



HAL
open science

Applications d'inégalités fonctionnelles à la mécanique statistique et au recuit simulé

Pierre-André Zitt

► **To cite this version:**

Pierre-André Zitt. Applications d'inégalités fonctionnelles à la mécanique statistique et au recuit simulé. Mathématiques [math]. Université de Nanterre - Paris X, 2006. Français. NNT: . tel-00114033

HAL Id: tel-00114033

<https://theses.hal.science/tel-00114033>

Submitted on 8 Jan 2007

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

UNIVERSITÉ PARIS X – NANTERRE
École doctorale Connaissance, Langage et Modélisation
Équipe Modal'X

Doctorat de Mathématiques

**Applications d'inégalités fonctionnelles à la
mécanique statistique et au recuit simulé**

Pierre-André ZITT

Thèse soutenue le 6 décembre 2006 devant le jury composé de :

Franck BARTHE	Examineur
Thierry BODINEAU	Rapporteur
Patrick CATTIAUX	Directeur de thèse
Christian LÉONARD	Examineur
Laurent MICLO	Rapporteur
Li Ming WU	Examineur

**Applications d'inégalités fonctionnelles à la mécanique statistique et au
recuit simulé**

Pierre-André ZITT
6 décembre 2006

Remerciements

Je tiens tout d'abord à remercier mon directeur de thèse, Patrick Cattiaux, qui a réussi à me rendre intelligible et intéressant le monde étrange des inégalités fonctionnelles. Ses nombreux conseils ont souvent débloqué des pistes qui me paraissaient sans issue, et ont aidé à améliorer la rédaction de ce mémoire.

Les explications de Thierry Bodineau sur de multiples résultats existant dans de nombreux modèles de mécanique statistique, les corrections attentives et les nombreuses suggestions de Laurent Miclo ont beaucoup apporté à cette thèse. Qu'ils en soient remerciés, ainsi que d'avoir accepté la tâche ingrate de rapporter cette thèse.

Je remercie également Franck Barthe, Christian Léonard et Li-Ming Wu d'avoir bien voulu faire partie de mon jury.

J'ai pu assister avant et pendant cette thèse à de nombreuses écoles d'été ou d'hiver, pendant lesquelles j'ai appris de nombreuses choses, proches de mon sujet ou plus lointaines, mais toujours très stimulantes. Je profite donc de ces lignes pour en remercier les organisateurs et intervenants, ainsi que l'équipe Modal'X qui m'a permis d'y participer.

Je remercie aussi cette équipe pour son accueil durant ces trois années — rendre accueillants les couloirs de l'entresol n'était pourtant pas une mince affaire. Les différentes générations d'habitants du mythique bureau E12 en particulier ont réussi à supporter mes plaisanteries, mes multiples preuves de l'inconsistance de ZFC, et n'écoutant que leur courage, m'ont même parfois invité à manger du chocolat et des salades à la fraise.

Merci aussi à mes proches, qui m'ont chacun à leur manière aidé et soutenu, pendant ces années, qui en rêve dans les escaliers de l'X, qui par des explications lumineuses sur le modèle M. standard, qui par son interprétation toute personnelle du recuit simulé, ...

Merci enfin à ma géographe préférée, pour tellement de choses qu'elles ne rentreraient pas entre ces deux marges...

★ ★ ★

La réalisation de ce mémoire a fait appel à de nombreux logiciels, tous disponibles librement ; merci donc sur le plan technique aux auteurs de $\text{T}_\text{E}\text{X}$ et $\text{L}_\text{A}\text{T}_\text{E}\text{X}$ pour le texte, de gnuplot, MetaPost, TikZ et ruby pour les dessins, et de svn et vim pour la gestion de versions et l'édition.

Chapitre 1

Introduction

I am looking over *Self-Control* again, & my opinion is confirmed of its being an excellently-meant, elegantly-written work, without anything of Nature or Probability in it.

Jane Austen, à propos de *Self-Control; a Novel* (M. Brunton, 1810).

There is no Algebraist nor Mathematician so expert in his science, as to place entire confidence in any truth immediately upon his discovery of it, or regard it as any thing, but a mere probability.

David Hume, *A Treatise on Human Nature*

L'objet de ce mémoire est de présenter quelques résultats concernant l'application d'inégalités fonctionnelles à deux modèles, l'un provenant de la mécanique statistique, l'autre d'un algorithme d'optimisation.

Dans cette introduction, nous présenterons d'abord les outils d'analyse fonctionnelle, et la façon dont ils sont apparus. Nous verrons ensuite pourquoi ces outils apparaissent naturellement dans l'étude de l'algorithme du recuit simulé. Enfin, nous esquisserons un historique des questions de mécanique statistique autour du modèle d'Ising pour motiver l'utilisation de techniques analytiques. Dans ces deux dernières sections, nous présenterons brièvement, puis tenterons de mettre en perspective nos principaux résultats.

Quelques notations

Tout au long de cette thèse, nous adopterons les notations suivantes. L'espérance sous une mesure de probabilité μ sera notée \mathbf{E}_μ , ou \mathbf{E} quand il n'y aura pas de confusion possible. Pour une fonction f μ -intégrable, on notera son intégrale sous l'une des formes suivantes :

$$\mathbf{E}_\mu[f] = \int f(x) d\mu(x) = \int f d\mu = \mu(f).$$

De même, pour $f \in L^2(\mu)$, on notera $\mathbf{Var}_\mu(f)$ la variance d'une fonction f sous μ : si une variable aléatoire X suit la loi μ , la variance de f est la variance de la variable aléatoire $f(X)$.

$$\mathbf{Var}_\mu(f) = \mathbf{E}_\mu \left[(f - \mathbf{E}_\mu[f])^2 \right] = \int (f - \mu f)^2 d\mu.$$

Pour f positive, on notera $\mathbf{Ent}_\mu(f)$ l'entropie de f sous la mesure μ , *i.e.* la quantité

$$\mathbf{Ent}_\mu(f) = \int f \log f d\mu - \left(\int f d\mu \right) \log \left(\int f d\mu \right),$$

où par convention, $\mathbf{Ent}_\mu(f) = \infty$ si $f \times \sup(\log(f), 0)$ n'est pas intégrable.

Le gradient et le laplacien usuel sur \mathbb{R}^d apparaîtront souvent, ils seront notés respectivement ∇ et Δ .

Enfin, dans la partie de mécanique statistique, les indices et exposants auront tendance à se multiplier, et il ne sera pas rare de croiser des objets comme $\mu_\Lambda^{z, \beta}$; nous omettrons parfois certains des indices, ou les passeront en paramètre de la fonction pour souligner la dépendance par rapport à une certaine quantité.

Les autres notations seront introduites au fur et à mesure.

1.1 À propos des inégalités fonctionnelles

Nature is unfair ? So much the better,
inequality is the only bearable thing, the
monotony of equality can only lead us to
boredom.

Francis Picabia

Ce sont les inégalités qu'on aime, tout en
prêchant l'égalité.

Louis de Bonald, Oeuvres complètes
tome 3, Pensées

Introduction Les inégalités fonctionnelles ne sont pas un sujet neuf : en témoigne le nom même de l'une des plus connues, l'inégalité de Poincaré. Si elles sont d'abord apparues dans des problèmes d'analyse — elles peuvent souvent s'interpréter comme preuves d'une injection continue entre deux espaces fonctionnels — elles ont trouvé une application naturelle dans le domaine des probabilités. Ce que nous présentons ici n'est qu'un bref aperçu du champ de ces applications ; pour une vue plus détaillée, on pourra consulter le panorama [ABC⁺00] et ses notes bibliographiques abondantes, auxquelles cette présentation doit beaucoup, et les nombreux cours disponibles sur le sujet (dans l'ordre chronologique, [Bak94, Roy99, GZ03]).

Pour mieux comprendre l'intérêt de ces inégalités en probabilités, on peut partir de la question naturelle du comportement en temps long d'un processus stochastique. Notre exemple fondamental de processus, que l'on verra réapparaître tout au long de cette thèse, est une diffusion avec dérive sur \mathbb{R}^d , appelée aussi *diffusion de Langevin*, et définie par l'équation différentielle stochastique (ÉDS) suivante :

$$dX_t = dB_t - \frac{1}{2} \nabla V(X_t) dt, \quad (1.1)$$

où V est une fonction réelle régulière appelée *potentiel*. Ce processus est donc gouverné par deux termes, un mouvement brownien et une dérive déterministe suivant le gradient d'un certain potentiel V .

Remarque. Cette équation est parfois écrite sans facteur $1/2$ devant le potentiel, ce qui rajoute un facteur 2 dans l'expression de la mesure invariante.

Cette diffusion admet des mesures invariantes, proportionnelles à $\exp(-V)$, et si ces mesures sont finies il y a donc une mesure de probabilité invariante ; en fait elle est même réversible pour le processus. Il est alors naturel d'étudier la convergence du processus vers cette mesure invariante. Partant d'un point x , on veut donc savoir si la loi de X_t est proche de μ , soit, en testant contre une fonction f , si $\mathbf{E}_x[f(X_t)] = P_t f(x)$ est proche de $\mu(f)$. On fait cette comparaison dans l'espace $L^2(d\mu)$, la quantité qui nous intéresse est donc finalement :

$$\mathbf{Var}_\mu(P_t f) = \mathbf{E}_\mu[(P_t f - \mu f)^2]. \quad (1.2)$$

La dérivation (au moins formelle) de cette quantité fait apparaître à droite le générateur L du semi-groupe ; le cadre étant simple, ce générateur peut être explicité ($L = (1/2) (\Delta \cdot - \nabla V \nabla \cdot)$). Un

rapide calcul d'intégration par parties donne, pour toute fonction g régulière et à croissance suffisamment lente à l'infini,

$$-\int gLg d\mu = \frac{1}{2} \int |\nabla g|^2 d\mu.$$

Si l'on dérive l'expression (1.2) de la variance, on obtient donc :

$$\frac{d}{dt} (\mathbf{Var}_\mu(P_t f)) = -2\mathbf{E}[P_t f L P_t f] = - \int |\nabla P_t f|^2 d\mu. \quad (1.3)$$

C'est à ce moment qu'intervient l'inégalité de Poincaré. Rappelons sa définition sans entrer dans les détails techniques.

Définition. *La mesure μ vérifie une inégalité de Poincaré si il existe une constante C_P telle que, pour toute fonction g ,*

$$\mathbf{Var}_\mu(g) \leq \frac{C_P}{2} \int |\nabla g|^2 d\mu.$$

Remarque. *Là encore, le facteur 1/2 est une convention, pour retrouver à droite la quantité $\frac{1}{2} \int |\nabla g|^2 = - \int gLg$, appelée énergie de g .*

Si l'on reprend la dérivation (1.3), et si la mesure invariante μ vérifie une inégalité de Poincaré, on peut l'appliquer avec $g = P_t f$ pour obtenir :

$$\frac{d}{dt} (\mathbf{Var}_\mu(P_t f)) \leq -\frac{2}{C_P} \mathbf{Var}_\mu(P_t f),$$

inégalité différentielle que l'on intègre entre 0 et t pour obtenir la *décroissance exponentielle du semi-groupe* dans $L^2(d\mu)$:

$$\mathbf{Var}_\mu(P_t f) \leq \exp\left(-\frac{2t}{C_P}\right) \mathbf{Var}_\mu(f). \quad (1.4)$$

L'inégalité fonctionnelle sur la *mesure invariante* du processus nous a donc donné des informations quantitatives sur le *comportement dynamique* du semi-groupe.

★ ★ ★

Cette étude dans L^2 a des avantages certains — en particulier, l'inégalité de Poincaré se traduit immédiatement en propriétés *spectrales* sur le générateur de la diffusion, et tous les outils de théorie spectrale sont alors disponibles (*cf.* par exemple [RS80], ou le cours en ligne [Hel03]).

Malheureusement, ce point de vue peut parfois être insuffisant, et on a parfois besoin de convergences plus fortes. On peut en particulier s'intéresser à l'éventuelle convergence en *variation totale* de la loi de X_t vers sa mesure d'équilibre. Rappelons la définition de la distance en variation totale :

1 Définition. *Soit (E, τ, \mathcal{B}) un espace topologique muni de sa tribu borélienne, et soit μ, ν deux mesures sur E . La distance en variation totale entre μ et ν est donnée par :*

$$\|\mu - \nu\|_{vt} = \sup_{A \in \mathcal{B}} |\mu(A) - \nu(A)|.$$

La notation vient du fait que $\|\cdot\|_{vt}$ définit une norme sur l'espace vectoriel des mesures signées sur E .

Cette distance peut être contrôlée par un terme de variance (ce qui permet d'utiliser les propriétés spectrales déjà évoquées), mais on peut également utiliser une entropie.

Proposition (Inégalité de Pinsker). *Pour tout couple de mesures μ, ν , l'entropie relative contrôle la distance en variation :*

$$\|\mu - \nu\|_{vt} \leq \sqrt{2H(\mu|\nu)} = \sqrt{2\mathbf{Ent}_\nu\left(\frac{d\mu}{d\nu}\right)}.$$

On peut alors se demander comment évolue l'entropie relative par rapport à la loi d'équilibre. Supposons que la loi du point d'origine admette la densité $f = f_0$ par rapport à μ . Il est alors facile de voir que la loi à l'instant t a la densité $f_t = P_t f$ par rapport à μ ; en effet, pour toute fonction g ,

$$\begin{aligned} \mathbf{E}[g(X_t)] &= \int \mathbf{E}_x[g(X_t)]f(x)\mu(dx) = \int P_t g(x)f(x)\mu(dx) \\ &= \int g(x)P_t f(x)\mu(dx) \end{aligned} \quad (\text{car } \mu \text{ est réversible})$$

La quantité intéressante à étudier devient donc l'entropie de $P_t f$, quand t varie. De la même façon qu'auparavant, la dérivation de cette entropie par rapport au temps fait intervenir naturellement une inégalité fonctionnelle. Ici, pour assurer la convergence exponentielle de notre quantité $\mathbf{Ent}(P_t f)$, nous avons besoin non plus de l'inégalité de Poincaré, mais d'une *inégalité de Sobolev logarithmique*.

2 Définition. *On dit que μ vérifie une inégalité de Sobolev logarithmique si, pour toute fonction f ,*

$$\mathbf{Ent}_\mu(f^2) \leq \frac{C_{LS}}{2} \int |\nabla f|^2 d\mu.$$

Pour une preuve détaillée du lien avec la convergence exponentielle de l'entropie, nous renvoyons à [ABC⁺00], section 2.6.

Cette inégalité est apparue plus récemment que la précédente; on la fait habituellement remonter à l'article de Gross [Gro75] (et on l'appelle parfois *inégalité de Gross*).

Ces deux inégalités ont donné naissance à une vaste littérature, et nous ne ferons ici qu'esquisser différentes directions de recherche.

Établir les inégalités, chercher les constantes optimales Comme les inégalités de Poincaré ou Sobolev logarithmiques peuvent être ou ne pas être valables, avec des constantes particulières, il est naturel de se demander quelles mesures les vérifient. Pour cela, on cherchera aussi des critères plus ou moins généraux permettant de les établir.

Si l'inégalité de Poincaré était déjà connue en 1975 pour certaines mesures (comme les mesures uniformes sur des ouverts réguliers et bornés de \mathbb{R}^d , dont nous reparlerons plus bas à propos du recuit simulé), ce n'était pas le cas pour l'inégalité de Gross. Les deux premiers exemples de mesure vérifiant cette inégalité sont les deux mesures les plus classiques de la théorie des probabilités.

Théorème. *La mesure de Bernoulli (pour un paramètre quelconque) et la mesure gaussienne standard sur \mathbb{R}^d vérifient une inégalité de Sobolev logarithmique; les constantes optimales sont connues. En particulier, pour la loi gaussienne et la loi de Bernoulli symétrique, la constante vaut 2.*

Notons que la valeur de la constante pour la loi de Bernoulli dépend en fait de la définition donnée à *l'énergie*, c'est-à-dire au choix que l'on fait pour définir le gradient discret dans un espace à deux points — pour plus de détails, les preuves et les références, cf. le chapitre 1 de [ABC⁺00].

Une fois que l'on connaît une mesure vérifiant l'une ou l'autre des inégalités fonctionnelles citées, on obtient « gratuitement » de nombreux autres exemples, en utilisant les propriétés suivantes.

Théorème. Tensorisation : *Si deux mesures μ et ν vérifient l'inégalité de Poincaré, alors $\mu \otimes \nu$ la vérifie également, avec la constante :*

$$C_{\mu \otimes \nu} = \max(C_\mu, C_\nu).$$

Perturbation : *Si la mesure $\mu = \exp(-V)$ sur \mathbb{R}^n vérifie l'inégalité de Poincaré, alors $\nu = \exp(-V + W)$ la vérifie également, dès que W est bornée.*

Le même résultat est valable pour l'inégalité de Sobolev logarithmique.

La propriété de tensorisation, déjà présente dans l'article de Gross, est fondamentale ; elle permet en particulier de voir ces inégalités comme infini-dimensionnelles (puisque si μ vérifie une des deux inégalités, $\mu^{\otimes n}$ la vérifie aussi, avec une constante indépendante de n).

Ces propriétés donnent déjà un catalogue de mesures vérifiant de bonnes inégalités, mais d'autres critères ont été développés. Citons en particulier le critère de courbure Γ_2 de Bakry et Émery ([BÉ85]), qui a permis d'obtenir des inégalités de Sobolev logarithmique dans de très nombreux cas, en particulier en mécanique statistique (cf. la suite de cette introduction et la partie I de la thèse).

Dans le cas unidimensionnel (des mesures sur \mathbb{R}), on peut trouver des critères explicites pour l'inégalité de Poincaré. Ces critères découlent des travaux de Muckenhoupt ([Muc72]) sur les inégalités de Hardy. Plus récemment, S. Bobkov et F. Götze ont établi dans [BG99] un critère analogue pour l'inégalité de Sobolev logarithmique. Une première approche de ces résultats se trouve dans [ABC⁺00], chapitre 6.

Des critères du même style ont ensuite été développés pour toutes sortes d'inégalités fonctionnelles : nous en verrons un exemple dans la partie II.

★ ★ ★

Établir une inégalité fonctionnelle est une chose, trouver la constante optimale (ou tout au moins une bonne approximation) en est une autre. Si les critères unidimensionnels évoqués précédemment donnent par exemple de bonnes constantes (*i.e.* qui ne s'éloignent de la valeur optimale qu'à un facteur universel près), d'autres techniques, comme la perturbation, n'ont pas cette qualité. Une bonne connaissance des constantes est pourtant souvent nécessaire en vue des applications, comme nous le verrons dans les deux parties de cette thèse.

