\sigma_x)\mu^+(\sigma_y) = m_{\text{sp}}^2$ et $\mu^-(\sigma_x \sigma_y) \rightarrow \mu^-(\sigma_x)\mu^-(\sigma_y) = m_{\text{sp}}^2$ à grande distance. Si on prend maintenant une combinaison linéaire $\mu = \alpha\mu^+ + (1 - \alpha)\mu^-$, l'aimantation moyenne et la fonction de corrélation à deux points pour μ sont :

$$\begin{aligned} \mu(\sigma_x) = \mu(\sigma_y) &= \alpha m_{\text{sp}} - (1 - \alpha)m_{\text{sp}} = (2\alpha - 1)m_{\text{sp}}, & \mu(\sigma_x)\mu(\sigma_y) &= m_{\text{sp}}^2(2\alpha - 1)^2, \\ \mu(\sigma_x \sigma_y) - \mu(\sigma_x)\mu(\sigma_y) &\rightarrow \alpha m_{\text{sp}}^2 + (1 - \alpha)m_{\text{sp}}^2 - m_{\text{sp}}^2(2\alpha - 1)^2 = 4\alpha(1 - \alpha)m_{\text{sp}}^2. \end{aligned} \quad (172)$$

Cette dernière quantité ne s'annule donc que pour $\alpha = 0$ ou 1 .

Extrémalité pour les mesures invariantes par translation

$\mathcal{G}_{\text{inv}}(\beta, \Phi)$ est lui aussi un espace convexe, on peut donc se demander comment sont caractérisés les éléments extrémaux de \mathcal{G}_{inv} . La notion de trivialité à l'infini doit être remplacé par la notion moins forte d'ergodicité.

Définition 8. *Une mesure invariante par translation $\mu \in \mathcal{M}_{\text{inv}}$ est dite ergodique si $\mu(A) \in \{0, 1\}$ pour tous les événements A invariants par translation.*

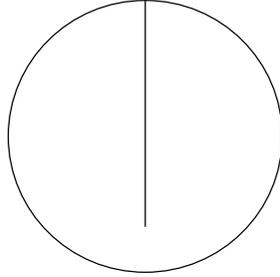
On peut montrer que si $\mu \in \mathcal{M}_{\text{inv}}$ est triviale à l'infini alors elle est ergodique (c'est ce que Jérémie a prouvé dans le cas d'une mesure produit, durant le cours sur la percolation).

On a le théorème suivant de caractérisation de l'extrémalité au sein de $\mathcal{G}_{\text{inv}}(\beta, \Phi)$:

Théorème 11. *Une mesure de Gibbs $\mu \in \mathcal{G}_{\text{inv}}(\beta, \Phi)$ est extrémale (dans $\mathcal{G}_{\text{inv}}(\beta, \Phi)$) si et seulement si elle est ergodique.*

Si μ est extrémale dans \mathcal{G} et invariante par translation alors elle est extrémale dans \mathcal{G}_{inv} : on ne peut pas l'écrire comme $\alpha\mu_1 + (1 - \alpha)\mu_2$ avec $\mu_{1,2} \in \mathcal{G}$, alors a fortiori on ne peut trouver une telle décomposition avec $\mu_{1,2} \in \mathcal{G}_{\text{inv}}$. Mais la réciproque est fautive : une mesure extrémale dans \mathcal{G}_{inv} peut admettre une décomposition non-triviale dans \mathcal{G} . Donc $\text{ext}(\mathcal{G}) \cap \mathcal{M}_{\text{inv}} \subset \text{ext}(\mathcal{G}_{\text{inv}})$ mais $\text{ext}(\mathcal{G}) \cap \mathcal{M}_{\text{inv}} \neq \text{ext}(\mathcal{G}_{\text{inv}})$ car l'inclusion réciproque est fautive.

Très schématiquement, sur l'exemple de la figure ci-dessous le disque représente un espace convexe (penser à \mathcal{G}) qui contient le segment, lui aussi convexe (penser à \mathcal{G}_{inv}). Les éléments extrêmes du segment sont ses deux extrémités, ceux du disque sa frontière. On voit que sur les deux éléments extrêmes par rapport au segment un seul des deux est extrême par rapport au disque.



On peut donner un exemple de la différence entre $\text{ext}(\mathcal{G}_{\text{inv}})$ et $\text{ext}(\mathcal{G}) \cap \mathcal{M}_{\text{inv}}$ avec un modèle antiferromagnétique (Ising avec $J < 0, h = 0$) à basse température, on construit μ_1 et μ_2 en prenant la limite thermodynamique des mesures avec des conditions aux bords fixées aux 2 configuration alternées, i.e. $\sigma_x = (-1)^{x_1 + \dots + x_d}$ et $\sigma_x = -(-1)^{x_1 + \dots + x_d}$. Alors μ_1 et μ_2 sont extrêmes dans \mathcal{G} , non-invariantes par translation. Considérons $\mu = \frac{1}{2}\mu_1 + \frac{1}{2}\mu_2$. Alors $\mu \in \mathcal{G}_{\text{inv}}$ et μ est extrême dans \mathcal{G}_{inv} , alors qu'elle ne l'est évidemment pas dans \mathcal{G} .

Classification pour le modèle d'Ising

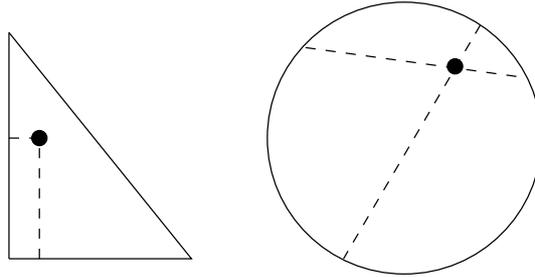
Résumons ici les résultats connus sur la structure de \mathcal{G} et \mathcal{G}_{inv} pour l'interaction d'Ising ferromagnétique sur \mathbb{Z}^d .

- En dimension 1 on a vu au premier TD qu'il n'y avait pas de transition de phase à température finie, on a donc toujours une unique mesure de Gibbs (et elle est invariante par translation).
- En champ non nul il y a unicité de la mesure de Gibbs à toute température (cf. examen de juin 2013). Le champ dans tout le système est suffisant pour briser la symétrie, il n'y a donc pas d'effet des conditions aux bords.
- On peut donc supposer à partir de maintenant $h = 0$ et $d \geq 2$. On a montré précédemment qu'à température suffisamment élevée il y avait une unique mesure de Gibbs, alors qu'à température suffisamment basse les mesures de Gibbs μ^+ et μ^- construites comme limite thermodynamique avec conditions aux bords + et - sont distinctes. Ce phénomène est monotone en β , ce que l'on peut montrer à partir d'inégalités de corrélation (cf. examen de juin 2014), et s'il y a non-unicité des mesures de Gibbs alors nécessairement μ_+ et μ_- sont distinctes (ces conditions aux bords encadrent toutes les conditions aux bords possibles). Il y a donc une unique température (inverse) de transition β_c telle que pour $\beta > \beta_c$ on a $\mathcal{G} = \mathcal{G}_{\text{inv}} = \{\mu^+\} = \{\mu^-\}$ (et par symétrie l'aimantation spontanée est nulle) alors que pour $\beta < \beta_c$ on a $\mu^+ \neq \mu^-$, et donc \mathcal{G} contient au moins le segment $[\mu^+, \mu^-]$ des combinaisons convexes de ces deux mesures.
- Plus précisément, il a été démontré [3, 4] qu'à toute température et en toute dimension $\mathcal{G}_{\text{inv}} = [\mu^+, \mu^-]$, les seules mesures de Gibbs extrémales et invariantes par translation sont donc μ^+ et μ^- .
- Par contre $\mathcal{G} = [\mu^+, \mu^-]$ n'est vrai qu'en dimension $d = 2$. Pour les dimensions $d \geq 3$ \mathcal{G} est strictement plus grand que \mathcal{G}_{inv} , il existe des conditions aux bords inhomogènes avec une paroi de domaine (par exemple $\underline{\tau}$ avec $\tau_x = +1$ si $x_1 > 0$, $\tau_x = -1$ si $x_1 \leq 0$) qui conduisent à la limite thermodynamique à des mesures de Gibbs extrémales dans \mathcal{G} , non invariantes par translation,

distinctes de μ^+ et μ^- . Ces dernières sont toutefois des points extrémaux de \mathcal{G} . La classification complète des éléments extrémaux de \mathcal{G} n'est pas achevée. En dimension 2 les parois de domaine (qui sont des lignes, de dimension 1) sont instables et ce phénomène ne se produit pas.

Autres aspects reliés à la convexité

Les ensembles convexes se divisent en deux catégories : ceux pour lesquels la décomposition de tout élément en combinaison linéaire d'éléments extrémaux est unique, que l'on appelle des simplexes (de Choquet), et les autres. Par exemple dans le plan, comme illustré ci-dessous, un triangle est un simplexe, alors qu'un carré ou un disque n'en sont pas. Le disque a un nombre infini d'éléments extrémaux, le carré que 4, mais dans les deux cas il y a non-unicité de la décomposition d'un point sur les éléments extrémaux.



On peut montrer (cf. [1]) que \mathcal{M} , \mathcal{M}_{inv} , \mathcal{G} et \mathcal{G}_{inv} sont des simplexes. Les points extrêmes dans \mathcal{M} sont les Dirac, dans \mathcal{M}_{inv} ce sont les mesures ergodiques.

4 Principe variationnel

Dans les deux chapitres précédents nous avons suivi deux approches complètement différentes pour l'étude d'un système infini muni d'une interaction Φ . D'une part on s'est intéressé seulement aux quantités thermodynamiques, et en particulier à l'existence et aux propriétés de la densité d'énergie libre $f(\beta, \Phi)$. D'autre part on a défini et discuté les propriétés de l'ensemble $\mathcal{G}(\beta, \Phi)$ des mesures de Gibbs, en considérant uniquement les probabilités des configurations microscopiques. On va voir dans ce chapitre qu'en fait ces deux approches sont intimement reliées par un principe variationnel. Ce chapitre s'appuiera largement sur les résultats établis au TD 2 (en volume fini) et sur les généralisations fonctionnelles des principes d'analyse convexe du TD 3.

4.1 En volume fini

Rappelons brièvement la démarche et les résultats du TD 2. Nous avons considéré des configurations discrètes $\underline{\sigma}$ dans un espace fini Σ , et des lois de probabilités $\mu \in \mathcal{M}$, définies de manière élémentaire par la donnée des probabilités $\mu(\underline{\sigma})$ pour chaque configuration. Nous avons défini l'entropie de Shannon d'une loi de probabilité selon

$$S(\mu) = - \sum_{\underline{\sigma} \in \Sigma} \mu(\underline{\sigma}) \ln \mu(\underline{\sigma}) , \quad (173)$$

et établi quelques unes de ses propriétés.

Etant donné un Hamiltonien H , i.e. une fonction $H(\underline{\sigma})$ définie sur Σ , l'énergie libre $F_c(\beta, H)$ et la loi de probabilité de l'ensemble canonique $\mu_c(\beta, H)$ sont données par

$$F_c(\beta, H) = -\frac{1}{\beta} \ln Z(\beta, H) , \quad Z(\beta, H) = \sum_{\underline{\sigma}} e^{-\beta H(\underline{\sigma})} , \quad \mu_c(\beta, H, \underline{\sigma}) = \frac{e^{-\beta H(\underline{\sigma})}}{Z(\beta, H)} . \quad (174)$$

On a finalement introduit une énergie libre dite de Gibbs, fonction d'un Hamiltonien et d'une loi de probabilité (et d'une température inverse) :

$$F_{\text{Gibbs}}(\beta, H, \mu) = \sum_{\underline{\sigma}} \mu(\underline{\sigma}) H(\underline{\sigma}) - \frac{1}{\beta} S(\mu) , \quad (175)$$

de la forme "énergie (moyenne) - température \times entropie".

On a pu alors démontrer deux principes variationnels :

- l'énergie libre canonique est donnée par la minimisation de l'énergie libre de Gibbs :

$$F_c(\beta, H) = \min_{\mu \in \mathcal{M}} F_{\text{Gibbs}}(\beta, H, \mu) . \quad (176)$$

- la loi de probabilité canonique est le minimiseur de F_{Gibbs} :

$$\mu_c(\beta, H) = \operatorname{argmin}_{\mu \in \mathcal{M}} F_{\text{Gibbs}}(\beta, H, \mu) . \quad (177)$$

L'objet de ce chapitre est la généralisation, en volume infini, de ces deux résultats. On doit donc commencer par construire l'équivalent de F_{Gibbs} (plus précisément de sa version intensive, l'énergie libre par spin) dans ce contexte-là, et donc définir ce que l'on entend par entropie d'une mesure sur $\Sigma = \chi^{\mathbb{Z}^d}$, ainsi que son énergie moyenne.

4.2 Caractérisation variationnelle de $f(\beta, \Phi)$

4.2.1 Entropie d'une mesure en volume infini

Définition 9. On définit la densité d'entropie (de Shannon) d'une mesure de probabilité $\mu \in \mathcal{M}$ dans le volume fini $\Lambda \subset \mathbb{Z}^d$ selon

$$s_\Lambda(\mu) = \frac{1}{|\Lambda|} S(\mu_\Lambda) = -\frac{1}{|\Lambda|} \sum_{\underline{\sigma}_\Lambda} \mu_\Lambda(\underline{\sigma}_\Lambda) \ln \mu_\Lambda(\underline{\sigma}_\Lambda) , \quad (178)$$

où μ_Λ est la loi marginale de μ , et l'on utilise la convention $0 \ln 0 = 0$.

Le lemme de Fekete admet une généralisation multidimensionnelle :

Lemme 22. Soit u_{n_1, \dots, n_d} sous-additive dans chacune de ses variables indépendamment : pour tout $i \in [1, d]$, on a $u_{n_1, \dots, n_{i-1}, n_i + m_i, n_{i+1}, \dots, n_d} \leq u_{n_1, \dots, n_{i-1}, n_i, n_{i+1}, \dots, n_d} + u_{n_1, \dots, n_{i-1}, m_i, n_{i+1}, \dots, n_d}$. Alors

$$\lim_{n_1, \dots, n_d \rightarrow \infty} \frac{u_{n_1, \dots, n_d}}{n_1 \dots n_d} \text{ existe et vaut } \inf_{n_1, \dots, n_d \geq 1} \frac{u_{n_1, \dots, n_d}}{n_1 \dots n_d}, \quad (179)$$

les d limites $n_i \rightarrow \infty$ pouvant être prises dans n'importe quel ordre.

Démonstration. Fixons $p_1, \dots, p_d \geq 1$, notons pour $i = 1, \dots, d$, $n_i = a_i p_i + b_i$, avec $a_i = \lfloor \frac{n_i}{p_i} \rfloor$ et $b_i \in [0, p_i - 1]$. En utilisant la sous-additivité successivement pour chaque coordonnée,

$$u_{n_1, \dots, n_d} \leq a_1 u_{p_1, n_2, \dots, n_d} + u_{b_1, n_2, \dots, n_d} \quad (180)$$

$$\leq a_1 a_2 u_{p_1, p_2, n_3, \dots, n_d} + a_1 u_{p_1, b_2, n_3, \dots, n_d} + a_2 u_{b_1, p_2, n_3, \dots, n_d} + u_{b_1, b_2, n_3, \dots, n_d} \quad (181)$$

$$\leq \dots \quad (182)$$

$$\leq a_1 \dots a_d u_{p_1, \dots, p_d} + \dots, \quad (183)$$

où les termes non-écrits dans la dernière ligne contiennent au maximum un produit de $d-1$ a_i multiplié par un terme borné si p_1, \dots, p_d est fixé. Il vient donc

$$\limsup_{n_1, \dots, n_d \rightarrow \infty} \frac{u_{n_1, \dots, n_d}}{n_1 \dots n_d} \leq \frac{u_{p_1, \dots, p_d}}{p_1 \dots p_d} \quad (184)$$

quand tous les n_i vont à l'infini, à n'importe quelle vitesse l'un par rapport à l'autre. Comme c'est vrai quelque soit (p_1, \dots, p_d) le limsup est plus petit que l'inf, donc la limite existe et est égale à l'inf. \square

Lemme 23. Pour une mesure de probabilité invariante par translation, $\mu \in \mathcal{M}_{\text{inv}}$, on peut définir une densité d'entropie par spin $s(\mu)$ par $s(\mu) = \lim_{\Lambda \rightarrow \infty} s_\Lambda(\mu)$ avec $\Lambda \rightarrow \infty$ le long d'une suite de rectangles : la limite existe et elle est indépendante de la suite $\Lambda \rightarrow \infty$. La fonction $s(\mu)$ vérifie les propriétés suivantes :

- $s(\mu) \in [0, \ln |\chi|]$
- la fonction $\mu \rightarrow s(\mu)$ de \mathcal{M}_{inv} dans \mathbb{R} est affine, $s(\alpha\mu + (1-\alpha)\nu) = \alpha s(\mu) + (1-\alpha)s(\nu)$, $\forall \mu, \nu \in \mathcal{M}_{\text{inv}}, \alpha \in [0, 1]$.

Démonstration. Prouvons d'abord l'existence de la limite. Pour $\mu \in \mathcal{M}_{\text{inv}}$, notons $S_{L_1, \dots, L_d} = S_\Lambda(\mu)$ où Λ est un rectangle de \mathbb{Z}^d de côtés L_1, \dots, L_d ; comme μ est invariante par translation la position du rectangle est arbitraire. Montrons que S_{L_1, \dots, L_d} est sous-additive. Pour une direction $i \in \{1, \dots, d\}$ on note Δ un rectangle de côtés $L_1, \dots, L_{i-1}, L_i + L'_i, L_{i+1}, \dots, L_d$, décomposé en deux rectangles disjoints Λ de côtés $L_1, \dots, L_{i-1}, L_i, L_{i+1}, \dots, L_d$ et Λ' de côtés $L_1, \dots, L_{i-1}, L'_i, L_{i+1}, \dots, L_d$. D'après les résultats du TD 2, $S_\Delta(\mu) \leq S_\Lambda(\mu) + S_{\Lambda'}(\mu)$ (il suffit de calculer la distance de Kullback-Leibler entre μ_Δ et $\mu_\Lambda \mu_{\Lambda'}$), ce qui est précisément la sous-additivité au sens du lemme de Fekete multidimensionnel, qui permet donc de conclure en l'existence de la limite.

L'indépendance de la limite par rapport au choix de la suite $\Lambda \rightarrow \infty$ est assurée par l'invariance sous les translations de μ (on peut donc bouger arbitrairement un coin des rectangles), et par l'énoncé du lemme de Fekete (la façon dont L_1, \dots, L_d divergent n'a pas d'influence sur la limite).

Signalons qu'en fait la limite existe pour toute suite $\Lambda \rightarrow \infty$ au sens de van Hove, pas seulement pour les rectangles. La preuve de cette généralisation repose sur une inégalité un peu plus forte sur $S[\mu]$,

$$S_{\Lambda \cup \Lambda'}(\mu) \leq S_\Lambda(\mu) + S_{\Lambda'}(\mu) - S_{\Lambda \cap \Lambda'}(\mu). \quad (185)$$

Pour prouver les deux propriétés :

- on a vu en TD que $S_\Lambda(\mu) \in [0, |\Lambda| \ln |\chi|]$, donc $s_\Lambda(\mu) \in [0, \ln |\chi|]$, et la limite quand $\Lambda \rightarrow \infty$ reste dans cet intervalle.

– on a aussi démontré que S_Λ était concave et “presque convexe”, i.e.

$$\alpha S_\Lambda(\mu) + (1-\alpha)S_\Lambda(\nu) \leq S_\Lambda(\alpha\mu + (1-\alpha)\nu) \leq \alpha S_\Lambda(\mu) + (1-\alpha)S_\Lambda(\nu) - \alpha \ln \alpha - (1-\alpha) \ln(1-\alpha) .$$

En divisant par $|\Lambda|$ les termes correctifs de la borne supérieure sont plus petits que $(\ln 2)/|\Lambda|$ quelque soit α , ils disparaissent donc pour $\Lambda \rightarrow \infty$, ce qui montre que la limite $s(\mu)$ est affine. \square

Notons que nous nous sommes restreints à des mesures invariantes par translation pour définir $s(\mu)$. En effet on pourrait avoir sinon des mesures μ qui sont déterministes sur des portions arbitrairement grandes du réseau (avec donc une densité d’entropie locale nulle) et uniformément aléatoires dans d’autres régions (où la densité d’entropie locale serait maximale). La limite thermodynamique de l’entropie serait alors mal définie, car fortement dépendante des détails microscopiques de la limite $\Lambda \rightarrow \infty$.

4.2.2 Energie moyenne d’une mesure en volume infini

Considérons maintenant la partie énergétique dans la définition de l’énergie libre de Gibbs, à savoir la généralisation en volume infini de l’expression $\sum_{\underline{\sigma}} \mu(\underline{\sigma}) H(\underline{\sigma})$. Désormais le modèle n’est plus défini directement par son Hamiltonien, mais par l’intermédiaire d’une interaction Φ .

Définition 10. *A une interaction $\Phi \in \mathcal{I}_g$ on associe la fonction de Σ dans \mathbb{R}*

$$u_\Phi = \sum_{X \ni 0} \frac{\phi_X}{|X|} . \quad (186)$$

Par définition de \mathcal{I}_g la somme converge normalement, u_Φ est donc continue. De plus $\Phi \rightarrow u_\Phi$ est une application linéaire de \mathcal{I}_g dans \mathcal{C} .

Lemme 24. *Pour $\Phi \in \mathcal{I}_g$ et $\mu \in \mathcal{M}_{\text{inv}}$,*

$$\lim_{\Lambda \rightarrow \infty} \frac{1}{|\Lambda|} \mu(H_\Lambda^\Phi) = \mu(u_\Phi) , \quad (187)$$

où la limite peut être prise au sens de Van Hove, et H_Λ^Φ est l’Hamiltonien aux conditions aux bords libres dans Λ pour l’interaction Φ .

On voit donc que u_Φ représente la contribution énergétique pour une variable (ici choisi à l’origine pour simplifier). Par exemple pour l’interaction d’Ising à proches voisins, $u(\underline{\sigma}) = -h\sigma_0 - \frac{J}{2} \sum_{x \in \partial 0} \sigma_x \sigma_0$, le facteur 1/2 permet d’éviter le double comptage des liens autour de l’origine.

Démonstration. Supposons pour commencer que Φ soit de portée finie r , et considérons un volume fini Λ . Alors

$$H_\Lambda^\Phi = \sum_{X \subset \Lambda} \phi_X = \sum_{X \subset \Lambda} \sum_{x \in X} \frac{\phi_X}{|X|} = \sum_{x \in \Lambda \setminus \partial_r \Lambda} \sum_{X \ni x} \frac{\phi_X}{|X|} + H_{\partial_r \Lambda}^\Phi , \quad (188)$$

où $\partial_r \Lambda$ contient les sommets de Λ à distance plus petite que r de Λ^c . En utilisant l’invariance par translation on a

$$\frac{1}{|\Lambda|} \mu(H_\Lambda^\Phi) = \mu(u_\Phi) - \frac{|\partial_r \Lambda|}{|\Lambda|} \mu(u_\Phi) + \frac{1}{|\Lambda|} \mu(H_{\partial_r \Lambda}^\Phi) . \quad (189)$$

Comme $\Lambda \rightarrow \infty$ au sens de van Hove $|\partial_r \Lambda| = o(|\Lambda|)$, les deux derniers termes disparaissent donc à la limite (rappelons que d’après le lemme 1 $\|H_{\partial_r \Lambda}\|_\infty \leq |\partial_r \Lambda| \|\Phi\|_g$).

On peut maintenant généraliser le résultat en utilisant la densité des interactions à portée finie dans \mathcal{I}_g , comme on l’avait fait pour prouver le théorème 3. Considérons donc $\Phi \in \mathcal{I}_g$, $\epsilon > 0$ et r

suffisamment grand pour que $\|\Phi - \Phi^{(r)}\|_g \leq \frac{\epsilon}{2}$, où $\Phi^{(r)}$ est la troncation de portée r de Φ définie dans la preuve du théorème 3. Alors

$$\|H_\Lambda^\Phi - H_\Lambda^{\Phi^{(r)}}\|_\infty = \|H_\Lambda^{\Phi - \Phi^{(r)}}\|_\infty \leq |\Lambda| \|\Phi - \Phi^{(r)}\|_g \leq |\Lambda| \frac{\epsilon}{2} \Rightarrow \left| \frac{1}{|\Lambda|} \mu(H_\Lambda^\Phi) - \frac{1}{|\Lambda|} \mu(H_\Lambda^{\Phi^{(r)}}) \right| \leq \frac{\epsilon}{2}, \quad (190)$$

et

$$\|u_\Phi - u_{\Phi^{(r)}}\|_\infty = \|u_{\Phi - \Phi^{(r)}}\|_\infty \leq \|\Phi - \Phi^{(r)}\|_g \leq \frac{\epsilon}{2} \Rightarrow |\mu(u_\Phi) - \mu(u_{\Phi^{(r)}})| \leq \frac{\epsilon}{2}. \quad (191)$$

Donc

$$\begin{aligned} \left| \frac{1}{|\Lambda|} \mu(H_\Lambda^\Phi) - \mu(u_\Phi) \right| &\leq \left| \frac{1}{|\Lambda|} \mu(H_\Lambda^\Phi) - \frac{1}{|\Lambda|} \mu(H_\Lambda^{\Phi^{(r)}}) \right| + \left| \frac{1}{|\Lambda|} \mu(H_\Lambda^{\Phi^{(r)}}) - \mu(u_{\Phi^{(r)}}) \right| + |\mu(u_\Phi) - \mu(u_{\Phi^{(r)}})| \\ &\leq \epsilon + \left| \frac{1}{|\Lambda|} \mu(H_\Lambda^{\Phi^{(r)}}) - \mu(u_{\Phi^{(r)}}) \right|. \end{aligned} \quad (192)$$

En passant à la limite quand $\Lambda \rightarrow \infty$ le dernier terme du membre de droite disparaît, comme ϵ est arbitraire on peut conclure la preuve. \square

4.2.3 Principe variationnel pour l'énergie libre

Des deux paragraphes précédents on est naturellement amené à définir la densité d'énergie libre de Gibbs, pour une interaction Φ et une mesure invariante par translation $\mu \in \mathcal{M}_{\text{inv}}$, selon $f_{\text{Gibbs}}(\beta, \Phi, \mu) = \mu(u_\Phi) - \frac{1}{\beta} s(\mu)$. On a alors une caractérisation variationnelle de l'énergie libre canonique en parfait analogue avec le résultat (176) précédemment établi en volume fini :

Théorème 12. *Pour tout $\Phi \in \mathcal{I}_g$ et pour tout β , la densité d'énergie libre peut s'exprimer comme*

$$f(\beta, \Phi) = \inf_{\mu \in \mathcal{M}_{\text{inv}}} \left[\mu(u_\Phi) - \frac{1}{\beta} s(\mu) \right]. \quad (193)$$

Démonstration. Prouvons d'abord que si $\mu \in \mathcal{M}_{\text{inv}}$, alors

$$f(\beta, \Phi) \leq \mu(u_\Phi) - \frac{1}{\beta} s(\mu). \quad (194)$$

C'est une simple généralisation du résultat obtenu en TD pour un volume fini. Reproduisons ici une des preuves possibles :

$$Z_\Lambda(\beta, \Phi) = \sum_{\underline{\sigma}_\Lambda} e^{-\beta H_\Lambda^\Phi(\underline{\sigma}_\Lambda)} = \sum_{\underline{\sigma}_\Lambda} \mu_\Lambda(\underline{\sigma}_\Lambda) e^{-\beta H_\Lambda^\Phi(\underline{\sigma}_\Lambda) - \ln \mu_\Lambda(\underline{\sigma}_\Lambda)} \quad (195)$$

$$\geq \exp \left[-\beta \sum_{\underline{\sigma}_\Lambda} \mu(\underline{\sigma}_\Lambda) H_\Lambda^\Phi(\underline{\sigma}_\Lambda) - \sum_{\underline{\sigma}_\Lambda} \mu_\Lambda(\underline{\sigma}_\Lambda) \ln \mu_\Lambda(\underline{\sigma}_\Lambda) \right], \quad (196)$$

en utilisant l'inégalité de Jensen sur l'exponentielle convexe. En prenant le logarithme et en divisant par $-\beta|\Lambda|$, il vient

$$f_\Lambda(\beta, \Phi) \leq \frac{1}{|\Lambda|} \mu(H_\Lambda^\Phi) - \frac{1}{\beta} s_\Lambda(\mu). \quad (197)$$

En passant à la limite $\Lambda \rightarrow \infty$ le long de rectangles et en utilisant les lemmes précédents on obtient bien (194).