Pour donner une idée de la difficulté du problème, considérons un ouvert connexe \mathcal{D} de \mathbb{R}^d , suffisamment régulier. Il est relativement facile de montrer que la mesure de Lebesgue sur \mathcal{D} vérifie l'inégalité de Poincaré (cf. par exemple [LD00], chapitre IV.7, proposition 2). Cependant la preuve classique, qui utilise l'injection compacte de l'espace de Sobolev H^1 dans L^2 (valable par exemple si \mathcal{D} vérifie la propriété de cône), ne dit rien sur la constante optimale. On peut également voir que cette constante dépend très fortement de la géométrie de \mathcal{D} — une modification infinitésimale de \mathcal{D} suffit à faire prendre n'importe quelle valeur à cette constante (cette question est déjà évoquée en 1937 par Courant et Hilbert — cf. le chapitre VI.2.6, « continuity of

eigenvalues » dans l'édition anglaise [CH53], et [Arr95] pour une présentation de résultats plus récents). Ce problème, ou plutôt une manière de le contourner, sera évoqué dans la partie II.

De même, une bonne approximation de la constante dans Sobolev logarithmique peut être aussi cruciale que difficile à établir : le problème *a priori* simple de recherche de la constante optimale pour la loi de Bernoulli asymétrique a résisté pendant plus de 25 ans, jusqu'aux travaux de Y. Higuchi et N. Yoshida (non publiés), et indépendamment de P. Diaconis et L. Saloff-Costes ([DSC96]). La dernière référence en particulier traite en détail l'application des inégalités de Poincaré ou Sobolev logarithmique aux problèmes de vitesse de convergence pour des chaînes de Markov à espace d'états finies.

Autres traductions des inégalités fonctionnelles Les inégalités fonctionnelles sont intimement liées à des propriétés géométriques de la mesure étudiée.

Le premier exemple de telles propriétés est la *concentration* — le fait qu'un petit ensemble (au sens de la distance sur l'espace) suffise pour rassembler presque toute la mesure. Pour préciser les choses et citer un résultat, introduisons la fonction de concentration α_μ correspondant à la mesure μ : c'est la meilleure fonction telle que, pour tout ensemble A de μ -masse supérieure à $1/2$,

$$\mu(A_r) \geq 1 - \alpha_\mu(r),$$

où A_r est l'ensemble épaissi $\{x, d(x, A) \leq r\}$. Une décroissance rapide de α signifie donc qu'un petit épaississement d'un ensemble de mesure $1/2$ concentre presque toute la mesure μ . On peut alors classer les mesures suivant leurs vitesses de concentration, et en particulier introduire la classe des mesures qui concentrent « au moins autant qu'une gaussienne » :

Définition. La mesure μ est dite à concentration gaussienne s'il existe deux constantes c, C telles que :

$$\alpha_\mu(r) \leq C \exp(-cr^2).$$

Cette concentration se traduit immédiatement par d'autres propriétés : elle permet par exemple de contrôler la déviation des fonctions lipschitziennes par rapport à leur moyenne. Ce genre de résultats trouve des applications aussi bien en statistique (*cf.* par exemple le panorama [LBB04]) que sur des modèles physiques comme les verres de spins (par exemple, dans la solution de Talagrand du modèle de Sherrington-Kirkpatrick ; on pourra consulter à ce sujet les chapitres 6 et 7 de [BS02]).

Quoiqu'il en soit, on peut relier cette propriété géométrique à une inégalité fonctionnelle :

Théorème. Si une mesure vérifie l'inégalité de Sobolev logarithmique, alors elle a la propriété de concentration gaussienne.

Pour plus de détails sur ces questions, des réciproques partielles, etc., on pourra consulter les notes de cours de M. Ledoux ([Led99]).

Signalons enfin, toujours en lien avec la géométrie, que l'on peut également étudier des questions d'*isopérimétrie* via des inégalités fonctionnelles — pour ce point de vue, on pourra consulter par exemple le chapitre 6.1 de [Vil03] et les références qui s'y trouvent.

Les semi-groupes Un des aspects sur lequel nous reviendrons le plus dans cette thèse est l'utilisation des inégalités fonctionnelles pour étudier un semi-groupe. Nous sommes partis de ce point de vue pour motiver l'introduction des inégalités fonctionnelles, et on le retrouvera au cœur de l'argumentation, aussi bien sur le recuit simulé que pour la mécanique statistique.

Il nous faut toutefois évoquer un autre aspect du lien avec les semi-groupes. L'inégalité de Sobolev logarithmique a en effet été introduite comme condition équivalente d'une propriété de semi-groupe, l'*hypercontractivité*. Cette propriété signifie que le semi-groupe P_t envoie l'espace L^2 dans des espaces plus petits, L^p (pour $p > 2$). En ce sens, il régularise les fonctions.

Cette notion d'hypercontractivité et ses dérivées ont également été largement étudiées ; on renvoie de nouveau aux cours de D. Bakry, d'A. Guionnet et de B. Zegarlinski pour une vue d'ensemble ([Bak94, GZ03]).

La grande famille des inégalités fonctionnelles Les deux inégalités fonctionnelles les plus connues, que nous avons déjà abondamment citées, ne sont pas indépendantes.

Théorème. *Si μ vérifie une inégalité de Sobolev logarithmique avec une constante C_{LS} , alors elle vérifie l'inégalité de Poincaré, et $C_P \leq \frac{1}{2}C_{LS}$.*

La preuve désormais classique peut se trouver par exemple dans [ABC⁺00] (section 1.2.6).

Heureusement, les inégalités fonctionnelles ne sont pas qu'au nombre de deux, et leur famille s'est beaucoup agrandie dans les dernières années : il devient alors nécessaire de voir qui implique qui, et avec quelle constantes. Une bonne référence pour ce paragraphe est l'article [BCR].

Pour motiver ces nouvelles inégalités, commençons par un exemple simple, sur \mathbb{R} . On l'obtient en considérant les mesures d'équilibre des processus de Langevin (gouvernés par l'É.D.S. (1.1)) pour le potentiel $V(x) = |x|^\alpha$, pour $\alpha > 0$ (en toute rigueur, on a besoin pour que l'É.D.S. soit valable de régularité sur V — on pourra remplacer dans l'énoncé suivant la valeur absolue $|x|$ par une version régularisée $\sqrt{1+x^2}$ sans changer les résultats).

3 Exemple. *Soit μ_α la mesure définie sur \mathbb{R} par :*

$$\mu_\alpha(dx) = Z_\alpha^{-1} \exp(-|x|^\alpha) dx,$$

où Z_α est une constante de normalisation.

Alors on a les comportements suivants :

- Si $\alpha \geq 2$, μ vérifie l'inégalité de Sobolev logarithmique ;
- Si $\alpha \geq 1$, μ vérifie l'inégalité de Poincaré.

Il y a donc un fossé entre $\alpha = 1$ et $\alpha = 2$: on peut penser par exemple que $\mu_{3/2}$ devrait, du point de vue de la convergence vers l'équilibre, ou de la concentration de la mesure, etc., se comporter « mieux » que μ_1 mais « moins bien » que la (quasi-)gaussienne μ_2 .

Pour identifier ces comportements intermédiaires, on peut chercher à généraliser les inégalités classiques, et plusieurs directions sont possibles.

Une première idée est de généraliser le membre de gauche de ces inégalités, pour tenter de trouver une formulation commune à Poincaré et Sobolev logarithmique. On peut remarquer par exemple que l'entropie d'une fonction, comme sa variance, peut s'écrire :

$$\int \phi(f) - \phi\left(\int f\right),$$

pour une fonction ϕ convexe (respectivement $x \mapsto x^2$ et $x \mapsto x \log(x)$). En prenant une fonction ϕ quelconque, on introduit des ϕ -entropies et les inégalités fonctionnelles correspondantes. Nous ne reviendrons pas sur ces considérations et renvoyons à [BCR] pour plus de renseignements.

Une autre idée est d'affaiblir l'inégalité en rajoutant un terme à droite. C'est l'idée de l'inégalité de Sobolev logarithmique *non tendue* (introduite dès l'article fondateur de Gross), puis plus tard de l'*inégalité faible de Poincaré* (définie par M. Röckner et F.-Y. Wang dans [RW01]). Celle-ci s'écrit :

$$\forall r > 0, \quad \mathbf{Var}_\mu(f) \leq \alpha(r) \int |\nabla f|^2 d\mu + r \|f\|_\infty^2, \quad (1.5)$$

où α est une fonction décroissante. La variance est donc presque contrôlée par l'énergie, mais on s'autorise un terme d'erreur. Celui-ci peut être rendu aussi petit que l'on veut, quitte à augmenter la constante devant l'énergie. M. Röckner et F.-Y. Wang donnent également ([RW01], théorème 3.1) un deuxième point de vue en montrant que si les restrictions de la mesure μ à des sous-ensembles vérifient des inégalités (fortes) de Poincaré, on peut en déduire une inégalité (faible) pour la mesure μ d'origine. Nous reviendrons sur ce point de vue, et sur ces inégalités faibles, dans la partie consacrée au recuit simulé.

Enfin, ces inégalités intermédiaires peuvent aussi s'exprimer grâce à des inégalités de Beckner généralisées, introduites par R. Latała et R. Oleszkiewicz dans [LO00] ; ce point de vue sera adopté dans la partie de mécanique statistique.

Les liens entre ces multiples inégalités — ainsi que les propriétés de concentration, d'isopérimétrie, de contrôle du semi-groupe qu'elles entraînent — sont maintenant mieux compris, par exemple grâce à leur reformulation quasi-systématique en inégalités *capacité-mesure* (cf. [BCR] pour cela, et la partie II pour un exemple d'utilisation de ces inégalités).

★ ★ ★

Après cette brève présentation du large champ des inégalités fonctionnelles, nous allons maintenant présenter les deux domaines dans lesquels nous avons pu les mettre à profit.

1.2 Le recuit simulé

Ne hâtez pas trop cependant la cuisson, mais conduisez votre feu de manière à diminuer graduellement la chaleur.

A. Dumas, *Mon dictionnaire de cuisine*, sur la manière de rôtir les viandes noires
Experiments that determine the low-temperature state of a material (...) are done by careful annealing, first melting the substance, then lowering the temperature slowly, and spending a long time at temperatures in the vicinity of the freezing point. If this is not done, and the substance is allowed to get out of equilibrium, (...) the substance may form a glass, with no crystalline order and only metastable, locally optimal structures.

S. Kirkpatrick, C.D. Gelatt et M.P. Vecchi, *Optimization by Simulated Annealing*, Science ([KGV83])

La deuxième partie de la thèse présente une application d'inégalités fonctionnelles à l'algorithme de « recuit simulé ».

Présentation Le recuit simulé est un algorithme stochastique de minimisation. Étant donné un espace E et une fonction V sur E , on cherche à localiser le ou les points où V atteint son minimum, et à estimer la valeur de V en ces points.

Dans un cadre continu, une des idées les plus classiques pour résoudre ce problème est de suivre la ligne de plus grande pente, *i.e.* d'effectuer une descente de gradient. Ceci correspond à résoudre l'équation différentielle :

$$dX_t = -\nabla V(X_t) dt. \quad (1.6)$$

Cette approche présente cependant plusieurs inconvénients — on peut par exemple remarquer qu'à l'approche du minimum, le processus ralentit de plus en plus — mais un des aspects le plus limitant est que le processus gouverné par (1.6) est piégé par les minima locaux.

Pour trouver le minimum global de V , il faut donc modifier le processus, et lui permettre de remonter les pentes. Une possibilité pour cela est de rajouter un terme aléatoire à la descente de gradient, et de définir X par :

$$dX_t = \sqrt{\sigma} dB_t - \nabla V(X_t) dt,$$

où σ est un réel positif, appelé *température*. Nous retrouvons ici l'équation de Langevin (1.1) évoquée précédemment. Malheureusement, comme indiqué précédemment à propos de cette équation, le nouveau processus ne converge plus vers un minimum : il se répartit suivant sa mesure invariante μ_σ , définie par

$$\mu_\sigma(dx) = \frac{1}{Z_\sigma} \exp\left(-\frac{V(x)}{\sigma}\right) dx.$$

Cette mesure n'est cependant pas sans intérêt pour trouver les minima de V : on voit facilement que, pour σ petit, elle se concentre sur les points x où V est minimal (plus précisément, elle converge faiblement vers une mesure dont le support est inclus dans l'ensemble des minima globaux). Fixer σ à une trop petite valeur risque cependant de faire perdre l'intérêt du terme aléatoire, en forçant le processus à rester longtemps dans les pièges locaux.

L'idée est alors de faire dépendre σ du temps, et de le diminuer graduellement. Le processus devient donc la diffusion inhomogène gouvernée par l'équation différentielle stochastique :

$$dX_t = \sqrt{\sigma(t)} dB_t - \nabla V(X_t) dt, \quad (1.7)$$

où σ est une fonction décroissante de t , de limite nulle. C'est cette « diffusion de recuit » que nous allons étudier. Avant cela, précisons les questions qui se posent pour toute application du recuit simulé.

Le problème du refroidissement L'algorithme de recuit simulé a été introduit par S. Kirkpatrick, C.D. Gelatt et M.P. Vecchi dans [KGV83], dans un cadre discret. Le but initial est de calculer des solutions approchées pour des problèmes « difficiles » comme celui du voyageur de commerce. Rappelons que ce problème consiste, étant donné n points dans l'espace \mathbb{R}^d , à trouver une trajectoire la plus courte possible passant par les n points. Dans cet exemple, un point x de l'espace E sera une trajectoire possible, $V(x)$ la longueur de cette trajectoire, et on cherchera à minimiser V . Pour appliquer l'idée du recuit simulé, il faut alors répondre à trois questions :

- Quelle notion de voisinage choisir, *i.e.* quelles sont les trajectoires « proches » d'une trajectoire donnée ?
- À quel moment arrêter l'algorithme ?
- Comment choisir la vitesse de refroidissement ?

La façon dont nous posons le problème élude la première question, puisqu'on se placera directement sur l'espace $E = \mathbb{R}^d$ pour la notion habituelle de voisinage. Pour le voyageur de commerce, on peut penser à diverses notions de voisinage (deux trajectoires peuvent être voisines si on peut passer de l'une à l'autre en intervertissant deux villes, par exemple), et le choix du voisinage change les caractéristiques de l'algorithme.

Nous ne nous attarderons pas non plus sur la deuxième question, même si la réponse est évidemment cruciale du point de vue pratique. Notons simplement que la méthode que nous adopterons pourrait théoriquement fournir des estimations de la vitesse de convergence (et donc du temps nécessaire pour obtenir une bonne approximation), avec les deux restrictions suivantes :

- l'explicitation de la vitesse paraît malaisée,
- elle dépendrait nécessairement du potentiel V .

Comme dans un cadre pratique, V est justement inconnu, il faudrait se résigner à des hypothèses *a priori*, où à des conditions d'arrêt heuristiques.

Pour plus d'éléments de réponse sur ces problèmes, y compris d'un point de vue pratique, on pourra consulter les synthèses [HJJ03, Loc00].

★ ★ ★

Nous nous pencherons donc sur la dernière question, celle du choix de σ en fonction de t . Nous présentons maintenant une façon intuitive de trouver la « bonne » vitesse de refroidissement. Pour cela, supposons que $\sigma(t)$ est pris décroissant par paliers, et qu'à chaque palier, on essaie d'attendre suffisamment pour se rapprocher de l'équilibre. Pour estimer ce temps d'attente, il nous faut étudier la convergence du processus à température fixée. Ceci rejoint les questions déjà posées dans la première partie de cette introduction, et fait intervenir les inégalités fonctionnelles. En particulier, si l'on connaît la constante de Poincaré C_σ pour la mesure d'équilibre μ_σ , le temps caractéristique de convergence (dans L^2) sera donné par C_σ (*cf.* l'équation (1.4), page 4).

Il nous reste alors à estimer C_σ . Il se trouve qu'il est possible de déterminer précisément le comportement asymptotique de cette quantité, et de la relier au profil du potentiel V . Le résultat est le suivant :

4 Proposition. *Il existe une constante d^* , dépendant explicitement du potentiel V , telle que :*

$$C_\sigma \approx \exp\left(\frac{d^*}{\sigma}\right).$$

Ce résultat demande des hypothèses sur V (*cf.* [FW98] pour le cas d'un domaine compact, et [Jac92] dans notre cas). Nous y reviendrons bientôt.

Supposons maintenant que l'on reste pendant un temps T_k au palier k , pour lequel $\sigma = 1/k$. D'après les remarques précédentes, il nous faut choisir $T_k = \exp(kd^*)$. Au temps

$$t = T_1 + T_2 + \dots + T_k \approx \exp(kd^*),$$

la valeur de σ est donc $1/k$. On en déduit donc la formule suivante pour σ en fonction du temps :

$$\sigma(t) = \frac{d^*}{\log(t)}. \tag{1.8}$$

Ces explications peuvent être rendues rigoureuses, au moins sous certaines hypothèses (c'est par cette approche que B. Hajek montre dans [Haj88] que, pour un espace d'états fini, cette vitesse de refroidissement est optimale). Nous verrons que le résultat est robuste : il est encore valable dans un cadre continu, et même dans des cas où la proposition 4 est mise en défaut.

Convergence dans le modèle continu La version continue du recuit simulé (*i.e.* le processus défini par l'équation (1.7)) est introduit par Chiang, Hwang et Sheu dans [HCS87]. Ces auteurs montrent un résultat de convergence en loi, pour un refroidissement en $c/\ln(t)$, sous les hypothèses :

$$V(x) \xrightarrow{|x| \rightarrow \infty} \infty \quad (1.9)$$

$$|\nabla V(x)| \xrightarrow{|x| \rightarrow \infty} \infty, \quad (1.10)$$

$$\exists C \in \mathbb{R}, \quad |\nabla V|^2 - \Delta V \geq C. \quad (1.11)$$

La preuve utilise des comparaisons avec les trajectoires de l'équation différentielle, et avec le processus à température fixée ; plusieurs arguments font intervenir directement des estimations de temps de transition entre différents ensembles. La constante c de l'hypothèse n'est pas la constante optimale : ce problème sera résolu dans [Roy89].

L. Miclo propose ensuite une approche plus analytique du problème. Il étudie directement le processus inhomogène défini par (1.7), via son entropie relative vis-à-vis de la mesure d'équilibre $\mu_{\sigma(t)}$. La preuve fait intervenir directement les inégalités fonctionnelles, et en particulier l'inégalité de Poincaré.