Il reste maintenant à montrer que l'on peut construire des μ qui approchent la borne inférieure arbitrairement près. Pour cela fixons Δ le cube de côté a avec un coin à l'origine, i.e. $\Delta = \{0, \dots, a-1\}^d$, et notons $\eta_\Delta(\underline{\sigma}_\Delta) = \frac{e^{-\beta H_\Delta^\Phi(\underline{\sigma}_\Delta)}}{Z_\Delta(\beta, \Phi)}$ la mesure aux conditions libres, définie sur \mathcal{F}_Δ . Pour la transformer en une mesure sur tout \mathcal{F} on considère la mesure produit des translatées de η_Δ , que l'on note

$$\nu_a = \bigotimes_{x \in a\mathbb{Z}^d} T_x \eta_\Delta. \quad (198)$$

Cette mesure est a -périodique dans toutes les directions mais pas invariante, on introduit donc finalement

$$\mu_a = \frac{1}{|\Delta|} \sum_{x \in \Delta} T_x \nu_a, \quad (199)$$

qui est bien invariante par translation. Pour conclure il reste à montrer que quand a tend vers l'infini la mesure μ_a sature la borne, autrement dit que

$$f(\beta, \Phi) = \lim_{a \rightarrow \infty} \left[\mu_a(u_\Phi) - \frac{1}{\beta} s(\mu_a) \right]. \quad (200)$$

Montrons cela en admettant temporairement :

$$(i) \quad s(\mu_a) = \frac{1}{|\Delta|} S(\eta_\Delta), \quad (201)$$

$$(ii) \quad \lim_{a \rightarrow \infty} \left[\mu_a(u_\Phi) - \frac{1}{|\Delta|} \eta_\Delta(H_\Delta^\Phi) \right] = 0. \quad (202)$$

En remarquant que

$$S(\eta_\Delta) = - \sum_{\underline{\sigma}_\Delta} \eta_\Delta(\underline{\sigma}_\Delta) \ln \left(\frac{e^{-\beta H_\Delta^\Phi(\underline{\sigma}_\Delta)}}{Z_\Delta(\beta, \Phi)} \right) = \ln Z_\Delta(\beta, \Phi) + \beta \eta_\Delta(H_\Delta^\Phi), \quad (203)$$

on a alors en utilisant (i)

$$\mu_a(u_\Phi) - \frac{1}{\beta} s(\mu_a) = \mu_a(u_\Phi) - \frac{1}{|\Delta|} \eta_\Delta(H_\Delta^\Phi) + f_\Delta(\beta, \Phi). \quad (204)$$

Quand $a \rightarrow \infty$ on sait grâce au théorème 3 que $f_\Delta \rightarrow f$, et en utilisant (ii) les deux premiers termes se compensent.

Prouvons maintenant (i). Notons Λ un rectangle fini, on a

$$S_\Lambda(\mu_a) = S_\Lambda \left(\frac{1}{|\Delta|} \sum_{x \in \Delta} T_x \nu_a \right), \quad (205)$$

en utilisant la concavité et la “presque convexité” de S_Λ on obtient

$$\frac{1}{|\Delta|} \sum_{x \in \Delta} S_\Lambda(T_x \nu_a) \leq S_\Lambda(\mu_a) \leq \frac{1}{|\Delta|} \sum_{x \in \Delta} S_\Lambda(T_x \nu_a) + \ln |\Delta|. \quad (206)$$

Comme $s(\mu_a) = \lim_{\Lambda \rightarrow \infty} \frac{1}{|\Lambda|} S_\Lambda(\mu_a)$, le résultat sera démontré si l'on obtient que pour tout $x \in \Delta$, $\lim_{\Lambda \rightarrow \infty} \frac{1}{|\Lambda|} S_\Lambda(T_x \nu_a) = \frac{1}{|\Delta|} S(\eta_\Delta)$. Considérons pour Λ un cube de côté L avec un coin à l'origine. Λ contient $\left(\lfloor \frac{L}{a} \rfloor\right)^d$ translatées complètes de $T_x \Delta$, et un reste que l'on notera Λ' . $T_x \nu_a$ est une mesure produit des translatées de η_Δ , d'où

$$\frac{1}{|\Lambda|} S_\Lambda(T_x \nu_a) = \frac{1}{|\Lambda|} \left(\left\lfloor \frac{L}{a} \right\rfloor \right)^d S(\eta_\Delta) + \frac{1}{|\Lambda|} S_{\Lambda'}(T_x \nu_a) \rightarrow \frac{1}{|\Delta|} S[\eta_\Delta]. \quad (207)$$

En effet l'entropie du reste est d'ordre $|\Lambda'|$, et $|\Lambda'|/|\Lambda| \rightarrow 0$ quand $L \rightarrow \infty$ avec a fixé.

Prouvons finalement (ii). D'une part,

$$\mu_a(u_\Phi) = \frac{1}{|\Delta|} \sum_{x \in \Delta} T_x \nu_a \left(\sum_{X \ni 0} \frac{\phi_X}{|X|} \right) = \frac{1}{|\Delta|} \sum_{x \in \Delta} \sum_{X \ni x} \frac{1}{|X|} \nu_a(\phi_X). \quad (208)$$

D'autre part,

$$\frac{1}{|\Delta|} \eta_\Delta(H_\Delta^\Phi) = \frac{1}{|\Delta|} \sum_{X \subset \Delta} \eta_\Delta(\phi_X) = \frac{1}{|\Delta|} \sum_{X \subset \Delta} \sum_{x \in X} \frac{1}{|X|} \nu_a(\phi_X). \quad (209)$$

La différence entre ces deux expressions est donc un terme de surface, dû aux interactions à cheval entre Δ et son extérieur. En prenant une interaction à portée finie et en prenant $a \rightarrow \infty$ on montre le résultat (ii), on conclut ensuite en utilisant la densité des interactions à portée finie dans \mathcal{I}_g . \square

4.3 Caractérisation variationnelle de $\mathcal{G}_{\text{inv}}(\beta, \Phi)$

On vient d'établir une expression variationnelle de l'énergie libre $f(\beta, \Phi)$, qui constitue une généralisation en volume infini de (176). On peut se demander maintenant si (177) admet aussi une extension en volume infini, autrement dit si les mesures de Gibbs associées à une interaction Φ sont complètement caractérisées par la valeur qu'elles donnent à l'énergie libre variationnelle f_{Gibbs} . La réponse est oui, si l'on se restreint aux mesures invariantes par translation ; on a en effet :

Théorème 13. *Si $\Phi \in \mathcal{I}_p$ les mesures de Gibbs invariantes par translation sont précisément celles qui sont optimales dans le théorème 12, i.e.*

$$\mu \in \mathcal{G}_{\text{inv}}(\beta, \Phi) \Leftrightarrow \mu \in \mathcal{M}_{\text{inv}} \text{ et } f(\beta, \Phi) = \mu(u_\Phi) - \frac{1}{\beta} s(\mu) . \quad (210)$$

On peut critiquer le résultat de cette approche qui ne conduit qu'à une caractérisation des mesures de Gibbs invariantes par translation, et pas de toutes les mesures de Gibbs. Mais d'une part les interactions dont on est parti sont invariantes par translation, donc les mesures dans \mathcal{G}_{inv} sont plus "naturelles" que celles de \mathcal{G} . D'autre part, dans le cas d'Ising les mesures de Gibbs non-invariantes par translation dont on a parlé précédemment sont celles qui contiennent une paroi de domaine entre des régions à aimantation positive et des régions à aimantation négative. Mais les deux phases homogènes ont la même densité d'énergie libre par spin, la paroi de domaine sur laquelle l'aimantation change est donc négligeable pour l'énergie libre, ce n'est qu'un effet de surface (et en effet on peut parler de tension de surface, comme pour les interfaces entre la phase vapeur et la phase liquide d'un fluide). Il est donc naturel que l'approche thermodynamique ne soit pas capable de détecter cette subtilité des mesures de Gibbs non-invariantes par translation, puisque les quantités intensives comme l'énergie libre par spin ne sont sensibles qu'aux effets de volume et pas à ceux de surface.

On ne prouvera pas ce théorème qui fait appel à des théorèmes d'analyse convexe assez techniques, et on renverra le lecteur vers [1] pour une discussion complète. Contentons-nous d'une esquisse du schéma de la preuve. Un résultat intermédiaire est la caractérisation des mesures de Gibbs invariantes par translation à partir des fonctionnelles tangentes de f (cf. le TD 3 pour une définition des fonctions tangentes) :

Définition 11. *Pour $\mu \in \mathcal{M}_{\text{inv}}$, on note v_μ la forme linéaire sur \mathcal{I}_g définie par $v_\mu(\Phi) = \mu(u_\Phi)$. C'est bien une forme linéaire à cause de la linéarité de $\Phi \rightarrow u_\Phi$ et de la linéarité de la moyenne.*

Théorème 14. *Si $\Phi \in \mathcal{I}_p$ et $\mu \in \mathcal{M}_{\text{inv}}$,*

$$\mu \in \mathcal{G}_{\text{inv}}(\beta, \Phi) \Leftrightarrow f(\beta, \Psi) \leq f(\beta, \Phi) + v_\mu(\Psi - \Phi) \quad \forall \Psi \in \mathcal{I}_g , \quad (211)$$

i.e. ssi v_μ est tangente à f en Φ .

Or le théorème 12 montre que $f(\Phi)$ est la transformée de Legendre de $s(\mu)$ (à des signes près et des détails techniques à cause de la dimension infinie des espaces \mathcal{M}_{inv} et \mathcal{I}_g qu'il faut restreindre pour avoir une paire duale). De plus au TD 3 on a montré, dans le cas évidemment beaucoup plus simple des fonctions réelles, que \hat{x} est tangente à f en x si et seulement si $f(x) + \hat{f}(\hat{x}) = \langle \hat{x}, x \rangle$. Si l'on admet la généralisation fonctionnelle de ce résultat cette condition devient ici précisément $f(\beta, \Phi) + \frac{1}{\beta} s(\mu) = \mu(u_\Phi)$. La combinaison du théorème 14 qui caractérise les mesures de Gibbs en termes des fonctionnelles tangentes de f , du théorème 12 qui dévoile la structure de transformée de Legendre de f , et de ce dernier résultat général d'analyse convexe conduit ainsi au théorème 13.

On peut justifier intuitivement, au vu des derniers résultats, le lien entre deux définitions a priori différentes d'une "transition de phase", à savoir la non-analyticité de l'énergie libre f comme fonction de β d'une part, et d'autre part l'apparition de mesures de Gibbs distinctes. En extrapolant à partir du cas des fonctions de \mathbb{R} , si f est dérivable elle admettra une seule fonctionnelle tangente, et donc une seule mesure de Gibbs.

5 Equivalence des ensembles

5.1 Discussion heuristique

On avait discuté brièvement dans l'introduction les deux "ensembles", canonique et microcanonique, qui peuvent être utilisés en physique statistique pour décrire deux situations physiques différentes. Dans l'ensemble microcanonique on suppose que le système étudié est isolé du reste de l'univers, son énergie E est donc une grandeur conservée constante; en l'absence d'autres informations sur le système on suppose que toutes les configurations avec cette énergie (à une petite constante additive près) sont équiprobables. La grandeur thermodynamique pertinente est l'entropie microcanonique $S(E)$, égale au logarithme du nombre de configurations avec l'énergie E (la constante de Boltzmann étant prise égale à 1). L'ensemble canonique modélise quant à lui un système à l'équilibre avec un thermostat de température $T = 1/\beta$, les échanges d'énergie sont permis entre le système et le thermostat, les probabilités des configurations sont $\exp[-\beta E(\underline{\sigma})]/Z$ (qui est en fait la loi maximisant l'entropie de Shannon avec une contrainte sur la valeur moyenne de l'énergie), la grandeur thermodynamique est l'énergie libre $F(T)$.

On s'attend à ce que ces deux ensembles soient équivalents dans la limite thermodynamique : en effet dans l'ensemble canonique, pour un système de taille $N \gg 1$ la valeur moyenne de l'énergie $\bar{E}(T)$ est d'ordre N , alors que les fluctuations de l'énergie autour de sa valeur moyenne ne sont que d'ordre \sqrt{N} . Puisque les fluctuations de l'énergie sont négligeables dans la limite $N \rightarrow \infty$ on peut remplacer l'ensemble canonique par l'ensemble microcanonique avec une valeur de l'énergie fixée à la valeur moyenne de l'ensemble canonique. Il y a de plus une structure de transformée de Legendre entre $F(T)$ et $S(E)$, comme on peut le voir avec ce raisonnement :

$$Z(\beta) = \sum_{\underline{\sigma}} e^{-\beta E(\underline{\sigma})} = \sum_E e^{S(E) - \beta E}, \quad F(T) = -\frac{1}{\beta} \ln Z(\beta) \approx \inf_E [E - TS(E)], \quad (212)$$

où l'on remplace la sommation sur E par la valeur maximale de l'intégrand, cette application de la méthode de Laplace étant justifiée heuristiquement par le caractère extensif dans la limite thermodynamique de E et de $S(E)$.

On va énoncer et prouver partiellement un résultat rigoureux de ce type-là, dans le cadre des modèles définis au chapitre 2.

5.2 Existence et propriétés de l'entropie microcanonique

La première chose à faire avant de définir l'entropie microcanonique est de déterminer les valeurs possibles, pour une interaction donnée, de la densité d'énergie. Rappelons que nous avons démontré que $\Phi \rightarrow f(\beta, \Phi)$ est une fonction concave des interactions, à β fixé. Si l'on considère Φ fixée, la fonction $\beta \rightarrow g(\beta, \Phi) = \beta f(\beta, \Phi) = f(1, \beta\Phi)$ est donc une fonction concave de β . Comme on l'a vu au TD 3, g admet donc des dérivées à gauche et à droite pour tout β , qui décroissent avec β , et g est dérivable en β partout sauf sur un ensemble au plus dénombrable de points.

Définition 12. *Pour une interaction $\Phi \in \mathcal{I}_g$, on note $e_{\max}(\Phi) = (D_+g)(0)$, $e_{\min}(\Phi) = \lim_{\beta \rightarrow +\infty} (D_-g)(\beta)$, et $\mathcal{E}_{\text{permis}}(\Phi) =]e_{\min}(\Phi), e_{\max}(\Phi)[$. Pour $e \in \mathcal{E}_{\text{permis}}(\Phi)$ on note $\beta(e, \Phi)$ une température inverse β telle que $e \in [(D_+g)(\beta), (D_-g)(\beta)]$, qui existe toujours d'après les propriétés des fonctions concaves. La fonction $\beta(e)$ est décroissante.*

On peut justifier ces définitions en rappelant qu'en volume fini,

$$\frac{d}{d\beta}(\beta f_\Lambda(\beta, \Phi)) = \frac{1}{|\Lambda|} \sum_{\underline{\sigma}_\Lambda} \eta_\Lambda(\underline{\sigma}_\Lambda) H_\Lambda^\Phi(\underline{\sigma}_\Lambda), \quad (213)$$

qui est donc la densité d'énergie moyenne pour des conditions aux bords libres.

Définition 13. *Pour $\Phi \in \mathcal{I}_g$, $e \in \mathcal{E}_{\text{permis}}(\Phi)$, $\delta > 0$, et $\Lambda \subset_{\text{f}} \mathbb{Z}^d$ on note*

$$s_{m, \Lambda, \delta}(e, \Phi) = \frac{1}{|\Lambda|} \ln N_{\Lambda, \delta}(e, \Phi), \quad N_{\Lambda, \delta}(e, \Phi) = |\{\underline{\sigma}_\Lambda : H_\Lambda^\Phi(\underline{\sigma}_\Lambda) \in [|\Lambda|(e - \delta), |\Lambda|(e + \delta)]\}| \quad (214)$$

la densité d'entropie microcanonique dans le volume fini Λ .

Cette quantité admet une limite thermodynamique, dans le sens suivant :

Théorème 15. Pour $\Phi \in \mathcal{I}_p$, $e \in \mathcal{E}_{\text{permis}}(\Phi)$,

$$\lim_{\delta \rightarrow 0} \liminf_{\Lambda \rightarrow \infty} s_{m,\Lambda,\delta}(e, \Phi) = \lim_{\delta \rightarrow 0} \limsup_{\Lambda \rightarrow \infty} s_{m,\Lambda,\delta}(e, \Phi) , \quad (215)$$

la limite $\Lambda \rightarrow \infty$ pouvant être prise au sens de Van Hove. On note $s_m(e, \Phi)$ la limite commune, et l'on a

$$s_m(e, \Phi) = \inf_{\beta \geq 0} [\beta e - \beta f(\beta, \Phi)] , \quad (216)$$

qui est donc une fonction concave de e (en tant qu'inf de fonctions affines et donc concaves).

La preuve complète de ce théorème nécessite des résultats de théorie ergodique plus avancés que le reste de ce cours, on se contentera donc d'en donner les lignes principales.

Démonstration (partielle). Considérons un volume fini Λ . Pour n'importe quelle valeur de β on peut écrire :

$$N_{\Lambda,\delta}(e, \Phi) = \sum_{\underline{\sigma}_\Lambda} \mathbb{I}(|H_\Lambda^\Phi(\underline{\sigma}_\Lambda) - |\Lambda|e| \leq \delta|\Lambda|) \quad (217)$$

$$= \sum_{\underline{\sigma}_\Lambda} \mathbb{I}(|H_\Lambda^\Phi(\underline{\sigma}_\Lambda) - |\Lambda|e| \leq \delta|\Lambda|) \frac{e^{-\beta H_\Lambda^\Phi(\underline{\sigma}_\Lambda)}}{Z_\Lambda(\beta, \Phi)} Z_\Lambda(\beta, \Phi) e^{\beta H_\Lambda^\Phi(\underline{\sigma}_\Lambda)} , \quad (218)$$

que l'on va maintenant borner inférieurement et supérieurement.

Comme dans la somme $H_\Lambda^\Phi(\underline{\sigma}_\Lambda) \leq |\Lambda|(e + \delta)$,

$$N_{\Lambda,\delta}(e, \Phi) \leq Z_\Lambda(\beta, \Phi) e^{\beta|\Lambda|(e+\delta)} \sum_{\underline{\sigma}_\Lambda} \mathbb{I}(|H_\Lambda^\Phi(\underline{\sigma}_\Lambda) - |\Lambda|e| \leq \delta|\Lambda|) \frac{e^{-\beta H_\Lambda^\Phi(\underline{\sigma}_\Lambda)}}{Z_\Lambda(\beta, \Phi)} \quad (219)$$

$$\leq Z_\Lambda(\beta, \Phi) e^{\beta|\Lambda|(e+\delta)} \sum_{\underline{\sigma}_\Lambda} \frac{e^{-\beta H_\Lambda^\Phi(\underline{\sigma}_\Lambda)}}{Z_\Lambda(\beta, \Phi)} = Z_\Lambda(\beta, \Phi) e^{\beta|\Lambda|(e+\delta)} , \quad (220)$$

où l'on a obtenu une borne supérieure de la première ligne en supprimant la fonction indicatrice. En prenant le log, en divisant par $|\Lambda|$ et en passant à la limite $\Lambda \rightarrow \infty$,

$$\limsup_{\Lambda \rightarrow \infty} s_{m,\Lambda,\delta}(e, \Phi) \leq -\beta f(\beta, \Phi) + \beta(e + \delta) , \quad (221)$$

$$\limsup_{\delta \rightarrow 0} \limsup_{\Lambda \rightarrow \infty} s_{m,\Lambda,\delta}(e, \Phi) \leq -\beta f(\beta, \Phi) + \beta e , \quad (222)$$

$$\limsup_{\delta \rightarrow 0} \limsup_{\Lambda \rightarrow \infty} s_{m,\Lambda,\delta}(e, \Phi) \leq \inf_{\beta \geq 0} [\beta e - \beta f(\beta, \Phi)] , \quad (223)$$

où dans la dernière étape on a utilisé la validité de l'inégalité précédente pour tout β . Cette partie de l'inégalité était donc élémentaire, la difficulté se trouve dans la borne inférieure. On peut écrire

$$N_{\Lambda,\delta}(e, \Phi) \geq Z_\Lambda(\beta, \Phi) e^{\beta|\Lambda|(e-\delta)} \eta_{\Lambda,\beta}(|H_\Lambda^\Phi(\underline{\sigma}_\Lambda) - |\Lambda|e| \leq \delta|\Lambda|) , \quad (224)$$

où $\eta_{\Lambda,\beta}(\underline{\sigma}_\Lambda) = \frac{e^{-\beta H_\Lambda^\Phi(\underline{\sigma}_\Lambda)}}{Z_\Lambda(\beta, \Phi)}$ est la mesure aux conditions aux bords libres dans le volume Λ ; on explicite ici la dépendance en β .

Supposons qu'il existe β_* tel que $e = \frac{dq}{d\beta}(\beta, \Phi)$ (alors nécessairement $\beta_* = \beta(e, \Phi)$). Selon l'idée intuitive expliquée précédemment, η_{Λ,β_*} devrait être concentrée sur des configurations de densité d'énergie proche de e , i.e. $\eta_{\Lambda,\beta_*}(|H_\Lambda^\Phi(\underline{\sigma}_\Lambda) - |\Lambda|e| \leq \delta|\Lambda|) \approx 1$. Supposons donc que l'on ait démontré

$$\liminf_{\delta \rightarrow 0} \liminf_{\Lambda \rightarrow \infty} \frac{1}{|\Lambda|} \ln \eta_{\Lambda,\beta(e,\Phi)}(|H_\Lambda^\Phi(\underline{\sigma}_\Lambda) - |\Lambda|e| \leq \delta|\Lambda|) = 0 ; \quad (225)$$

alors le théorème serait prouvé. En effet, on aurait alors

$$\liminf_{\delta \rightarrow 0} \liminf_{\Lambda \rightarrow \infty} s_{m,\Lambda,\delta}(e, \Phi) \geq [\beta(e, \Phi)e - \beta(e, \Phi)f(\beta(e, \Phi), \Phi)] \geq \inf_{\beta \geq 0} [\beta e - \beta f(\beta, \Phi)] , \quad (226)$$

ce qui permet de conclure vu la borne supérieure démontrée en (223).

Le point crucial dans la démonstration est donc la preuve de (225) ; celle-ci, que l'on admettra ici, repose sur l'idée d'ergodicité. On a introduit cette notion dans la partie 3.6, les mesures ergodiques étant par définition celles invariantes par translation et qui donnent une probabilité soit 0 soit 1 aux évènements invariants par translation. Un des résultats centraux de la théorie ergodique est le « théorème ergodique », qui énonce que pour μ une mesure ergodique, et f une fonction (continue),

$$\lim_{\Lambda \rightarrow \infty} \frac{1}{|\Lambda|} \sum_{x \in \Lambda} T_x f = \mu(f) , \quad (227)$$

μ -presque sûrement et dans L_1 . Autrement dit l'ergodicité traduit une équivalence entre une moyenne « spatiale » (dans le membre de gauche) et une moyenne « d'ensemble » (dans celui de droite). Or on peut voir $\frac{1}{|\Lambda|} H_\Lambda^\Phi$ comme une moyenne spatiale de $|\Lambda|$ translatées de u_Φ (la fonction définie en (186), à des termes de bords près), cette moyenne spatiale est donc, presque sûrement dans la limite $\Lambda \rightarrow \infty$, égale à sa moyenne d'ensemble pour une mesure ergodique. Comme on l'a vu précédemment les mesures de Gibbs invariantes par translation et extrémales ont cette propriété, en prenant donc un élément μ de $\text{ext}\mathcal{G}_{\text{inv}}(\beta(e, \Phi), \Phi)$ on a un chemin naturel pour prouver (225). On passe ici sous silence un certain nombre de difficultés, des preuves complètes peuvent se trouver dans [1, 2]. □

5.3 Discussion

D'après ce qu'on a vu au TD 3 on peut inverser la transformée de Legendre (216) en

$$f(\beta, \Phi) = \inf_{e \in \mathcal{E}_{\text{permis}}} \left[e - \frac{1}{\beta} s_m(e, \Phi) \right] . \quad (228)$$

Le fait que l'énergie libre soit la transformée de Legendre de l'entropie microcanonique est naturel au vu du calcul rudimentaire de l'équation (212). Ce qui est hautement non-trivial c'est que l'entropie microcanonique soit une fonction concave de l'énergie, et donc que la transformée de Legendre entre les $f(\beta)$ et $s_m(e)$ soit inversible. Ce fait repose crucialement sur la dimension finie du réseau sur lequel on a défini le modèle (et peut devenir faux en champ moyen).

On peut justifier intuitivement la concavité de l'entropie microcanonique de la façon suivante. Considérons e_1 et e_2 deux densités d'énergie permises, et un paramètre $\alpha \in]0, 1[$. Prenons un (grand) volume Λ , par exemple un cube de côté L , et divisons-le en deux rectangles disjoints, $\Lambda = \Lambda_1 \cup \Lambda_2$, avec $|\Lambda_1| = \alpha|\Lambda|$ et $|\Lambda_2| = (1 - \alpha)|\Lambda|$, en prenant pour Λ_1 un rectangle de côtés $\alpha L \times L \times \dots \times L$ et pour Λ_2 des côtés $(1 - \alpha)L \times L \times \dots \times L$. Construisons maintenant toutes les configurations dans Λ_1 de densité d'énergie e_1 (à un petit δ près), et toutes les configurations dans Λ_2 de densité d'énergie proche de e_2 . Il y a environ $e^{|\Lambda_1|s_m(e_1)} \times e^{|\Lambda_2|s_m(e_2)}$ paires de telles configurations ; leur juxtaposition donne une configuration dans Λ de densité d'énergie proche de $\alpha e_1 + (1 - \alpha)e_2$, en supposant négligeable les effets de bord à l'interface entre Λ_1 et Λ_2 . Les configurations ainsi construites par produit de configurations dans Λ_1 et Λ_2 de densités d'énergie respectivement e_1 et e_2 n'est qu'un sous-ensemble de toutes les configurations dans Λ de densité d'énergie $\alpha e_1 + (1 - \alpha)e_2$, on a donc

$$e^{|\Lambda|s_m(\alpha e_1 + (1 - \alpha)e_2)} \geq e^{|\Lambda_1|s_m(e_1)} \times e^{|\Lambda_2|s_m(e_2)} = e^{|\Lambda|(\alpha s_m(e_1) + (1 - \alpha)s_m(e_2))} , \quad (229)$$

soit

$$s_m(\alpha e_1 + (1 - \alpha)e_2) \geq \alpha s_m(e_1) + (1 - \alpha)s_m(e_2) . \quad (230)$$

Cet argument repose sur l'existence d'une tension de surface finie et d'une surface négligeable devant le volume, ce qui est faux en champ moyen (par exemple dans le cadre du modèle de Curie-Weiss étudié au TD 1).

5.4 Autres paires d'ensembles

L'équivalence d'ensembles considérée ci-dessus traite du passage entre l'ensemble microcanonique où l'énergie E est fixée strictement, à l'ensemble canonique où elle n'est fixée qu'en moyenne, par l'intermédiaire de la température inverse β ; E et β sont deux quantités thermodynamiquement conjuguées, dans le sens où l'une est intensive, l'autre extensive, et c'est le produit des deux qui apparaît dans l'exponentielle donnant le poids d'une configuration canonique. Il existe bien d'autres paires de grandeurs conjuguées : pour un système de particules notamment, on a la conjugaison $p \leftrightarrow V$ entre pression et volume, et $\mu \leftrightarrow N$ entre potentiel chimique et nombre de particules. On peut donc construire un grand nombre d'ensembles, en choisissant au sein de chaque paire conjuguée si l'on prend la grandeur intensive ou l'extensive comme variable. Pour les raisons exposées ci-dessus on a l'équivalence entre tous ces ensembles dans la limite thermodynamique, pour les systèmes en dimension finie (et en l'absence de transition de phase), avec des structures de transformée de Legendre entre les différentes fonctions thermodynamiques.

Dans le cadre des systèmes sur réseau traités dans ce cours, la construction la plus générale consiste à considérer n interactions Φ_1, \dots, Φ_n , parmi lesquelles p sont traitées de manière canonique, avec des « températures inverses » β_1, \dots, β_p , les autres étant traités de manière microcanonique, en fixant leurs densités d'énergie par spin e_{p+1}, \dots, e_n . La fonction thermodynamique correspondante est alors de la forme :

$$g(\beta_1, \Phi_1, \dots, \beta_p, \Phi_p, e_{p+1}, \Phi_{p+1}, \dots, e_n, \Phi_n) \quad (231)$$

$$= \lim_{\delta \rightarrow 0} \limsup_{\Lambda \rightarrow \infty} \frac{1}{|\Lambda|} \ln \left(\sum_{\underline{\sigma}_\Lambda} e^{-\sum_{i=1}^p \beta_i H_\Lambda^{\Phi_i}(\underline{\sigma}_\Lambda)} \prod_{i=p+1}^n \mathbb{I}(H_\Lambda^{\Phi_i}(\underline{\sigma}_\Lambda) \in [|\Lambda|(e_i - \delta), |\Lambda|(e_i + \delta)]) \right),$$

et on a des structures de transformée de Legendre pour chaque couple de variables $\beta_i \leftrightarrow e_i$. Un cas particulier que l'on va utiliser dans la partie 6 concerne les systèmes de spin d'Ising, $\chi = \{-1, +1\}$, pour une interaction Φ arbitraire. Une observable naturelle d'un tel système est l'aimantation (la somme des spins dans un volume fini), conjuguée à un champ magnétique. Introduisons donc deux énergies libres selon que l'on traite cette variable de manière canonique ou microcanonique (l'interaction Φ étant elle toujours en canonique à la température inverse β) :

$$f(\beta, \Phi, h) = -\frac{1}{\beta} \lim_{\Lambda \rightarrow \infty} \frac{1}{|\Lambda|} \ln \left(\sum_{\underline{\sigma}_\Lambda} e^{-\beta H_\Lambda^\Phi(\underline{\sigma}_\Lambda) + \beta h \sum_{x \in \Lambda} \sigma_x} \right), \quad (232)$$

$$\hat{f}(\beta, \Phi, m) = -\frac{1}{\beta} \lim_{\delta \rightarrow 0} \limsup_{\Lambda \rightarrow \infty} \frac{1}{|\Lambda|} \ln \left(\sum_{\underline{\sigma}_\Lambda} e^{-\beta H_\Lambda^\Phi(\underline{\sigma}_\Lambda)} \mathbb{I} \left(\sum_{x \in \Lambda} \sigma_x \in [|\Lambda|(m - \delta), |\Lambda|(m + \delta)] \right) \right),$$

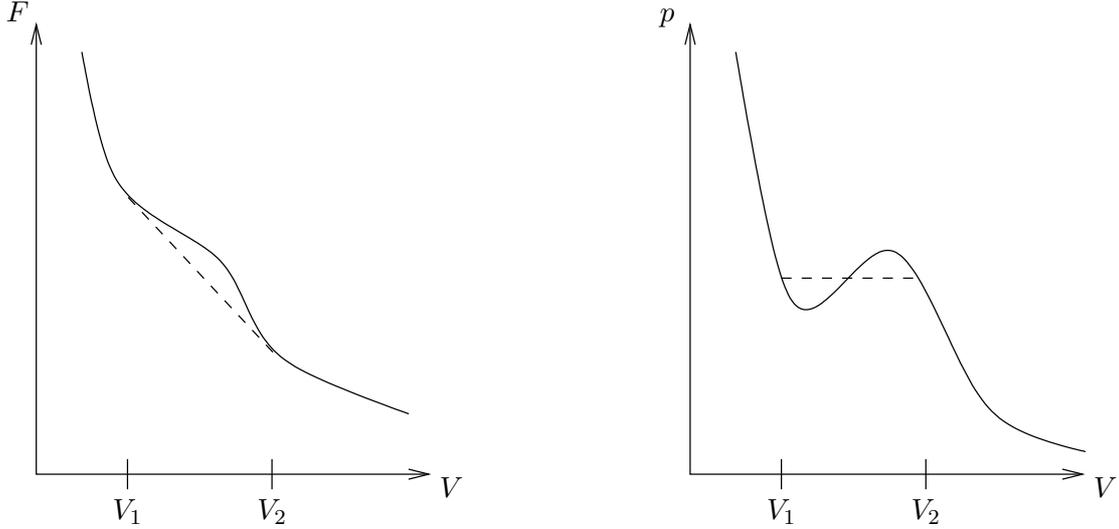
où m est la densité d'aimantation par spin. On peut montrer, avec un schéma de preuve semblable à celui partiellement exposé ci-dessus, qu'à β, Φ fixés $f(h)$ est concave, $\hat{f}(m)$ est convexe, et elles sont (aux signes près) transformées de Legendre l'une de l'autre,

$$f(h) = \inf_{m \in [-1, 1]} [\hat{f}(m) - h m], \quad \hat{f}(m) = \sup_h [f(h) + h m]. \quad (233)$$

6 Limite de Kac

6.1 Motivation - Construction de Maxwell

On a vu dans le chapitre précédente que l'énergie libre des modèles définis en dimension finie doit vérifier des propriétés de convexité. Les approximations (incontrôlées) de type champ moyen peuvent toutefois conduire à des énergies libres non-convexes. L'exemple physique le plus révélateur est celui du gaz de van der Waals, pour lequel à basse température on obtient une énergie libre $F(T, V, N)$ non-convexe dans le volume V , autrement dit des isothermes dans le plan (V, p) non-monotones, ce qui n'est pas acceptable physiquement. La résolution habituelle de ce problème consiste à effectuer la construction de Maxwell, autrement dit à remplacer $F(V)$ par son enveloppe convexe, ou, de manière équivalente, à remplacer la partie non-monotone de l'isotherme par un palier qui vérifie la loi des aires :



Justifions brièvement l'équivalence entre la prise de l'enveloppe convexe de $F(V)$ et la construction du palier de Maxwell sur $p(V)$. Rappelons que la pression et l'énergie libre sont reliées par $p = -\frac{\partial F}{\partial V}|_{T,N}$ d'après les relations générales de thermodynamique. Notons $[V_1, V_2]$ l'intervalle sur lequel $F(V)$ diffère de son enveloppe convexe, qui est une droite de pente notée $-p_0$. Par construction $F'(V_1) = F'(V_2) = -p_0$, la partie affine de l'enveloppe convexe étant tangente à F en ces deux points. De plus $p_0 = -\frac{F(V_2) - F(V_1)}{V_2 - V_1}$. Calculons maintenant l'aire entre les isothermes avant et après l'introduction du palier en p_0 :

$$(V_2 - V_1) p_0 - \int_{V_1}^{V_2} dV p(V) = (V_2 - V_1) p_0 + F(V_2) - F(V_1) = 0 . \quad (234)$$

La construction de Maxwell correspond à imposer l'égalité des pressions et des potentiels chimiques entre les deux phases en coexistence ; ces derniers se lisent comme l'ordonnée à l'origine de la tangente en un point de la courbe.

L'objectif de ce chapitre est de justifier rigoureusement la construction de Maxwell dans le cadre d'une limite bien contrôlée. On utilisera pour simplifier le langage d'un modèle de spins d'Ising sur réseau, une approche similaire peut être suivie pour les systèmes de particules.

Rappelons tout d'abord les résultats du TD 1 pour le modèle de champ moyen de Curie-Weiss, défini par un Hamiltonien agissant sur N spins d'Ising sans géométrie,

$$H(\underline{\sigma}) = -\frac{1}{2} \sum_{i,j=1}^N \sigma_i \sigma_j - h \sum_{i=1}^N \sigma_i . \quad (235)$$

On a montré que son énergie libre par spin vaut, dans la limite thermodynamique,

$$f_{\text{cm}}(\beta, h) = \inf_{m \in [-1,1]} [\widehat{f}_{\text{cm}}(\beta, m) - h m] , \quad (236)$$

$$\widehat{f}_{\text{cm}}(\beta, m) = -\frac{1}{2} m^2 - \frac{1}{\beta} s(m) , \quad s(m) = -\frac{1+m}{2} \ln \left(\frac{1+m}{2} \right) - \frac{1-m}{2} \ln \left(\frac{1-m}{2} \right) . \quad (237)$$

Insistons sur le fait que ce modèle ne rentre absolument pas dans le cadre des modèles de dimension finie traités dans les chapitres précédents ; en effet à basse température $\widehat{f}_{\text{cm}}(\beta, m)$ n'est pas convexe en m , ce qui serait contradictoire pour un modèle en dimension finie. L'objectif de la suite est d'étudier un système de dimension finie, qui dans une certaine limite conduise aux résultats de champ moyen corrigés par la construction de Maxwell.

6.2 Interaction de Kac

Considérons une fonction φ de \mathbb{R}^d dans \mathbb{R} , continue et à dérivées continues, à support inclus dans la boule unité de \mathbb{R}^d , non-négative. Elle est donc intégrable au sens de Riemann (une fonction définie sur un compact est Riemann intégrable ssi elle est bornée et presque partout continue), et on la suppose normalisée avec $\int dr \varphi(r) = 1$. On supposera aussi qu'elle est symétrique autour de l'origine, i.e. $\varphi(r) = \varphi(-r)$. Définissons l'interaction de Kac à l'échelle $R \in \mathbb{R}_+$, notée $\Phi(R)$, selon

$$\phi_X(\underline{\sigma}) = \begin{cases} -h \sigma_x & \text{si } X = \{x\} \\ -J_{x,y} \sigma_x \sigma_y & \text{si } X = \{x, y\} \\ 0 & \text{sinon} \end{cases}, \quad \text{avec } J_{x,y} = \frac{1}{R^d} \varphi\left(\frac{x-y}{R}\right) \quad (238)$$

Les spins à distance plus grande que R n'interagissent pas entre eux, tous les couplages pour des distances nettement plus petites que R sont par contre du même ordre. Pour $R \approx 1$ on a une interaction à proches voisins, pour $R \approx \text{diam}(\Lambda)$ on retrouve approximativement le modèle de Curie-Weiss. La limite de Kac, étudiée notamment par Lebowitz et Penrose en 1966, consiste à prendre $R \rightarrow \infty$, mais seulement après la limite thermodynamique, ce qui va permettre d'interpoler entre un modèle de dimension finie à portée finie et un modèle de champ moyen. Notons en effet que pour tout R fini l'interaction définie ci-dessus est à portée finie (φ étant de support borné), donc la limite thermodynamique de l'énergie libre existe, notons-la $f_{\text{Kac}}(\beta, h, R, d, \varphi) = \lim_{\Lambda \rightarrow \infty} f_{\Lambda}(\beta, \Phi(R))$.

Le choix de la normalisation des $J_{x,y}$ et de φ est tel que $\sum_{y \in \mathbb{Z}^d \setminus x} J_{x,y} \approx 1$ (quand R est grand), ce qui correspond à la même normalisation pour le couplage total d'un spin avec les autres spins que pour le modèle de Curie-Weiss.

6.3 Limite de Kac pour $f(h)$

Le premier résultat que l'on veut prouver est le suivant :

Théorème 16. *Dans la limite des grandes portées l'interaction de Kac a l'énergie libre de champ moyen :*

$$\lim_{R \rightarrow \infty} f_{\text{Kac}}(\beta, h, R, d, \varphi) = f_{\text{cm}}(\beta, h) . \quad (239)$$

Remarquons que dans la limite de Kac l'énergie libre devient indépendant des détails de φ (du moment que son intégrale est normalisée à 1), mais aussi, ce qui est a priori plus surprenant, de la dimension d de l'espace. En particulier pour $d = 1$ la limite $R \rightarrow \infty$ est nécessairement singulière : pour tout R fini l'interaction est à portée finie, et f_{Kac} est une fonction analytique de h (comme on l'a vu au TD 1 il n'y a pas de transition de phase en dimension 1), alors que sa limite f_{cm} est singulière en $h = 0$ (à basse température). Il existe aussi des résultats en $d \geq 2$ qui montrent l'existence d'une transition de phase pour R grand mais fini [2].

Démonstration. Le schéma de la preuve consistera à prouver des bornes supérieures et inférieures coïncidentes sur les \limsup et \liminf de f_{Kac} quand $R \rightarrow \infty$.

Une des inégalités vient de la méthode variationnelle vue au TD3. On a en effet, dans un volume fini Λ ,

$$H_{\Lambda}^{\Phi}(\underline{\sigma}) = -\frac{1}{2} \sum_{\substack{x,y \in \Lambda \\ x \neq y}} J_{x,y} \sigma_x \sigma_y - h \sum_{x \in \Lambda} \sigma_x . \quad (240)$$

En appliquant la méthode variationnelle du TD 3 avec une mesure produit de spins indépendants d'aimantation moyenne m arbitraire, il vient

$$f_\Lambda(\beta, \Phi) \leq -\frac{1}{2}m^2 \left(\frac{1}{|\Lambda|} \sum_{\substack{x,y \in \Lambda \\ x \neq y}} J_{x,y} \right) - hm - \frac{1}{\beta} s(m) . \quad (241)$$

En prenant la limite thermodynamique $\Lambda \rightarrow \infty$ avec R fixé, comme les effets de bords disparaissent il vient

$$f_{\text{Kac}}(\beta, h, R, d, \varphi) \leq -\frac{1}{2}J(R)m^2 - hm - \frac{1}{\beta} s(m) , \text{ avec } J(R) = \sum_{x \in \mathbb{Z}^d \setminus 0} J_{0,x} = \sum_{x \in \mathbb{Z}^d \setminus 0} \frac{1}{R^d} \varphi\left(\frac{x}{R}\right) . \quad (242)$$

Or quand $R \rightarrow \infty$ cette dernière somme de Riemann devient une intégrale,

$$J(R) \xrightarrow{R \rightarrow \infty} \int dr \varphi(r) = 1 , \quad (243)$$

on a donc, en notant que m était arbitraire jusque-là,

$$\limsup_{R \rightarrow \infty} f_{\text{Kac}}(\beta, h, R, d, \varphi) \leq \inf_{m \in [-1,1]} \left[-\frac{1}{2}m^2 - hm - \frac{1}{\beta} s(m) \right] = f_{\text{cm}}(\beta, h) . \quad (244)$$

Il faut maintenant trouver une borne inférieure pour l'énergie libre. L'idée dans la suite est que pour $R \gg 1$, tous les spins dans une zone de taille ℓ avec $1 \ll \ell \ll R$ interagissent entre eux, on peut donc voir le système total comme une juxtaposition de modèles de Curie-Weiss interagissant les uns avec les autres.

Plus formellement, fixons un entier positif ℓ avec $\ell < R$, et prenons pour le système $\Lambda \subset_f \mathbb{Z}^d$ un cube de côté L , avec pour simplifier L multiple de ℓ . Λ sera donc vu comme l'union de cubes C_α de côtés ℓ , l'indice α allant de 1 à $N = (\frac{L}{\ell})^d$. On notera $r_\alpha \in \mathbb{Z}^d$ la position du centre du cube C_α (si ℓ est pair on prendra la partie entière de ses coordonnées).

Pour une configuration $\underline{\sigma}_\Lambda$ des spins d'Ising dans Λ on note $m_\alpha(\underline{\sigma})$ l'aimantation par spin dans le cube α , i.e.

$$m_\alpha(\underline{\sigma}) = \frac{1}{\ell^d} \sum_{x \in C_\alpha} \sigma_x , \quad (245)$$

qui appartient à $\mathcal{M}_\ell = \{-1, -1 + \frac{2}{\ell^d}, \dots, 1 - \frac{2}{\ell^d}, 1\}$.

On va choisir dans la suite ℓ petit devant l'échelle R sur laquelle varient sensiblement les couplages $J_{x,y}$. Pour exploiter cette idée nous allons réécrire la fonction de partition

$$Z_\Lambda = \sum_{\underline{\sigma}_\Lambda} \exp \left[\beta \frac{1}{2} \sum_{\substack{x,y \in \Lambda \\ x \neq y}} J_{x,y} \sigma_x \sigma_y + \beta h \sum_{x \in \Lambda} \sigma_x \right] \quad (246)$$

en termes des aimantations de bloc m_α , au prix d'une erreur qui deviendra négligeable dans la limite de Kac. On a

$$\frac{1}{2} \sum_{\substack{x,y \in \Lambda \\ x \neq y}} J_{x,y} \sigma_x \sigma_y = \frac{1}{2} \sum_{\alpha=1}^N \sum_{x,y \in C_\alpha} J_{x,y} \sigma_x \sigma_y - \frac{1}{2} N \ell^d J_{0,0} + \frac{1}{2} \sum_{\substack{\alpha, \alpha'=1 \\ \alpha \neq \alpha'}}^N \sum_{x \in C_\alpha, y \in C_{\alpha'}} J_{x,y} \sigma_x \sigma_y , \quad (247)$$

où l'on a utilisé l'invariance par translation de J pour traiter à part les termes diagonaux. En remplaçant maintenant dans les $J_{x,y}$ les coordonnées de x, y par celles du centre du bloc auquel ils appartiennent,

$$\frac{1}{2} \sum_{\substack{x,y \in \Lambda \\ x \neq y}} J_{x,y} \sigma_x \sigma_y = \frac{1}{2} \ell^{2d} \sum_{\alpha, \alpha'=1}^N J_{r_\alpha, r_{\alpha'}} m_\alpha(\underline{\sigma}) m_{\alpha'}(\underline{\sigma}) + E , \quad (248)$$

où l'erreur E commise dans ce remplacement vaut

$$E = -\frac{1}{2}N\ell^d J_{0,0} + \frac{1}{2} \sum_{\alpha=1}^N \sum_{x,y \in C_\alpha} [J_{x,y} - J_{r_\alpha, r_\alpha}] \sigma_x \sigma_y + \frac{1}{2} \sum_{\substack{\alpha, \alpha'=1 \\ \alpha \neq \alpha'}}^N \sum_{x \in C_\alpha, y \in C_{\alpha'}} [J_{x,y} - J_{r_\alpha, r_{\alpha'}}] \sigma_x \sigma_y \quad (249)$$

Dans la première somme x et y sont au plus à distance ℓ du centre r_α du bloc, donc

$$|J_{x,y} - J_{r_\alpha, r_\alpha}| = \frac{1}{R^d} \left| \varphi \left(\frac{x-y}{R} \right) - \varphi(0) \right| \leq C \frac{1}{R^d} \frac{\ell}{R}, \quad (250)$$

où C est une constante qui ne dépend que de φ et de la dimension, que l'on pourrait expliciter en terme d'une borne supérieure sur les dérivées premières de φ (qui sont finies par les hypothèses faites). Dans la deuxième somme on peut écrire similairement

$$|J_{x,y} - J_{r_\alpha, r_{\alpha'}}| \leq C' \frac{1}{R^d} \frac{\ell}{R} \mathbb{I}(\|r_\alpha - r_{\alpha'}\| \leq 2R), \quad (251)$$

la dernière fonction indicatrice étant justifiée car le support de φ est inclus dans la boule unité (le facteur 2, non-optimal, assure qu'en cas contraire aucun des spins de C_α ne peut interagir avec un spin de $C_{\alpha'}$). On peut donc borner l'erreur selon

$$|E| \leq \frac{1}{2} L^d \frac{\varphi(0)}{R^d} + \frac{1}{2} \frac{L^d}{\ell^d} \ell^{2d} C \frac{1}{R^d} \frac{\ell}{R} + \frac{1}{2} \frac{L^d}{\ell^d} \ell^{2d} C'' \frac{1}{R^d} \frac{\ell}{R} \frac{R^d}{\ell^d}, \quad (252)$$

le dernier facteur comptant le nombre de blocs qui peuvent interagir avec un bloc donné. On écrira finalement

$$|E| \leq L^d b(\ell, R, \varphi, d), \quad \text{avec} \quad b(\ell, R, \varphi, d) = O\left(\frac{1}{R^d}\right) + O\left(\left(\frac{\ell}{R}\right)^{d+1}\right) + O\left(\frac{\ell}{R}\right). \quad (253)$$

Revenant maintenant au calcul de la fonction de partition,

$$Z_\Lambda \leq e^{L^d b(\ell, R, \varphi, d)} \sum_{\underline{\sigma}_\Lambda} \exp \left[\beta \frac{1}{2} \ell^{2d} \sum_{\alpha, \alpha'=1}^N J_{r_\alpha, r_{\alpha'}} m_\alpha(\underline{\sigma}) m_{\alpha'}(\underline{\sigma}) + \beta h \ell^d \sum_{\alpha=1}^N m_\alpha(\underline{\sigma}) \right]. \quad (254)$$

Comme dans cette dernière expression toute la dépendance dans les configurations passe par les aimantations des blocs m_α , la somme vaut

$$\sum_{\{m_\alpha \in \mathcal{M}_\ell\}_{\alpha=1}^N} Z(m_1, \dots, m_N), \quad (255)$$

$$\text{où} \quad Z(m_1, \dots, m_N) = \exp \left[\sum_{\alpha=1}^N \ell^d s_\ell(m_\alpha) + \beta \frac{1}{2} \ell^{2d} \sum_{\alpha, \alpha'=1}^N J_{r_\alpha, r_{\alpha'}} m_\alpha m_{\alpha'} + \beta h \ell^d \sum_{\alpha=1}^N m_\alpha \right], \quad (256)$$

où l'on a introduit l'entropie $s_\ell(m)$ pour $m \in \mathcal{M}_\ell$ selon

$$e^{\ell^d s_\ell(m)} = \binom{\ell^d}{\ell^d \frac{1+m}{2}}. \quad (257)$$

On va maintenant utiliser le fait que les interactions sont ferromagnétiques (i.e. que $J_{x,y} \geq 0$) et qu'ainsi les configurations des $\{m_\alpha\}$ avec m_α constante dans l'espace sont privilégiées. En effet,

$$\sum_{\alpha, \alpha'=1}^N J_{r_\alpha, r_{\alpha'}} m_\alpha m_{\alpha'} = -\frac{1}{2} \sum_{\alpha, \alpha'=1}^N J_{r_\alpha, r_{\alpha'}} (m_\alpha - m_{\alpha'})^2 + \sum_{\alpha=1}^N m_\alpha^2 \sum_{\alpha'=1}^N J_{r_\alpha, r_{\alpha'}} \leq \sum_{\alpha=1}^N m_\alpha^2 \sum_{\alpha'=1}^N J_{r_\alpha, r_{\alpha'}}, \quad (258)$$

où l'on a utilisé la symétrie $J_{x,y} = J_{y,x}$. En majorant la somme des $Z(m_1, \dots, m_N)$ par le produit du nombre de termes et d'une borne supérieure sur chacun des termes, il vient

$$\begin{aligned}
\sum_{\{m_\alpha \in \mathcal{M}_\ell\}_{\alpha=1}^N} Z(m_1, \dots, m_N) &\leq |\mathcal{M}_\ell|^N \sup_{m_1, \dots, m_N} Z(m_1, \dots, m_N) \\
&\leq |\mathcal{M}_\ell|^N \sup_{m_1, \dots, m_N} \exp \left[\sum_{\alpha=1}^N \ell^d s_\ell(m_\alpha) + \beta \frac{1}{2} \ell^{2d} \sum_{\alpha=1}^N m_\alpha^2 \sum_{\alpha'=1}^N J_{r_\alpha, r_{\alpha'}} + \beta h \ell^d \sum_{\alpha=1}^N m_\alpha \right] \\
&\leq |\mathcal{M}_\ell|^N \exp \left[N \sup_{m \in \mathcal{M}_\ell} \left\{ \ell^d s_\ell(m) + \beta \frac{1}{2} \ell^{2d} m^2 \sup_{\alpha} \left(\sum_{\alpha'=1}^N J_{r_\alpha, r_{\alpha'}} \right) + \beta h \ell^d m \right\} \right]
\end{aligned} \tag{259}$$

où l'on a utilisé l'équation (258) pour passer de la première à la deuxième ligne.