Ceci permet de mieux comprendre l'intérêt des hypothèses (1.10) et (1.11). En effet, on a le résultat suivant ([Kun02], proposition 3.7, cité par [Cat05], section 5) :

5 Proposition. *Soit V une fonction \mathcal{C}^2 , et $\mu_\sigma = Z_\sigma^{-1} \exp(-V/\sigma)$ Si*

$$\liminf_{|x| \rightarrow \infty} \left(\frac{1}{2} |\nabla V|^2 - \sigma \Delta V \right) > 0, \quad (1.12)$$

alors μ_σ vérifie une inégalité de trou spectral.

Si les propriétés (1.10) et (1.11) sont vérifiées, (1.12) est vraie dès que σ est suffisamment petit, et les mesures μ_σ vérifient donc bien une inégalité de trou spectral.

Un nouveau résultat de convergence Le résultat de la partie II est une généralisation du théorème de convergence de [HCS87]. Il répond aux questions suivantes :

- L'algorithme converge-t-il si la propriété de trou spectral n'est plus vérifiée ? En particulier, que se passe-t-il dans le cas simple ou le potentiel V se comporte à l'infini comme une puissance $|x|^\alpha$ de la distance, avec $\alpha \in]1, 2[$, de sorte que (1.10) n'est plus vérifiée ?
- Si il y a tout de même convergence, doit-on ralentir σ pour tenir compte de l'affaiblissement des hypothèses « à l'infini », ou est-ce le comportement « local » (près du minimum global) qui importe ?

Théorème (Th. 71, page 110). *Si le potentiel V croît à l'infini plus rapidement que la distance, si son gradient est borné, et si ses minimas locaux ne sont pas « pathologiques », on peut toujours définir une constante d^* , qui ne dépend que de la structure « locale » (près de 0) des minima, et telle que le processus du recuit converge pour une vitesse du type*

$$\sigma(t) = \frac{c}{\ln(t)},$$

pour $c > d^*$.

Une fois de plus, nous renvoyons au corps de la thèse pour un énoncé plus précis des hypothèses et résultats.

Pour contourner le problème posé par l'absence de trou spectral, nous avons utilisé l'inégalité faible de Poincaré introduite précédemment (équation (1.5)). Cette dernière est en effet vraie sous des hypothèses très faibles sur la mesure μ .

Une des difficultés a été de contrôler le comportement de cette inégalité sur μ_σ lorsque σ varie : pour ce faire, nous avons utilisé la caractérisation de l'inégalité faible en termes de relation capacité/mesure. À cette occasion, nous avons pu réinterpréter la caractérisation de l'inégalité faible comme collection d'inégalités fortes « locales » (cf. la discussion qui suit l'équation (1.5)) : l'inégalité faible de Poincaré peut s'étudier par un critère capacité-mesure restreint aux ensembles *de mesure suffisamment grande*. C'est ce qui nous permet de montrer que le comportement « local » prime sur le comportement à l'infini : les ensembles de mesure importante sont en effet concentrés près du minimum global du potentiel V .

★ ★ ★

Évoquons brièvement les perspectives. Beaucoup de questions restent en suspens. Déjà, le cas limite où $V(x) = |x|$ à l'infini n'est pas traité. Nous demandons également qu'il n'y ait aucun piège (minimum local) de V en dehors d'un compact ; on pourrait penser que des pièges suffisamment « petits » ne devraient pas perturber le comportement global du système. Une étude poussée demanderait cependant probablement de revenir à l'étude trajectorielle du processus.

1.3 Mécanique statistique et modèle d'Ising

Magnetism, as you recall from physics class, is a powerful force that causes certain items to be attracted to refrigerators.

Dave Barry

Although here we expect to find a more favorable result than from a general three-dimensional model, we recognize here too a disappearance of J with H. Thus in the model chosen here, whose essential characteristic is the restriction of the mutual effect on the one between neighboring elements, no ferromagnetism appears.

E. Ising, à propos de son modèle en dimension trois

L'autre partie de cette thèse est consacrée à l'étude d'un modèle de physique statistique, le modèle d'Ising. Ce modèle est, avec la percolation, l'un des plus connus et les plus étudiés de la physique statistique, et c'est pour le décrire qu'ont été développées de nombreuses théories qui ont ensuite pu être appliquées dans d'autres domaines.

La présentation faite ici est donc naturellement partielle et partielle ; pour plus de détails sur la théorie générale, nous renvoyons aux ouvrages [Geo88, Pru86, Lig85], ainsi qu'au plus

récent [GHM01] (qui s'intéresse en particulier aux liens avec la percolation et aux aspects « géométriques » des phases). D'autres références plus spécifiques seront données au cours du texte.

Cette introduction s'articule en trois temps. Nous esquissons d'abord un historique du modèle, et évoquons les résultats classiques, obtenus dans leur majorité avant les années 70. Nous détaillons ensuite un aspect de ce modèle, à savoir son comportement en l'absence de transition de phase, et les liens que l'on a pu faire avec des modèles dynamiques. Enfin, nous introduisons la variante du modèle que nous étudierons, où l'on considère des spins à valeurs dans un espace non-borné, pour laquelle nous rappelons les résultats connus et la façon dont notre résultat principal s'y inscrit.

Premiers résultats

Naissance du modèle Le « modèle d'Ising » a en fait été inventé en 1920 par E. Lenz. C'est un modèle de physique statistique : il vise à expliquer un phénomène macroscopique complexe par un mécanisme microscopique simple (éventuellement modélisé comme aléatoire).

Plus précisément, on essaie d'expliquer le phénomène suivant. À température fixée, on soumet une substance ferromagnétique (un morceau de fer, de cobalt, ...) à un champ magnétique : la substance s'aimante, dans le sens du champ. Si on baisse maintenant le champ magnétique jusqu'à le faire disparaître, il y a deux comportements en fonction de la température. Au-dessus d'une certaine température (la température critique, ou température de Curie) la substance ne garde aucune mémoire du champ magnétique lorsque celui-ci a disparu. En revanche, si l'on se place à une température inférieure, la substance reste aimantée alors que le champ n'est plus là : c'est ce phénomène d'hystérésis, appelé *magnétisation spontanée*, que l'on cherche à expliquer.

Le mécanisme microscopique proposé pour cela est le suivant. On modélise la substance par un ensemble de spins (moments magnétiques, que l'on peut voir comme des petits aimants), répartis sur un réseau (\mathbb{Z}^d ou un sous-ensemble fini de \mathbb{Z}^d), et qui ne peuvent prendre que deux valeurs, +1 ou -1, (les aimants pointent soit vers le haut soit vers le bas).

1 Terminologie. *Nous aurons à manipuler plusieurs types d'ensembles de sites, i.e. de sous-ensembles de \mathbb{Z}^d . Nous adoptons la terminologie suivante :*

- Un volume fini sera un sous-ensemble fini de \mathbb{Z}^d .
- Une boîte finie sera un sous-ensemble fini parallélépipédique.
- Les termes boîte infinie, volume infini signifieront \mathbb{Z}^d tout entier.
- Enfin, on dira qu'une suite croissante de volumes Λ_n tend vers \mathbb{Z}^d si la condition suivante est vérifiée :

$$\forall i \in \mathbb{Z}^d, \exists n, i \in \Lambda_n.$$

Trois effets s'appliquent au système :

- La température, qui désorganise aléatoirement les spins : plus elle est élevée, plus ils se retournent aléatoirement,
- le champ magnétique, qui les aligne (par exemple vers le haut),
- l'interaction, à portée microscopique : un spin dont tous les voisins immédiats sont (par exemple) vers le haut aura tendance à s'orienter dans le même sens.

Pour formaliser tout cela, on doit réfléchir à ce que peut être un état macroscopique du système. Un tel état ne peut pas être une simple configuration microscopique : intuitivement en effet, la modification d'un seul spin ne change pas fondamentalement l'état global du système. Cependant, on ne peut pas identifier deux configurations qui ne diffèrent qu'en un seul point, sinon de proche en proche, toutes les configurations devraient être confondues.

L'idée est alors d'attribuer à chaque configuration un *poids*, ce qui amène naturellement (après normalisation des poids) à voir un état macroscopique, aussi appelé *phase*, comme une *mesure de probabilité* sur les configurations microscopiques.

Ceci étant précisé, on peut définir les états qui correspondent aux contraintes évoquées plus haut (effets du magnétisme, de la température et de l'interaction). Suivant un formalisme classique en physique statistique, on attribue une *énergie* à chaque configuration. Le choix typique de cette énergie, appelée hamiltonien et notée H , est le suivant :

$$H(\mathbf{x}) = - \sum_{i \sim j} x_i x_j - h \sum_i x_i, \quad (1.13)$$

où la notation $i \sim j$ signifie que i est voisin de j , et le paramètre h dénote l'intensité du champ magnétique extérieur. Les configurations favorisées seront celles de *basse énergie*, i.e. celles où les spins voisins sont *alignés*, et où ils *suivent le champ magnétique*.

On définit alors les mesures de Gibbs :

Définition. Une mesure de Gibbs à la température $T = 1/\beta$ est une mesure de probabilité qui, à une configuration \mathbf{x} , attribue le poids suivant :

$$\mu_\beta(\mathbf{x}) = \frac{1}{Z_\beta} \exp(-\beta H(\mathbf{x})),$$

où la constante $Z_{\beta,h}$ est appelée fonction de partition et est définie par :

$$Z_{\beta,h} = \sum_{\mathbf{x}} \exp(-\beta H(\mathbf{x})).$$

Cette définition n'est valable *stricto sensu* que pour des volumes finis, à condition de ne considérer dans la somme (1.13) que les sites i, j appartenant au volume considéré. Pour de tels volumes, la mesure de Gibbs est alors unique.

Si l'on veut interpréter la transition de phase comme la possible coexistence de plusieurs mesures de Gibbs, il est alors nécessaire de donner un sens à la définition précédente *en volume infini*. Ceci est possible de plusieurs manières, la plus élégante étant sans doute la propriété de compatibilité dite *DLR* (pour Dobrushin, Lanford et Ruelle).

Une fois cette définition posée, la question de la transition de phase se traduit alors ainsi :

Question. *Quelles sont les mesures de Gibbs en volume infini ?*

Si il y a plusieurs mesures possibles, les conditions physiques (température, champ magnétique) ne suffisent pas à décrire l'objet macroscopique, et il y a *transition de phase*.

La question de la présence de transition de phase est donc fondamentale, et a donné lieu à énormément de développements.

Les premières réponses Si nous avons présenté le modèle à partir des mesures de Gibbs infinie, il faut savoir que cette approche ne date en fait que des années 1970. Les réflexions précédentes s'appuyaient sur des calculs d'espérance de certaines fonctions.

Plus précisément, nous avons vu plus haut que l'on pouvait, dans une boîte finie, définir pour chaque couple (T, h) une mesure de Gibbs. En réalité, on peut rajouter d'autres paramètres, des *conditions aux bords*. Ceci revient à fixer les valeurs des spins aux bords de la boîte. Deux choix particulièrement intéressants sont les conditions aux bords uniformes, entièrement constituées de spins orientés dans le même sens (il y en a donc deux). Dans une

boîte finie, disons le cube $[-n, n]^d$, on a donc défini deux mesures particulières, que l'on note $\mu_n(\beta, +)$ et $\mu_n(\beta, -)$. La fonction la plus simple sur laquelle on peut tester ces mesures est la valeur du spin à l'origine, et si l'on peut montrer que

$$\mu_n(\beta, +)(x_0) \qquad \text{et} \qquad \mu_n(\beta, -)(x_0)$$

admettent des limites différentes pour n tendant vers l'infini, on pourra en déduire la présence d'une transition de phase.

Ceci permet d'expliquer le caractère prépondérant de la fonction de partition, grâce au petit calcul suivant. Appelons *énergie libre* le logarithme renormalisé de la fonction de partition (on supposera la condition aux bords nulle) :

$$F_\Lambda(\beta, h) = \frac{1}{\beta \text{Card}(\Lambda)} \log(Z(\beta, h, \Lambda)).$$

Dans une boîte finie, dérivons cette énergie libre par rapport au champ magnétique h :

$$\begin{aligned} \frac{\partial F_\Lambda(\beta, h)}{\partial h} &= \frac{1}{\beta \text{Card}(\Lambda) Z(\beta, h, \Lambda)} \frac{\partial}{\partial h} \left(\sum_{\mathbf{x}} \exp(\beta(\sum x_i x_j + h \sum x_i)) \right) \\ &= \frac{1}{\text{Card}(\Lambda) Z(\beta, h, \Lambda)} \sum_{\mathbf{x}} (\sum x_i) \exp(-H(\mathbf{x})) \\ &= \mu_{\beta, h, \Lambda}(m_\Lambda(\mathbf{x})), \end{aligned}$$

où $m(\mathbf{x})$ est la *magnétisation* sur la boîte Λ (définie par $m_\Lambda(\mathbf{x}) = \frac{1}{\text{Card}(\Lambda)} \sum x_i$). La dérivée de l'énergie libre donne donc la moyenne (sous $\mu_{\Lambda, \beta, h}$) de la magnétisation.

Supposons alors que :

- l'énergie libre converge, quand Λ tend vers \mathbb{Z}^d vers une fonction $F(\beta, h)$, indépendante de la suite de boîtes et de la condition aux bords :

$$F_\Lambda(\beta, h) \xrightarrow{\Lambda \rightarrow \mathbb{Z}^d} F(\beta, h), \tag{1.14}$$

- cette fonction soit dérivable,
- on puisse intervertir limite et dérivation ;
- la moyenne de la magnétisation ne diffère pas trop de la moyenne du spin à l'origine.

Alors la dérivée de l'énergie libre (en volume infini) devrait donner la valeur moyenne du spin en 0, qui ne pourrait donc prendre qu'une seule valeur : ainsi, $\mu_{\beta, +}$ et $\mu_{\beta, -}$ coïncideraient. En fait, tous ces arguments peuvent être rendus rigoureux ([LML72, Rue72]), et il y a équivalence entre :

- La fonction $F(\beta, h)$ est dérivable en h au point (β_0, h_0) ,
- Les moyennes $\mu(\beta_0, h_0, +)(x_0)$ et $\mu(\beta_0, h_0, -)(x_0)$ coïncident.

Ceci justifie que l'on s'intéresse à l'énergie libre en volume infini et à ses propriétés de régularité.

★ ★ ★

On a donc cherché à expliciter la forme de cette énergie libre ; ceci a été fait en dimension 1 dès l'introduction du modèle, par Ising, qui a montré l'absence de transition de phase, ce qui l'a convaincu du manque d'intérêt du modèle... Plus tard, L. Onsager ([Ons44]) a pu calculer la solution en dimension 2, uniquement pour un champ magnétique nul, et a trouvé une singularité en une valeur critique explicite (de la température) ; on a pu ultérieurement justifier que cette singularité correspondait bien à la transition de phase (cf. par exemple les

notes bibliographiques de [Geo88], pp. 450–451). Ce calcul, même après simplification, reste très technique et ne semble guère se prêter à une généralisation en dimension supérieure; on a donc naturellement cherché d'autres voies d'approche du problème.

La première, introduite par Peierls en 1936, est la méthode des contours. Elle s'appuie sur une remarque simple : ce qui compte pour calculer l'énergie d'une configuration, ce sont les frontières entre les zones de + et les zones de -. Ceci peut aider à estimer les probabilités d'événements du type « le spin en 0 est un - », sous des conditions aux bords + : en effet, si cet événement est réalisé, il y a nécessairement un contour (une interface) entre le spin en 0 et le bord de la boîte, et ce contour a un certain coût énergétique. On peut montrer que, si la température est suffisamment faible, ce coût est assez important pour garantir que l'événement a une probabilité qui reste strictement inférieure à 1/2 même quand la boîte grandit. On obtient ainsi la présence d'une transition de phase en dimension supérieure à 2. Notons que l'objectif ici est toujours de montrer qu'il y a une transition de phase; c'est en effet ce phénomène qui a motivé l'introduction du modèle, et qui a longtemps été considéré comme son principal intérêt.

L'essor et la mathématisation Dans les années 70, la théorie sur les modèles de physique statistique se développe et se mathématise. La notion de mesure de Gibbs, en particulier, trouve une définition claire comme « mesure définie par ses probabilités conditionnelles », avec les équations DLR (pour Dobrushin, [Dob70], Lanford et Ruelle [LR69]). Au même moment apparaissent des outils fondamentaux, les inégalités de corrélation, aux premiers rang desquelles figure l'inégalité FKG (du nom de Kasteleyn, Fortuin et Ginibre ([FKG71])). Elle s'appuie sur le fait qu'on peut munir les configurations d'un ordre partiel :

$$\mathbf{x} \leq \mathbf{y} \iff \forall i, x_i \leq y_i.$$

On peut alors définir des fonctions *croissantes*, et des événements croissants (les événements dont l'indicatrice est croissante). Par exemple, les événements « les deux points A et B sont reliés par une chaîne de spins + » ou « la magnétisation moyenne est positive » sont croissants. On peut également en déduire un ordre (partiel) sur les mesures : on dira que μ est inférieure à ν si

$$\forall f, f \text{ croissante}, \quad \mu(f) \leq \nu(f).$$

L'inégalité traduit alors le caractère attractif du modèle (les spins ont tendance à s'aligner avec leurs voisins) par une propriété rigoureuse : deux événements croissants (ou plus généralement, deux fonctions croissantes) sont toujours positivement corrélés.

Théorème (Inégalité FKG). *Pour tout volume fini, toute condition aux bords \mathbf{z} , et toutes valeurs de h et β , et tout couple (f, g) de fonctions croissantes :*

$$\mu_{\Lambda, \mathbf{z}}(f; g) \geq 0.$$

La conséquence principale de ce fait, qui est propre aux systèmes attractifs, est que les mesures se comportent bien vis-à-vis de cet ordre. On peut par exemple montrer qu'une augmentation du champ magnétique, ou des conditions aux bords, fait augmenter la mesure au sens FKG.

Sous de bonnes conditions (réalisées pour le modèle d'Ising standard), on peut alors en déduire que toute mesure de Gibbs est comprise entre deux extrêmes, que l'on peut construire comme mesures limites quand les conditions aux bords sont extrémales (constituées uniquement de + ou de -). En particulier, la mesure de Gibbs est unique si et seulement si les mesures

extrémales coïncident, donc la question de la transition de phase se ramène à montrer l'identité ou la différence de deux mesures bien définies.

Remarque. Notons que l'extrémalité dont il est question est l'extrémalité au sens FKG. En particulier, les deux mesures μ_+, μ_- n'ont pas de raison d'être les seuls points extrémaux du convexe formé par les mesures de Gibbs.

Ceci donne naissance à d'autres questions ; on peut par exemple se demander si l'ensemble des mesures de Gibbs n'est constitué que du segment $[\mu_-, \mu_+]$, ou si la situation est plus complexe (en fait elle l'est : si en dimension 2, toute mesure de Gibbs appartient bien à ce segment, ce n'est vrai en dimension supérieure que pour les mesures de Gibbs invariantes par translation, cf. les travaux d'Aizenmann, Higuchi, Dobrushin ([Aiz80, Hig81, Dob72]) et le récent [Bod05]).