On peut passer maintenant à la limite thermodynamique $\Lambda \rightarrow \infty$, ici en prenant $L \rightarrow \infty$. On sait d'après les résultats généraux que la limite existe, on peut donc écrire

$$\lim_{\Lambda \rightarrow \infty} \frac{1}{L^d} \ln Z_\Lambda \leq b(\ell, R, \varphi, d) + \frac{1}{\ell^d} \ln(\ell^d + 1) + \sup_{m \in \mathcal{M}_\ell} \left\{ s_\ell(m) + \beta \frac{1}{2} m^2 \sum_{x \in \ell \mathbb{Z}^d} \ell^d J_{0,x} + \beta h m \right\}, \tag{260}$$

autrement dit

$$f_{\text{Kac}}(\beta, h, R, d, \varphi) \geq \inf_{m \in \mathcal{M}_\ell} \left\{ -\frac{1}{2} J(R, \ell) m^2 - h m - \frac{1}{\beta} s_\ell(m) \right\} - \frac{1}{\beta} b(\ell, R, \varphi, d) - \frac{1}{\beta} \frac{1}{\ell^d} \ln(\ell^d + 1) \tag{261}$$

avec

$$J(R, \ell) = \sum_{x \in \ell \mathbb{Z}^d} \ell^d J_{0,x} = \sum_{x \in \ell \mathbb{Z}^d} \frac{\ell^d}{R^d} \varphi \left(\frac{x}{R} \right) = \sum_{x \in \mathbb{Z}^d} \left(\frac{\ell}{R} \right)^d \varphi \left(\frac{\ell}{R} x \right) \tag{262}$$

Considérons maintenant la limite $R \rightarrow \infty$. On est libre de choisir dans la borne ci-dessus une dépendance arbitraire de ℓ par rapport à R , on va donc prendre $\ell(R) \rightarrow \infty$ quand $R \rightarrow \infty$, mais avec $\ell(R)/R \rightarrow 0$. On peut par exemple prendre $\ell(R) = \lfloor \sqrt{R} \rfloor$. Alors les deux derniers termes de l'équation (262) vont à 0 (d'après la caractérisation de b donnée en (253)), $J(R, \ell(R)) \rightarrow 1$, $s_\ell(m) \rightarrow s(m)$ (en utilisant le développement de Stirling des factorielles), et la minimisation sur \mathcal{M}_ℓ peut être remplacé par une minimisation sur $[-1, 1]$ (l'erreur ainsi faite est d'ordre $(\ln \ell)/\ell$ qui est bien négligeable). On obtient finalement

$$\liminf_{R \rightarrow \infty} f_{\text{Kac}}(\beta, h, R, d, \varphi) \geq f_{\text{cm}}(\beta, h), \tag{263}$$

qui combinée avec la borne (244) achève la preuve du théorème. \square

Notons que l'on peut exploiter de manière beaucoup plus poussée le passage ("coarse-graining" en anglais) des variables microscopiques σ_i aux variables "mésoscopiques" m_α . Dans la limite de Kac on peut en effet étudier les probabilités de profils d'aimantation mésoscopique $m(x)$, avec des coordonnées d'espace x désormais continues, qui sont données par une fonctionnelle d'énergie libre fait d'un terme local et d'une interaction de couplage φ . Ce genre d'études, largement développées dans le livre de Presutti, permet notamment de définir la tension de surface entre deux phases à une transition du premier ordre.

6.4 Limite de Kac pour $\hat{f}(m)$

Revenons finalement sur l'objectif initial du chapitre, à savoir la justification de la construction de Maxwell. Dans le cadre des systèmes de spin on définirait $\hat{f}_{\text{Kac}}(\beta, m, R, d, \varphi)$, l'énergie libre microcanonique pour la paire de variables conjuguées $m \leftrightarrow h$, en contraignant la fonction de partition Z_Λ aux configurations d'aimantation par spin m (à un δ près, pris à 0 après la limite thermodynamique). D'après les résultats généraux énoncés et partiellement prouvés au chapitre précédent :

- $f_{\text{Kac}}(\beta, h, R, d, \varphi)$ est une fonction concave de h (c'est facile à prouver en calculant sa dérivée seconde en volume fini)

- $\widehat{f}_{\text{Kac}}(\beta, m, R, d, \varphi)$ est une fonction convexe de m (on l'admet ici, même schéma que la preuve partielle donnée pour la concavité de l'entropie)
- $\widehat{f}_{\text{Kac}}(\beta, m, R, d, \varphi)$ et $f_{\text{Kac}}(\beta, h, R, d, \varphi)$ sont transformées de Legendre l'une de l'autre (à des signes près par rapport à la définition standard des transformées de Legendre) :

$$f_{\text{Kac}}(h) = \inf_{m \in [-1, 1]} [\widehat{f}_{\text{Kac}}(m) - h m] , \quad \widehat{f}_{\text{Kac}}(m) = \sup_{h \in \mathbb{R}} [f_{\text{Kac}}(h) + h m] , \quad (264)$$

en gardant implicite les autres arguments (égaux) dans ces deux fonctions.

Dans la limite de Kac $R \rightarrow \infty$ on a vu que f_{Kac} avait pour limite $f_{\text{cm}}(\beta, h)$. On va finalement montrer un théorème similaire sur les énergies libres fonctions de m , qui justifie rigoureusement la construction de Maxwell :

Théorème 17. *Dans la limite des grandes portées l'interaction de Kac a pour énergie libre fonction de l'aimantation l'enveloppe convexe de celle de champ moyen :*

$$\lim_{R \rightarrow \infty} \widehat{f}_{\text{Kac}}(\beta, m, R, d, \varphi) = E.C. \widehat{f}_{\text{cm}}(\beta, m) \quad (265)$$

Démonstration. D'après le résultat admis rappelé ci-dessus, $\widehat{f}_{\text{Kac}}(m, R) = \sup_{h \in \mathbb{R}} [f_{\text{Kac}}(h, R) + h m]$. Donc pour tout h , $\widehat{f}_{\text{Kac}}(m, R) \geq f_{\text{Kac}}(h, R) + h m$. En passant à la limite $R \rightarrow \infty$ et en utilisant le théorème précédent, il vient

$$\liminf_{R \rightarrow \infty} \widehat{f}_{\text{Kac}}(\beta, m, R, d, \varphi) \geq f_{\text{cm}}(\beta, h) + h m , \quad (266)$$

et comme c'est vrai quelque soit h ,

$$\liminf_{R \rightarrow \infty} \widehat{f}_{\text{Kac}}(\beta, m, R, d, \varphi) \geq \sup_{h \in \mathbb{R}} [f_{\text{cm}}(\beta, h) + h m] . \quad (267)$$

Par définition $f_{\text{cm}}(\beta, h) = \inf_m [f_{\text{cm}}(\beta, m) - h m]$, donc par les propriétés générales des transformées de Legendre vues au TD 4, $\sup_{h \in \mathbb{R}} [f_{\text{cm}}(\beta, h) + h m] = E.C. \widehat{f}_{\text{cm}}(\beta, m)$. On a donc prouvé pour l'instant

$$\liminf_{R \rightarrow \infty} \widehat{f}_{\text{Kac}}(\beta, m, R, d, \varphi) \geq E.C. \widehat{f}_{\text{cm}}(\beta, m) , \quad (268)$$

il reste donc à prouver une borne supérieure sur $\widehat{f}_{\text{Kac}}(\beta, m, R, d, \varphi)$ dans la limite $R \rightarrow \infty$. Pour cela on peut utiliser l'équation (242) pour écrire

$$f_{\text{Kac}}(\beta, h, R, d, \varphi) \leq -\frac{1}{2}J(R)m'^2 - h m' - \frac{1}{\beta}s(m') \quad \forall m' \in [-1, 1] . \quad (269)$$

Alors

$$\widehat{f}_{\text{Kac}}(\beta, m, R, d, \varphi) = \sup_h [f_{\text{Kac}}(\beta, h, R, d, \varphi) + h m] \leq \sup_h \left[-\frac{1}{2}J(R)m'^2 + h(m - m') - \frac{1}{\beta}s(m') \right] , \quad (270)$$

cette inégalité étant valable pour tout m' . En prenant $m' = m$ il vient

$$\widehat{f}_{\text{Kac}}(\beta, m, R, d, \varphi) \leq -\frac{1}{2}J(R)m^2 - \frac{1}{\beta}s(m) , \quad (271)$$

et en prenant la limite $R \rightarrow \infty$, pour laquelle $J(R) \rightarrow 1$, on obtient

$$\limsup_{R \rightarrow \infty} \widehat{f}_{\text{Kac}}(\beta, m, R, d, \varphi) \leq \widehat{f}_{\text{cm}}(\beta, m) . \quad (272)$$

En regroupant les deux inégalités il vient

$$E.C. \widehat{f}_{\text{cm}}(\beta, m) \leq \liminf_{R \rightarrow \infty} \widehat{f}_{\text{Kac}}(\beta, m, R, d, \varphi) \leq \limsup_{R \rightarrow \infty} \widehat{f}_{\text{Kac}}(\beta, m, R, d, \varphi) \leq \widehat{f}_{\text{cm}}(\beta, m) . \quad (273)$$

Or $\limsup_{R \rightarrow \infty} \widehat{f}_{\text{Kac}}(\beta, m, R, d, \varphi)$ est une fonctions convexe de m (en tant que \limsup de telles fonctions pour tout R finie), comprise entre \widehat{f}_{cm} et son enveloppe convexe, elle est donc nécessairement égale à cette dernière, ce qui conclut la preuve du théorème. \square

Références

- [1] B. Simon. *The statistical mechanics of lattice gases*, volume 1. Princeton University Press, 1993.
- [2] E. Presutti. *Scaling limits in statistical mechanics and microstructures in continuum mechanics*. Springer Science & Business Media, 2008.
- [3] T. Bodineau. Translation invariant Gibbs states for the Ising model. *Probability Theory and Related Fields*, **135**(2), 153–168 (2006).
- [4] M. Aizenman, H. Duminil-Copin, and V. Sidoravicius. Random Currents and Continuity of Ising Model's Spontaneous Magnetization. *Communications in Mathematical Physics*, **334**(2), 719–742 (2015).

TD 1 : Modèles unidimensionnels et de champ moyen

Modèles unidimensionnels

La chaîne d'Ising

On considère une chaîne de N spins d'Ising $\sigma_i = \pm 1$, $i \in [1, N]$, avec conditions aux bords périodiques $\sigma_{N+1} = \sigma_1$, en équilibre avec un thermostat. La probabilité d'une configuration est donc $e^{-\beta H}/Z$, où Z est la fonction de partition qui normalise cette distribution. Les moyennes selon cette loi de probabilité seront notées $\langle \bullet \rangle$. On prendra comme Hamiltonien :

$$H = -J \sum_{i=1}^N \sigma_i \sigma_{i+1} - h \sum_{i=1}^N \sigma_i, \quad (274)$$

où $J > 0$ correspond à un couplage ferromagnétique entre plus proches voisins, et h représente l'effet d'un champ magnétique extérieur.

1. Mettre la fonction de partition sous la forme :

$$Z = \sum_{\underline{\sigma}} \prod_{i=1}^N T(\sigma_i, \sigma_{i+1}), \quad (275)$$

avec T symétrique (i.e. $T(\sigma, \sigma') = T(\sigma', \sigma)$).

On notera \mathbb{T} la matrice d'ordre 2 telle que $\mathbb{T}_{\sigma\sigma'} = T(\sigma, \sigma')$.

2. Exprimer Z en fonction de \mathbb{T} .
3. Trouver les valeurs propres λ_{\pm} de \mathbb{T} (avec par convention $\lambda_+ > \lambda_-$).
4. Exprimer la densité d'énergie libre par spin $f = -\frac{1}{N\beta} \ln Z$ en fonction des λ_{\pm} , simplifier le résultat dans la limite thermodynamique $N \rightarrow \infty$.
5. On s'intéresse à l'aimantation moyenne par spin $m = \langle \sigma_i \rangle$. Exprimer m en fonction de \mathbb{T} et de la matrice $\hat{\sigma}$:

$$\hat{\sigma} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (276)$$

6. Trouver une autre expression de m en fonction d'une dérivée de f , montrer que dans la limite thermodynamique :

$$m = \frac{\text{sh}(\beta h)}{\sqrt{\text{sh}^2(\beta h) + e^{-4\beta J}}}. \quad (277)$$

Tracer l'allure de cette courbe en fonction de h à différentes températures. Y a-t-il une transition de phase dans ce modèle ?

La chaîne d'Ising avec interactions à seconds voisins

On considère maintenant le modèle d'Ising unidimensionnel avec des interactions à premiers et seconds voisins :

$$H = -J_1 \sum_{i=1}^N \sigma_i \sigma_{i+1} - J_2 \sum_{i=1}^N \sigma_i \sigma_{i+2}, \quad (278)$$

et on étend la condition aux limites périodiques à $\sigma_{N+2} = \sigma_2$. On supposera N pair.

7. Mettre la fonction de partition sous la forme :

$$Z = \sum_{\underline{\sigma}} \prod_{i=0}^{(N/2)-1} T((\sigma_{2i+1}, \sigma_{2i+2}), (\sigma_{2i+3}, \sigma_{2i+4})). \quad (279)$$

8. Donner une matrice \mathbb{T} d'ordre 4 telle que $Z = \text{Tr } \mathbb{T}^{(N/2)}$.

On admettra les deux résultats suivants :

- Le théorème de Perron-Frobenius affirme qu'une matrice réelle M dont tous les éléments sont strictement positifs admet une valeur propre λ_0 non-dégénérée et strictement positive, et telle que le module de toutes les autres valeurs propres est strictement inférieur à λ_0 .
- Soit $P(\lambda, \beta)$ un polynôme unitaire en λ dont les coefficients sont des fonctions analytiques de β . Si λ_0 est une racine simple de $P(\cdot, \beta_0)$, alors il existe une fonction $\lambda(\beta)$, analytique dans un voisinage de β_0 avec $\lambda(\beta_0) = \lambda_0$, telle que $P(\lambda(\beta), \beta) = 0$.

9. En utilisant ces deux résultats, montrer qu'il n'y a pas de transition de phase dans ce modèle, et plus généralement dans tout modèle unidimensionnel à variables discrètes et à portée d'interaction finie.

Le modèle de Curie-Weiss

On considère un système de N spins d'Ising, $\underline{\sigma} = (\sigma_1, \dots, \sigma_N) \in \{-1, +1\}^N$, interagissant selon l'Hamiltonien

$$H(\underline{\sigma}) = -\frac{J}{2N} \sum_{i,j=1}^N \sigma_i \sigma_j - h \sum_{i=1}^N \sigma_i, \quad (280)$$

avec $J > 0$. Ce modèle est dit de champ moyen car chaque degré de liberté interagit avec tous les autres, il n'y a donc pas de structure géométrique sous-jacente.

1. Pourquoi a-t-on divisé la constante de couplage J par le facteur $1/N$?
2. On suppose le système à l'équilibre avec un thermostat à la température T . En remarquant que H ne dépend de $\underline{\sigma}$ que par l'intermédiaire de l'aimantation $\sum_i \sigma_i$, mettre la fonction de partition sous la forme :

$$Z = \sum_{m \in \mathcal{M}_N} \mathcal{N}_m^N e^{-\beta N [-\frac{J}{2} m^2 - hm]}. \quad (281)$$

On explicitera l'ensemble \mathcal{M}_N dans lequel varie m , ainsi que l'expression de \mathcal{N}_m^N .

3. On définit l'énergie libre par spin dans la limite thermodynamique selon

$$f(\beta, h) = \lim_{N \rightarrow \infty} -\frac{1}{\beta N} \ln Z. \quad (282)$$

On rappelle que

$$\frac{1}{N} \ln \binom{N}{\alpha N} = -\alpha \ln \alpha - (1 - \alpha) \ln(1 - \alpha) + O\left(\frac{\ln N}{N}\right) \quad (283)$$

dans la limite $N \rightarrow \infty$. En déduire que

$$f(\beta, h) = \inf_{m \in [-1, 1]} \widehat{f}(m; \beta, h), \quad (284)$$

où l'on explicitera la fonction $\widehat{f}(m; \beta, h)$.

4. Tracer l'allure de $\widehat{f}(m; \beta, h = 0)$ en fonction de m , pour différentes températures. Quelle est la valeur de la température critique T_c où leur comportement change qualitativement ?
5. Quel est l'effet d'un champ $h > 0$ sur ces courbes ?
6. On définit $m_*(\beta, h)$ comme le point où le minimum est atteint dans l'équation (284). Donner l'équation implicite vérifiée par cette quantité.
7. Tracer l'allure de $m_*(\beta, h)$ en fonction de h pour deux températures, supérieure et inférieure à T_c .
8. On définit l'aimantation spontanée $m_{\text{sp}}(\beta) = \lim_{h \rightarrow 0^+} m_*(\beta, h)$. Tracer son allure en fonction de la température. Calculer l'exposant β_{cm} qui régit le comportement de m_{sp} au voisinage de T_c , i.e. tel que $m_{\text{sp}}(T) \propto (T_c - T)^{\beta_{\text{cm}}}$, où \propto désigne un équivalent à une constante multiplicative près.

Corrigé du TD 1

Modèles unidimensionnels

La méthode présentée ici, dite de la matrice de transfert, permet de résoudre très efficacement tous les modèles unidimensionnels. D'un point de vue computationnel elle effectue la somme sur le nombre exponentiel de configurations des poids de Boltzmann en un nombre polynomial d'opérations, en organisant de manière convenable le calcul.

La chaîne d'Ising

1.

$$Z = \sum_{\underline{\sigma}} e^{\beta J \sum_{i=1}^N \sigma_i \sigma_{i+1} + \beta h \sum_{i=1}^N \frac{\sigma_i + \sigma_{i+1}}{2}} = \sum_{\underline{\sigma}} \prod_{i=1}^N T(\sigma_i, \sigma_{i+1}), \quad \text{avec } T(\sigma, \sigma') = e^{\beta J \sigma \sigma' + \beta h \frac{\sigma + \sigma'}{2}}.$$

En introduisant une notation matricielle, où la première ligne (resp. colonne) correspond à $\sigma = +1$ (resp. $\sigma' = +1$), la deuxième à -1 , on a

$$\mathbb{T} = \begin{pmatrix} e^{\beta J + \beta h} & e^{-\beta J} \\ e^{-\beta J} & e^{\beta J - \beta h} \end{pmatrix}.$$

2. On reconnaît les règles de multiplication matricielle dans la sommation sur les configurations de Z ,

$$\begin{aligned} Z &= \sum_{\sigma_1, \dots, \sigma_N} \mathbb{T}_{\sigma_1, \sigma_2} \mathbb{T}_{\sigma_2, \sigma_3} \cdots \mathbb{T}_{\sigma_N, \sigma_1} = \sum_{\sigma_1, \sigma_3, \dots, \sigma_N} (\mathbb{T}^2)_{\sigma_1, \sigma_3} \mathbb{T}_{\sigma_3, \sigma_4} \cdots \mathbb{T}_{\sigma_N, \sigma_1} \\ &= \sum_{\sigma_1} (\mathbb{T}^N)_{\sigma_1, \sigma_1} = \text{Tr}(\mathbb{T}^N). \end{aligned}$$

3. En résolvant l'équation quadratique $\lambda^2 - \text{Tr}(\mathbb{T})\lambda + \det(\mathbb{T}) = 0$ on obtient

$$\lambda_{\pm} = e^{\beta J} \cosh(\beta h) \pm \sqrt{e^{2\beta J} \sinh^2(\beta h) + e^{-2\beta J}}.$$

4.

$$f = -\frac{1}{N\beta} \ln Z = -\frac{1}{N\beta} \ln(\lambda_+^N + \lambda_-^N) \xrightarrow{N \rightarrow \infty} -\frac{1}{\beta} \ln(\lambda_+),$$

puisque $\lambda_+ > \lambda_-$.

5.

$$\langle \sigma_i \rangle = \frac{1}{Z} \sum_{\sigma_1, \dots, \sigma_N} \mathbb{T}_{\sigma_1, \sigma_2} \cdots \mathbb{T}_{\sigma_{i-1}, \sigma_i} \sigma_i \mathbb{T}_{\sigma_i, \sigma_{i+1}} \cdots \mathbb{T}_{\sigma_N, \sigma_1} = \frac{1}{Z} \text{Tr}(\mathbb{T}^i \hat{\sigma} \mathbb{T}^{N-i}) = \frac{\text{Tr}(\hat{\sigma} \mathbb{T}^N)}{\text{Tr}(\mathbb{T}^N)}.$$

L'indépendance en i vient de l'invariance par translation due aux conditions aux bords périodiques, qui se traduisent ici par l'invariance cyclique de la trace. On pourrait aussi exprimer de manière similaire des fonctions de corrélation, par exemple $\langle \sigma_i \sigma_{i+r} \rangle = \frac{\text{Tr}(\hat{\sigma} \mathbb{T}^r \hat{\sigma} \mathbb{T}^{N-r})}{\text{Tr}(\mathbb{T}^N)}$.

6. On a aussi

$$m = -\frac{\partial f}{\partial h} = \frac{1}{\beta} \frac{(\partial_h \lambda_+) \lambda_+^{N-1} + (\partial_h \lambda_-) \lambda_-^{N-1}}{\lambda_+^N + \lambda_-^N} \xrightarrow{N \rightarrow \infty} \frac{1}{\beta} \frac{\partial_h \lambda_+}{\lambda_+} = \frac{1}{\beta} \frac{\partial}{\partial h} \ln \lambda_+.$$

On constate au passage que la dérivation par rapport à h et la limite $N \rightarrow \infty$ commutent ici ; c'est loin d'être toujours le cas, et c'est une des manifestations de l'absence de transition de phase dans ce modèle. A partir de cette dernière expression de m , et en simplifiant les expressions, on arrive à l'expression de l'énoncé. C'est une fonction impaire, croissante, saturant à ± 1 pour $h \rightarrow \pm \infty$.

Quand la température diminue elle est de plus en plus abrupte, mais elle reste régulière à toute température positive ; en particulier la limite $m(h \rightarrow 0^+)$, qui définit l'aimantation spontanée, est nulle à toute température, il n'y a pas de transition de phase. La pente en 0, qui est la susceptibilité magnétique, vaut $\beta e^{2\beta J}$, elle diverge (exponentiellement) seulement à température nulle.

La chaîne d'Ising avec interactions à seconds voisins

7. A cause des interactions à seconds voisins on ne peut plus écrire le poids de Boltzmann comme un produit entre des facteurs faisant intervenir seulement des spins plus proches voisins ; pour contourner cette difficulté on considère des blocs formés de 2 spins adjacents, de sorte que les interactions sont seulement entre blocs plus proches voisins. Il faut faire attention à éviter le double comptage d'interactions ; il n'y a pas unicité de la fonction T , un choix possible est

$$T((\sigma_1, \sigma_2), (\sigma_3, \sigma_4)) = e^{\beta J_1(\frac{\sigma_1\sigma_2}{2} + \sigma_2\sigma_3 + \frac{\sigma_3\sigma_4}{2}) + \beta J_2(\sigma_1\sigma_3 + \sigma_2\sigma_4)} .$$

8. En prenant pour l'ordre des lignes et des colonnes les configurations $(+, +), (+, -), (-, +), (-, -)$, il vient pour la matrice \mathbb{T} :

$$\mathbb{T} = \begin{pmatrix} e^{2\beta J_1 + 2\beta J_2} & e^{\beta J_1} & e^{-\beta J_1} & e^{-2\beta J_2} \\ e^{-\beta J_1} & e^{-2\beta J_1 + 2\beta J_2} & e^{-2\beta J_2} & e^{\beta J_1} \\ e^{\beta J_1} & e^{-2\beta J_2} & e^{-2\beta J_1 + 2\beta J_2} & e^{-\beta J_1} \\ e^{-2\beta J_2} & e^{-\beta J_1} & e^{\beta J_1} & e^{2\beta J_1 + 2\beta J_2} \end{pmatrix} .$$

9. On peut se convaincre que pour tout modèle unidimensionnel à variables discrètes et à portée d'interaction finie on peut construire une matrice \mathbb{T} de taille indépendante de N , telle que la fonction de partition s'écrive $Z = \text{Tr}(\mathbb{T}^{(N/p)})$, où p est indépendant de N . Les éléments de matrice de \mathbb{T} sont des poids de Boltzmann $e^{-\beta H}$, donc strictement positifs, d'après le théorème de Perron-Frobenius le rayon spectral λ_0 est donc strictement positif et correspond à une valeur propre non-dégénérée. Ceci donne pour la limite thermodynamique de l'énergie libre $f = -\ln(\lambda_0)/(\beta p)$. De plus les éléments de matrice sont analytique en β , donc aussi les coefficients du polynôme caractéristique dont λ_0 est racine ; par Perron-Frobenius cette racine est simple pour tout β , d'après l'autre résultat rappelé dans l'énoncé (qui est une forme du théorème des fonctions implicites) λ_0 est donc analytique en β . Comme elle est strictement positive le point de branchement en 0 du logarithme n'est jamais atteint, donc f est analytique en β , il n'y a donc pas de transition de phase dans toute cette classe de modèles.

Le modèle de Curie-Weiss

1. Cette normalisation permet d'avoir un Hamiltonien extensif, puisque la première somme sur les paires de sites contient N^2 termes. Comme chaque spin interagit avec tous les autres il faut que cette interaction soit faible pour avoir une limite thermodynamique sensée.
2. En notant $m(\underline{\sigma}) = \frac{1}{N} \sum_{i=1}^N \sigma_i$ l'aimantation par spin d'une configuration, on voit que

$$Z = \sum_{\underline{\sigma}} e^{-\beta N(-\frac{J}{2}m(\underline{\sigma})^2 - hm(\underline{\sigma}))} .$$

Le fait que l'énergie ne dépende de la configuration microscopique que par l'intermédiaire de son aimantation est une situation très particulière, due au caractère champ moyen du modèle : en dimension finie on peut construire des configurations de même aimantation moyenne avec des énergies très différentes. On note maintenant que $m(\underline{\sigma}) \in \mathcal{M}_N = \{-1, -1 + \frac{2}{N}, \dots, 1 - \frac{2}{N}, 1\}$; en comptant le nombre de spins $+$ et $-$ dans une configuration d'aimantation par spin $m \in \mathcal{M}_N$ on se rend compte que le facteur combinatoire \mathcal{N}_m^N qui compte le nombre de configurations avec cette valeur de m est

$$\mathcal{N}_m^N = \binom{N}{N\frac{1+m}{2}} ,$$

correspondant au choix des indices des spins + (ou -).

3. On doit calculer

$$f(\beta, h) = \lim_{N \rightarrow \infty} -\frac{1}{\beta N} \ln \left(\sum_{m \in \mathcal{M}_N} \binom{N}{N \frac{1+m}{2}} e^{-\beta N (-\frac{J}{2} m^2 - hm)} \right). \quad (285)$$

Dans cette limite la somme est sur un nombre linéaire (en N) de termes, chacun de ces termes étant exponentiels. On peut donc appliquer la méthode de Laplace, qui à cet ordre-là consiste à ne garder que l'ordre exponentiel du terme dominant,

$$\begin{aligned} f(\beta, h) &= \inf_{m \in [-1, 1]} \widehat{f}(m; \beta, h) \quad \text{avec} \\ \widehat{f}(m; \beta, h) &= e(m) - \frac{1}{\beta} s(m) \\ e(m) &= -\frac{J}{2} m^2 - hm, \\ s(m) &= -\frac{1+m}{2} \ln \left(\frac{1+m}{2} \right) - \frac{1-m}{2} \ln \left(\frac{1-m}{2} \right). \end{aligned}$$

On reconnaît un terme énergétique et entropique dans cette énergie libre, ce dernier prenant la forme d'une entropie de Shannon.