Parallèlement, S.A. Pirogov et Y.G. Sinai ([PS75, Sin82]) ont cherché à généraliser l'argument des contours de Peierls. Pour cela, ils ont défini puis étudié des modèles auxiliaires portant directement sur les contours, en oubliant les configurations de spins sous-jacentes. Ils ont ainsi obtenu des résultats sur le comportement à basse température du modèle d'origine — en particulier, leur théorie permet de décrire, en fonction des paramètres du hamiltonien (ici la température et le champ extérieur), quelles phases sont susceptibles d'apparaître, *i.e.* d'établir le *diagramme de phases* du modèle (cf. le chapitre 6 de [Pru86] pour une courte introduction, et les notes des sections 19.1, 19.2 de [Geo88] pour plus de références). Cette approche est toujours utilisée (voir par exemple le cours de M. Zahradník, [Zah98]).

La région d'unicité et le modèle d'Ising dynamique

Les critères de Dobrushin et Shlosman Parallèlement à ces recherches sur la coexistence de phases, un nouvel intérêt naît pour la région où il n'y a pas transition de phase. R.L. Dobrushin établit en 1970 ([Dob70]) un critère simple, entraînant l'unicité de la mesure de Gibbs. Pour exprimer ce critère, nous avons besoin de comparer des mesures entre elles.

Définition. Soit (E, d) un espace métrique séparable, et soit μ et ν deux mesures sur E . Soit Π l'ensemble des mesures sur $E \times E$ qui ont pour marginales μ et ν . La distance de Wasserstein entre μ et ν est donnée par :

$$d_W(\mu, \nu) = \inf \int d(x, y) d\pi(x, y).$$

Pour une preuve que ceci définit bien une distance, cf. par exemple [Vil03].

Le critère de Dobrushin dit intuitivement la chose suivante. Considérons le seul spin à l'origine, conditionnellement à tous ses voisins (d'après la définition d'une mesure de Gibbs, cela revient à conditionner uniquement par ses voisins immédiats). Faisons maintenant varier les conditions aux bords (*i.e.* les spins autres que 0). Si la mesure correspondante sur le spin à l'origine ne dépend pas trop des variations dans les conditions aux bords, on peut imaginer que chaque spin se comportera à peu près indépendamment des autres, et qu'il n'y aura qu'une seule mesure de Gibbs, proche d'une mesure produit.

Plus formellement, le critère est le suivant. Notons \mathcal{C}_i l'ensemble des couples de configurations (\mathbf{x}, \mathbf{y}) qui ne diffèrent qu'en i : $(\mathbf{x}, \mathbf{y}) \in \mathcal{C}_i \iff (\forall j \neq i, x_j = y_j)$. Soit d_S une distance sur l'espace « mono-spin » $\{-1, +1\}$ pour le modèle d'Ising standard), et d_W la distance de Wasserstein associée.

6 Théorème. Soit ρ_i la quantité

$$\rho_i = \sup \left\{ \frac{d_W(\mu_{\{0\}}^{\mathbf{x}}, \nu_{\{0\}}^{\mathbf{y}})}{d_S(x_i, y_i)}; (\mathbf{x}, \mathbf{y}) \in \mathcal{C}_i \right\}.$$

Alors, si

$$\sum_i \rho_i < 1,$$

la mesure de Gibbs en volume infini est unique.

Remarque. Ce critère a été majoritairement utilisé pour des modèles où les spins sont discrets. Dans ce cas, on peut remplacer la distance de Wasserstein par la distance en variation totale (qui n'est qu'un cas particulier de distance de Wasserstein, où l'on prend la distance $d_S(x, y) = \mathbf{1}_{x \neq y}$ comme référence sur l'espace de départ). Nous reviendrons plus loin au cas des spins continus non-bornés.

Ce critère est ensuite très largement utilisé. Quand il est vérifié, il donne de plus de nombreuses indications sur l'unique mesure de Gibbs (ses corrélations, sa régularité en fonction des paramètres, etc.). Pour plus de renseignements sur ces points, on pourra consulter le chapitre 8 de [Geo88] (des exemples de modèles sont donnés au point 8.9 et 8.13, et des conséquences dans la section 8.2).

★ ★ ★

Malgré le succès du critère, il est apparu qu'il ne recouvrait pas toute la région d'unicité. R.L. Dobrushin et S.B. Shlosman publient donc ultérieurement ([DS85b]) un second critère (que nous noterons **DSU**), analogue du premier mais sur des ensembles de sites plus grands (non nécessairement réduits à un point). Nous y reviendrons dans la partie I (section 2.5). La région d'application est *a priori* plus vaste, mais les propriétés de régularité ne sont plus nécessairement vérifiées.

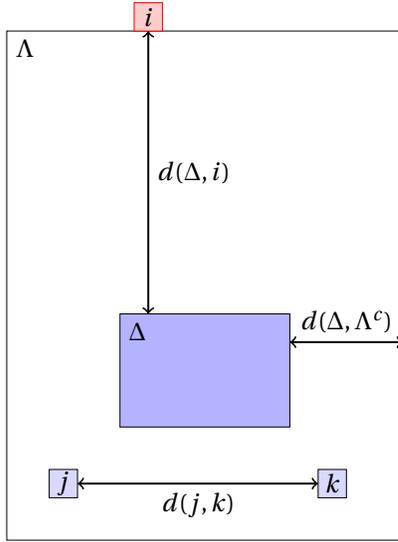
Les deux mêmes auteurs publient alors une série d'articles ([DS85a, DS85b, DS87]) concernant justement ce problème de régularité. Pour expliquer leur résultats, donnons d'abord une définition informelle.

Définition. On dit qu'une mesure vérifie une propriété de mélange quand ses marginales sur les sous-boîtes éloignées du bord ne dépendent que faiblement des conditions aux bords (cf. le schéma 1.1).

On dit qu'il y a décroissance des corrélations quand on peut contrôler la covariance entre des spins distants par une fonction (décroissante, tendant vers 0) de la distance entre les sites concernés.

Enfin, les propriétés d'analyticité étudient la régularité de l'énergie libre en fonction du paramètre de magnétisation h , ou plus généralement en fonction du hamiltonien.

Dobrushin et Shlosman montrent que des formes fortes de ces propriétés (mélange fort exponentiel, décroissance exponentielle des corrélations dans tous les volumes, etc.) sont équivalentes — ainsi, dans une certaine région de l'espace des phases, il y a une unique mesure de Gibbs qui vérifie de multiples propriétés. Toutes ces propriétés équivalentes sont appelées « conditions de mélange de Dobrushin–Shlosman », ou « conditions d'analyticité complète » (que nous abrègerons en **CA**).



On note $\mu_{\Lambda, \Delta}^x$ la marginale de μ_{Λ}^x sur la boîte Δ .
Il y a *mélange fort* pour la distance de Wasserstein s'il existe C telle que pour tout Λ, Δ , tout i à la frontière de Δ , et tout couple $(x, y) \in \mathcal{C}_i$:

$$d_W(\mu_{\Lambda, \Delta}^x, \mu_{\Lambda, \Delta}^y) \leq C \exp\left(-\frac{d(\Delta, i)}{C}\right) d_S(x_i, y_i).$$

Il y a *mélange faible* s'il existe C telle :

$$d_W(\mu_{\Lambda, \Delta}^x, \mu_{\Lambda, \Delta}^y) \leq C \sum_{y \in \partial \Delta} \exp\left(-\frac{d(\Delta, y)}{C}\right) d_S(x_i, y_i).$$

Il y a *décroissance exponentielle des corrélations* s'il existe C tel que pour tout Λ , toute condition aux bords y et tous sites j, k dans Λ ,

$$\mu_{\Lambda, y}(x_j, x_k) \leq C \exp\left(-\frac{d(j, k)}{C}\right).$$

FIG. 1.1: Des exemples de mélange et de décroissance des corrélations

Remarque. Le terme de « mélange » est à rapprocher de son utilisation pour les processus. Un processus est mélangant quand il « oublie sa condition initiale », une mesure de Gibbs l'est quand elle « oublie sa condition aux bords ». Le point de vue de « chaîne de Markov spatiale » pour étudier les mesures de Gibbs est classique, cf. par exemple [Geo88]. Pour une récapitulation des diverses notions de mélange pour les processus, on pourra consulter [Bra05].

Ces auteurs donnent de plus un « critère constructif », intermédiaire entre le critère original de Dobrushin et le critère **DSU**, qui garantit les conditions **CA**.

Remarque. Tous ces critères et propriétés ont en réalité de multiples expressions et variantes, suivant le choix des distances, les hypothèses sur les potentiels d'interaction, etc. Une bonne référence pour comprendre les liens entre ces différents critères dans le cas des spins compacts est [MO94]. Notons de plus que, comme certains de ces critères entraînent des propriétés de mélange, ils sont parfois désignés sous le nom de « conditions de mélange », ce qui peut prêter à confusion.

Le modèle dynamique Parallèlement à ces études « statiques », plusieurs auteurs ([Hol85, AH87, SZ92b]) s'intéressent aux liens avec un modèle dynamique, appelé *dynamique de Glauber* (ou modèle d'Ising stochastique). Ce modèle d'évolution est très simple à décrire : partant d'une configuration donnée, on la fait évoluer à chaque pas de temps suivant la règle suivante :

- on choisit un site au hasard dans la boîte, puis
- avec une probabilité dépendant uniquement du champ extérieur, des spins voisins et de la température, on le retourne.

Ceci définit une chaîne de Markov sur l'ensemble des configurations, et pour un bon choix de la probabilité, cette chaîne admet comme unique mesure réversible la mesure de Gibbs (statique) dans la boîte. On peut également définir une dynamique analogue directement dans la boîte infinie \mathbb{Z}^d , auquel cas toute mesure de Gibbs (infinie) est réversible par rapport à la dynamique.

Étudier cette dynamique amène naturellement à s'intéresser à son éventuelle convergence. Notons au passage que cette dynamique n'est pas la plus adaptée pour simuler la mesure d'équilibre (*cf.* à ce sujet l'algorithme de Swendsen et Wang [SW87]), mais qu'elle est la plus étudiée. On peut attendre intuitivement les phénomènes suivants :

- Si la mesure de Gibbs est unique, la dynamique devrait la « trouver » rapidement.
- Si plusieurs mesures d'équilibre sont possibles, la convergence devrait ralentir, le système se mettant à hésiter entre les limites possibles.

On a donc cherché à voir si ces comportements étaient vérifiés. En particulier, on a voulu savoir si, au moins dans certaines régions « loin de la transition de phase », la convergence vers l'unique mesure invariante était rapide.

C'est évidemment ici que les inégalités fonctionnelles rentrent en jeu : la connaissance des constantes de Poincaré et de Sobolev logarithmiques pour les mesures de Gibbs devient une clé pour étudier ces convergences.

Remarque. *L'idée précédente est compliquée par le phénomène de métastabilité. Expliquons brièvement en quoi il consiste. Supposons que l'on se place dans une boîte, avec une température faible, et un champ magnétique très faible (mais non-nul), et des conditions aux bords libres (i.e. le bord n'influe pas sur la boîte). La présence du champ magnétique fait que la mesure de Gibbs devrait favoriser les configurations qui suivent ce champ (disons les configurations de magnétisation strictement positives). Supposons maintenant que l'on parte d'une configuration composée entièrement de spins $-$. Comme le champ a été pris très petit devant la température, les spins $-$ n'auront que très peu de chances de se retourner si tous leurs voisins sont aussi des spins $-$. Il faudra attendre très longtemps avant que le système ne « comprenne » que le champ magnétique extérieur favorise les spins $+$. C'est ce genre de phénomène qui, même dans une boîte finie, peut ralentir le processus.*

Inversement, quand la température est élevée, chaque spin se comporte quasiment indépendamment des autres, et on s'attend à une convergence beaucoup plus rapide.

Aux propriétés étudiées auparavant (mélange, analyticité etc.) se sont donc rajoutées les questions de convergence du modèle dynamique, pour certains volumes et conditions aux bords, à certaines vitesses, et pour diverses notions de convergence.

Dans cette direction, D.W. Stroock et B. Zegarlinski montrent dans [SZ92a] le résultat suivant.

7 Théorème. *Si il existe une constante C telle que toute mesure de Gibbs en volume fini (et pour des conditions aux bords quelconques) vérifie une inégalité de Poincaré de constante C , alors toutes ces mesures vérifient en fait une inégalité de Sobolev logarithmique, pour une même constante. De plus ces conditions sont équivalentes aux conditions **CA** de Dobrushin et Shlosman.*

Notons enfin pour finir que l'article déjà cité de Martinelli et Olivieri ([MO94]) fait un point détaillé sur toutes les conditions équivalentes, et met en évidence une limitation de l'approche, en discutant en particulier du type de boîte pour lesquelles on peut espérer des propriétés de mélange ou de décroissance des corrélations.

Le cas des spins non-bornés

Définition du modèle Tous les travaux que nous venons d'évoquer se limitent à un cadre relativement simple, dans le sens où les spins considérés sont à valeurs dans $\{-1, 1\}$, ou au pire dans une variété *compacte*. Or la physique, et en particulier la théorie quantique des champs, a

aussi donné naissance à des modèles très similaires à celui d'Ising, mais faisant intervenir des spins *non-bornés* (en pratique à valeurs dans \mathbb{R}). Cette modification a de lourdes conséquences techniques, et de nombreuses preuves valables dans les modèles classiques doivent être revues pour s'en accommoder.

Ceci se comprend facilement : même si l'on ne considère qu'un volume fini, un hamiltonien qui ne prendrait en compte que les interactions entre les spins n'aurait par exemple que peu de chances d'être intégrable. Cependant, le fait de se placer dans un espace continu permet aussi certaines simplifications ; le modèle dynamique devient en effet une diffusion avec dérive, sur laquelle on sait beaucoup de choses.

Pour empêcher que les spins prennent de trop grandes valeurs, on pénalise ces fortes valeurs par une auto-interaction : on rajoute au hamiltonien un terme du type

$$\sum_i V(x_i),$$

où V tend vers l'infini à l'infini — les grandes valeurs de x_i donnent donc un hamiltonien important, et ont un poids faible dans la mesure de Gibbs.

Le problème suivant est de donner une bonne notion de température. On a en effet, même en ne retenant que les définitions les plus simples, au moins 4 choix de paramètres :

$$\begin{aligned} H(\mathbf{x}) &= \sum V(x_i) - J \sum x_i x_j, & Z &= \int \exp(-\beta H), \\ H(\mathbf{x}) &= \sum V(x_i) + J \sum (x_i - x_j)^2, & Z &= \int \exp(-\beta H), \\ H(\mathbf{x}) &= \sum V(x_i) - J \sum x_i x_j, & Z &= \int \exp(-H). \\ H(\mathbf{x}) &= \sum V(x_i) + J \sum (x_i - x_j)^2, & Z &= \int \exp(-H). \end{aligned}$$

On peut par exemple trouver la seconde notation dans [DS81, Mat04], la troisième dans les articles de N. Yoshida ([Yos99], par exemple), la quatrième chez beaucoup d'auteurs. Nous adopterons selon les cas la seconde ou la troisième formulation.

Remarque. *Le « bon choix » de paramètres ne semble pas clair au vu des articles que nous avons consulté.*

Notons tout d'abord que pour des valeurs finies des paramètres, ce choix ne semble avoir que peu d'importance : n'importe quelle notation peut être utilisée, puisque l'on peut passer de l'une à l'autre en changeant les définitions de V, β, J .

Le problème est que de nombreux résultats sont annoncés comme « perturbatifs », i.e. valables pour des petites valeurs des paramètres. La bonne définition de ces paramètres est alors importante.

En particulier, on cherche souvent à voir si des résultats valables pour le modèle d'Ising standard ont un analogue dans le cas non-borné. On peut alors se demander à quoi correspondent les résultats de haute ou basse température.

Concrètement, la plupart des résultats sont annoncés pour des interactions faibles ($J \rightarrow 0$, dans la notation IV). L'autre asymptotique ($J \rightarrow \infty$) ne ressemble alors pas au modèle d'Ising classique (on obtiendrait une loi où tous les spins seraient identiques, et l'unique valeur serait distribuée suivant la loi libre de densité $\exp(-V)$).

Inversement, nous n'avons trouvé que peu d'articles traitant un cas de « basse température ». Parmi eux, celui de O. Matte, qui conclut à l'existence de plusieurs mesures de Gibbs (et à des

estimations de trou spectral) est exprimé dans le cadre II, avec β grand (ce qui est logique) et J fixé, mais suffisamment petit. Dans ces notations, c'est le régime de haute température qui dégénère (la mesure en un seul site dégénère pour $\beta \rightarrow 0$). Toutefois, nous verrons qu'il est bien possible d'utiliser dans ce cadre le critère de Dobrushin pour montrer l'unicité pour β non-nul suffisamment petit (cf. le chapitre 2.4).

Nous verrons que plusieurs hypothèses sur l'auto-interaction sont nécessaires, en particulier pour montrer des inégalités fonctionnelles. Les plus centrales sont les suivantes.

Hypothèse 1. « Convexité à l'infini » V se décompose en une somme $V_1 + V_2$, où V_1 est strictement convexe, et V_2 est bornée.

Hypothèse 2. « Super-convexité » Pour tout $m > 0$, il existe une décomposition de V en $V_1 + V_2$, où $V_1'' \geq m$, et V_2 est bornée.

Deux exemples fondamentaux Les deux exemples le plus souvent cités sont les suivants :

- le cas d'une auto-interaction convexe. Dans ce cas, la mesure de Gibbs est unique, cf. infra.
- le cas du *double puits*. La forme la plus simple est de prendre pour V le polynôme :

$$V(x) = ax^4 - bx^2$$

où a et b sont strictement positifs. Ce qui fait intuitivement l'intérêt du modèle est que chaque spin (indépendamment des autres) a deux positions privilégiées, dans chacun des puits de potentiel. On espère ainsi qu'une transition de phase a plus de chances d'apparaître, chaque phase se localisant préférentiellement dans un des puits.

Ils sont représentés dans la figure 1.2. Notons que le deuxième exemple vérifie la condition de super-convexité.

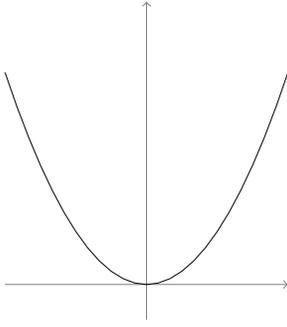
Adaptation des résultats classiques Avant de généraliser les équivalences entre inégalités fonctionnelles, mélange et décroissance des corrélations, on a naturellement cherché à adapter les résultats « classiques » du modèle d'Ising. Pour plusieurs raisons, on ne doit considérer dans ces études que certaines mesures de Gibbs, appelées *mesures tempérées*. Pour un récapitulatif des différentes définitions que nous avons pu trouver, et la preuve qu'elles sont en fait équivalentes, nous renvoyons à la section 2.2.