4. La fonction $s(m)$ est positive, paire, concave, s'annule en $m = \pm 1$ avec des tangentes infinies, et est maximale en $m = 0$ où elle vaut $\ln 2$. La contribution énergétique, à $h = 0$, est une parabole maximale en 0. Selon la température il y a une compétition entre les effets énergétiques (qui poussent à $m \approx \pm 1$ pour minimiser l'énergie) et entropiques (qui favorisent $m \approx 0$ qui correspond au plus grand nombre de configurations). A haute température $\widehat{f}(m; \beta, h = 0)$ est convexe avec un unique minimum en $m = 0$. A basse température $m = 0$ devient un maximum local, $\widehat{f}(m; \beta, h = 0)$ a deux minima locaux en $\pm m_*$ avec $m_* > 0$, c'est l'aimantation spontanée. La température critique séparant ces deux régimes correspond à l'annulation de la dérivée seconde de $\widehat{f}(m)$ en $m = 0$. Après un petit calcul on trouve $T_c = J$.
5. L'ajout d'un champ magnétique induit un terme supplémentaire linéaire dans l'énergie libre. A basse température un champ infinitésimal brise la dégénérescence entre les deux minima locaux ; pour des valeurs plus grandes du champ il n'existe plus qu'un seul minimum local (cette transition est dite spinodale).
6. En dérivant l'expression de \widehat{f} , et en utilisant l'identité $\frac{1}{2} \ln \left(\frac{1+x}{1-x} \right) = \operatorname{arctanh} x$, on obtient

$$m_*(\beta, h) \quad \text{solution de} \quad m = \tanh(\beta(Jm + h)). \quad (286)$$

7. En tant que fonction de h m_* est croissante, impaire, et sature à ± 1 pour $h \rightarrow \pm \infty$. A haute température ($T > T_c$) $m_*(h)$ est continue, donc elle s'annule pour $h = 0$. Au contraire à basse température il y a une aimantation spontanée, les limites en $h \rightarrow 0^\pm$ sont opposées et différentes de zéro.
8. L'aimantation spontanée augmente quand la température diminue, continûment à partir de $T = T_c$, jusqu'à atteindre 1 à température nulle. Pour déterminer l'exposant critique on développe l'équation implicite (286) au voisinage de la température critique, en posant $\beta J = 1 + \epsilon$. Au plus bas ordre non-trivial en m et ϵ il vient

$$m = m + \epsilon m - \frac{1}{3} m^3 + \dots, \quad \text{soit} \quad m = \sqrt{3\epsilon}. \quad (287)$$

L'exposant de champ moyen β_{cm} vaut donc $1/2$.

TD 2 : Entropie et principe variationnel

Entropie de Shannon

On considère un ensemble dénombrable d'évènements élémentaires, et l'ensemble des lois de probabilités p qui donnent une probabilité non-nulle à un nombre fini d'évènements. Pour simplifier on notera $p = (p_1, \dots, p_M)$, avec les $p_i \in [0, 1]$ et $\sum_{i=1}^M p_i = 1$.

En 1948 Shannon a introduit une fonction entropie $S(p)$, qui quantifie le "manque d'information a priori" sur un tirage d'un évènement aléatoire avec la loi p . Il a imposé les conditions suivantes sur S :

(P1) la fonction S est une fonction continue des probabilités p_1, \dots, p_M

(P2) $S(p_1, \dots, p_M) \leq S(1/M, \dots, 1/M)$

(P3) $S(p_1, \dots, p_M, 0) = S(p_1, \dots, p_M)$

(P4) propriété d'additivité : soient A l'évènement global correspondant aux évènements de 1 à m de probabilité $q_A = \sum_{i=1}^m p_i$ et B l'évènement global correspondant aux évènements de $m+1$ à M de probabilité $q_B = \sum_{i=m+1}^M p_i$. Alors on a :

$$S(p_1, \dots, p_M) = S(q_A, q_B) + q_A S\left(\frac{p_1}{q_A}, \dots, \frac{p_m}{q_A}\right) + q_B S\left(\frac{p_{m+1}}{q_B}, \dots, \frac{p_M}{q_B}\right)$$

On va montrer que l'expression $S(p_1, \dots, p_M) = -\sum_{i=1}^M p_i \ln p_i$ est, à une constante multiplicative près, la seule fonction vérifiant ces propriétés. On utilise la convention $0 \ln 0 = 0$. On notera $\sigma(M) = S(1/M, \dots, 1/M)$ l'entropie du cas équiprobable.

1. Montrer que $\sigma(M)$ est une fonction croissante de M .
2. Montrer que $S(1) = 0$. Commenter.
3. Généraliser la quatrième propriété pour montrer que

$$S(p_1, \dots, p_{m_1}, \dots, p_{m_2}, \dots, p_{m_\alpha}) = S(q_1, \dots, q_\alpha) + \sum_{i=1}^{\alpha} q_i S\left(\frac{p_{m_{i-1}+1}}{q_i}, \dots, \frac{p_{m_i}}{q_i}\right)$$

avec $q_i = p_{m_{i-1}+1} + \dots + p_{m_i}$ ($m_0 = 0$).

4. En déduire que $\sigma(MN) = \sigma(M) + \sigma(N)$.
5. En déduire que $\sigma(M) = k \ln M$, où k est une constante. Quel doit être son signe ?
6. En déduire que, pour $(p_1, \dots, p_M) \in \mathbb{Q}^M$, $S(p_1, \dots, p_M) = -k \sum_{i=1}^M p_i \ln p_i$.
7. En déduire l'expression de l'entropie de Shannon pour des probabilités quelconques.

L'entropie de Shannon joue un rôle fondamental en physique statistique, en probabilité (en particulier pour la théorie des grandes déviations) et en théorie de l'information. Dans ce dernier contexte elle apparaît notamment dans les théorèmes de Shannon, qui montrent (i) que le taux de compression possible sans perte d'information d'une suite de symboles générés par une source aléatoire en un fichier binaire est précisément l'entropie de la source, (ii) et que la vitesse de communication au travers d'un canal bruité est bornée par l'information mutuelle (une notion reliée à l'entropie que l'on va définir ci-dessous) entre l'entrée et la sortie du canal. Pour ces aspects de théorie de l'information on pourra consulter : *Elements of Information Theory* de T. Cover et J. Thomas, ainsi que *Information Theory, Inference, and Learning Algorithms* de D. MacKay, disponible gratuitement sur le site <http://www.inference.phy.cam.ac.uk/mackay/itila/book.html>

Propriétés de l'entropie

Pour faire le lien avec le cours on note à partir de maintenant $\underline{\sigma} \in \Sigma = \chi^\Delta$ les configurations dans une partie Δ finie de \mathbb{Z}^d , et μ une loi de probabilité sur ces configurations, donnée de manière élémentaire par les $\mu(\underline{\sigma}) \in [0, 1]$ avec $\sum_{\underline{\sigma}} \mu(\underline{\sigma}) = 1$. On note alors l'entropie de cette loi

$$S(\mu) = - \sum_{\underline{\sigma}} \mu(\underline{\sigma}) \ln \mu(\underline{\sigma}) . \quad (288)$$

1. Montrer que S est concave, i.e. $S(\alpha\mu + (1 - \alpha)\nu) \geq \alpha S(\mu) + (1 - \alpha)S(\nu)$ si μ, ν sont deux lois de probabilité sur le même espace de configurations, et $\alpha \in [0, 1]$. *Indication* : montrer d'abord que $g(x) = -x \ln x$ est une fonction concave sur $[0, 1]$.
2. Montrer que S est "presque convexe", dans le sens que

$$S(\alpha\mu + (1 - \alpha)\nu) \leq \alpha S(\mu) + (1 - \alpha)S(\nu) - \alpha \ln \alpha - (1 - \alpha) \ln(1 - \alpha) . \quad (289)$$

Indication : montrer que g est sous-additive, i.e. $g(a + b) \leq g(a) + g(b)$, et que $g(uv) = ug(v) + vg(u)$.

3. On introduit la "distance" de Küllback-Leibler entre deux mesures de probabilité par

$$D(\mu||\nu) = \sum_{\underline{\sigma}} \mu(\underline{\sigma}) \ln \left(\frac{\mu(\underline{\sigma})}{\nu(\underline{\sigma})} \right) , \quad (290)$$

qui n'est pas une distance puisqu'elle n'est pas symétrique. Montrer que $D(\mu||\nu) \geq 0$. Dans quel cas a-t-on $D(\mu||\nu) = 0$?

4. Montrer que $S(\mu) \in [0, |\Delta| \ln |\chi|]$. Dans quels cas ces bornes sont-elles atteintes ?
5. On considère un sous-ensemble de variables $\Lambda \subset \Delta$ et son complémentaire $\Lambda^c = \Delta \setminus \Lambda$, et les lois marginales de μ

$$\mu_\Lambda(\underline{\sigma}_\Lambda) = \sum_{\underline{\sigma}_{\Lambda^c}} \mu((\underline{\sigma}_\Lambda, \underline{\sigma}_{\Lambda^c})) , \quad \mu_{\Lambda^c}(\underline{\sigma}_{\Lambda^c}) = \sum_{\underline{\sigma}_\Lambda} \mu((\underline{\sigma}_\Lambda, \underline{\sigma}_{\Lambda^c})) . \quad (291)$$

On note $S_\Lambda(\mu)$ et $S_{\Lambda^c}(\mu)$ les entropies de ces deux lois. Montrer que

$$S(\mu) \leq S_\Lambda(\mu) + S_{\Lambda^c}(\mu) . \quad (292)$$

Dans quel cas la borne est-elle atteinte ? On définit l'information mutuelle entre Λ et Λ^c comme $I_{\Lambda, \Lambda^c}(\mu) = S_\Lambda(\mu) + S_{\Lambda^c}(\mu) - S(\mu) = I_{\Lambda^c, \Lambda}(\mu)$, qui est donc positive d'après l'inégalité ci-dessus.

6. La loi conditionnelle dans Λ sachant Λ^c est définie par $\mu(\underline{\sigma}_\Lambda | \underline{\sigma}_{\Lambda^c}) = \frac{\mu((\underline{\sigma}_\Lambda, \underline{\sigma}_{\Lambda^c}))}{\mu_{\Lambda^c}(\underline{\sigma}_{\Lambda^c})}$, et l'entropie conditionnelle selon

$$S_{\Lambda|\Lambda^c}(\mu) = \sum_{\underline{\sigma}_{\Lambda^c}} \mu_{\Lambda^c}(\underline{\sigma}_{\Lambda^c}) \left[- \sum_{\underline{\sigma}_\Lambda} \mu(\underline{\sigma}_\Lambda | \underline{\sigma}_{\Lambda^c}) \ln \mu(\underline{\sigma}_\Lambda | \underline{\sigma}_{\Lambda^c}) \right] . \quad (293)$$

Interpréter cette définition, exprimer l'entropie conditionnelle à l'aide des entropies marginales et de l'information mutuelle, et justifier le moyen mnémotechnique "le conditionnement réduit l'entropie".

Principe variationnel de Gibbs

Toujours dans le cadre d'un espace de configurations fini on considère maintenant un Hamiltonien $H(\underline{\sigma})$ et on définit l'énergie libre $F_c(\beta, H)$ et la loi de probabilité de l'ensemble canonique $\mu_c(\beta, H)$ selon

$$F_c(\beta, H) = -\frac{1}{\beta} \ln Z(\beta, H) , \quad Z(\beta, H) = \sum_{\underline{\sigma}} e^{-\beta H(\underline{\sigma})} , \quad \mu_c(\beta, H, \underline{\sigma}) = \frac{e^{-\beta H(\underline{\sigma})}}{Z(\beta, H)} . \quad (294)$$

On définit par ailleurs l'énergie libre de Gibbs $F_{\text{Gibbs}}(\beta, H, \mu)$ qui associe un réel à la donnée d'une température inverse β , d'un Hamiltonien H et d'une mesure de probabilité μ :

$$F_{\text{Gibbs}}(\beta, H, \mu) = \sum_{\underline{\sigma}} \mu(\underline{\sigma}) H(\underline{\sigma}) - \frac{1}{\beta} S(\mu) . \quad (295)$$

On notera \mathcal{M} l'ensemble des mesures de probabilités sur cet espace de configurations.

1. Montrer que $F_{\text{Gibbs}}(\beta, H, \mu) \geq F_c(\beta, H)$ pour tout $\mu \in \mathcal{M}$, et que $F_{\text{Gibbs}}(\beta, H, \mu_c(\beta, H)) = F_c(\beta, H)$. On a donc la caractérisation variationnelle de l'énergie libre et de la loi canonique,

$$F_c(\beta, H) = \min_{\mu \in \mathcal{M}} F_{\text{Gibbs}}(\beta, H, \mu) , \quad \mu_c(\beta, H) = \operatorname{argmin}_{\mu \in \mathcal{M}} F_{\text{Gibbs}}(\beta, H, \mu) . \quad (296)$$

2. Montrer que la loi canonique est celle qui maximise l'entropie $S(\mu)$ lorsque l'on contraint l'énergie moyenne $\sum_{\underline{\sigma}} \mu(\underline{\sigma}) H(\underline{\sigma})$ à prendre une valeur donnée. Quelle est l'interprétation de la température dans ce calcul ?
3. Montrer que pour une loi de probabilité arbitraire $\mu \in \mathcal{M}$ on a l'expression variationnelle suivante de l'entropie :

$$S(\mu) = \min_H \left[\beta \sum_{\underline{\sigma}} \mu(\underline{\sigma}) H(\underline{\sigma}) - \beta F_c(\beta, H) \right] . \quad (297)$$

On pourra commencer par supposer pour simplifier que $\mu(\underline{\sigma}) > 0$ pour toutes les configurations.

4. On considère un modèle d'Ising sur une portion rectangulaire Δ de \mathbb{Z}^d , avec des conditions aux bords périodiques, interagissant selon l'Hamiltonien

$$H(\underline{\sigma}) = -J \sum_{\langle ij \rangle} \sigma_i \sigma_j - h \sum_{i \in \Delta} \sigma_i , \quad (298)$$

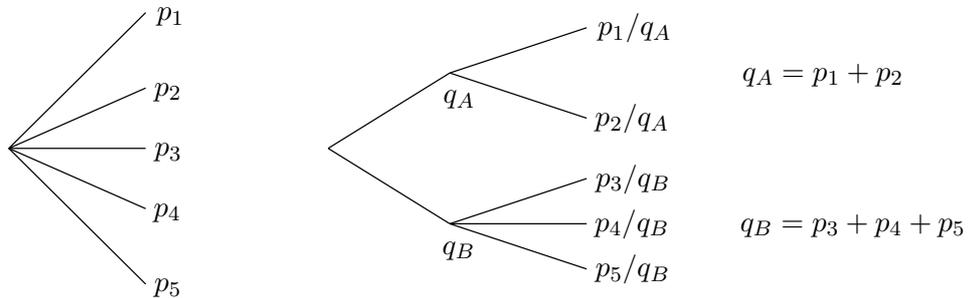
où la première somme porte sur les liens de Δ . Donner une borne supérieure à l'énergie libre par spin $F_c(\beta, H)/|\Delta|$, en considérant pour μ une mesure produit de spins indépendants.

Corrigé du TD 2

Entropie de Shannon

Commentons les hypothèses faites sur S :

- (P1) La continuité est “naturelle”, changer infinitésimalement une probabilité ne peut provoquer de grand changement dans le manque d’information.
- (P2) Le manque d’information est maximal si les évènements susceptibles de se produire sont équiprobables, il n’y a pas de stratégie possible pour deviner l’évènement qui va se produire.
- (P3) Rajouter un évènement qui ne se produira jamais ne change pas le manque d’information sur le tirage.
- (P4) Tirer un évènement dans $[1, M]$ est équivalent à tirer A ou B avec les probabilités q_A et q_B , puis, conditionné à ce premier tirage, choisir l’évènement élémentaire à l’intérieur du groupe choisi avec la probabilité conditionnelle. Le manque d’information est donc additif sous cette décomposition.



1.

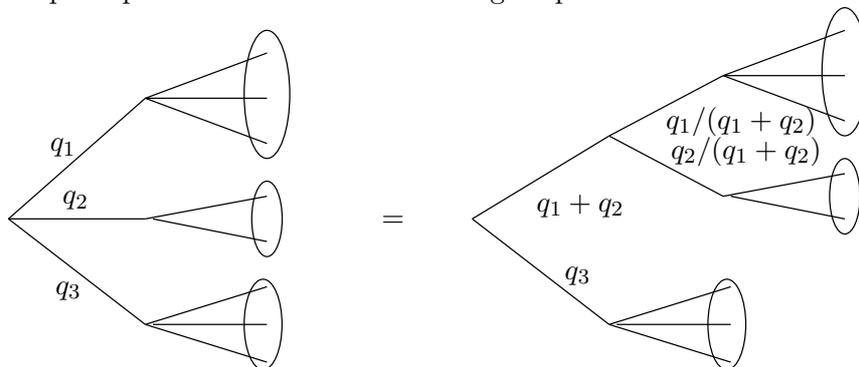
$$\sigma(M) = S\left(\frac{1}{M}, \dots, \frac{1}{M}\right) = S\left(\frac{1}{M}, \dots, \frac{1}{M}, 0\right) \leq S\left(\frac{1}{M+1}, \dots, \frac{1}{M+1}\right) = \sigma(M+1) \quad (299)$$

2. En appliquant la règle de décomposition avec $M = 2$, $q_A = p_1$, $q_B = p_2$, il vient

$$S(p_1, p_2) = S(p_1, p_2) + (p_1 + p_2)S(1) = S(p_1, p_2) + S(1) . \quad (300)$$

Il n’y a en effet aucun manque d’information sur un évènement dont le tirage est déterministe.

3. Comme l’hypothèse (P4) correspond à cette formule pour $\alpha = 2$ on peut faire une récurrence sur α , dont le principe est schématisé sur cette figure pour $\alpha = 3$:



Explicitement,

$$\begin{aligned} S(p_1, \dots, p_{m_1}, \dots, p_{m_2}, \dots, p_{m_\alpha}) &= S(q_1 + \dots + q_{\alpha-1}, q_\alpha) \\ &+ (q_1 + \dots + q_{\alpha-1})S\left(\frac{p_1}{q_1 + \dots + q_{\alpha-1}}, \dots, \frac{p_{m_{\alpha-1}}}{q_1 + \dots + q_{\alpha-1}}\right) \\ &+ q_\alpha S\left(\frac{p_{m_{\alpha-1}+1}}{q_\alpha}, \dots, \frac{p_{m_\alpha}}{q_\alpha}\right) \end{aligned} \quad (301)$$

en utilisant (P4). On utilise ensuite l'hypothèse de récurrence au rang $\alpha - 1$ pour transformer le S de la deuxième ligne, ce qui donne

$$\begin{aligned}
& S(q_1 + \dots + q_{\alpha-1}, q_\alpha) \\
& + (q_1 + \dots + q_{\alpha-1}) \left[S\left(\frac{q_1}{q_1 + \dots + q_{\alpha-1}}, \dots, \frac{q_{\alpha-1}}{q_1 + \dots + q_{\alpha-1}}\right) \right. \\
& \quad \left. + \sum_{i=1}^{\alpha-1} \frac{q_i}{q_1 + \dots + q_{\alpha-1}} S\left(\frac{p_{m_{i-1}+1}}{q_i}, \dots, \frac{p_{m_i}}{q_i}\right) \right] \\
& + q_\alpha S\left(\frac{p_{m_{\alpha-1}+1}}{q_\alpha}, \dots, \frac{p_{m_\alpha}}{q_\alpha}\right) \\
= & S(q_1 + \dots + q_{\alpha-1}, q_\alpha) + (q_1 + \dots + q_{\alpha-1}) S\left(\frac{q_1}{q_1 + \dots + q_{\alpha-1}}, \dots, \frac{q_{\alpha-1}}{q_1 + \dots + q_{\alpha-1}}\right) + q_\alpha S\left(\frac{q_\alpha}{q_\alpha}\right) \\
& + \sum_{i=1}^{\alpha} q_i S\left(\frac{p_{m_{i-1}+1}}{q_i}, \dots, \frac{p_{m_i}}{q_i}\right)
\end{aligned}$$

En utilisant (P4) on voit finalement que la première ligne est $S(q_1, \dots, q_\alpha)$, ce qui conclut la preuve.

4. Il suffit de considérer l'entropie de MN évènements équiprobables, regroupés en $\alpha = M$ blocs de N évènements.
5. En itérant cette relation il vient $\sigma(l^n) = n\sigma(l)$. On veut exprimer $\sigma(l)$ en terme de $\sigma(2)$, on définit donc $m(n) = \lfloor n \frac{\ln l}{\ln 2} \rfloor$, de sorte que $2^m \leq l^n < 2^{m+1}$. En utilisant aussi la croissance de σ on obtient

$$\sigma(2^m) \leq \sigma(l^n) \leq \sigma(2^{m+1}) \quad (302)$$

$$m\sigma(2) \leq n\sigma(l) \leq (m+1)\sigma(2) \quad (303)$$

$$\frac{1}{n} \left\lfloor n \frac{\ln l}{\ln 2} \right\rfloor \sigma(2) \leq \sigma(l) \leq \left(\frac{1}{n} \left\lfloor n \frac{\ln l}{\ln 2} \right\rfloor + \frac{1}{n} \right) \sigma(2). \quad (304)$$

En prenant la limite $n \rightarrow \infty$ à l fixé il vient $\sigma(l) = \frac{\sigma(2)}{\ln 2} \ln l = k \ln l$. Il faut que k soit positive pour que l'entropie du cas équiprobable soit croissante.

6. En réduisant au même dénominateur on peut écrire $p_i = N_i/D$, où D et les N_i sont des entiers. En utilisant la décomposition de l'entropie pour D évènements équiprobables groupés en M blocs contenant chacun N_i évènements élémentaires, il vient

$$S\left(\frac{1}{D}, \dots, \frac{1}{D}\right) = S\left(\frac{N_1}{D}, \dots, \frac{N_M}{D}\right) + \sum_{i=1}^M \frac{N_i}{D} S\left(\frac{1}{N_i}, \dots, \frac{1}{N_i}\right), \quad (305)$$

soit

$$S\left(\frac{N_1}{D}, \dots, \frac{N_M}{D}\right) = \sigma(D) - \sum_{i=1}^M \frac{N_i}{D} \sigma(N_i) = -k \sum_{i=1}^M \frac{N_i}{D} \ln\left(\frac{N_i}{D}\right) = -k \sum_{i=1}^M p_i \ln p_i, \quad (306)$$

en utilisant l'expression précédemment établie pour σ .

7. Comme \mathbb{Q}^M est dense dans \mathbb{R}^M et que S est supposée continue cette expression reste valable pour des probabilités arbitraires.

Propriétés de l'entropie

1. On peut remarquer que $g(x) = -x \ln x$ est une fonction concave sur $[0, 1]$, sa dérivée seconde $-1/x$ étant négative. Pour une valeur de $\underline{\sigma}$ on a donc

$$-(\alpha\mu(\underline{\sigma}) + (1-\alpha)\nu(\underline{\sigma})) \ln(\alpha\mu(\underline{\sigma}) + (1-\alpha)\nu(\underline{\sigma})) \geq -\alpha\mu(\underline{\sigma}) \ln \mu(\underline{\sigma}) - (1-\alpha)\nu(\underline{\sigma}) \ln \nu(\underline{\sigma}), \quad (307)$$

qu'il suffit ensuite de sommer sur $\underline{\sigma}$ pour obtenir $S(\alpha\mu + (1-\alpha)\nu) \geq \alpha S(\mu) + (1-\alpha)S(\nu)$.

2. Notons que pour $a, b \geq 0$

$$g(a+b) = -(a+b) \ln(a+b) = -a \ln(a+b) - b \ln(a+b) \leq -a \ln a - b \ln b = g(a) + g(b), \quad (308)$$

car le logarithme est croissant. De plus

$$g(uv) = -uv \ln(uv) = -uv \ln u - uv \ln v = ug(v) + vg(u), \quad (309)$$

donc

$$g(\alpha x + (1-\alpha)y) \leq g(\alpha x) + g((1-\alpha)y) = \alpha g(x) + (1-\alpha)g(y) + xg(\alpha) + yg(1-\alpha). \quad (310)$$

pour $\alpha, x, y \in [0, 1]$.

En appliquant cette inégalité avec $x = \mu(\underline{\sigma})$, $y = \nu(\underline{\sigma})$ et en sommant sur $\underline{\sigma}$ il vient bien le résultat demandé.

3. Le logarithme étant une fonction concave, pour toute fonction positive $f(\underline{\sigma})$ et toute mesure de probabilité $\mu(\underline{\sigma})$ l'inégalité de Jensen conduit à

$$\sum_{\underline{\sigma}} \mu(\underline{\sigma}) \ln(f(\underline{\sigma})) \leq \ln \left(\sum_{\underline{\sigma}} \mu(\underline{\sigma}) f(\underline{\sigma}) \right). \quad (311)$$

En appliquant cette inégalité à $f(\underline{\sigma}) = \frac{\nu(\underline{\sigma})}{\mu(\underline{\sigma})}$,

$$-D(\mu||\nu) \leq \ln \left(\sum_{\underline{\sigma}} \mu(\underline{\sigma}) \frac{\nu(\underline{\sigma})}{\mu(\underline{\sigma})} \right) = 0. \quad (312)$$

Le logarithme étant strictement concave il n'y a égalité que si f est constante, autrement dit $D(\mu||\nu) = 0 \Rightarrow \mu = \nu$.

Notons qu'on peut aussi prouver l'inégalité à partir de $\ln x \leq x - 1$:

$$-D(\mu||\nu) = \sum_{\underline{\sigma}} \mu(\underline{\sigma}) \ln \left(\frac{\nu(\underline{\sigma})}{\mu(\underline{\sigma})} \right) \leq \sum_{\underline{\sigma}} (\nu(\underline{\sigma}) - \mu(\underline{\sigma})) = 0 \quad (313)$$

4. Comme $\mu(\underline{\sigma}) \in [0, 1]$ les termes de la somme sont non-négatifs, d'où $S(\mu) \geq 0$. Pour que $S(\mu)$ soit égale à 0 il faut que $\mu(\underline{\sigma}) \ln(\mu(\underline{\sigma})) = 0$ pour tout $\underline{\sigma}$, i.e. que $\mu(\underline{\sigma}) \in \{0, 1\}$, comme μ est normalisée cela implique que $\mu(\underline{\sigma}) = \delta_{\underline{\sigma}, \underline{\tau}}$, la mesure est supportée par une seule configuration $\underline{\tau}$. Notons ν la mesure uniforme sur les $|\chi|^{|\Delta|}$ configurations, avec $\nu(\underline{\sigma}) = \frac{1}{|\chi|^{|\Delta|}}$. Alors

$$D(\mu||\nu) = -S(\mu) + |\Delta| \ln |\chi| \geq 0 \Rightarrow S(\mu) \leq |\Delta| \ln |\chi|. \quad (314)$$

On peut aussi le voir comme une conséquence de l'hypothèse (P2) de l'exercice précédent. La borne n'est atteinte que pour $\mu = \nu$, la mesure uniforme.

5. Comme $\mu_{\Lambda} \mu_{\Lambda^c}$ est une mesure de probabilité sur le même espace de configurations que μ on peut calculer

$$\begin{aligned} D(\mu||\mu_{\Lambda} \mu_{\Lambda^c}) &= -S(\mu) - \sum_{\underline{\sigma}_{\Lambda}, \underline{\sigma}_{\Lambda^c}} \mu((\underline{\sigma}_{\Lambda}, \underline{\sigma}_{\Lambda^c})) \ln \mu_{\Lambda}(\underline{\sigma}_{\Lambda}) - \sum_{\underline{\sigma}_{\Lambda}, \underline{\sigma}_{\Lambda^c}} \mu((\underline{\sigma}_{\Lambda}, \underline{\sigma}_{\Lambda^c})) \ln \mu_{\Lambda^c}(\underline{\sigma}_{\Lambda^c}) \\ &= -S(\mu) - \sum_{\underline{\sigma}_{\Lambda}} \mu_{\Lambda}(\underline{\sigma}_{\Lambda}) \ln \mu_{\Lambda}(\underline{\sigma}_{\Lambda}) - \sum_{\underline{\sigma}_{\Lambda^c}} \mu(\underline{\sigma}_{\Lambda^c}) \ln \mu_{\Lambda^c}(\underline{\sigma}_{\Lambda^c}) \\ &= -S(\mu) + S_{\Lambda}(\mu) + S_{\Lambda^c}(\mu) \geq 0 \end{aligned} \quad (315)$$

La borne est atteinte si $\mu = \mu_{\Lambda} \mu_{\Lambda^c}$, c'est-à-dire si les variables dans Λ sont indépendantes de celles dans Λ^c , l'information mutuelle est alors nulle.

6. L'entropie conditionnelle est la moyenne, sur la variable conditionnante, de l'entropie de la loi conditionnelle. D'après la définition,

$$S_{\Lambda|\Lambda^c}(\mu) = - \sum_{\underline{\sigma}} \mu(\underline{\sigma}) \ln \left(\frac{\mu(\underline{\sigma})}{\mu_{\Lambda^c}(\underline{\sigma}_{\Lambda^c})} \right) = S(\mu) - S_{\Lambda^c}(\mu) = S_{\Lambda}(\mu) - I_{\Lambda, \Lambda^c}(\mu) \leq S_{\Lambda}(\mu) \quad (316)$$

puisque l'information mutuelle est positive. On peut ainsi interpréter

$$I_{\Lambda, \Lambda^c}(\mu) = S_{\Lambda}(\mu) - S_{\Lambda|\Lambda^c}(\mu) = S_{\Lambda^c}(\mu) - S_{\Lambda^c|\Lambda}(\mu) \quad (317)$$

comme la réduction (moyenne) du manque d'information sur une des deux variables aléatoires quand l'autre est connue.

Principe variationnel de Gibbs

1. Notons que l'énergie libre de Gibbs $F_{\text{Gibbs}}(\beta, H, \mu)$ a bien la forme "énergie (moyenne) - température \times entropie". Calculons la distance de Kullback-Leibler entre μ et $\mu_c(\beta, H)$:

$$D(\mu||\mu_c(\beta, H)) = \sum_{\underline{\sigma}} \mu(\underline{\sigma}) \ln \left(\mu(\underline{\sigma}) \frac{Z(\beta, H)}{e^{-\beta H(\underline{\sigma})}} \right) = -S(\mu) + \ln Z(\beta, H) + \beta \sum_{\underline{\sigma}} \mu(\underline{\sigma}) H(\underline{\sigma}) , \quad (318)$$

soit

$$F_{\text{Gibbs}}(\beta, H, \mu) = F_c(\beta, H) + \frac{1}{\beta} D(\mu||\mu_c(\beta, H)) . \quad (319)$$

Comme $D \geq 0$ on a $F_{\text{Gibbs}}(\beta, H, \mu) \geq F_c(\beta, H)$, et comme $D(\mu_c(\beta, H)||\mu_c(\beta, H)) = 0$ la borne est atteinte pour $\mu = \mu_c(\beta, H)$. Comme $D(\mu||\nu) = 0$ seulement si $\mu = \nu$, $\mu_c(\beta, H)$ est l'unique point où le minimum est atteint.

Une preuve alternative de l'inégalité consiste à écrire

$$Z(\beta, H) = \sum_{\underline{\sigma}} e^{-\beta H(\underline{\sigma})} = \sum_{\underline{\sigma}} \mu(\underline{\sigma}) e^{-\beta H(\underline{\sigma}) - \ln \mu(\underline{\sigma})} \geq e^{-\beta \sum_{\underline{\sigma}} \mu(\underline{\sigma}) H(\underline{\sigma}) - \sum_{\underline{\sigma}} \mu(\underline{\sigma}) \ln \mu(\underline{\sigma})} , \quad (320)$$

en appliquant l'inégalité de Jensen à la fonction exponentielle qui est convexe. Il suffit ensuite de prendre le logarithme et de diviser par $-\beta$.

Pour vérifier que $F_{\text{Gibbs}}(\beta, H, \mu_c(\beta, H)) = F_c(\beta, H)$ on aurait aussi pu calculer l'entropie de la loi canonique,

$$S(\mu_c(\beta, H)) = - \sum_{\underline{\sigma}} \mu_c(\beta, H, \underline{\sigma}) \ln \left(\frac{e^{-\beta H(\underline{\sigma})}}{Z(\beta, H)} \right) = \ln Z(\beta, H) + \beta \sum_{\underline{\sigma}} \mu_c(\beta, H, \underline{\sigma}) H(\underline{\sigma}) . \quad (321)$$

2. Pour maximiser une fonction en imposant des contraintes on utilise la méthode des multiplicateurs de Lagrange. Ici on doit maximiser $S(\mu)$ par rapport à μ , en imposant la normalisation $\sum_{\underline{\sigma}} \mu(\underline{\sigma}) = 1$ et $\sum_{\underline{\sigma}} \mu(\underline{\sigma}) H(\underline{\sigma}) = E$, une énergie donnée. La méthode des multiplicateurs de Lagrange consiste donc à chercher un point stationnaire de $S(\mu) + A \sum_{\underline{\sigma}} \mu(\underline{\sigma}) + B \sum_{\underline{\sigma}} \mu(\underline{\sigma}) H(\underline{\sigma})$, où A et B sont les multiplicateurs de Lagrange. En annulant la dérivée par rapport à $\mu(\underline{\sigma})$ de cette fonction on obtient

$$\mu(\underline{\sigma}) = e^{A+B H(\underline{\sigma})} . \quad (322)$$

Les multiplicateurs de Lagrange A et B sont ensuite fixés pour satisfaire les contraintes. On retrouve alors la loi canonique, avec $B = -\beta$, et $e^A = 1/Z$. On peut voir aussi l'énergie libre de Gibbs, à un facteur $-1/\beta$ près, comme la fonction de Lagrange à maximiser. Le fait que l'entropie soit maximale (et non minimale) à cet extrémum vient de sa concavité.

3. La fonction qui à H associe

$$\beta \sum_{\underline{\sigma}} \mu(\underline{\sigma}) H(\underline{\sigma}) - \beta F_c(\beta, H) \quad (323)$$

est convexe et dérivable (c'est évident pour le premier terme, et cela a été démontrée en cours pour le deuxième). Le minimum est donc atteint au point \hat{H} où ses dérivées par rapport aux $H(\underline{\sigma})$ s'annulent, autrement dit $\hat{H}(\underline{\sigma}) = -\frac{1}{\beta} \ln \mu(\underline{\sigma}) + C$, où C est une constante indépendante de $\underline{\sigma}$. On a donc un Hamiltonien \hat{H} tel que $\mu_c(\beta, \hat{H}) = \mu$. Comme $F_{\text{Gibbs}}(\beta, \hat{H}, \mu_c(\beta, \hat{H}) = \mu) = F_c(\beta, \hat{H})$, il vient

$$\sum_{\underline{\sigma}} \mu(\underline{\sigma}) \hat{H}(\underline{\sigma}) - \frac{1}{\beta} S(\mu) = F_c(\beta, \hat{H}), \quad \text{soit } S(\mu) = \beta \sum_{\underline{\sigma}} \mu(\underline{\sigma}) \hat{H}(\underline{\sigma}) - \beta F_c(\beta, \hat{H}). \quad (324)$$

4. Notons μ la mesure produit dans Δ , où chaque spin a indépendamment l'aimantation m , $\mu(\underline{\sigma}) = \prod_{i \in \Delta} \frac{1+m\sigma_i}{2}$. On a alors

$$\frac{1}{|\Delta|} F_{\text{Gibbs}}(\beta, H, \mu) = -Jdm^2 - hm - \frac{1}{\beta} s(m), \quad s(m) = -\frac{1+m}{2} \ln \left(\frac{1+m}{2} \right) - \frac{1-m}{2} \ln \left(\frac{1-m}{2} \right),$$

car il y a $2d$ liens autour de chaque site mais il ne faut les compter qu'une fois pour chacune des deux extrémités. La borne étant vraie pour tout m , on peut donc conclure

$$\frac{1}{|\Delta|} F_c(\beta, H) \leq \inf_{m \in [-1,1]} f_{\text{cm}}(m), \quad f_{\text{cm}}(m) = -Jdm^2 - hm - \frac{1}{\beta} s(m), \quad (325)$$

où l'on a reconnu l'expression de l'énergie libre de champ moyen établi au TD 1.

Plus généralement, le principe variationnel de Gibbs est souvent exploité en mécanique statistique, sous le nom de "méthodes variationnelles". En effet en général on est incapable de calculer exactement $F_c(\beta, H)$ pour un Hamiltonien $H(\underline{\sigma})$ un peu compliqué. On peut néanmoins calculer $F_{\text{Gibbs}}[\beta, H, \mu]$ pour des mesures μ suffisamment simples. En cherchant le minimum de cette fonction dans un sous-espace de \mathcal{M} on obtient donc une borne supérieure de la quantité inaccessible F_c . Quand μ se cantonne aux mesures produits on obtient ainsi l'approximation de champ moyen.

Dans ce contexte le principe variationnel est souvent exprimé sous la forme

$$F_c(\beta, H) \leq F_0 + \langle H - H_0 \rangle_0, \quad (326)$$

où la mesure variationnelle est prise de la forme $\mu = e^{-\beta H_0} / Z_0$, les moyennes avec cette mesure étant notées $\langle \bullet \rangle_0$, et F_0 étant l'énergie libre associée au Hamiltonien variationnel H_0 .

TD 3 : Transformées de Legendre

Fonctions convexes

On considère une fonction f de \mathbb{R} dans \mathbb{R} , que l'on suppose convexe :

$$f(\alpha x + (1 - \alpha)y) \leq \alpha f(x) + (1 - \alpha)f(y) \quad \text{pour } x, y \in \mathbb{R} \text{ et } \alpha \in [0, 1]. \quad (327)$$

1. Pour trois points $x_1 < x_2 < x_3$, montrer que

$$\frac{f(x_2) - f(x_1)}{x_2 - x_1} \leq \frac{f(x_3) - f(x_1)}{x_3 - x_1} \leq \frac{f(x_3) - f(x_2)}{x_3 - x_2}. \quad (328)$$

2. Montrer que les dérivées à gauche et à droite de f

$$(D_-f)(x) = \lim_{h \rightarrow 0^+} \frac{f(x) - f(x-h)}{h}, \quad (D_+f)(x) = \lim_{h \rightarrow 0^+} \frac{f(x+h) - f(x)}{h}, \quad (329)$$

existent pour tous les x et vérifient, si $x < y$

$$(D_-f)(x) \leq (D_+f)(x) \leq (D_-f)(y) \leq (D_+f)(y). \quad (330)$$

3. En déduire que f est continue sur \mathbb{R} et dérivable sur \mathbb{R} privé d'un nombre au plus dénombrable de points.

4. Montrer que si \mathcal{G} est une famille de fonctions convexes, la fonction f définie par $f(x) = \sup_{g \in \mathcal{G}} g(x)$ est convexe. On dit que $f : \mathbb{R} \rightarrow \mathbb{R} \cup \{+\infty\}$ est convexe si elle vérifie (327) au sens de l'arithmétique sur la droite réelle étendue.

5. On dit que $v_\gamma(x) = \gamma x$ est une fonction linéaire tangente au point x à la fonction f convexe ssi

$$f(y) \geq f(x) + v_\gamma(y - x) \quad \forall y. \quad (331)$$

Montrer qu'en tout point une fonction convexe de \mathbb{R} dans \mathbb{R} admet au moins une fonction linéaire tangente. Quelle est la condition d'unicité de la fonction linéaire tangente en un point ?

6. Pour une fonction f pas forcément convexe on définit son enveloppe convexe $E.C. f$ comme à la question 4 en prenant pour \mathcal{G} la famille $\{g \text{ convexe et } g(x) \leq f(x) \forall x\}$; l'enveloppe convexe est donc la plus grande fonction convexe qui minore f , et si f est convexe $E.C. f = f$. Montrer que l'on peut définir de manière équivalente l'enveloppe convexe à partir de la famille des fonctions affines minorant f , i.e. avec $\mathcal{G}' = \{g_{a,b} : g_{a,b}(x) = ax + b \leq f(x) \forall x\}$.

Transformées de Legendre

Pour une fonction f définie sur \mathbb{R} , pas forcément convexe, on définit sa transformée de Legendre

$$\hat{f}(\hat{x}) = \sup_x [\hat{x}x - f(x)]. \quad (332)$$

1. Interpréter graphiquement cette définition.

2. Calculer les transformées de Legendre de :

$$f(x) = \beta \frac{1}{2} x^2, \quad f(x) = \beta |x|, \quad f(x) = \beta ||x| - 1|. \quad (333)$$

On supposera $\beta > 0$ dans ces trois cas.

3. Montrer que \hat{f} est convexe.

4. Montrer que $\hat{\hat{f}}(x) \leq f(x)$.

5. Montrer que si g est une fonction affine minorant f , i.e. si $g(x) = ax + b \leq f(x) \forall x$, alors $g(x) \leq \hat{f}(x) \forall x$.
6. En déduire que $\hat{\hat{f}} = E.C.f$. Si f est convexe de \mathbb{R} dans \mathbb{R} on a donc $\hat{\hat{f}} = f$, et on peut alors inverser la transformée de Legendre avec $f(x) = \sup_{\hat{x}} [\hat{x}x - \hat{f}(\hat{x})]$.
7. Montrer que v_γ est tangente au point x à la fonction f si et seulement si $f(x) + \hat{f}(\gamma) = v_\gamma(x)$.

On peut assez facilement généraliser les résultats ci-dessus au cas des fonctions de \mathbb{R}^n dans \mathbb{R} : la définition de convexité (327) reste valable si f est une fonction de X dans \mathbb{R} . On peut aussi généraliser la transformée de Legendre avec

$$\hat{f}(\hat{x}) = \sup_{x \in \mathbb{R}^n} [\langle \hat{x}, x \rangle - f(x)] , \quad (334)$$

où \hat{x} est aussi un vecteur de \mathbb{R}^n , en prenant pour $\langle \hat{x}, x \rangle$ le produit scalaire dans \mathbb{R}^n .

Il existe aussi une généralisation fonctionnelle (plus compliquée). La définition (327) reste valable si f est une fonction de X dans \mathbb{R} avec X un espace vectoriel arbitraire. En notant X^* le dual de X (i.e. l'espace vectoriel des formes linéaires de X vers \mathbb{R}), il semble naturel de définir la transformée de Legendre \hat{f} comme une fonction de X^* dans \mathbb{R} selon

$$\hat{f}(\hat{x}) = \sup_{x \in X} [\langle \hat{x}, x \rangle - f(x)] , \quad (335)$$

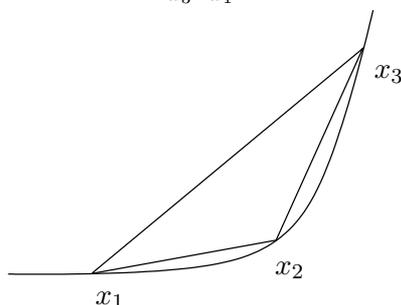
où l'on a noté $\langle \hat{x}, x \rangle = \hat{x}(x)$ l'action de la forme linéaire \hat{x} . De même une fonctionnelle linéaire tangente à f en x sera un élément \hat{x} du dual tel que $f(y) \geq f(x) + \langle \hat{x}, y - x \rangle$ pour tout $y \in X$.

Cependant la généralisation de certains théorèmes, en particulier la caractérisation des fonctionnelles tangentes par $f(x) + \hat{f}(\hat{x}) = \langle \hat{x}, x \rangle$, n'est valide qu'avec des hypothèses supplémentaires, discutées notamment par Fenchel (dans le cas général on parle de transformée de Legendre-Fenchel). On se restreint en particulier au sous-espace vectoriel \hat{X} de X^* tel que $\forall x \in X \setminus \{0\}$, il existe $\hat{x} \in \hat{X}$ avec $\langle \hat{x}, x \rangle \neq 0$; on appelle alors $\langle X, \hat{X} \rangle$ une paire duale. Il faut aussi restreindre la maximisation dans (335) à un ensemble convexe vérifiant des « bonnes » propriétés.

Corrigé du TD 3

Fonctions convexes

1. Ecrivons $x_2 = \alpha x_1 + (1 - \alpha)x_3$, avec $\alpha = \frac{x_3 - x_2}{x_3 - x_1} \in]0, 1[$.



De l'inégalité de convexité

$$f(x_2) \leq \alpha f(x_1) + (1 - \alpha)f(x_3) \quad (336)$$

on tire sans difficultés les deux inégalités, qui portent sur les pentes des trois droites tracées sur cette figure.

2. A x fixé, avec $h > 0$, notons $g_+(h) = \frac{f(x+h) - f(x)}{h}$. En prenant $x_1 = x$, $x_2 = x + h$ et $x_3 = x + h'$ dans l'inégalité précédente on montre que g_+ est une fonction croissante, elle admet donc une limite $(\Delta_+ f)(x)$ quand $h \rightarrow 0^+$.

De même $g_-(h) = \frac{f(x) - f(x-h)}{h}$ est décroissante (prendre $x_3 = x$, $x_2 = x - h$, $x_1 = x - h'$), d'où l'existence de sa limite $(\Delta_- f)(x)$ quand $h \rightarrow 0^+$.

Comme $g_-(h) \leq g_+(h)$ (prendre $x_2 = x$, $x_1 = x - h$, $x_3 = x + h$), les limites vérifient $(D_- f)(x) \leq (D_+ f)(x)$.

Finalement si $x < y$, en prenant $z \in]x, y[$,

$$(D_+ f)(x) \leq \frac{f(z) - f(x)}{z - x} \leq \frac{f(y) - f(z)}{y - z} \leq (D_- f)(y). \quad (337)$$

3. Une fonction qui est dérivable à gauche et à droite en x est continue en x , donc f est continue.

Par ailleurs une fonction monotone est continue partout sauf en un nombre au plus dénombrable de points : considérons g croissante, les limites à gauche et à droite de g existent partout, notons $X_\epsilon(I)$ les points d'un intervalle borné I où la discontinuité de g est $> \epsilon$. Sur I g est bornée, $X_\epsilon(I)$ est donc fini pour tout $\epsilon > 0$. Donc l'ensemble des points de discontinuité de g sur \mathbb{R} est $\bigcup_{n>0, p>0} X_{1/n}([-p, p])$, au plus dénombrable.

On peut appliquer ce résultat à $\Delta_+ f$ qui est bien monotone. En un point de continuité x de $\Delta_+ f$ on a $(\Delta_+ f)(x) = (\Delta_- f)(x)$ (en prenant $y \rightarrow x^+$ dans l'inégalité (4) de l'énoncé), donc f est dérivable en x .

- 4.

$$f(\alpha x + (1 - \alpha)y) = \sup_{g \in \mathcal{G}} g(\alpha x + (1 - \alpha)y) \leq \sup_{g \in \mathcal{G}} [\alpha g(x) + (1 - \alpha)g(y)] \quad (338)$$

$$\leq \alpha \sup_{g \in \mathcal{G}} [g(x)] + (1 - \alpha) \sup_{g \in \mathcal{G}} [g(y)] = \alpha f(x) + (1 - \alpha)f(y). \quad (339)$$

5. On a vu qu'en tout point x une fonction convexe f admet des dérivées à gauche et à droite $(D_- f)(x)$ et $(D_+ f)(x)$, avec $(D_- f)(x) \leq (D_+ f)(x)$. Si $\gamma \in [(D_- f)(x), (D_+ f)(x)]$ alors v_γ vérifie la définition, car pour $y > x$

$$\frac{f(y) - f(x)}{y - x} \geq (D_+ f)(x) \geq \gamma \Rightarrow f(y) \geq f(x) + \gamma(y - x) = f(x) + v_\gamma(y - x), \quad (340)$$

et pour $y < x$

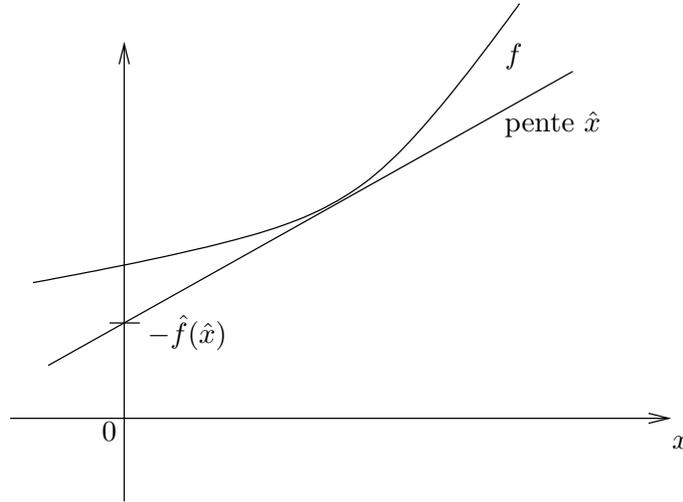
$$\frac{f(x) - f(y)}{x - y} \leq (D_- f)(x) \leq \gamma \Rightarrow f(y) \geq f(x) + \gamma(y - x) = f(x) + v_\gamma(y - x). \quad (341)$$

Il y a donc unicité si et seulement si $(D_- f)(x) = (D_+ f)(x)$, autrement dit si f est dérivable au point x considéré.

6. Notons h la fonction obtenue à partir de la famille \mathcal{G}' , i.e. $h(x) = \sup_{g \in \mathcal{G}'} g(x)$. Comme $\mathcal{G}' \subset \mathcal{G}$, on a $h(x) \leq (E.C. f)(x)$. Par ailleurs $E.C. f$ est une fonction convexe, elle admet donc au moins une fonction linéaire tangente en tout point. Considérons un point x_0 arbitraire, et une fonction affine $g_{a,b}$ tangente à $E.C. f$ en x_0 . Comme $g_{a,b}$ minore $E.C. f$ elle minore aussi f , donc $g_{a,b} \in \mathcal{G}'$, donc $h(x_0) \geq g_{a,b}(x_0) = (E.C. f)(x_0)$. Comme c'est vrai quelque soit x_0 , et comme on a aussi l'égalité inverse, on en conclut $h = E.C. f$.

Transformées de Legendre

1. L'ordonnée à l'origine d'une droite de pente \hat{x} intersectant f en x est $f(x) - \hat{x}x$. Il faut donc chercher la droite de pente \hat{x} la plus basse possible qui touche f , son ordonnée à l'origine est $-\hat{f}(\hat{x})$.



Si f est dérivable les points stationnaires de la fonction dont on cherche le sup sont solutions de $\hat{x} = f'(x)$.

2. Dans le premier cas la relation entre x et \hat{x} est donnée par $\hat{x} = \beta x$, on trouve $\hat{f}(\hat{x}) = \frac{1}{\beta} \frac{1}{2} \hat{x}^2$.
Pour la transformée de Legendre de la valeur absolue il vient

$$\hat{f}(\hat{x}) = \begin{cases} 0 & \text{si } \hat{x} \in [-\beta, \beta] \\ +\infty & \text{sinon} \end{cases} \quad (342)$$

Dans le dernier cas on trouve

$$\hat{f}(\hat{x}) = \begin{cases} |\hat{x}| & \text{si } \hat{x} \in [-\beta, \beta] \\ +\infty & \text{sinon} \end{cases} \quad (343)$$

3. \hat{f} est le sup de fonctions convexes de \hat{x} (car affines), elle est donc convexe.
4. D'après la définition,

$$\hat{f}(x) = \sup_{\hat{x}} [x\hat{x} - \sup_{x'} [\hat{x}x' - f(x')]] = \sup_{\hat{x}} \inf_{x'} [f(x') + \hat{x}(x - x')]. \quad (344)$$

Quelque soit \hat{x} l'inf sur x' est plus petit que sa valeur en x qui vaut $f(x)$, d'où $\hat{f}(x) \leq f(x)$.

5.

$$\inf_{x'} [f(x') + \hat{x}(x - x')] \geq \inf_{x'} [ax' + b + \hat{x}(x - x')] = \begin{cases} ax + b & \text{si } \hat{x} = a \\ -\infty & \text{sinon} \end{cases}, \quad (345)$$

d'où $\hat{f}(x) \geq ax + b = g(x)$.

6. On a donc d'après les deux dernières questions $(E.C. f)(x) \leq \hat{f}(x) \leq f(x)$ pour tout x . Or \hat{f} est convexe et minore f implique $\hat{f}(x) \leq (E.C. f)(x)$ pour tout x , ce qui permet de conclure $\hat{f} = E.C. f$. Si f est convexe elle est égale à son enveloppe convexe, et donc à sa double transformée de Legendre. On peut vérifier explicitement la relation $\hat{f} = f$ pour f convexe sur les deux premiers exemples traités ci-dessus. Dans le troisième cas on trouve

$$\hat{f}(x) = \begin{cases} 0 & \text{si } x \in [-1, 1] \\ \beta(|x| - 1) & \text{sinon} \end{cases}, \quad (346)$$

qui est bien l'enveloppe convexe de f .