Citons quelques étapes de l'étude du modèle de spins réels :

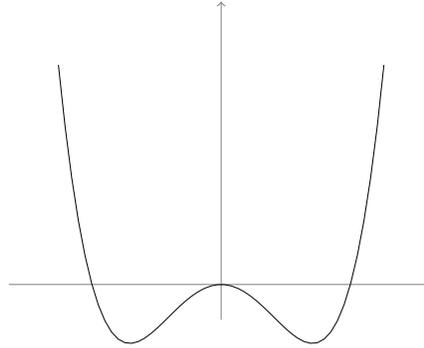
- J.L. Lebowitz et E. Presutti montrent dans [LP76, LP81] que l'on peut définir l'énergie libre en boîte infinie (de la même façon qu'en (1.14)), qu'il existe au moins une mesure de Gibbs en volume infini, et que la différentiabilité de l'énergie libre (en fonction du champ magnétique) entraîne l'unicité de la mesure de Gibbs.
- M. Cassandro *et al.* ([COPP78]) établissent l'analogie de la condition d'unicité de Dobrushin.
- J. Bellissard et R. Høegh-Krohn ([BHK82]) généralisent les résultats précédents, et étudient les propriétés topologiques de l'ensemble des mesures de Gibbs.

Viennent ensuite les questions de transition de phase.

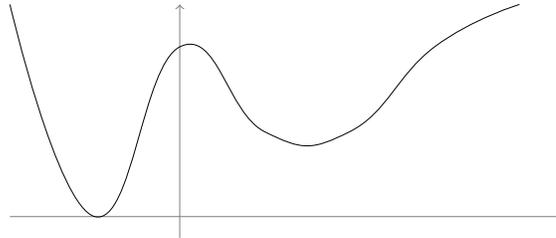
Pour les preuves de l'*existence* d'une transition de phase, nous n'avons trouvé dans la littérature que deux arguments. Le premier se trouve dans un article de Dobrushin et Shlosman, [DS81], et concerne des potentiels asymétriques (cf. le troisième exemple de la figure 1.2) ; il y est montré que si les deux puits vérifient certaines conditions, il existe une valeur critique



Un potentiel convexe, pour lequel la mesure de Gibbs est unique.



L'exemple du double puits $U(x) = ax^4 - bx^2$. La mesure de Gibbs est unique pour certaines valeurs de β et J . Pour certains choix de a et b , et certaines valeurs de β, J , il y a deux mesures de Gibbs.



Un exemple de potentiel asymétrique. Sous certaines conditions sur la forme de U et la valeur de J , il existe un β critique qui « équilibre les puits » et pour lequel deux mesures de Gibbs coexistent.

FIG. 1.2: Trois exemples de potentiels d'auto-interaction

de β qui les « équilibre », et crée une transition de phase. Notons que le cas du double puits y est considéré comme acquis (pour β grand), avec un renvoi à [Mal75], qui ne semble pourtant traiter de d'espaces de spins compacts.

Le second est, toujours pour un double puits, une adaptation de l'argument de Peierls ; les hypothèses et le résultat précis peuvent être trouvés dans la thèse [Mat04] : notons simplement que l'argument s'applique dans le cadre II, avec $V = ax^4 - bx^2$, J petit (ce qui semble surprenant — une interaction forte devrait favoriser la transition de phase) et β suffisamment grand (la valeur critique pouvant dépendre de J).

Pour montrer l'absence de transition de phase, la première idée est d'utiliser la condition de Dobrushin. Le régime où elle s'applique n'est cependant pas évident à déterminer, en raison du flou qui règne sur le choix des paramètres. Nous reviendrons sur cette question dans la section 2.4 (partie I). Nous renvoyons également ici à l'article de G. Royer ([Roy79]) qui utilise une itération dans l'esprit de la condition de Dobrushin pour montrer l'unicité pour des interactions polynomiales convexes, ou des petites perturbations d'interactions convexes.

Une autre voie, traitée dans [BHK82], est d'utiliser la régularité de l'énergie libre. En particulier, on trouve dans cet article le résultat suivant :

8 Théorème. *Si l'énergie libre $F(\beta, h)$ est dérivable par rapport à h en un certain point (β_0, h_0) , alors la mesure de Gibbs est unique en ce point.*

Cette technique peut être utilisée pour le modèle discret : un théorème classique de Lee et Yang assure que la fonction F ne peut avoir des points singuliers que si le champ magnétique h vaut 0.

On pourrait espérer que, pour un potentiel symétrique, et pour h non nul, ce théorème s'applique, et montre la dérivabilité. Malheureusement ce n'est pas le cas en toute généralité : ce théorème est vrai pour certains potentiels (comme $V = ax^4 - bx^2$), mais faux pour d'autres. Pour plus de détails, on pourra consulter la section 4.6 de [GJ87].

Une troisième approche est d'utiliser la convergence des modèles d'évolution. Cette approche, due à H. Doss et G. Royer ([DR79, Roy99]) est intimement liée aux inégalités fonctionnelles, et nous y reviendrons en détail.

Mélange, décroissance des corrélations et convergence des modèles d'évolution L'adaptation des résultats sur les liens entre propriétés statiques et dynamique pour les spins non bornés a lieu à la fin des années 1990. B. Zegarlinski ([Zeg96]) s'intéresse à la preuve de l'inégalité de Sobolev logarithmique uniforme en dimension 1, grâce à des propriétés de décroissance des corrélations.

N. Yoshida publie ensuite une série d'articles sur le sujet, et montre en particulier dans [Yos99] une inégalité de Sobolev logarithmique uniforme, en régime perturbatif (*i.e.* pour J petit, dans le cadre II évoqué ci-dessus).

Parallèlement, T. Bodineau et B. Helffer ([BH99b]) obtiennent un résultat similaire, en relâchant les hypothèses de N. Yoshida (et en particulier celle de super-convexité). Leur approche est plus analytique, et s'appuie notamment sur les travaux d'Helffer et Sjöstrand ([HS94, Hel99]) sur les représentations des corrélations en termes de Laplaciens de Witten. Ce point de vue sera également réinterprété par M. Ledoux ([Led01]) en termes de semi-groupe (voir aussi [GR01], qui relâche encore les hypothèses sur l'auto-interaction et sortent du cadre de convexité à l'infini).

Tous ces articles s'attachent à montrer la *validité d'une inégalité de Sobolev logarithmique*, dans un cadre où les techniques classiques (et en particulier, le critère de Bakry et Émery) ne s'applique pas directement. L'utilisation des corrélations apparaît donc plus ici comme un outil que comme un résultat en soi. Résumons l'approche de ces différents articles (et plus particulièrement de [BH99b, Led01]) :

- On montre une inégalité de trou spectral, sur un spin, uniformément par rapport aux conditions aux bords.
- Pour une interaction (J) faible, on en déduit une inégalité de *trou spectral* uniforme en la boîte et la condition aux bords (*cf.* on peut remarquer que la valeur correspondante de J est la même que celle qui garantit l'application du critère d'unicité de Dobrushin, *cf.* la remarque 2.4, page 49).
- On montre une inégalité de Sobolev logarithmique sur les fonctions ne dépendant que d'un spin, en utilisant un argument de Bakry et Émery, en utilisant l'hypothèse de super-convexité, ou l'hypothèse de convexité et la faible valeur de J .
- Par un argument d'itération, qui utilise l'inégalité mono-spin précédente et la décroissance des corrélations, on déduit l'inégalité de Sobolev logarithmique uniforme cherchée. Suivant l'argument utilisé, la constante est inconnue ([Yos99],[BH99b]) ou explicite ([Led01]).

Ces preuves passent donc toutes par l'inégalité de Poincaré uniforme; à ce sujet, nous renvoyons aussi à [Wu06], qui montre une telle inégalité en s'appuyant sur le critère de Dobrushin (dont nous verrons qu'il est vérifié pour J ou β petit). Nous mentionnons également l'article [Mar06], qui sous une condition de faibles interactions (malheureusement peu explicite) donne une borne sur la constante de Sobolev logarithmique.

★ ★ ★

N. Yoshida cherche ensuite à vérifier si les différentes propriétés (mélange, décroissance des corrélations, inégalités uniformes) sont toujours aussi fortement reliées. Dans [Yos01], il établit, sous l'hypothèse de super-convexité, l'équivalence entre le trou spectral uniforme, Sobolev logarithmique uniforme, des propriétés de mélange exponentiel et la décroissance des corrélations. Il établit également un critère constructif, dans l'esprit des critères de Dobrushin et Shlosman. L'article [Yos00] fait également le lien avec des propriétés dynamiques de convergence. Précisons que ces résultats sont valables en régime *quelconque* (non-perturbatif).

C'est en nous appuyant sur les techniques développées dans ces derniers articles que nous prouvons notre premier résultat. C'est une adaptation au cas des spins non-bornés d'un autre article de N. Yoshida ([Yos97]), et il peut s'exprimer ainsi.

9 Théorème. *Soit c_N une famille de constantes telles que, dans tout volume de diamètre inférieur à N , et pour toute condition aux bords \mathbf{z} , la mesure $\mu_\Lambda^{\mathbf{z}}$ vérifie une inégalité de trou spectral avec :*

$$C_P(\mu_\Lambda^{\mathbf{z}}) \leq c_N.$$

Supposons que c_N vérifie :

$$c_N \leq C \frac{N}{\log(N)}.$$

Alors, si C est suffisamment petit, les constantes $C_P(\mu_\Lambda^{\mathbf{z}})$ sont majorées uniformément en Λ, \mathbf{z} .

De même, si les mesures $\mu_\Lambda^{\mathbf{z}}$ vérifient des propriétés de mélange polynomiales suffisamment fortes, alors le mélange est en réalité exponentiel.

Pour les hypothèses précises et la preuve, cf. la partie I, et en particulier les sections 3.3, 3.4. Ce résultat est à rapprocher de nombreux autres, sur l'absence d'un « régime intermédiaire » entre les comportements polynomiaux et exponentiels (cf. [Yos97], remarque 3.3). Comme le résultat d'origine, il est également valable si l'on remplace, dans l'hypothèse et la conclusion, « tout volume fini » par « toute boîte finie », ou « toute boîte de côtés multiples d'un certain m ». Il est également très proche d'un résultat de T. Bodineau et B. Helffer (cf. le théorème 56 et la discussion qui le suit pour une comparaison de ces résultats).

★ ★ ★

Parallèlement, G. Royer publie en 1999 ([Roy99]) les notes d'un cours sur les inégalités de Sobolev logarithmiques et leur application aux systèmes de spins réels. En plus des résultats déjà cités, on y trouve le théorème suivant :

10 Théorème. *Soit Λ_n la suite croissante de boîtes $[-n, n]^d$, qui tend vers \mathbb{Z}^d . Si les mesures de Gibbs μ_n sur Λ_n avec condition aux bords nulle vérifient une inégalité de Sobolev logarithmique pour une constante C indépendante de n ,*

alors la mesure de Gibbs tempérée en volume infini est unique.

Remarque. *On peut prendre pour Λ_n toute suite de volumes suffisamment régulière — par exemple, le théorème est valable dès que les Λ_n sont des boîtes et que le volume $\text{Card}(\Lambda_n)$ est majoré par un polynôme en la distance $d(0, \Lambda_n^c)$.*

Ce changement de point de vue est intéressant : plutôt que d'étudier les conditions fortes sur les constantes de Sobolev logarithmiques — qui entraînent une régularité forte de la phase unique, cf. supra, on se demande *quel contrôle des constantes suffit pour entraîner l'absence de transition de phase*. Notons aussi que les contrôles ne sont demandés que pour *une* condition au bord ; ceci pourrait sembler intéressant pour opérer une vérification directe des inégalités, grâce à des critères explicites comme ceux disponibles dans [Cat04, Cat05]. Nos efforts dans cette direction n'ont malheureusement pas abouti.

Au vu des liens précédemment évoqués entre Sobolev logarithmique et Poincaré pour le modèle étudié, il devient naturel de se demander si le théorème précédent pourrait se généraliser au cas où l'on ne suppose qu'une inégalité de Poincaré uniforme (toujours pour une seule condition aux bords). Si nous n'avons pas pu montrer ce résultat, nous avons tout de même réussi à affaiblir l'hypothèse en ne demandant qu'une inégalité de Beckner, intermédiaire entre Poincaré et log-Sobolev. Cette inégalité (définie précisément page 69) dépend d'un paramètre a : pour $a = 0$ c'est l'inégalité de Poincaré, pour $a = 1$ on retrouve Sobolev logarithmique.

11 Théorème. *Soit Λ_n une suite croissante de volumes finis, tendant vers \mathbb{Z}^d . Si V vérifie certaines hypothèses,*

si les mesures μ_n sur Λ_n avec condition aux bords nulle vérifient une inégalité de Beckner pour un paramètre a et une constante C indépendante de n ,

et si a est suffisamment grand (proche de 1),

alors la mesure de Gibbs tempérée en volume infini est unique.

Les hypothèses sur V et la preuve de ce théorème peuvent être trouvés dans la partie I (théorème 44, page 69).

Perspectives

Nous énumérons ici plusieurs questions encore sans réponse à notre connaissance.

Quel est le diagramme de phases du modèle, même dans le cas simple du double puits ? Si le potentiel est symétrique et que l'on rajoute un champ magnétique, la mesure de Gibbs est-elle unique (ceci pourrait être infirmé par [DS81]) ? Rajouter un terme en hx dans le hamiltonien correspond-il d'ailleurs à imposer un champ magnétique ?

Peut-on expliciter des potentiels sur lesquels l'inégalité de Sobolev logarithmique (éventuellement, pour la seule condition aux bords nulle) pourrait être montrée directement, et pour quelles valeurs de β ou de J ? En particulier, existe-t-il une région où la mesure est unique mais où les conditions de Dobrushin et Shlosman (**CA** ou même **DSU**) sont mises en défaut ?

Première partie

Mécanique statistique

L'objectif de cette partie est d'étudier un modèle de mécanique statistique, analogue du modèle d'Ising mais où les spins prennent leurs valeurs dans \mathbb{R} .

Elle est organisée en deux chapitres. Dans le premier, nous introduisons le modèle et les hypothèses générales. Nous définissons les mesures de Gibbs tempérées, et vérifions que les différentes notions trouvées dans la littérature coïncident bien. Nous rappelons ensuite (essentiellement d'après Dobrushin, Dobrushin-Shlosman) des critères d'existence et d'unicité. Nous introduisons aussi, en suivant G. Royer, un modèle dynamique particulièrement adapté au problème, dans le sens où ses mesures d'équilibre sont précisément les mesures de Gibbs tempérées que l'on cherche à étudier.

Dans le second chapitre, nous nous inspirons des multiples travaux de la fin des années 1990 sur les inégalités fonctionnelles pour ce modèle ([BH99b, Yos99, Led01, Roy99], etc.). Nous obtenons une preuve de l'unicité avec une hypothèse d'inégalités fonctionnelles uniformes ou « presque uniformes », dans la lignée d'un résultat de G. Royer, mais avec des hypothèses plus faibles (théorèmes 40 et 44, respectivement p. 61 et 69). L'idée est ici d'utiliser le lien dynamique/statique, en montrant que si la dynamique converge suffisamment rapidement (grâce aux inégalités fonctionnelles), elle ne peut avoir qu'une seule mesure limite.

Enfin, nous améliorons des résultats de N. Yoshida sur le régime de « régularité » des mesures de Gibbs, en montrant que, si des inégalités fonctionnelles sont vérifiées sur toutes les boîtes et conditions aux bords, et que les constantes n'augmentent pas trop vite, elles sont en fait uniformes — ce sont les théorèmes 55 et 60, pages 82 et 92. Nous montrons au passage qu'un mélange polynomial suffisamment rapide est nécessairement exponentiel.

Chapitre 2

Un modèle d'Ising à spins réels

2.1 Notations et hypothèses

On considère un système de spins réels sur le réseau \mathbb{Z}^d . On note \mathcal{S} l'ensemble des sites, Ω l'ensemble des configurations. Pour \mathbf{x} une configuration, on note \mathbf{x}_Λ la projection de \mathbf{x} sur la boîte Λ , et $(\mathbf{x}_\Lambda \cdot \mathbf{z}_{\Lambda^c})$ la configuration hybride égale à \mathbf{x} sur Λ et \mathbf{z} en dehors.

On note $U_{\Lambda, \mathbf{z}}$ le hamiltonien sur la boîte Λ avec conditions aux bords \mathbf{z} ; c'est la spécification de cet hamiltonien qui détermine le modèle.

Dans le cas le plus général, on se donne une famille de potentiels d'interaction $(V_L)_{L \subset \subset \mathbb{Z}^d}$, fonctions de Ω dans \mathbb{R} tels que $V_\Lambda(\mathbf{x})$ ne dépende que de \mathbf{x}_Λ . Le hamiltonien du système en volume fini s'écrit :

$$U_{\Lambda, \mathbf{z}}(\mathbf{x}) = \sum_{L, L \cap \Lambda \neq \emptyset} V_L(\mathbf{x}_\Lambda \cdot \mathbf{z}_{\Lambda^c}).$$

Remarque (Température, mesure d'équilibre et potentiel). *Suivant les auteurs, les conventions diffèrent légèrement sur le potentiel et la mesure d'équilibre associée, et la littérature physique rajoute le paramètre température au problème. Pour clarifier les choses, les notations sont répertoriées dans le tableau 2.1. On choisit ici les notations de G. Royer, et la mesure d'équilibre est donc proportionnelle à $\exp(-2U)$.*

Notons que, comme indiqué dans l'introduction, nous ne paramétrons pas le hamiltonien par la température inverse β . Ceci n'a que peu d'incidence puisque nos résultats ne sont pas perturbatifs. Ce problème sera rediscuté dans la section 2.4.

Le choix des V_L donne naissance aux différents modèles. Dans la plupart des modèles considérés habituellement, les seules interactions présentes sont les auto-interactions et les interactions entre deux spins :

$$V_L = 0 \text{ dès que } \text{Card} L > 2.$$

Les auto-interactions, aussi appelées « phases », sont les fonctions $V_{\{i\}}$. Les interactions par paires sont les $V_{\{i, j\}}$.