7. Pour tout x et γ on a $f(x) + \hat{f}(\gamma) \geq v_\gamma(x)$, par définition de la transformée de Legendre. De plus

$$f(x) + \hat{f}(\gamma) \leq v_\gamma(x) \quad \Leftrightarrow \quad \hat{f}(\gamma) = \sup_y [\gamma y - f(y)] \leq v_\gamma(x) - f(x) \quad (347)$$

$$\Leftrightarrow \quad \gamma y - f(y) \leq v_\gamma(x) - f(x) \quad \forall y \quad (348)$$

$$\Leftrightarrow \quad f(y) \geq f(x) + v_\gamma(y - x) \quad \forall y. \quad (349)$$

Examen de juin 2013 : Les zéros de Lee et Yang

Il est parfois instructif d'étendre le domaine de variation des paramètres d'un problème de leurs valeurs physiques, i.e. réelles, à tout le plan complexe, et d'étudier ainsi leurs propriétés d'analyticité. Considérons par exemple un modèle ferromagnétique et le champ magnétique h comme paramètre. Par définition la fonction de partition Z du modèle n'est jamais nulle quand $h \in \mathbb{R}$, car c'est une somme de termes positifs; en 1952 Lee et Yang ont étudié les points d'annulation de $Z(h)$ en considérant $h \in \mathbb{C}$ et ont découvert que ces zéros correspondaient nécessairement à des valeurs purement imaginaires du champ magnétique. Dans cet exercice on va démontrer ce résultat et présenter une de ses applications.

Le théorème de Lee et Yang pour un graphe fini

On considère un graphe $G = (V, E)$, i.e. un ensemble fini V de sommets et un ensemble de liens $e \in E$ reliant chacun deux sommets de V distincts, notés i_e et j_e . Chaque lien est muni d'une constante de couplage J_e réelle (mais pour l'instant de signe arbitraire). On place un spin d'Ising $\sigma_i \in \{-1, +1\}$ sur chaque sommet, leur configuration collective est notée $\underline{\sigma} = \{\sigma_i, i \in V\} \in \{-1, +1\}^{|V|}$. Ils subissent l'effet de champ magnétiques h_i a priori complexes et différents pour chaque site, et les interactions selon les liens du graphe. L'énergie d'une configuration est donc

$$H(\underline{\sigma}) = - \sum_{e \in E} J_e \sigma_{i_e} \sigma_{j_e} - \sum_{i \in V} h_i \sigma_i . \quad (350)$$

Chaque fois que cela sera possible sans ambiguïté on numérotera les sommets par des entiers, i.e. on supposera $V = \{1, \dots, |V|\}$.

1. Ecrire la fonction de partition $\hat{Z}_G(h_1, \dots, h_{|V|})$ du modèle en contact avec un thermostat de température inverse β . On notera $Z_G(h) = \hat{Z}_G(h, \dots, h)$ sa valeur quand les champs magnétiques sont tous égaux à h .
2. Mettre la fonction de partition sous la forme

$$\hat{Z}_G(h_1, \dots, h_{|V|}) = e^{-\sum_{i \in V} \beta h_i + \sum_{e \in E} \beta J_e} \hat{P}_G(z_1, \dots, z_{|V|}) , \quad (351)$$

où $z_i = e^{2\beta h_i}$. Ici et dans toute la suite les z et z_i désignent des variables complexes. Comme \hat{P}_G est un polynôme de degré 1 dans chacune de ses variables on peut l'écrire

$$\hat{P}_G(z_1, \dots, z_{|V|}) = \sum_{U \subset V} c_{G,U} \prod_{i \in U} z_i , \quad (352)$$

où la somme porte sur les $2^{|V|}$ parties de V .

Exprimer les coefficients $c_{G,U}$ en termes des $\tau_e = e^{-2\beta J_e}$, et interpréter.

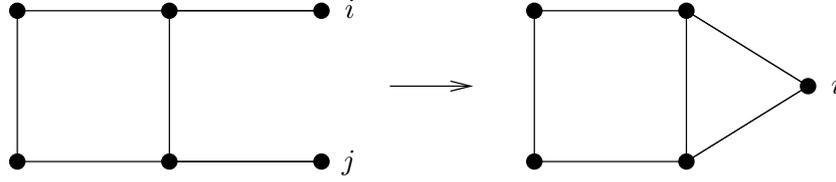
Comme ci-dessus on notera $P_G(z) = \hat{P}_G(z, \dots, z)$ le polynôme obtenu quand tous les champs magnétiques sont égaux.

3. Que vaut $\hat{P}_G(z_1, \dots, z_{|V|})$ lorsque $E = \emptyset$, i.e. pour un système de spins indépendants? Quelles sont alors les racines de $P_G(z)$? Pour quelle valeur du champ magnétique h la fonction de partition $Z(h)$ s'annule-t-elle?
4. Considérons maintenant $V = \{1, 2\}$ et un graphe G fait d'un seul lien entre les deux sommets, avec un couplage J . Exprimer alors $\hat{P}_G(z_1, z_2)$ et $P_G(z)$ (on notera $\tau = e^{-2\beta J}$). Quelles sont les racines de $P_G(z)$, et les valeurs du champ magnétique associé? On distinguera les cas $J \geq 0$ et $J < 0$. Montrer que si $J \geq 0$,

$$|z_1| < 1 \text{ et } |z_2| < 1 \Rightarrow \hat{P}_G(z_1, z_2) \neq 0 . \quad (353)$$

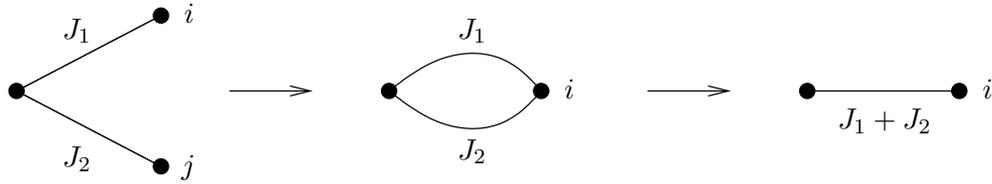
Indication : on pourra étudier les propriétés de l'homographie $z \rightarrow -\frac{1+\tau z}{\tau+z}$, en particulier l'image du cercle unité par cette application.

5. Supposons qu'un graphe $G = (V, E)$ puisse se décomposer en deux graphes disjoints $G_1 = (V_1, E_1)$ et $G_2 = (V_2, E_2)$, i.e. V est l'union disjointe de V_1 et V_2 , les liens de E_1 (resp. E_2) ne sont adjacents qu'à des sommets de V_1 (resp. V_2), et $E = E_1 \cup E_2$. Exprimer \widehat{P}_G à l'aide de \widehat{P}_{G_1} et \widehat{P}_{G_2} .
6. On définit maintenant une opération de contraction sur les graphes, schématisé ainsi sur un exemple :



La définition formelle de la contraction est la suivante : pour un graphe $G = (V, E)$ arbitraire sur $|V|$ sommets et $i \neq j$ deux sommets distincts non liés dans E on note $G^{(i,j)}$ le graphe sur les $|V| - 1$ sommets $V \setminus j$ obtenu en identifiant le sommet j au sommet i (i.e. j est remplacé par i dans toutes les arêtes de E).

Il peut arriver dans certains cas particuliers que cette opération de contraction conduise à des arêtes multiples dans le graphe contracté, par exemple :



Par convention on remplace alors les arêtes multiples par une unique arête portant la somme des couplages.

Isolons la dépendance en z_i, z_j du polynôme \widehat{P}_G d'un graphe G arbitraire :

$$\widehat{P}_G(z_i, z_j, \underline{z}') = A(\underline{z}') + B(\underline{z}')z_i + C(\underline{z}')z_j + D(\underline{z}')z_i z_j, \quad (354)$$

où A, B, C et D sont des polynômes de degré 1 dans les $\underline{z}' = \{z_k, k \in V \setminus \{i, j\}\}$.

Que vaut le polynôme $\widehat{P}_{G^{(i,j)}}(z_i, \underline{z}')$ obtenu après contraction ?

7. Supposons que a, b, c, d soient quatre nombres complexes tels que

$$|z_1| < 1 \text{ et } |z_2| < 1 \Rightarrow a + bz_1 + cz_2 + dz_1 z_2 \neq 0. \quad (355)$$

Montrer qu'alors

$$|z| < 1 \Rightarrow a + dz \neq 0. \quad (356)$$

8. On dit qu'un graphe $G = (V, E)$ vérifie la propriété de Lee-Yang si son polynôme associé est tel que

$$|z_1| < 1 \text{ et } |z_2| < 1 \text{ et } \dots \text{ et } |z_{|V|}| < 1 \Rightarrow \widehat{P}_G(z_1, \dots, z_{|V|}) \neq 0. \quad (357)$$

Montrer que si un graphe vérifie la propriété de Lee-Yang alors le graphe obtenue par contraction de deux sommets la vérifie aussi.

9. En déduire que la propriété de Lee-Yang est vérifiée pour n'importe quel graphe $G = (V, E)$ dont tous les couplages sont ferromagnétiques (i.e. $J_e \geq 0 \forall e \in E$).

Indication : on pourra construire un graphe \tilde{G} pour lequel la propriété est facilement vérifiée et qui conduise à G par une série de contractions.

10. En utilisant la propriété de symétrie $Z_G(h) = Z_G(-h)$ et ses conséquences sur le polynôme $P_G(z)$ terminer la preuve du théorème de Lee et Yang :

Pour tout graphe G dont les couplages sont ferromagnétiques, $Z_G(h) = 0 \Rightarrow h \in i\mathbb{R}$.

Analyticité de l'énergie libre

On considère maintenant le modèle d'Ising avec interactions à plus proches voisins sur le réseau \mathbb{Z}^d , défini par l'interaction

$$\phi_X(\underline{\sigma}) = \begin{cases} -h\sigma_x & \text{si } X = \{x\} \\ -J\sigma_x\sigma_y & \text{si } X = \{x, y\} \text{ avec } \text{dist}(x, y) = 1 \\ 0 & \text{sinon} \end{cases}, \quad (358)$$

avec J une constante de couplage réelle positive.

On note $f(\beta, h)$ la limite thermodynamique de l'énergie libre par spin de ce modèle, dont l'existence a été démontrée en cours pour $h \in \mathbb{R}$. On rappelle que

$$f = \lim_{\Lambda \rightarrow \infty} f_\Lambda, \quad f_\Lambda = -\frac{1}{\beta} \frac{1}{|\Lambda|} \ln Z_\Lambda, \quad (359)$$

où Λ est une partie finie de \mathbb{Z}^d , Z_Λ la fonction de partition du modèle avec conditions aux bords libres dans Λ , et la limite $\Lambda \rightarrow \infty$ est prise au sens de Van Hove.

11. Proposer une ébauche de preuve, sans trop de détails, de l'affirmation suivante : f est une fonction analytique de h sur $\mathbb{C} \setminus i\mathbb{R}$.
12. En déduire que la fonction $h \in \mathbb{R} \rightarrow f(\beta, h)$ ne peut être singulière qu'en $h = 0$. Montrer que les dérivées à gauche et à droite en 0 de cette fonction sont toujours bien définies. Expliquer qualitativement dans quelles conditions ces deux dérivées peuvent être différentes, et interpréter alors leurs valeurs.
13. Relier qualitativement cette éventuelle singularité au comportement, dans la limite thermodynamique, du lieu des zéros de Lee et Yang.
14. On rappelle le résultat du TD 1 pour ce modèle en dimension $d = 1$: la fonction de partition d'une chaîne de N spins avec des conditions aux bords périodiques s'écrit

$$Z_N(\beta, h) = \lambda_+(\beta, h)^N + \lambda_-(\beta, h)^N, \quad \text{avec } \lambda_\pm \text{ les racines de } \lambda^2 - x(y + y^{-1})\lambda + (x^2 - x^{-2}) = 0,$$

où $x = e^{\beta J}$ et $y = e^{\beta h}$.

Commenter votre réponse à la question précédente au vu de cette expression.

Corrigé de l'examen de juin 2013

Le théorème de Lee et Yang pour un graphe fini

1. Par définition,

$$\widehat{Z}_G(h_1, \dots, h_{|V|}) = \sum_{\underline{\sigma} \in \{-1, +1\}^V} e^{\sum_{e \in E} \beta J_e \sigma_{i_e} \sigma_{j_e} + \sum_{i \in V} \beta h_i \sigma_i} \quad (360)$$

2. En factorisant le terme de l'énoncé,

$$\widehat{Z}_G(h_1, \dots, h_{|V|}) = e^{-\sum_{i \in V} \beta h_i + \sum_{e \in E} \beta J_e} \sum_{\underline{\sigma} \in \{-1, +1\}^V} e^{\sum_{e \in E} \beta J_e (\sigma_{i_e} \sigma_{j_e} - 1) + \sum_{i \in V} \beta h_i (\sigma_i + 1)}. \quad (361)$$

On peut voir la somme sur les $2^{|V|}$ configurations comme une somme sur les $2^{|V|}$ ensembles $U \subset V$ des spins égaux à $+1$ ("Up"). On a alors $\sigma_i + 1 = 2$ si $i \in U$ et $\sigma_i + 1 = 0$ si $i \notin U$, d'où le facteur $\prod_{i \in U} z_i$. Par ailleurs $\sigma_{i_e} \sigma_{j_e} - 1$ vaut 0 si i_e et j_e appartiennent tous les deux à U ou à $V \setminus U$ (i.e. si les deux spins sont de même signe), -2 sinon. On a donc la forme de l'énoncé pour \widehat{P}_G avec

$$c_{G,U} = \prod_{e \in E \cap \partial U} \tau_e = e^{-2\beta \sum_{e \in E \cap \partial U} J_e}, \quad (362)$$

où ∂U désigne les paires de sommets dont un élément est dans U , l'autre dans $V \setminus U$. $c_{G,U}$ est donc le facteur de Boltzmann pour une configuration où les spins de U et de $V \setminus U$ sont de signes opposés, sans compter l'énergie des champs, et en ayant soustrait l'énergie du fondamental.

3. On a alors

$$\widehat{P}_G(z_1, \dots, z_{|V|}) = \prod_{i \in V} (1 + z_i), \quad P_G(z) = (1 + z)^{|V|}. \quad (363)$$

Ce dernier polynôme a une racine en -1 , dégénérée $|V|$ fois. Cela correspond à $h \in \frac{i}{2\beta}(\pi + 2\pi\mathbb{Z})$, un champ imaginaire pur.

4. En appliquant la formule générale il vient

$$\widehat{P}_G(z_1, z_2) = 1 + \tau z_1 + \tau z_2 + z_1 z_2, \quad P_G(z) = 1 + 2\tau z + z^2. \quad (364)$$

– Dans le cas antiferromagnétique, $J < 0$, $\tau > 1$, les racines de P_G sont $-\tau \pm \sqrt{\tau^2 - 1}$. Ces deux racines sont négatives et de modules différents de 1 , les champs magnétiques correspondants ont donc une partie imaginaire $\pi/(2\beta)$ (modulo π/β) et une partie réelle non-nulle.

– Dans le cas ferromagnétique, $J \geq 0$, $\tau \in [0, 1]$, les racines de P_G sont $-\tau \pm i\sqrt{1 - \tau^2}$, toutes les deux de module 1 , les champs correspondants sont donc imaginaires purs.

Pour une valeur de z_1 donnée le polynôme \widehat{P}_G s'annule si $z_2 = f(z_1)$ avec $f(z) = -\frac{1+\tau z}{\tau+z}$. Comme $f(e^{i\theta}) = -\frac{1+\tau e^{i\theta}}{\tau+e^{i\theta}} = -e^{i\theta} \frac{\tau+e^{-i\theta}}{\tau+e^{i\theta}}$, l'homographie f envoie le cercle unité sur lui-même. L'origine est envoyée en $-\frac{1}{\tau}$, qui dans le cas ferromagnétique est de module plus grand que 1 , l'infini est envoyé en $-\tau$, donc "par continuité" pour $J \geq 0$ cette application échange extérieur et intérieur du cercle unité. Plus formellement on peut calculer

$$|f(z)|^2 = 1 - \frac{1 - \tau^2}{|\tau + z|^2} (|z|^2 - 1), \quad (365)$$

telle que $|z| < 1$ conduit à $|f(z)| \geq 1$. Autrement dit si $|z_1| < 1$ il faudrait prendre $|z_2| \geq 1$ pour annuler \widehat{P}_G , ce qui prouve le résultat de l'énoncé.

5. En notant $V_1 = \{1, \dots, |V_1|\}$ et $V_2 = \{|V_1| + 1, \dots, |V_1| + |V_2|\}$, il vient $\widehat{Z}_G(h_1, \dots, h_{|V_1|+|V_2|}) = \widehat{Z}_{G_1}(h_1, \dots, h_{|V_1|})\widehat{Z}_{G_2}(h_{|V_1|+1}, \dots, h_{|V_1|+|V_2|})$ et donc

$$\widehat{P}_G(z_1, \dots, z_{|V_1|+|V_2|}) = \widehat{P}_{G_1}(z_1, \dots, z_{|V_1|})\widehat{P}_{G_2}(z_{|V_1|+1}, \dots, z_{|V_1|+|V_2|}) \quad (366)$$

6. Le poids d'une configuration avec $\sigma_i = \sigma$ dans $G^{(i,j)}$ est celui de la configuration correspondante avec $\sigma_i = \sigma_j = \sigma$ dans G . En termes des polynômes il ne reste donc que les termes proportionnels à A (pour $\sigma = -1$) et à D (pour $\sigma = +1$) :

$$\widehat{P}_{G^{(i,j)}}(z_i, \underline{z}') = A(\underline{z}') + D(\underline{z}')z_i . \quad (367)$$

7. Supposons d'abord $d \neq 0$. En prenant $z_1 = z_2 = z$, l'hypothèse implique que le polynôme $a + (b+c)z + dz^2$ a deux racines de module ≥ 1 , donc le produit des racines a/d est de module ≥ 1 . Or $a + dz$ s'annule justement pour $-a/d$, d'où le résultat. Si $d = 0$ on a forcément $a \neq 0$ pour que l'hypothèse soit vérifiée, et alors le résultat est trivial.
8. Soit G un graphe vérifiant la propriété de Lee-Yang, i, j les sommets sur lesquels on effectue la contraction. Choisissons arbitrairement \underline{z}' avec $|z_k| < 1$ pour tout $k \in V \setminus i, j$, et notons $a = A(\underline{z}')$, $b = B(\underline{z}')$, $c = c(\underline{z}')$ et $d = D(\underline{z}')$. Comme G vérifie la propriété de Lee-Yang ces coefficients vérifient l'hypothèse de la question précédente, donc pour tout z_i avec $|z_i| < 1$ on a $0 \neq a + dz_i = A(\underline{z}') + D(\underline{z}')z_i = \widehat{P}_{G^{(i,j)}}(z_i, \underline{z}')$. Le graphe contracté vérifie donc la propriété de Lee-Yang.
9. On construit \widetilde{G} en dupliquant chaque sommet de V un nombre de fois égal à son degré dans G (sauf les sommets isolés qui sont laissés tels quels), désignons avec un léger abus de notation i_e les différentes copies du sommet i dans les arêtes e où il est présent, qui sont maintenant des sommets différents dans \widetilde{G} . Les arêtes de \widetilde{G} reproduisent celles de G , chaque sommet de \widetilde{G} étant utilisé au plus une fois. Illustrons la définition par un exemple :



Comme \widetilde{G} se décompose en graphes disjoints (arêtes isolées et sommets isolés) son polynôme se factorise selon :

$$\widehat{P}_{\widetilde{G}} = \prod_{e \in E} (1 + \tau_e z_{i_e} + \tau_e z_{j_e} + z_{i_e} z_{j_e}) \prod_{i \in I} (1 + z_i) , \quad (368)$$

où I désigne les sommets isolés dans G . C'est un produit de termes qui vérifient chacun la propriété de Lee-Yang (c'est évident pour les sommets isolés, pour les arêtes c'est le résultat de la question 4 en supposant les couplages ferromagnétiques), donc \widetilde{G} vérifie cette propriété. Or G peut être obtenu de \widetilde{G} en contractant successivement sur les sommets de \widetilde{G} qui proviennent du même sommet dans G , la propriété étant conservée par contraction elle est donc vraie pour G .

10. La symétrie $Z_G(h) = Z_G(-h)$ provient de l'invariance des couplages à deux corps sous la transformation $\underline{\sigma} \rightarrow -\underline{\sigma}$. En terme des polynômes il est facile de voir qu'elle se traduit par

$$P_G(z) = z^{|V|} P_G\left(\frac{1}{z}\right) . \quad (369)$$

D'après la propriété de Lee et Yang les V racines de P_G sont de module ≥ 1 (si les couplages sont ferromagnétiques). En utilisant la symétrie ci-dessus elles doivent aussi être de module ≤ 1 , la seule possibilité est qu'elles soient de module 1. Comme $z = e^{2\beta h}$ cela correspond à des champs magnétiques imaginaires purs. Finalement Z_G ne peut s'annuler qu'aux racines de P_G (l'exponentielle en facteur n'est jamais nulle), ce qui conclut la preuve.

Analyticité de l'énergie libre

11. La fonction de partition en volume fini Z_Λ est clairement une fonction analytique de h dans tout le plan complexe. D'après le théorème de Lee et Yang elle ne s'annule que sur un ensemble dénombrable de points (les $|\Lambda|$ racines du polynôme associé à ce graphe fini et leurs images avec une périodicité $i\pi/\beta$) inclus dans $i\mathbb{R}$, donc f_Λ est une fonction analytique sur \mathbb{C} privé de ces points (il faudrait aussi vérifier que Z_Λ n'est jamais un réel négatif, de telle sorte que l'on puisse utiliser la branche principale du logarithme), a fortiori sur $\mathbb{C} \setminus i\mathbb{R}$. La convergence simple de $f_\Lambda(\beta, h)$ vers $f(\beta, h)$ peut se démontrer en adaptant la méthode du cours présentée pour $h \in \mathbb{R}$. On montrerait ensuite que f_Λ est bornée uniformément en Λ sur tout compact de $\mathbb{C} \setminus i\mathbb{R}$, ce qui implique que la convergence est uniforme sur tout compact, et donc que la limite est analytique (avec le théorème de Morera).
12. En tant que fonction de h réel f est donc analytique sur $\mathbb{R} \setminus 0$, la seule singularité peut être en $h = 0$. Comme l'énergie libre est de manière générale concave dans les interactions, ici f est concave en h . On peut le retrouver en calculant les dérivées de f_Λ en volume fini :

$$\frac{\partial}{\partial h} f_\Lambda(\beta, h) = -\langle m(\underline{\sigma}) \rangle_\Lambda, \quad \frac{\partial^2}{\partial h^2} f_\Lambda(\beta, h) = -\beta(\langle m(\underline{\sigma})^2 \rangle_\Lambda - \langle m(\underline{\sigma}) \rangle_\Lambda^2) \leq 0, \quad (370)$$

où les moyennes sont prises avec la loi de Gibbs-Boltzmann dans le volume Λ , et $m(\underline{\sigma})$ est l'aimantation par spin. La concavité est préservée par passage à la limite, f est donc une fonction concave de $h \in \mathbb{R}$, ses dérivées à gauche et à droite existent donc en tout point. En présence d'une transition de phase, c'est-à-dire pour $d \geq 2$ et à suffisamment basse température, on a $(D_- f)(0) = m_{\text{sp}}$ et $(D_+ f)(0) = -m_{\text{sp}}$, avec $m_{\text{sp}} > 0$ l'aimantation spontanée.

13. On a vu que dans un volume fini Λ les zéros de la fonction de partition (donc les singularités de l'énergie libre) sont sur l'axe des champs imaginaires. De plus $h = 0$ ne peut annuler Z_Λ , on peut donc définir $\delta_\Lambda > 0$ comme la partie imaginaire du zéro de Lee et Yang le plus proche de l'axe réel. Dans la limite thermodynamique deux comportements distincts sont possibles : si $\delta_\Lambda > \delta$ uniformément en Λ , la limite thermodynamique $f(\beta, h)$ est analytique pour $|\text{Im } h| < \delta$, elle est donc analytique sur l'axe réel et il n'y a pas de transition de phase. Au contraire si $\delta_\Lambda \rightarrow 0$ dans la limite thermodynamique on ne peut relier analytiquement $h > 0$ et $h < 0$, les zéros de Lee et Yang se rapprochant de plus en plus de l'axe réel, ce qui traduit l'apparition d'une aimantation spontanée.
14. Considérons une valeur de h qui annule $Z_N(\beta, h)$. Il faut que les racines λ_\pm vérifient $\left(\frac{\lambda_+}{\lambda_-}\right)^N = -1$. Elles doivent donc être de même module r , on notera $\lambda_\pm = r e^{i\alpha_\pm}$. D'après l'équation de l'énoncé on constate que le produit des racines, $(x^2 - x^{-2})$, est réel, on a donc $\alpha_- = -\alpha_+$. Sans perte de généralité on peut supposer $\alpha_+ \in [0, \pi]$, et de l'équation $e^{2iN\alpha_+} = -1$ on tire $\alpha_+ = \frac{\phi_j}{2}$ avec $\phi_j = \frac{\pi}{N}(2j - 1)$, pour $j \in \{1, \dots, N\}$. En identifiant le produit et la somme des racines d'après l'équation de l'énoncé on obtient

$$\begin{cases} \lambda_+ \lambda_- = r^2 = x^2 - x^{-2} \\ \lambda_+ + \lambda_- = 2r \cos\left(\frac{\phi_j}{2}\right) = x(y + y^{-1}) \end{cases} \Rightarrow y + y^{-1} = 2\sqrt{1 - x^{-4}} \cos\left(\frac{\phi_j}{2}\right) \in \mathbb{R}. \quad (371)$$

D'après le théorème de Lee et Yang $h = i\tilde{h}$ est imaginaire pur, donc $y + y^{-1} = 2\cos(\beta\tilde{h})$. Pour chaque valeur de j on a une solution \tilde{h}_j dans $[0, \pi/\beta]$, selon

$$\cos(\beta\tilde{h}_j) = \sqrt{1 - x^{-4}} \cos\left(\frac{\phi_j}{2}\right), \quad (372)$$

Le zéro le plus proche de l'axe réel correspond à $j = 1$, dans la limite thermodynamique $N \rightarrow \infty$ sa partie imaginaire est donc solution de $\cos(\beta\tilde{h}) = \sqrt{1 - x^{-4}}$. A toute température strictement positive les zéros de Lee et Yang restent donc uniformément éloignés de l'axe réel dans la limite thermodynamique, ce qui est une traduction de l'absence de transition de phase en dimension 1.

Examen de juin 2014 : Inégalités de corrélation

Systèmes de taille finie

On considère un ensemble Λ fini de points, sur lesquels on place des spins d'Ising $\sigma_x \in \{-1, +1\}$. On note $\underline{\sigma} = \{\sigma_x, x \in \Lambda\}$ la configuration globale du système, et l'on définit l'énergie d'une configuration selon

$$H(\underline{\sigma}) = -\frac{1}{2} \sum_{\substack{x,y \in \Lambda \\ x \neq y}} J_{x,y} \sigma_x \sigma_y - \sum_{x \in \Lambda} h_x \sigma_x . \quad (373)$$

Pour un sous-ensemble $X \subset \Lambda$ on notera $\sigma_X = \prod_{x \in X} \sigma_x$ le produit des spins dans ce sous-ensemble de points. Quand X est le singleton $\{x\}$ on écrira indifféremment $\sigma_{\{x\}} = \sigma_x$.