	Notation « hamiltonienne »	1 ^{re} notation ([Roy99])	2 ^e notation
Mesure invariante	$Z_\beta^{-1} e^{-\beta U}$	$Z^{-1} e^{-2U}$	$Z^{-1} e^{-U}$
Générateur	$\frac{1}{2}\Delta - \frac{\beta}{2}\nabla U \nabla$	$\frac{1}{2}\Delta - \nabla U \nabla$	$\frac{1}{2}\Delta - \frac{1}{2}\nabla U \nabla$

TAB. 2.1: Différentes conventions pour l'énergie

On précise maintenant les hypothèses que l'on fera dans la suite sur les différents potentiels d'interaction.

Hypothèse 1 ((forme de l'auto-interaction)). *On suppose que la phase est la même en tout point,*

$$V_{\{i\}}(\mathbf{x}) = V(x_i),$$

où V vérifie les conditions suivantes :

- *Convexité à l'infini* — V se décompose en une somme $V_1 + V_2$, où V_1 est uniformément strictement convexe, et V_2 est une fonction \mathcal{C}^2 à support compact.
- *Croissance polynomiale à l'infini* — il existe des constantes a'_V, b'_V , et un $d_V > 0$ tels que pour tout x , $|V(x)| \leq a'_V |x|^{d_V} + b'_V$.
- *Il existe une constante $a''_V \in]0, 1[$ telle que la fonction $x \mapsto a''_V V'(x)^2 - V''(x)$ soit bornée inférieurement.*

Hypothèse 2 (forme des interactions par paire). *Les interactions par paire sont du type*

$$V_{\{i,j\}}(\mathbf{x}) = J(j-i)x_i x_j,$$

où J est une fonction à support fini de \mathbb{Z}^d dans \mathbb{R} , symétrique par rapport à 0. La fonction J n'est pas nécessairement positive.

On notera $p(i) = |J(i)|$ et $\sigma = \sum_{i \in \mathbb{Z}^d} p(i)$.

On supposera $J(0) = 0$, ce qui revient à dire que toute l'auto-interaction est prise en compte dans la fonction V .

Hypothèse 3 (Lien entre auto-interaction et interactions entre spins). *L'auto-interaction est suffisamment forte par rapport aux interactions :*

$$\sigma < \inf V''_1.$$

Faisons maintenant quelques remarques sur ces hypothèses. Elles correspondent à un modèle vu comme généralisation du modèle d'Ising. Le premier point restrictif est l'hypothèse sur la décomposition de V . Cette hypothèse est assez forte. Elle implique en particulier que $|V'|$ tend vers l'infini en l'infini, et interdit donc les puits de potentiel à l'infini.

Si on relâche l'hypothèse pour permettre une perturbation bornée lipschitzienne (dont on ne contrôle donc pas la dérivée seconde), on peut toujours montrer l'existence d'une mesure de Gibbs et construire les dynamiques en volume fini, mais il paraît difficile de construire la dynamique infinie (faute de contrôler de façon adéquate les dynamiques dans deux boîtes différentes).

Ces hypothèses sur les dérivées de la perturbation sont en décalage avec les cadres fixés dans les articles [BH99a, GR01], dans lesquels n'importe quelle perturbation bornée est acceptée (et, en particulier, les minima locaux de l'auto-interaction ne sont pas nécessairement localisés dans un compact). Si le contrôle de la dérivée seconde, qui sert à construire la dynamique infini-dimensionnelle, n'a pas d'influence sur les considérations de ces articles, il semble difficile de se passer d'une hypothèse du type « coercivité » pour montrer l'existence d'une mesure de Gibbs tempérée.

L'interaction par paire est directement copiée de celle du modèle d'Ising compact ; une autre approche (qui correspond à des modèles d'interface) est de pénaliser les configurations pour lesquelles $x_i - x_j$ est grand en introduisant un terme $W(x_i - x_j)$, avec une fonction W

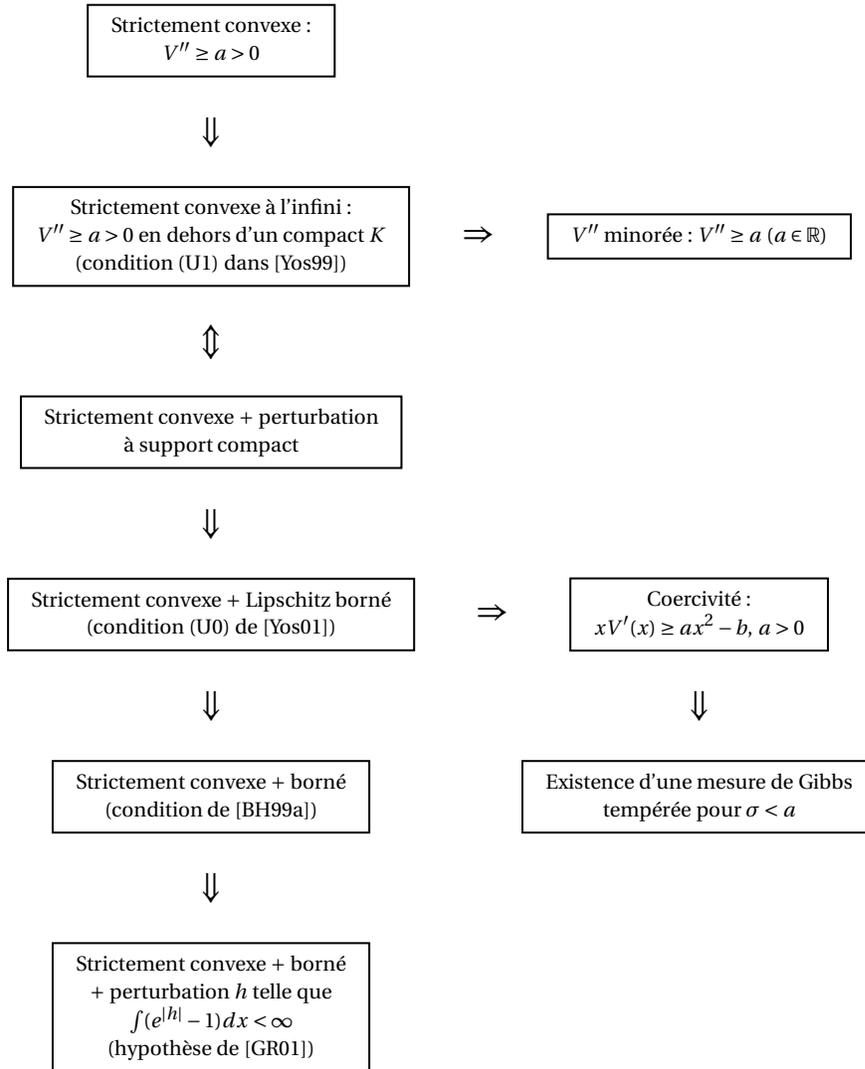


FIG. 2.1: Les multiples hypothèses sur l'auto interaction

Notons que l'article [Yos01] introduit une autre condition (notée (U1)), qui stipule que l'on peut trouver pour tout m une décomposition de la phase en une somme $V_1 + V_2$, où $V_1'' \geq m$ et V_2 est lipschitzienne et bornée.

tendant vers l'infini à l'infini. On peut remarquer que dans le cas simple $W(x) = x^2$, et quitte à rajouter des termes quadratiques à l'auto-interaction, on retrouve le même modèle que ci-dessus (cf. cependant la discussion faite dans l'introduction sur le choix des paramètres, page 22).

Enfin, la dernière hypothèse, qui pourrait sembler très restrictive, ne l'est pas vraiment. Si l'on considère par exemple le cas simple d'une auto-interaction polynomiale, paire, de degré supérieur à 4 (par exemple, $V(x) = x^4 - x^2$), on peut, pour tout σ , trouver une décomposition $V = V_1 + V_2$ avec $\inf V_1'' > \sigma$, et V_2 bornée et lipschitzienne, donc l'hypothèse ne réduit pas la généralité.

De même, dans le cas limite d'une auto-interaction en ax^2 (modèle Gaussien), la condition $\sigma < \inf V''$ est déjà nécessaire pour que les mesures de Gibbs en volume fini existent (et donc pour que le modèle ait un sens).

Avant de continuer, on montre ici la propriété de coercivité, qui servira entre autres pour justifier l'existence d'une mesure de Gibbs en volume infini.

12 Proposition (Coercivité et propriétés reliées). *Il existe des constantes $a_V > 0$, $b_V \in \mathbb{R}$ telles que :*

$$\forall x \in \mathbb{R}, \quad xV'(x) \geq a_V x^2 - b_V.$$

On peut prendre pour a_V tout réel vérifiant $a_V < \inf V_1''$.

De même, pour tout $a < \inf V_1''$, il existe une constante b telle que :

$$\forall x, y, \quad (x - y)(V'(x) - V'(y)) \geq a(x - y)^2 - b.$$

Enfin, il existe une constante c_V , éventuellement négative, telle que :

$$\forall x, y, \quad (x - y)(V'(x) - V'(y)) \geq c_V(x - y)^2$$

Démonstration. On peut supposer que V_1 atteint sa borne inférieure en 0. Soit $a < \inf V_1''$. L'application du théorème des accroissements finis entre x et y donne :

$$(x - y)(V_1'(x) - V_1'(y)) = V''(c_{x,y})(x - y)^2 \geq \inf(V_1'')(x - y)^2.$$

On en déduit :

$$\begin{aligned} (x - y)(V'(x) - V'(y)) &\geq \inf(V_1'')(x - y)^2 - 2\|V_2'\|_\infty \times |x - y| \\ &\geq a(x - y)^2 + (\inf V'' - a)(x - y)^2 - 2\|V_2'\|_\infty \times |x - y|. \end{aligned}$$

En notant $(-b)$ la borne inférieure du terme entre parenthèses, on obtient la deuxième inégalité annoncée.

Pour démontrer la première, on raisonne de la même façon, en ayant fixé y .

La dernière inégalité est une conséquence du fait que V'' est bornée inférieurement. \square

Enfin, on rassemble ici quelques estimations sur le hamiltonien et ses dérivées.

13 Proposition. *Soit \mathbf{z} une configuration, L une boîte et i un site de L . Alors la dérivée du hamiltonien par rapport à la valeur du spin en i est donnée par :*

$$\partial_i U_{L,\mathbf{z}} = \frac{\partial U_{L,\mathbf{z}}(\mathbf{x})}{\partial x_i} = V'(x_i) + \sum_{j \in \mathbb{Z}^d} J(i - j)(\mathbf{x}^L \cdot \mathbf{z})_j,$$

et on a, pour les constantes a_V, b_V, c_V de la proposition précédente, les inégalités :

$$x_i \partial_i U_{L,\mathbf{z}} \geq a_V x_i^2 - b_V - \frac{1}{2} \sum_j p(i-j) \left(x_i^2 + (\mathbf{x}^L \cdot \mathbf{z})_j^2 \right) \quad (2.1)$$

$$\geq (a_V - \frac{\sigma}{2}) x_i^2 - b_V - \frac{1}{2} \sum_j p(i-j) (\mathbf{x}^L \cdot \mathbf{z})_j^2, \quad (2.2)$$

$$(x_i - y_i) (\partial_i U_{L,\mathbf{z}}(\mathbf{x}) - \partial_i U_{L,\mathbf{z}}(\mathbf{y})) \geq (c_V - \frac{\sigma}{2}) (x_i - y_i)^2 - \frac{1}{2} \sum_{j \in L} p(i-j) (x_j - y_j)^2. \quad (2.3)$$

Remarque. Dans la dernière inégalité, rappelons que c_V n'est pas nécessairement positif, et que $c_V - \sigma/2$ peut donc très bien être négatif.

Démonstration. L'expression de la dérivée est immédiate, l'inégalité provient de la coercivité et de la minoration :

$$J(i-j)xy \geq -p(i-j) \left(\frac{x^2 + y^2}{2} \right).$$

La seconde inégalité se prouve de la même manière, en utilisant la proposition précédente. \square

2.2 Mesures de Gibbs tempérées

On veut étudier les mesures de Gibbs en volume infini pour les spins non bornés. Une approche naturelle est de les définir comme les mesures sur l'ensemble des configurations qui vérifient la propriété de compatibilité de Dobrushin, Lanford et Ruelle. Rappelons sa définition. Pour cela, notons que l'on peut considérer les mesures en volume fini comme un noyau de probabilités :

$$\begin{aligned} \mu_L : \mathbb{R}^{\mathbb{Z}^d} &\rightarrow \text{mesures sur } \mathbb{R}^{\mathbb{Z}^d} \\ \mathbf{z} &\rightarrow \mu_{L,\mathbf{z}} \otimes \delta_{z_{\Lambda^c}} \end{aligned}$$

14 Définition. On dit qu'une mesure μ sur $\mathbb{R}^{\mathbb{Z}^d}$ vérifie la propriété DLR si pour tout volume fini L , on a la relation de compatibilité :

$$\mu \mu_L = \mu.$$

Autrement dit, pour toute fonction f ,

$$\int \int f((\mathbf{y}^L \cdot \mathbf{z})) d\mu_{\Lambda,\mathbf{z}}(\mathbf{y}) d\mu(\mathbf{z}) = \int f(\mathbf{z}) d\mu(\mathbf{z}).$$

On note \mathcal{G} l'ensembles des mesures μ vérifiant cette propriété.

Malheureusement, l'ensemble \mathcal{G} n'est pas vraiment adapté au problème. Pour des raisons techniques, mais aussi physiques (cf. [Roy99], remarque 4.2.19, p. 82 pour un exemple de mesure DLR que l'on ne veut pas considérer), on doit restreindre l'ensemble des mesures en contraignant les configurations à ne pas augmenter trop vite.

On définit donc l'ensemble des configurations tempérées :

15 Définition. Une configuration de spins \mathbf{x} est dite tempérée si elle ne croît que polynomialement à l'infini :

$$\exists c_{\mathbf{x}}, d_{\mathbf{x}} \quad \forall i \in \mathbb{Z}^d, \quad |x_i| \leq c_{\mathbf{x}} (1 + |i|)^{d_{\mathbf{x}}}.$$

L'ensemble des configurations tempérées est noté \mathcal{S}' . On utilisera également les sous ensembles $\mathcal{P}_{p,C}$ et \mathcal{P}_p définis ainsi :

$$\begin{aligned}\mathcal{P}_{p,C} &= \{\mathbf{x}, \forall i, |x_i| \leq C(1+|i|)^p\} \\ \mathcal{P}_p &= \bigcap_{C>0} \mathcal{P}_{p,C} = \{\mathbf{x} \in \mathcal{S}', d_{\mathbf{x}} \leq p\}\end{aligned}$$

Pour passer des configurations tempérées aux mesures tempérées, plusieurs idées sont possibles. La première est d'étudier la configuration « moyenne » : pour une mesure μ , c'est la configuration $\mathbf{h}_\mu \stackrel{\text{déf}}{=} (\mu(|x_i|))_{i \in \mathbb{Z}^d}$. Cette idée se traduit par les deux définitions ([BHK82, Yos99] pour la première, [COPP78] pour la seconde) :

$$\begin{aligned}\mathcal{G}_t &\stackrel{\text{déf}}{=} \{\mu \in \mathcal{G}, \mathbf{h}_\mu \in \mathcal{S}'\}, \\ \mathcal{G}_t^{\text{Copp}} &\stackrel{\text{déf}}{=} \{\mu \in \mathcal{G}, \|\mathbf{h}_\mu\|_\infty < \infty\}.\end{aligned}$$

La deuxième idée est de considérer directement les ensembles chargés par les mesures considérées. C'est ce qui est fait dans [Roy99] et dans la thèse de O. Matte (qui cite [AKRT00]) :

$$\begin{aligned}\mathcal{G}_t^{\text{Roy}} &\stackrel{\text{déf}}{=} \{\mu \in \mathcal{G}, \mu(\mathcal{S}') = 1\}, \\ \mathcal{G}_t^{\text{Mat}} &\stackrel{\text{déf}}{=} \{\mu \in \mathcal{G}, \exists p, \mu(\mathcal{P}_p) = 1\}.\end{aligned}$$

En réalité ces quatre définitions concordent :

16 Théorème. *Les ensembles $\mathcal{G}_t^{\text{Roy}}$, $\mathcal{G}_t^{\text{Copp}}$, $\mathcal{G}_t^{\text{Mat}}$ et \mathcal{G}_t sont égaux.*

De plus, pour tout $p > d$ (la dimension du réseau), toute mesure de Gibbs tempérée vérifie :

$$\mu(\mathcal{P}_p) = 1.$$

Avant la preuve, introduisons, en suivant [Roy99], un outil très utile pour étudier les configurations tempérées. Rappelons que l'ensemble des configurations $\mathbb{R}^{\mathbb{Z}^d}$, muni de la topologie produit, est métrisable (on peut le munir de la distance $d(\mathbf{x}, \mathbf{y}) = \sum_k 2^{-k} \frac{d_0(x_k, y_k)}{1+d_0(x_k, y_k)}$, où d_0 est la distance habituelle sur \mathbb{R} et où on énumère les sites de \mathbb{Z}^d dans un ordre quelconque). Le problème de cette distance est qu'elle ne tient absolument pas compte de la géométrie du réseau, qui gouverne les interactions par paires dans la mesure de Gibbs. En particulier, lorsque l'on voudra comparer l'évolution suivant la dynamique de Langevin de deux configurations, ce choix de distance ne sera pas adapté.

C'est ce qui motive l'introduction d'une autre distance : plutôt que de donner au sites des poids indépendants de la géométrie du réseau, on définit un poids α à partir de l'interaction. Plus précisément, on note :

$$\begin{aligned}p(i) &= |J(i)|, \\ \sigma &= \sum_{i \in \mathbb{Z}^d} p(i),\end{aligned}$$

et pour tout $\sigma' > \sigma$,

$$\alpha = \sum_{k=0}^{\infty} \frac{1}{\sigma'^k} p^{*k}.$$

Ce poids dépend de σ' . En fait on fixera σ' de manière à ce que :

$$0 < \sigma < \sigma' < a_V < \inf V_1'' , \quad (2.4)$$

où a_V , le coefficient de coercivité, peut bien être choisi entre σ et $\inf V_1''$ (cf. l'hypothèse 3 et la proposition 12).

Ce poids a de bonnes propriétés :

17 Proposition. *Le poids α est surharmonique par rapport au noyau p , c'est-à-dire :*

$$\alpha * p \leq \sigma' \alpha .$$

De plus, il décroît exponentiellement par rapport à la distance à l'origine :

$$\exists c_\alpha, d_\alpha, \forall i \in \mathbb{Z}^d, \quad \alpha(i) \leq c_\alpha \exp(-d_\alpha |i|).$$

La première propriété montre que α est bien adapté à l'interaction. La seconde permet d'inclure les configurations tempérées dans un ensemble muni d'une structure agréable, à savoir $l^2(\alpha)$.

18 Théorème. *Soit $l^2(\alpha)$ l'ensemble des configurations \mathbf{x} telles que la quantité suivante soit finie :*

$$\|\mathbf{x}\|_\alpha = \left(\sum_{i \in \mathbb{Z}^d} \alpha(i) x_i^2 \right)^{\frac{1}{2}} < \infty .$$

Alors $l^2(\alpha)$ est un espace de Hilbert (réel), et on a l'inclusion :

$$\mathcal{S}' \subset l^2(\alpha) .$$

Cette norme permet de contrôler l'influence des spins aux bords d'une boîte sur ceux de l'intérieur, c'est l'objet du lemme suivant.

19 Lemme. *Il existe des constantes A et B telles que, pour toute configuration \mathbf{z} et pour toute boîte L , les spins dans la boîte soient contrôlés par les spins extérieurs de la façon suivante :*

$$\mu_{L,\mathbf{z}}(\|\mathbf{x}\|_\alpha^2) \leq A + B \sum_{i \notin L} \alpha(i) z_i^2 .$$

Démonstration. On suit l'argument de Royer ([Roy99], lemme 4.2.6). On utilise les inégalités obtenues sur le hamiltonien (proposition 13) :

$$x_i \partial_i U_{L,\mathbf{z}} \geq a_V x_i^2 - b_V - \frac{1}{2} \sum_j p(i-j) \left(x_i^2 + (\mathbf{x}^L \cdot \mathbf{z})_j^2 \right) .$$

On multiplie par $\frac{2}{Z_{L,\mathbf{z}}} \exp(-2U_{L,\mathbf{z}})$, et on intègre :

$$\frac{2}{Z_{L,\mathbf{z}}} \int x_i \partial_i U_{L,\mathbf{z}} e^{-2U_{L,\mathbf{z}}} d\mathbf{x} \geq 2a_V \mu_{L,\mathbf{z}}(x_i^2) - 2b_V - \left(\sum_j p(i-j) \right) \mu_{L,\mathbf{z}}(x_i^2) - \sum_j p(i-j) \mu_{L,\mathbf{z}}\left((\mathbf{x}^L \cdot \mathbf{z})_j^2 \right) .$$

On remarque que le terme de gauche vaut 1 (il suffit d'intégrer par parties par rapport à x_i), et on utilise le fait que $\sum_j p(i-j) = \sigma$:

$$1 \geq (2a_V - \sigma) \mu_{L,\mathbf{z}}(x_i^2) - 2b_V - \sum_j p(i-j) \mu_{L,\mathbf{z}}\left((\mathbf{x}^L \cdot \mathbf{z})_j^2 \right) .$$

On fait passer $2b_V$ à gauche, et on somme sur les sites i , en pondérant par $\alpha(i)$:

$$(1 + 2b_V) \sum_i \alpha(i) \geq (2a_V - \sigma) \sum_i \alpha(i) \mu_{L,z}(x_i^2) - \sum_j \left(\sum_i \alpha(i) p(i-j) \right) \mu_{L,z} \left((\mathbf{x}^L \cdot \mathbf{z})^2_j \right).$$

Comme $\alpha * p \leq \sigma' \alpha$, il vient :

$$(1 + 2b_V) |\alpha| \geq (2a_V - \sigma) \mu_{L,z}(\|\mathbf{x}\|_\alpha^2) - \sigma' \sum_j \alpha(j) \left((\mathbf{x}^L \cdot \mathbf{z})^2_j \right).$$

Enfin, comme la dernière somme est majorée par $\sum_j \alpha(j) x_j^2 + \sum_{j \notin L} z_j^2$, on a :

$$(2a_V - \sigma - \sigma') \mu_{L,z}(\|\mathbf{x}\|_\alpha^2) \leq (1 + 2b_V) |\alpha| + \sigma' \sum_{i \notin L} \alpha(i) z_i^2.$$

Par choix de a_V et de σ' (équation (2.4)), $2a_V - \sigma - \sigma' > 0$. On peut donc diviser par cette quantité pour obtenir le résultat annoncé. \square

Tous les outils sont réunis pour prouver le théorème 16. Il s'agit donc de montrer que les quatre définitions des mesures tempérées coïncident. Certaines inclusions sont évidentes d'après les définitions :

$$\begin{aligned} \mathcal{G}_t^{\text{Copp.}} &\subset \mathcal{G}_t \\ \mathcal{G}_t^{\text{Mat.}} &\subset \mathcal{G}_t^{\text{Roy.}} \end{aligned}$$

Il suffit donc de montrer deux inclusions pour conclure.

1^{re} inclusion : $\mathcal{G}_t \subset \mathcal{G}_t^{\text{Mat.}}$. On suit la preuve qui figure dans [BHK82] (qui elle-même cite la thèse [Pic79]).

Soit μ un élément de \mathcal{G}_t . Le vecteur \mathbf{h}_μ est donc tempéré :

$$\exists c_\mu, p_\mu, \quad \forall i \in \mathbb{Z}^d, \quad (\mathbf{h}_\mu)_i = \mu(|x_i|) \leq c_\mu (1 + |i|)^{p_\mu}.$$

On cherche à montrer que μ ne charge qu'un sous ensemble de \mathcal{S}' . Soit p un entier, que l'on fixera plus tard. On va établir :

$$\mu(\mathcal{P}_p) = 1.$$

Comme $\mathcal{P}_p = \bigcup_{C>0} \mathcal{P}_{p,C}$, son complémentaire (dans l'ensemble des configurations) vérifie :

$$\mu(\mathcal{P}_p^c) = \mu\left(\bigcap_{C>0} \mathcal{P}_{p,C}^c\right) = \inf_{C>0} \mu(\mathcal{P}_{p,C}^c).$$

Il suffit donc de montrer que cette borne inférieure est nulle pour conclure que μ ne charge que \mathcal{P}_p .

On regarde donc le poids que μ donne à $\mathcal{P}_{p,C}^c$.

$$\begin{aligned} \mu(\mathcal{P}_{p,C}^c) &= \mu(\{\mathbf{x}, \exists i, |x_i| > C(1 + |i|)^p\}) \\ &= \mu\left(\bigcup_{i \in \mathbb{Z}^d} \{\mathbf{x}, |x_i| > C(1 + |i|)^p\}\right) \\ &\leq \sum_{i \in \mathbb{Z}^d} \mu(\{\mathbf{x}, |x_i| > C(1 + |i|)^p\}) \\ &\leq \sum_{i \in \mathbb{Z}^d} \frac{1}{C(1 + |i|)^p} \mu(|x_i|), \end{aligned}$$

où la dernière ligne découle de l'inégalité de Markov. D'après l'hypothèse sur \mathbf{h}_μ , on peut majorer $\mu(|x_i|)$:

$$\begin{aligned}\mu(\mathcal{D}_{p,C}^c) &\leq \sum_{i \in \mathbb{Z}^d} \frac{c_\mu}{C(1+|i|)^p} (1+|i|)^{p_\mu} \\ &\leq \frac{c_\mu}{C} \sum_i (1+|i|)^{p_\mu-p}.\end{aligned}$$

Choisissons p tel que $p > p_\mu + d$. Ce choix garantit que la somme qui apparaît ci-dessus est finie, et il existe donc une constante c_d dépendant de la dimension d du réseau, telle que :

$$\mu(\mathcal{D}_{p,C}^c) \leq \frac{c_\mu c_d}{C}.$$

En prenant la borne inférieure sur les valeurs de C , comme C est arbitrairement grand, on voit que le poids donné par μ aux configurations croissant plus vite que $|i|^p$ est nul :

$$\mu(\mathcal{D}_p^c) = \inf_{C>0} \mu(\mathcal{D}_{p,C}^c) = 0.$$

2^e étape : $\mathcal{G}_t^{\text{Roy.}} \subset \mathcal{G}_t^{\text{Copp.}}$. Soit μ une mesure dans $\mathcal{G}_t^{\text{Roy.}}$, qui ne charge donc que les configurations tempérées. L'idée est d'utiliser un contrôle sur les mesures de Gibbs en boîte finies, et de l'injecter dans la propriété de compatibilité *DLR* pour montrer l'intégrabilité du spin en 0. L'inégalité en boîte finie est donnée par le lemme 19. On veut donc maintenant utiliser la propriété *DLR*. Pour ce faire, fixons un réel positif h et une mesure de Gibbs $\mu \in \mathcal{G}_t^{\text{Roy.}}$.

Sur toute boîte finie Λ , et pour toute configuration tempérée z , on peut appliquer le lemme 19 :

$$\begin{aligned}\int (x_0^2 \wedge h) d\mu_{\Lambda,z}(\mathbf{x}) &\leq \frac{1}{\alpha(0)} \int \alpha(0) x_0^2 d\mu_{\Lambda,z}(\mathbf{x}) \\ &\leq \frac{1}{\alpha(0)} \mu_{\Lambda,z}(\|\mathbf{x}\|_\alpha^2) \\ &\leq \frac{1}{\alpha(0)} \left(A + B \sum_{i \notin \Lambda} \alpha(i) z_i^2 \right).\end{aligned}$$

Si l'on note

$$f_n(\mathbf{z}) = \int (x_0^2 \wedge h) d\mu_{\Lambda_n, \mathbf{z}}(\mathbf{x}),$$

on a donc, pour tout z , une majoration de $f_n(\mathbf{z})$. Prenons la limite supérieure pour n tendant vers l'infini : comme \mathbf{z} est tempérée et α à décroissance exponentielle, $\limsup \sum_{\Lambda_n^c} \alpha(i) z_i^2 = 0$, et on obtient la majoration indépendante de z :

$$\limsup_n f_n(\mathbf{z}) \leq \frac{A}{\alpha_0}.$$

Comme les f_n sont uniformément bornées (par h), on peut appliquer le lemme de Fatou à la suite de fonctions positives $g_n = h - f_n$ pour l'intégration par rapport à μ . On obtient $\int \liminf g_n \leq \liminf \int g_n$, et en repassant aux fonctions f_n ,

$$\limsup \int f_n(\mathbf{z}) d\mu(\mathbf{z}) \leq \int \limsup f_n(\mathbf{z}) d\mu(\mathbf{z}).$$

La suite qui apparaît à gauche est en fait constante : grâce aux équations *DLR*,

$$\int f_n(\mathbf{z}) d\mu(\mathbf{z}) = \int \int (x_0^2 \wedge h) d\mu_{\Lambda_n, \mathbf{z}}(\mathbf{x}) d\mu(\mathbf{z}) = \int (x_0^2 \wedge h) d\mu(\mathbf{x}).$$

Quant au terme de droite, il est majoré par A/α_0 . Finalement,

$$\int (x_0^2 \wedge h) d\mu \leq \frac{A}{\alpha_0}.$$

Comme ceci est valable pour tout h et que le membre de droite est indépendant de h , on a par convergence monotone pour $h \rightarrow \infty$:

$$\int x_0^2 d\mu \leq \frac{A}{\alpha_0}.$$

Pour finir, on remarque que la même démonstration marche pour n'importe quel site à la place de 0 (pour le prouver en i , on applique la même technique avec un poids translaté $\tilde{\alpha}(\cdot) = \alpha(\cdot - i)$, pour lequel le lemme 19 est toujours valable, avec les mêmes constantes A et B , grâce à l'invariance par translation des interactions). On peut donc écrire, grâce à l'inégalité de Jensen :

$$\|\mathbf{h}_\mu\|_\infty = \sup_i \mu(|x_i|) \leq \sup_i \mu(x_i^2)^{1/2} < \infty$$

La mesure μ est donc bien dans $\mathcal{G}_t^{\text{Copp}}$, ce que nous voulions démontrer.

Troisième étape : les mesures tempérées ne chargent que certaines configurations Il s'agit de montrer qu'il existe un p tel que toute mesure tempérée ne charge en fait que des configurations de \mathcal{P}_p . En réalité, c'est une conséquence des preuves des inclusions précédentes. En effet, si μ est une mesure tempérée, on sait que le vecteur h_μ est borné. En particulier h_μ est dans \mathcal{P}_0 . D'après la preuve de $\mathcal{G}_t \subset \mathcal{G}_t^{\text{Mat}}$, pour tout $p > d$, on a $\mu(\mathcal{P}_p) = 1$: c'est la propriété cherchée, et tout $p > d$ convient.

Ceci achève la preuve du théorème 16.

2.3 Existence de mesures de Gibbs tempérées

La première question à poser, avant d'étudier si l'état de Gibbs en volume infini est unique, est celle de l'existence de mesures de Gibbs tempérées. Ce type de résultat est souvent obtenu par un argument de compacité, et nous ne dérogerons pas à la règle. Le point problématique est justement de montrer qu'un certain ensemble de mesures est compact, alors que chaque spin évolue dans un espace non-compact. Il va donc falloir utiliser de la tension ; le contrôle dans $l^2(\alpha)$ obtenu dans la section précédente (dans le lemme 19) va nous y aider.

La mesure de Gibbs en volume infini sera donc obtenue comme point limite d'une suite. Un choix naturel pour cette suite est de considérer les $\mu_{\Lambda_n, \mathbf{z}}$, où \mathbf{z} est une configuration tempérée fixée, et Λ_n est la boîte $[-n, n]^d$. Pour simplifier, on fixe $\mathbf{z} = 0$ et on abrège $\mu_{\Lambda_n, 0}$ en μ_n .

Soit $\varepsilon > 0$, il nous faut trouver un compact K_ε de $\mathbb{R}^\mathcal{S}$ (pour la topologie produit) tel que :

$$\forall n, \quad \mu_n(K_\varepsilon) \geq 1 - \varepsilon.$$

Comme $\mu_n(\|\mathbf{x}\|_\alpha) \leq \sqrt{\mu_n(\|\mathbf{x}\|_\alpha^2)}$, le lemme 19 nous donne directement une constante A telle que :

$$\mu_n(\|\mathbf{x}\|_\alpha^2) \leq A^2, \tag{2.5}$$

$$\mu_n(\|\mathbf{x}\|_\alpha) \leq A. \tag{2.6}$$

Soit donc K_ε l'ensemble de configurations suivant :

$$K_\varepsilon = \left\{ \mathbf{x}, \|\mathbf{x}\|_\alpha \leq \frac{A}{\varepsilon} \right\}.$$

Par l'inégalité de Markov,

$$\mu_n(K_\varepsilon^c) = \mu_n \left(\left\{ \|\mathbf{x}\|_\alpha > \frac{A}{\varepsilon} \right\} \right) \leq \frac{\varepsilon}{A} \mu_n(\|\mathbf{x}\|_\alpha \leq \varepsilon).$$

De plus K_ε est compact. En effet, il est inclus dans le compact

$$\prod \left[-\frac{A^2}{\varepsilon^2 \alpha(i)}, \frac{A^2}{\varepsilon^2 \alpha(i)} \right]$$

(puisque $\alpha(i)x_i^2 \leq A^2$ dès que \mathbf{x} est dans K_ε); et il est fermé (la fonction $\mathbf{x} \mapsto \|\mathbf{x}\|_\alpha$ est semi-continue inférieurement, par exemple comme borne supérieure des fonctions continues $\mathbf{x} \mapsto (\sum_{|i| \leq n} \alpha(i)x_i^2)^{1/2}$).

La suite μ_n est donc tendue, et d'après le théorème de Prohorov, elle admet une sous-suite qui converge faiblement vers une mesure de probabilités μ .

Cette mesure limite est une mesure de Gibbs. Pour le voir, il suffit de montrer, pour toute boîte finie L , la propriété :

$$\mu \mu_L = \mu,$$

où μ_L est le noyau $\mathbf{z} \mapsto \mu_{L,\mathbf{z}}$. Soit f une fonction bornée, cylindrique, lipschitzienne pour la norme l^1 sur son support :

$$|f(\mathbf{x}) - f(\mathbf{y})| \leq C_f \sum_{i \in \text{supp } f} |x_i - y_i|.$$

Pour n suffisamment grand (pour garantir $L \subset \Lambda_n$ et f à support dans Λ_n), on a :

$$\mu_n \mu_L f = \mu_n f,$$

par la propriété de compatibilité en boîte finie. Comme $\mu_L f$ est continue (cf. l'annexe A) et bornée, la convergence faible de μ_n le long de la sous-suite assure :

$$\mu \mu_L f = \mu f.$$

D'après la proposition 66 (annexe A), ceci implique $\mu \mu_L = \mu$, et la mesure μ est donc bien de Gibbs.

Montrons enfin qu'elle est tempérée. On réutilise l'inégalité (2.5), dont on déduit :

$$\mu_n(x_0^2) \leq \frac{A^2}{\alpha(0)}.$$

Pour tout h positif, en notant $A' = A^2 / \alpha(0)$, on a :

$$\mu_n(h \wedge x_0^2) \leq A'.$$

On utilise la convergence faible pour obtenir $\mu(h \wedge x_0^2) \leq A'$, puis la convergence monotone pour $h \rightarrow \infty$. Finalement, $\mu(|x_0|) \leq A'$. On peut faire le même raisonnement au site i avec un poids translaté. Il existe donc une constante A' , indépendante de i , telle que :

$$\mu(|x_i|) \leq A'.$$

Ceci montre que la mesure μ est tempérée (cf. la section précédente).

Nous avons donc montré, pour $\mathbf{z} = 0$, le résultat :

20 Théorème. *Il existe une mesure de Gibbs tempérée en volume infini.*

De plus, pour toute configuration tempérée \mathbf{z} , la suite $\mu_{\Lambda_n, \mathbf{z}}$ des mesures de Gibbs dans la boîte Λ_n avec des conditions aux bords \mathbf{z} admet une sous-suite qui converge vers une mesure de Gibbs tempérée.

Le cas où \mathbf{z} est quelconque se traite exactement de la même manière, seule la constante A doit être modifiée pour tenir compte de \mathbf{z} .

2.4 Unicité : le critère de Dobrushin

Nous rappelons dans cette section une preuve classique d'unicité de la mesure de Gibbs, adaptée au cadre des spins réels.

Ce cas est en réalité prévu dès l'article fondateur de Dobrushin ([Dob70]), mais les présentations simplifiées des références plus récentes se concentrent souvent sur le cas particulier des spins compacts, et abandonnent en particulier les distances de Wasserstein pour se limiter à l'utilisation de la distance en variation. Une exception à cette règle est l'article de A.Val. Antoniouk et A.Vict. Antoniouk, [AA96], dont nous nous inspirons directement. La principale modification est que nous explicitons l'usage des distances de Wasserstein, sous-jacentes dans [AA96].

Dans un second temps, nous tentons de clarifier la région de l'espace des paramètres où ce critère s'applique.