On notera $\langle \bullet \rangle$ les moyennes sur l'ensemble des configurations avec le poids de Boltzmann pour une température inverse β , i.e. pour une fonction $f(\underline{\sigma})$ on écrit

$$\langle f(\underline{\sigma}) \rangle = \sum_{\underline{\sigma} \in \{-1,1\}^\Lambda} f(\underline{\sigma}) \frac{e^{-\beta H(\underline{\sigma})}}{Z}, \quad Z = \sum_{\underline{\sigma} \in \{-1,1\}^\Lambda} e^{-\beta H(\underline{\sigma})} . \quad (374)$$

1. Donner l'interprétation physique des constantes $J_{x,y}$ et h_x . Dans toute la suite on supposera $J_{x,y} \geq 0 \forall x, y$; interpréter cette condition.
2. On suppose $h_x \geq 0 \forall x \in \Lambda$. Montrer que pour tout sous-ensemble $X \subset \Lambda$,

$$\langle \sigma_X \rangle \geq 0 . \quad (375)$$

Indication : on pourra développer en série les exponentielles dans le numérateur de $\langle \sigma_X \rangle$.

3. En utilisant le résultat de la question précédente (supposant donc $h_x \geq 0 \forall x \in \Lambda$), montrer que pour toute paire X, Y de sous-ensembles ($X \subset \Lambda, Y \subset \Lambda$) on a l'inégalité suivante, dite de Griffiths :

$$\langle \sigma_X \sigma_Y \rangle \geq \langle \sigma_X \rangle \langle \sigma_Y \rangle . \quad (376)$$

Indication : on pourra noter que

$$\langle \sigma_X \sigma_Y \rangle - \langle \sigma_X \rangle \langle \sigma_Y \rangle = \frac{1}{Z^2} \sum_{\underline{\sigma}, \underline{\sigma}'} (\sigma_X \sigma_Y - \sigma_X \sigma_Y') e^{-\beta(H(\underline{\sigma}) + H(\underline{\sigma}'))} \quad (377)$$

et faire un changement de variable en introduisant $\tau_x = \sigma_x \sigma_x'$.

4. Commenter les similitudes et les différences entre l'inégalité de Griffiths (376) et celle de Harris vue dans le cours sur la percolation.
5. On considère deux choix $\{h_x^{(1)}, x \in \Lambda\}$ et $\{h_x^{(2)}, x \in \Lambda\}$ des paramètres h_x , et l'on note $\langle \bullet \rangle_1$ et $\langle \bullet \rangle_2$ les moyennes avec ces deux choix (les $J_{x,y}$ étant les mêmes dans les deux cas). Montrer (en utilisant l'inégalité (376)) que

$$0 \leq h_x^{(1)} \leq h_x^{(2)} \quad \forall x \in \Lambda \quad \Rightarrow \quad \langle \sigma_X \rangle_1 \leq \langle \sigma_X \rangle_2 , \quad (378)$$

quelque soit le choix de $X \subset \Lambda$.

6. Par passage à la limite ce résultat est vrai si certains des h_x valent $+\infty$; interpréter cette situation.
7. Comment varie $\langle \sigma_X \rangle$ avec β , pour une valeur donnée des $J_{x,y} \geq 0$ et des $h_x \geq 0$?

On notera \mathcal{C}_X^\uparrow les fonctions dépendants des variables dans un sous-ensemble $X \subset \Lambda$, croissantes pour l'ordre partiel, i.e. $f \in \mathcal{C}_X^\uparrow$ signifie

$$\sigma_x \leq \sigma'_x \quad \forall x \in X \quad \Rightarrow \quad f(\{\sigma_x, x \in X\}) \leq f(\{\sigma'_x, x \in X\}) . \quad (379)$$

On admettra dans la suite l'inégalité suivante, dite FKG (pour Fortuyn, Kasteleyn, Ginibre),

$$f, g \in \mathcal{C}_X^\uparrow \quad \Rightarrow \quad \langle fg \rangle \geq \langle f \rangle \langle g \rangle , \quad (380)$$

valable pour les modèles d'Ising avec $J_{x,y} \geq 0$, et *quelque soit le signe des h_x* . On admettra aussi le corollaire suivant (qui s'obtient à partir de (380) de la même façon que (378) s'obtient à partir de (376)) :

$$f \in \mathcal{C}_X^\uparrow \quad \text{et} \quad h_x^{(1)} \leq h_x^{(2)} \quad \forall x \in \Lambda \quad \Rightarrow \quad \langle f \rangle_1 \leq \langle f \rangle_2 . \quad (381)$$

8. A quelle condition σ_X est-elle une fonction croissante dans le sens défini ci-dessus ?

Applications à la théorie des mesures de Gibbs

On se place maintenant dans le cadre du cours sur les mesures de Gibbs, et l'on considère des mesures de probabilité sur $\Sigma = \{-1, +1\}^{\mathbb{Z}^d}$, les configurations d'un système infini en dimension d . L'interaction sera celle d'Ising avec un couplage $J \geq 0$ entre plus proches voisins et en champ extérieur nul. Rappelons quelques notations et résultats du cours que vous pourrez admettre dans vos réponses :

- Pour une configuration $\underline{\tau} \in \Sigma$ et un volume fini $\Lambda \subset \mathbb{Z}^d$ on note μ_Λ^τ la mesure dans Λ avec condition aux bords $\underline{\tau}$, c'est-à-dire

$$\mu_\Lambda^\tau(\underline{\sigma}) = \begin{cases} \frac{e^{-\beta H_\Lambda^\tau(\underline{\sigma})}}{Z_\Lambda^\tau} & \text{si } \sigma_x = \tau_x \quad \forall x \in \mathbb{Z}^d \setminus \Lambda , \\ 0 & \text{sinon} \end{cases} , \quad (382)$$

avec l'Hamiltonien dans Λ pour la condition aux bords $\underline{\tau}$ et la fonction de partition associée

$$H_\Lambda^\tau(\underline{\sigma}) = -J \sum_{\substack{\langle x,y \rangle \\ x \in \Lambda, y \in \Lambda}} \sigma_x \sigma_y - J \sum_{\substack{\langle x,y \rangle \\ x \in \Lambda, y \notin \Lambda}} \sigma_x \tau_y , \quad Z_\Lambda^\tau = \sum_{\underline{\sigma}_\Lambda} e^{-\beta H_\Lambda^\tau(\underline{\sigma})} , \quad (383)$$

où $\langle x, y \rangle$ dénote les liens entre plus proches voisins sur le réseau.

- On notera μ_Λ^+ cette mesure quand la condition aux bords $\underline{\tau}$ est $\tau_x = +1 \quad \forall x$.
- Par définition de la topologie employée sur l'ensemble \mathcal{M} des mesures de probabilité sur Σ , une suite $\mu_n \in \mathcal{M}$ converge vers $\mu \in \mathcal{M}$ si et seulement si $\mu_n(f)$ converge vers $\mu(f)$ pour toute fonction f de Σ dans \mathbb{R} qui dépend d'un nombre fini de variables.
- On note $\Lambda_n \nearrow \mathbb{Z}^d$ si Λ_n est une suite de parties finies de \mathbb{Z}^d , qui est croissante pour l'inclusion ($\Lambda_n \subset \Lambda_{n+1}$), et qui envahit \mathbb{Z}^d (i.e. $\forall x \in \mathbb{Z}^d \exists N$ tel que $\forall n \geq N$ on a $x \in \Lambda_n$).
- Par des arguments généraux de compacité on avait vu que si $\Lambda_n \nearrow \mathbb{Z}^d$, de toute suite $\mu_{\Lambda_n}^\tau$ on pouvait extraire une sous-suite convergente vers une mesure de Gibbs (pour toute condition aux bords et tout type d'interactions) .

On va voir que les inégalités de corrélation permettent d'obtenir des résultats plus forts dans le cadre du modèle d'Ising.

9. (a) Montrer en utilisant le corollaire (378) de l'inégalité de Griffiths que si $\Lambda' \subset \Lambda$,

$$\mu_\Lambda^+(\sigma_X) \leq \mu_{\Lambda'}^+(\sigma_X) . \quad (384)$$

pour toute partie finie X de \mathbb{Z}^d .

- (b) En déduire que si Λ_n est une suite arbitraire avec $\Lambda_n \nearrow \mathbb{Z}^d$, la suite $\mu_{\Lambda_n}^+$ converge vers une mesure de Gibbs (sans avoir besoin d'extraire une sous-suite).
- (c) Montrer que si Λ_n et Λ'_n sont deux suites de parties finies de \mathbb{Z}^d avec $\Lambda_n \nearrow \mathbb{Z}^d$ et $\Lambda'_n \nearrow \mathbb{Z}^d$, alors $\mu_{\Lambda_n}^+$ et $\mu_{\Lambda'_n}^+$ convergent vers la même limite.

On notera μ^+ la limite des $\mu_{\Lambda_n}^+$ quand $\Lambda_n \nearrow \mathbb{Z}^d$, dont on vient de montrer l'existence et l'indépendance par rapport au choix de Λ_n ; cette dernière propriété implique que μ^+ est invariante par translation (on ne demande pas de preuve de cette affirmation). Avec le même type de raisonnement on montrerait que si la condition aux bords est $\tau_x = -1 \forall x$ les suites $\mu_{\Lambda_n}^-$ convergent vers une mesure μ^- invariante par translation, indépendamment de la suite Λ_n .

10. (a) On note $m_{\text{sp}}(\beta) = \mu^+(\sigma_0)$ l'aimantation moyenne d'un spin sous la mesure μ^+ (en explicitant la dépendance en β). Montrer que $m_{\text{sp}}(\beta)$ est une fonction croissante de β .
- (b) Que vaut $\mu^-(\sigma_0)$?
- (c) Rappeler brièvement les résultats du cours obtenus par les méthodes de Dobrushin et de Peierls (on supposera $d \geq 2$), et leurs conséquences sur la fonction $m_{\text{sp}}(\beta)$.
11. On va maintenant montrer, à partir du corollaire (381) de l'inégalité FKG, que μ^+ est une mesure extrême.

- (a) Montrer que si Λ est un volume fini de \mathbb{Z}^d , $X \subset \Lambda$, $f \in \mathcal{C}_X^\uparrow$ et $\underline{\tau}, \underline{\tau}'$ sont deux configurations avec $\tau_x \leq \tau'_x \forall x \in \mathbb{Z}^d$, alors

$$\mu_{\Lambda}^{\underline{\tau}}(f) \leq \mu_{\Lambda}^{\underline{\tau}'}(f). \quad (385)$$

- (b) Soit ν une mesure de Gibbs pour l'interaction d'Ising, et X une partie finie de \mathbb{Z}^d . Montrer que pour toute fonction $f \in \mathcal{C}_X^\uparrow$,

$$\nu(f) \leq \mu^+(f). \quad (386)$$

- (c) En déduire que μ^+ est extrême au sein des mesures de Gibbs, i.e. que l'on ne peut écrire $\mu^+ = \alpha\mu_1 + (1 - \alpha)\mu_2$ avec $\alpha \in]0, 1[$ et $\mu_1 \neq \mu_2$ deux mesures de Gibbs distinctes.

De la même façon on montre que μ^- est extrême et que pour n'importe quelle mesure de Gibbs ν on a la généralisation naturelle de (386) pour toute fonction $f \in \mathcal{C}_X^\uparrow$:

$$\mu^-(f) \leq \nu(f) \leq \mu^+(f). \quad (387)$$

12. On va finalement montrer que $m_{\text{sp}}(\beta) = 0$ est une condition suffisante pour assurer l'unicité des mesures de Gibbs (la réciproque étant évidemment vraie).

- (a) On note $\rho_x = \frac{1+\sigma_x}{2}$ et $\rho_X = \prod_{x \in X} \rho_x$ pour une partie finie X de \mathbb{Z}^d . Montrer que la fonction

$$f = \sum_{x \in X} \rho_x - \rho_X \quad (388)$$

est croissante.

- (b) En déduire, en utilisant (387), un encadrement de $\mu^+(\rho_X) - \mu^-(\rho_X)$.
- (c) Conclure que si $m_{\text{sp}}(\beta) = 0$ alors il n'existe qu'une mesure de Gibbs.

Corrigé de l'examen de juin 2014

Systèmes de taille finie

1. $J_{x,y}$ représente l'intensité de l'interaction entre les spins σ_x et σ_y , h_x celle du champ magnétique extérieur que subit le spin σ_x . La condition $J_{x,y} \geq 0$ signifie que les interactions sont ferromagnétiques, les spins ont tendance à s'aligner dans la même direction pour minimiser l'énergie.
2. La fonction de partition étant évidemment positive, il suffit de montrer que

$$\sum_{\underline{\sigma}} \sigma_X \prod_{\substack{x,y \in \Lambda \\ x \neq y}} e^{\frac{1}{2} \beta J_{x,y} \sigma_x \sigma_y} \prod_{x \in \Lambda} e^{\beta h_x \sigma_x} \geq 0. \quad (389)$$

En développant en série les exponentielles cette expression se décompose comme une somme de termes proportionnels (à des facteurs combinatoires positifs près) à

$$\prod_{\substack{x,y \in \Lambda \\ x \neq y}} \left(\frac{1}{2} \beta J_{x,y} \right)^{n_{x,y}} \prod_{x \in \Lambda} (\beta h_x)^{n_x} \sum_{\underline{\sigma}} \prod_{x \in X} \sigma_x \prod_{\substack{x,y \in \Lambda \\ x \neq y}} (\sigma_x \sigma_y)^{n_{x,y}} \prod_{x \in \Lambda} \sigma_x^{n_x}, \quad (390)$$

où les n sont des entiers positifs ou nuls. Comme les $\beta J_{x,y}$ et βh_x sont positifs, le facteur devant la somme l'est aussi. Par ailleurs la sommation sur $\underline{\sigma}$ est positive : elle vaut $2^{|\Lambda|}$ si chacun des spins σ_x apparaît un nombre pair de fois, 0 sinon.

3. En suivant l'indication de l'énoncé,

$$\begin{aligned} \langle \sigma_X \sigma_Y \rangle - \langle \sigma_X \rangle \langle \sigma_Y \rangle &= \frac{1}{Z^2} \sum_{\underline{\sigma}, \underline{\sigma}'} (\sigma_X \sigma_Y - \sigma_X \sigma'_Y) e^{\frac{1}{2} \sum_{x,y} \beta J_{x,y} (\sigma_x \sigma_y + \sigma'_x \sigma'_y) + \sum_x \beta h_x (\sigma_x + \sigma'_x)} \\ &= \frac{1}{Z^2} \sum_{\underline{\sigma}, \underline{\sigma}'} \sigma_X \sigma_Y (1 - \sigma_Y \sigma'_Y) e^{\frac{1}{2} \sum_{x,y} \beta J_{x,y} \sigma_x \sigma_y (1 + \sigma_x \sigma'_x \sigma_y \sigma'_y) + \sum_x \beta h_x \sigma_x (1 + \sigma_x \sigma'_x)} \\ &= \frac{1}{Z^2} \sum_{\underline{\sigma}, \underline{\tau}} \sigma_X \sigma_Y (1 - \tau_Y) e^{\frac{1}{2} \sum_{x,y} \beta J_{x,y} \sigma_x \sigma_y (1 + \tau_x \tau_y) + \sum_x \beta h_x \sigma_x (1 + \tau_x)}. \end{aligned} \quad (391)$$

En séparant les deux sommations on note que cette expression vaut

$$\frac{1}{Z^2} \sum_{\underline{\tau}} (1 - \tau_Y) \sum_{\underline{\sigma}} \sigma_X \sigma_Y e^{\frac{1}{2} \sum_{x,y} \beta J_{x,y} (1 + \tau_x \tau_y) \sigma_x \sigma_y + \sum_x \beta h_x (1 + \tau_x) \sigma_x}. \quad (392)$$

A $\underline{\tau}$ fixé la somme sur $\underline{\sigma}$ est positive d'après la question précédente : elle est en effet proportionnelle à $\langle \sigma_{X \Delta Y} \rangle$ (ici Δ désigne la différence symétrique) pour un modèle où les couplages sont $J_{x,y}(1 + \tau_x \tau_y)$ et les champs $h_x(1 + \tau_x)$, bien positifs quelque soit $\underline{\tau}$. Comme $1 - \tau_Y$ est aussi positif, c'est bien le cas de toute l'expression.

4. L'inégalité de Harris s'écrit $\langle fg \rangle \geq \langle f \rangle \langle g \rangle$, elle a donc la même forme que l'inégalité de Griffiths et exprime que deux variables aléatoires sont corrélées positivement. Cependant les hypothèses de validité sont différentes dans les deux cas : pour Harris on suppose les variables élémentaires (ici les spins) indépendantes, ce qui n'est pas le cas pour Griffiths dès que les couplages $J_{x,y}$ sont non-nuls. Par ailleurs l'inégalité de Harris suppose les fonctions f et g croissantes, ici on considère une famille particulière de fonctions σ_X .
5. On peut remarquer que

$$\frac{\partial}{\partial h_x} \langle \sigma_X \rangle = \beta (\langle \sigma_X \sigma_x \rangle - \langle \sigma_X \rangle \langle \sigma_x \rangle) \geq 0, \quad (393)$$

l'inégalité provenant du résultat de la question 3. Autrement dit augmenter un champ magnétique fait augmenter les moyennes des produits de spin. En modifiant un champ à la fois pour passer des champs $h_x^{(1)}$ aux champs $h_x^{(2)}$ les moyennes $\langle \sigma_X \rangle$ augmentent à chaque étape, d'où l'inégalité.

6. Un champ h_x égal à $+\infty$ correspond à fixer strictement $\sigma_x = +1$ dans la loi de probabilité de Gibbs-Boltzmann, les configurations avec $\sigma_x = -1$ ayant un poids nul.

7. Comme ci-dessus

$$\frac{\partial}{\partial J_{x,y}} \langle \sigma_X \rangle = \beta (\langle \sigma_X \sigma_x \sigma_y \rangle - \langle \sigma_X \rangle \langle \sigma_x \sigma_y \rangle) \geq 0, \quad (394)$$

i.e. augmenter un couplage fait augmenter les moyennes des produits de spin. Comme la dépendance en β de $\langle \sigma_X \rangle$ est entièrement par l'intermédiaire des produits $\beta J_{x,y}$ et βh_x , on conclut que $\langle \sigma_X \rangle$ est croissante en β .

8. σ_X est croissante si et seulement si $|X| = 1$. En effet σ_x est bien croissante, et pour $|X| \geq 2$ on trouve facilement des configurations qui violent l'inégalité, par exemple $\underline{\sigma} = (-1, -1, +1, \dots, +1)$ et $\underline{\sigma}' = (-1, +1, +1, \dots, +1)$.

Applications à la théorie des mesures de Gibbs

9. (a) Les spins hors de Λ sont fixés à $+1$ dans les deux membres de l'équation, on peut donc sans perte de généralité supposer $X \subset \Lambda$. On est donc exactement dans le cadre de la première partie du problème, avec $\mu_\Lambda^+(\sigma_X) = \langle \sigma_X \rangle_1$ et $\mu_{\Lambda'}^+(\sigma_X) = \langle \sigma_X \rangle_2$ et des champs $h_x^{(1)}$ égaux à J multiplié par le nombre de voisins de x dans $\mathbb{Z}^d \setminus \Lambda$, alors que $h_x^{(2)} = +\infty$ si $x \in \Lambda \setminus \Lambda'$ et $h_x^{(2)} = h_x^{(1)}$ si $x \in \Lambda'$. On peut donc utiliser le résultat de la question 5 pour conclure.

(b) Considérons $\Lambda_n \nearrow \mathbb{Z}^d$, une partie finie X de \mathbb{Z}^d , et la suite $\mu_{\Lambda_n}^+(\sigma_X)$. D'après la réponse à la question précédente cette suite est décroissante, elle admet donc une limite. Comme toute fonction f dépendant d'un nombre fini de spins d'Ising peut s'écrire comme une combinaison linéaire de produits σ_X d'un nombre fini de spins, cela implique que $\mu_{\Lambda_n}^+(f)$ converge pour toute fonction f dépendant d'un nombre fini de variables. D'après la caractérisation de la convergence des mesures rappelée dans l'énoncé ceci est suffisant pour conclure à la convergence de la suite $\mu_{\Lambda_n}^+$.

(c) Comme $\Lambda_n \nearrow \mathbb{Z}^d$ et $\Lambda'_n \nearrow \mathbb{Z}^d$ on peut construire deux suites croissantes n_p et n'_p telles que

$$\Lambda_{n_1} \subset \Lambda'_{n'_1} \subset \Lambda_{n_2} \subset \Lambda'_{n'_2} \subset \Lambda_{n_3} \subset \dots \quad (395)$$

Pour une partie finie X de \mathbb{Z}^d notons $b_p = \mu_{\Lambda_{n_p}}^+(\sigma_X)$ et $b'_p = \mu_{\Lambda'_{n'_p}}^+(\sigma_X)$. D'après (395) et la réponse à la question 9a on a

$$b_1 \geq b'_1 \geq b_2 \geq b'_2 \geq b_3 \geq \dots, \quad (396)$$

les deux suites b_p et b'_p ont donc la même limite. Comme ci-dessus on étend le résultat par combinaison linéaire : pour toute fonction dépendant d'un nombre fini de variables $\lim \mu_{\Lambda_n}^+(f) = \lim \mu_{\Lambda'_n}^+(f)$, et comme c'est vrai quelque soit f on a bien $\lim \mu_{\Lambda_n}^+ = \lim \mu_{\Lambda'_n}^+$.

10. (a) $\mu_\Lambda^+(\sigma_0)$ est une fonction croissante de β d'après la question 7, en prenant $X = \{0\}$. La monotonie est préservée dans la limite $\Lambda \nearrow \mathbb{Z}^d$.

(b) Par symétrie $\mu^-(\sigma_0) = -m_{\text{sp}}(\beta)$.

(c) D'après Dobrushin il existe une température inverse β_1 telle que pour $\beta < \beta_1$ il n'existe qu'une mesure de Gibbs. Dans ce cas $\mu^+ = \mu^-$, on a donc $m_{\text{sp}}(\beta) = 0$ pour $\beta < \beta_1$. A l'inverse le résultat de Peierls a montré l'existence (en dimension $d \geq 2$) d'une température inverse β_2 telle que $m_{\text{sp}}(\beta) > 0$ pour $\beta > \beta_2$. Avec en plus la monotonie qui vient de Griffiths, il ne peut y avoir qu'une seule température critique qui sépare deux régimes avec ou sans aimantation spontanée.

11. (a) C'est une application du résultat de l'énoncé, les champs imposés sur les spins dans Λ par la condition aux bords $\underline{\tau}'$ étant supérieurs à ceux de la condition aux bords $\underline{\tau}$.

(b) Comme ν est une mesure de Gibbs, $\nu(f_\Lambda) = \nu(f)$, où comme dans le cours $f_\Lambda(\underline{\tau})$ est la fonction moyennée dans Λ avec la condition aux bords $\underline{\tau}$. Si l'on choisit Λ suffisamment grand pour que $X \subset \Lambda$, on a $f_\Lambda(\underline{\tau}) = \mu_\Lambda^\tau(f) \leq \mu_\Lambda^+(f)$ quelque soit $\underline{\tau}$, d'après la question précédente. Donc $\nu(f) = \nu(f_\Lambda) \leq \mu_\Lambda^+(f)$. Ceci est vrai quelque soit Λ incluant X , on peut prendre la limite $\Lambda \nearrow \mathbb{Z}^d$ pour conclure.

(c) D'après la question précédente on a $\mu_1(f) \leq \mu^+(f)$ et $\mu_2(f) \leq \mu^+(f)$. Pour que $\alpha\mu_1(f) + (1 - \alpha)\mu_2(f)$ soit égale à $\mu^+(f)$ avec $\alpha \in]0, 1[$ il faut donc que $\mu_1(f) = \mu_2(f) = \mu^+(f)$. Comme c'est vrai quelque soit X et $f \in \mathcal{C}_X^\uparrow$ cela implique $\mu_1 = \mu_2 = \mu^+$ (car toute fonction dépendant d'un nombre fini de variables peut se décomposer comme une combinaison linéaire de fonctions croissantes), il n'existe donc pas de décomposition non-triviale de μ^+ .

On a vu en cours que les mesures de Gibbs extrémales sont celles qui sont triviales à l'infini. On peut comprendre intuitivement que μ^+ soit triviale à l'infini, car elle donne un poids 1 seulement aux évènements contenant les configurations qui sont « égales à +1 à l'infini », et 0 sinon.

12. (a) Les ρ_x sont évidemment croissantes, ainsi que leur somme, la difficulté pourrait donc provenir de la soustraction de ρ_X elle-même croissante. Mais pour que ρ_X augmente strictement il faut qu'au moins un des ρ_x ait strictement augmenté, ce qui compense le premier effet. Plus explicitement, considérons deux configurations $\underline{\sigma}, \underline{\sigma}'$ avec $\sigma_x \leq \sigma'_x$ pour tout $x \in X$, notons ρ_x, ρ_X et f pour la première configuration, et ρ'_x, ρ'_X et f' pour la deuxième. Deux cas sont à considérer : soit $\rho_X = \rho'_X$, et alors $f \leq f'$. Soit $\rho_X < \rho'_X$, et alors la seule possibilité est $\rho_X = 0$ (il existe donc un $\rho_x = 0$) et $\rho'_X = 1$ (et donc $\rho'_x = 1$ pour tous les x), on vérifie donc $f \leq f'$ aussi dans ce cas-là.

(b) Comme f est croissante on a $\mu^+(f) \geq \mu^-(f)$, soit

$$\mu^+ \left(\sum_{x \in X} \rho_x \right) - \mu^- \left(\sum_{x \in X} \rho_x \right) \geq \mu^+(\rho_X) - \mu^-(\rho_X) . \quad (397)$$

Comme ρ_X est elle-même croissante, le membre de droite de cette inégalité est positif. Par ailleurs, en utilisant l'invariance par translation de μ^+ et μ^- , ainsi que la définition de m_{sp} , il vient

$$|X|m_{\text{sp}}(\beta) \geq \mu^+(\rho_X) - \mu^-(\rho_X) \geq 0 . \quad (398)$$

(c) Si $m_{\text{sp}}(\beta) = 0$ on a donc $\mu^+(\rho_X) = \mu^-(\rho_X)$, et donc $\nu(\rho_X) = \mu_+(\rho_X)$ pour toute mesure de Gibbs ν . Comme toute fonction dépendant d'un nombre fini de variables peut se décomposer comme une combinaison linéaire de ρ_X , cela implique $\nu = \mu_+$, il n'existe donc qu'une mesure de Gibbs si l'aimantation spontanée s'annule.