Le critère

Pour cette section seulement, nous changeons légèrement les notations, en réintroduisant le paramètre β de température inverse. Le but principal est en effet de comprendre pour quels paramètres (hautes températures, ou interactions faibles) le critère de Dobrushin peut s'appliquer. Nous remplaçons donc le hamiltonien de la section 2.1 par l'expression (formelle) suivante :

$$U(\mathbf{x}) = \sum V(x_i) + J \sum_{i \sim j} (x_i - x_j)^2, \quad (2.7)$$

et nous étudions les mesures de densité $Z_{\beta, J}^{-1} \exp(-\beta H)$ (qui dépendent donc de β). Notons que nous avons également simplifié le terme d'interaction en nous limitant aux interactions aux plus proches voisins.

Nous aurons besoin de l'estimation *a priori* suivante.

21 Théorème. *Pour toute mesure de Gibbs tempérée μ , il existe c_μ telle qu'en tout site i ,*

$$\mu(|x_i|) \leq c_\mu.$$

Il s'agit en réalité de l'une des définitions équivalentes des mesures tempérées (*cf.* le théorème 16).

Le critère de Dobrushin a déjà été énoncé dans l'introduction, nous le rappelons ici, sous une forme légèrement différente.

Rappelons que \mathcal{C}_j est l'ensemble des couples de configurations (\mathbf{x}, \mathbf{y}) qui ne diffèrent qu'en j : $(\mathbf{x}, \mathbf{y}) \in \mathcal{C}_j \iff \forall i \neq j, x_i = y_i$, et que d_W est la distance de Wasserstein entre deux mesures sur \mathbb{R} .

22 Théorème. *Supposons qu'il existe une famille C_{ij} telle que, pour tout couple (i, j) , et pour tout couple $(\mathbf{x}, \mathbf{y}) \in \mathcal{C}_j$,*

$$d_W(\mu_{(i)}^{\mathbf{x}}, \mu_{(i)}^{\mathbf{y}}) \leq C_{i,j} |x_j - y_j|.$$

La quantité C_{ij} quantifie l'influence du site j sur i .

Alors, si

$$\sup_i \sum_j C_{ij} = \gamma < 1,$$

il ne peut y avoir qu'au plus une mesure de Gibbs tempérée.

Pour le voir, considérons deux mesures de Gibbs tempérées μ et ν . Pour montrer qu'elles sont égales, il suffit de montrer :

$$|\mu(f) - \nu(f)| = 0,$$

pour toute fonction f dans une classe suffisamment grande. On choisira la classe \mathcal{C} des fonctions cylindriques, \mathcal{C}^1 , de norme triple finie, où la norme triple vaut :

$$\|f\| = \sum_i \|\nabla_i f\|_\infty.$$

Nous allons établir une suite de « bonnes estimations » de la différence entre μ et ν , qui à la limite nous donnera l'égalité voulue.

23 Définition. *Une famille de réels $(a_i)_{i \in \mathbb{Z}^d}$ est une « bonne estimation » de la différence entre μ et ν si, pour toute fonction f de \mathcal{C} ,*

$$|\mu f - \nu f| \leq \sum_i a_i \|\nabla_i f\|_\infty. \quad (2.8)$$

La preuve se réduit alors à montrer le lemme suivant.

24 Lemme. – *Il existe un réel a tel que la famille constante \bar{a} définie par*

$$\forall i, \quad \bar{a}_i = a$$

soit une bonne estimation.

– *Si (a_i) est une bonne estimation, il en est de même pour la famille (Ca) , définie par*

$$(Ca)_i = \sum_j C_{ij} a_j.$$

Admettons pour l'instant le lemme. Une récurrence immédiate montre que $(C^n \bar{a})$ est une bonne estimation. Une autre récurrence nous donne alors $\|C^n\|_\infty \leq \gamma^n a$. En reportant ceci dans la définition d'une bonne estimation, il vient :

$$\begin{aligned} |\mu f - \nu f| &\leq \sum_i (C^n \bar{a})_i \|\nabla_i f\|_\infty \\ &\leq a \gamma^n \sum_i \|\nabla_i f\|_\infty \\ &\leq a \gamma^n \|f\|. \end{aligned}$$

Comme ceci est valable pour tout n et que γ est strictement inférieur à 1, $\mu(f) = \nu(f)$.

Il nous reste donc à montrer le lemme.

Preuve du lemme 24 Montrons d'abord que (μ, ν) est bien estimée par une famille constante. Pour cela, on utilise le fait que f est cylindrique. Si l'on note $\text{supp}(f)$ son support, et i_1, \dots, i_m une énumération de ce support, on peut donc décomposer $f(\mathbf{x}) - f(0)$ en construisant une suite de configurations qui interpolent entre 0 et \mathbf{x} :

$$\mathbf{x}^l = (\mathbf{x}^{(i_1, i_2, \dots, i_l)} 0).$$

On a donc $\mathbf{x}^0 = 0$, et \mathbf{x}^m coïncide avec \mathbf{x} sur le support de f .

On décompose $f(\mathbf{x}) - f(0)$:

$$\begin{aligned} |f(\mathbf{x}) - f(0)| &= \left| \sum_{l=1}^m f(\mathbf{x}^l) - f(\mathbf{x}^{l-1}) \right| \\ &\leq \sum_{k \in \text{supp}(f)} \|\nabla_k f\|_\infty x_k. \end{aligned}$$

On en déduit :

$$\begin{aligned} |\mu f - \nu f| &= |\mu(f - f(0)) - \nu(f - f(0))| \\ &\leq \sum_{k \in \text{supp}(f)} (\|\nabla_k f\|_\infty \mu(|x_k|)) + \sum_{k \in \text{supp}(f)} (\|\nabla_k f\|_\infty \nu(|x_k|)) \\ &\leq (c_\mu + c_\nu) \|f\|, \end{aligned}$$

où c_μ, c_ν sont donnés par le théorème 21. Le premier point du lemme est prouvé.

Soit maintenant a une bonne estimation ; montrons que (Ca) en est une. Pour cela, définissons pour tout volume fini J un poids a^J par :

$$a_i^J = \begin{cases} \min(a_i, (Ca)_i), & \text{si } i \in J, \\ a_i & \text{sinon.} \end{cases} \quad (2.9)$$

Si a^J est, pour tout J fini, une bonne estimation, alors (Ca) en est une : pour n'importe quelle fonction f dans \mathcal{C} , il suffit de prendre $J = \text{supp}(f)$ pour montrer l'équation (2.8). On s'est donc ramené à montrer, pour tout volume fini J et tout site $k \notin J$,

$$\text{Si } a^J \text{ est une bonne estimation, alors } a^{J \cup \{k\}} \text{ en est une.} \quad (2.10)$$

Notons f_k la fonction de deux variables

$$f_k(\mathbf{x}, \mathbf{y}) = \int_{\mathbb{R}} f((z^{(k)} \mathbf{x})) d\mu_{\{k\}}^{\mathbf{y}}(z),$$

et

$$g_k(\mathbf{x}) = f_k(\mathbf{x}, \mathbf{x}).$$

La relation DLR nous dit :

$$\int f(\mathbf{x}) d\mu(\mathbf{x}) = \int g_k(\mathbf{x}) d\mu(\mathbf{x}).$$

Majorons $\mu(f) - \nu(f)$ en utilisant ce fait, et la bonne estimation a^J .

$$\begin{aligned} \left| \int f d\mu - \int f d\nu \right| &= \left| \int g_k(\mathbf{x}) d\mu - \int g_k(\mathbf{x}) d\nu \right| \\ &\leq \sum_j a_j^J \|\nabla_j g_k\|_\infty. \end{aligned} \quad (2.11)$$

Pour étudier le terme $\|\nabla_j g_k\|_\infty$, on utilise la formulation équivalente :

$$\|\nabla_j g_k\|_\infty = \sup \left\{ \left| \frac{g_k(\mathbf{x}) - g_k(\mathbf{y})}{x_j - y_j} \right| ; (\mathbf{x}, \mathbf{y}) \in \mathcal{C}_j \right\}. \quad (2.12)$$

Majorons maintenant le terme $|g_k(\mathbf{x}) - g_k(\mathbf{y})|$.

$$\begin{aligned} |g_k(\mathbf{x}) - g_k(\mathbf{y})| &= |f_k(\mathbf{x}, \mathbf{x}) - f_k(\mathbf{x}, \mathbf{y}) + f_k(\mathbf{x}, \mathbf{y}) - f_k(\mathbf{y}, \mathbf{y})| \\ &\leq |f_k(\mathbf{x}, \mathbf{x}) - f_k(\mathbf{x}, \mathbf{y})| + |f_k(\mathbf{x}, \mathbf{y}) - f_k(\mathbf{y}, \mathbf{y})|. \end{aligned} \quad (2.13)$$

Pour majorer le premier terme, on revient à la définition de f_k . Introduisons la fonction $h_k^{\mathbf{x}}$

$$\begin{aligned} h_k^{\mathbf{x}} : \mathbb{R} &\rightarrow \mathbb{R}, \\ z &\mapsto f((z \stackrel{(k)}{\cdot} \mathbf{x})) \end{aligned}$$

La différence cherchée est donc :

$$|f_k(\mathbf{x}, \mathbf{x}) - f_k(\mathbf{x}, \mathbf{y})| = \left| \mu_{\{k\}}^{\mathbf{x}}(h_k^{\mathbf{x}}) - \mu_{\{k\}}^{\mathbf{y}}(h_k^{\mathbf{x}}) \right|.$$

Comme $h_k^{\mathbf{x}}$ est lipschitzienne, et comme on contrôle la distance de Wasserstein entre les deux mesures considérées (c'est l'hypothèse du critère), on a (cf. par exemple le théorème 11.8.2 de [Dud02]) :

$$|f_k(\mathbf{x}, \mathbf{x}) - f_k(\mathbf{x}, \mathbf{y})| \leq C_{kj} |x_j - y_j| \|h_k^{\mathbf{x}}\|_{Lip} \quad (2.14)$$

Par définition de $h_k^{\mathbf{x}}$, sa norme de Lipschitz est bornée, uniformément en \mathbf{x} , par $\|\nabla_k f\|_\infty$, ce qui donne :

$$|f_k(\mathbf{x}, \mathbf{x}) - f_k(\mathbf{x}, \mathbf{y})| \leq C_{kj} |x_j - y_j| \|\nabla_k f\|_\infty. \quad (2.15)$$

Pour estimer le second terme de (2.13), on écrit directement :

$$\begin{aligned} |f_k(\mathbf{x}, \mathbf{y}) - f_k(\mathbf{y}, \mathbf{y})| &\leq \int_{\mathbb{R}} \left| f((z \stackrel{(k)}{\cdot} \mathbf{x})) - f((z \stackrel{(k)}{\cdot} \mathbf{y})) \right| d\mu_{\{k\}}^{\mathbf{y}}(z) \\ &\leq |x_j - y_j| \|\nabla_j f\|_\infty. \end{aligned} \quad (2.16)$$

En insérant (2.15) et (2.16) dans (2.13), on a donc :

$$|g_k(\mathbf{x}) - g_k(\mathbf{y})| \leq |x_j - y_j| (C_{kj} \|\nabla_k f\|_\infty + \|\nabla_j f\|_\infty).$$

L'équation (2.12) donne alors, pour $j \neq k$ (le terme correspondant pour $j = k$ est nul) :

$$\|\nabla_j g_k\|_\infty \leq C_{kj} \|\nabla_k f\|_\infty + \|\nabla_j f\|_\infty. \quad (2.17)$$

Revenons alors à (2.11) :

$$\begin{aligned} \left| \int f d\mu - \int f dv \right| &\leq \sum_j a_j^I \|\nabla_j g_k\|_\infty && \text{(par (2.11))} \\ &\leq \left(\sum_{j \neq k} a_j^I C_{kj} \right) \|\nabla_k f\|_\infty + \sum_{j \neq k} a_j^I \|\nabla_j f\|_\infty && \text{(par (2.17))} \\ &\leq (Ca)_k \|\nabla_k f\|_\infty + \sum_{j \neq k} a_j^I \|\nabla_j f\|_\infty && \text{(car } Ca^J \leq Ca) \end{aligned}$$

Par ailleurs, a^J est une bonne estimation, donc :

$$\left| \int f d\mu - \int f d\nu \right| \leq (a^J)_k \|\nabla_k f\|_\infty + \sum_{j \neq k} a_j^J \|\nabla_j f\|_\infty.$$

On peut donc choisir la meilleure de ces deux majorations. Par définition de $a^{J \cup \{k\}}$ (équation (2.9)), on obtient :

$$\left| \int f d\mu - \int f d\nu \right| \leq \sum_j a_j^{J \cup \{k\}} \|\nabla_j f\|_\infty.$$

Ceci achève la preuve du lemme 24, et du critère de Dobrushin.

Pour quels paramètres a-t-on unicité ?

Une fois le critère établi, il est naturel de se demander s'il permet, dans le cadre des spins réels, de montrer des résultats analogues à ceux valables pour le modèle classique. Ceux-ci sont multiples (cf. la section 8.1 de [Geo88], et particulièrement les exemples 8.9, 8.13). Citons en deux.

25 Théorème. *Dans le modèle d'Ising classique, le critère de Dobrushin s'applique si β est suffisamment petit, ou, à β fixé, si le champ magnétique h est suffisamment important (en valeur absolue).*

Le principal problème est que les résultats précédents sont montrés dans un cas où la distance de Wasserstein devient une distance en variation, ce qui facilite les calculs. Il est cependant possible d'estimer les distances de Wasserstein, car les mesures à comparer sont des mesures sur \mathbb{R} . Ceci permet d'établir le résultat suivant :

26 Théorème ([COPP78], Théorème 2.3). *Pour le modèle défini par (2.7), les coefficients de la matrice de dépendance C_{ij} vérifient :*

$$C_{ij} = \beta J \mathbf{1}_{i \sim j} \sup_t \int \left(x - \int y d\mu_t(y) \right)^2 d\mu_t(x),$$

où μ_t est définie par :

$$d\mu_t(x) = \frac{1}{Z_t} \exp \left(-\beta V(x) - \beta J \left(\sum_j \mathbf{1}_{i \sim j} \right) x^2 + tx \right).$$

Pour pouvoir appliquer le critère de Dobrushin, il nous faut donc borner la quantité :

$$\int \left(x - \int y d\mu_t(y) \right)^2 d\mu_t(x)$$

indépendamment de t . Pour cela, il suffit de contrôler la constante de Poincaré de μ_t :

$$\int \left(x - \int y d\mu_t(y) \right)^2 d\mu_t(x) = \mathbf{Var}_{\mu_t}(x \mapsto x) \leq C_P(\mu_t).$$

On se ramène donc à montrer une propriété de trou spectral uniforme sur une famille de mesures perturbées.

Remarque. Cette propriété est étudiée en particulier dans [GR01], où elle est baptisée « uniform spectral gap inequality ». Elle peut aussi être vue comme un trou spectral indépendant des conditions aux bord, mais restreint aux boîtes de taille 1.

Cette propriété unidimensionnelle est essentielle dans la preuve des inégalités uniformes sur toutes les boîtes et conditions aux bord, telle qu'elle est développée dans [BH99b, Led01, GR01] (cf. l'introduction).

À ce propos, on peut remarquer que la région de l'espace des paramètres (β, J) où [GR01, Led01] montrent un trou spectral uniforme (sur les boîtes et les conditions aux bords) est la même que celle où le critère d'unicité de Dobrushin est valable (cf. les valeurs explicites de β, J dans le théorème suivant et dans [GR01]). Autrement dit, il n'est pas prouvé qu'il existe des valeurs de β, J pour lesquelles on ait analyticit  compl te sans pour autant v rifier le crit re de Dobrushin.

Pour expliciter les conditions sur β, J , nous utiliserons les deux propri t s classiques suivantes :

27 Th or me. Si la fonction ψ v rifie $\psi'' \geq c_\psi > 0$, alors la mesure μ_ψ   densit  $1/Z_\psi \exp(-\psi)$ v rifie une in galit  de Poincar  de constante $C_\psi = \frac{1}{c_\psi}$.

Si la mesure μ , de densit  $Z_V^{-1} \exp(-V)$, v rifie une in galit  de Poincar  de constante C_μ , alors pour toute fonction born e W , la mesure $\nu = Z_{V,W} \exp(-V - W)$ en v rifie une pour la constante

$$C_\nu = C_\mu \exp(\text{osc } W).$$

On pourra trouver une preuve du premier r sultat dans [Led01], corollaire 1.4, et du second dans [Roy99], proposition 3.1.18.

On obtient alors les r sultats suivants.

28 Th or me. Cas strictement convexe. Si V est strictement convexe, alors pour tous β et J , la mesure de Gibbs temp r e est unique. En particulier, si le mod le est gaussien ($V = ax^2$), il n'y a qu'une mesure de Gibbs temp r e.

Cas Convexe   l'infini. Si V est convexe   l'infini, $V = V_1 + V_2$, avec $V_1'' \geq c_1$, et si

$$\frac{2dJ}{c_1 + 2dJ} \exp(\beta \text{osc}(V_2)) < 1,$$

alors la mesure de Gibbs temp r e est unique. En particulier, ceci est vrai :

-   J fix , pour β suffisamment petit (hautes temp ratures),
-   β fix , pour J suffisamment petit (interactions faibles).

Terminons cette section par une courte discussion. Nous avons retrouv  sous deux formes (haute temp rature ou interactions faibles) des r sultats valables dans le mod le d'Ising habituel (pour β petit). En revanche, ce crit re ne peut en aucun cas montrer l'unicit  si l'on rajoute un champ magn tique h au mod le : la quantit 

$$\sup_t \int (x - \int y d\mu_t)^2 d\mu_t(x)$$

est invariante si l'on change $V(x)$ en $V(x) + hx$. Comme les spins sont quelconques, on ne peut pas garantir que le champ magn tique ait plus d'influence sur un site que tous ses voisins r unis. Il faudrait pour contourner cette difficult  trouver un  quivalent du crit re de Dobrushin qui n'aurait besoin d' tre v rifi  que pour des valeurs « cr dibles » des spins du bord.

Notons aussi que, contrairement au cas standard, il n'y a pas à notre connaissance de preuve de l'unicité de la mesure de Gibbs en dehors du régime d'application du critère de Dobrushin. D'autres conditions suffisantes d'unicité peuvent être données (cf. la section suivante, qui montre l'unicité sous une condition faible de mélange, et les théorèmes 40, 44 inspirés des travaux de G. Royer), mais elles ne donnent pas lieu à des estimations explicites en terme des paramètres du hamiltonien.

2.5 Unicité : le critère de Dobrushin–Shlosman

Nous mettons au clair dans cette section la preuve de l'unicité de la mesure de Gibbs sous la condition de Dobrushin–Shlosman, et ses liens avec la propriété de mélange faible.

Rappelons d'abord les définitions des deux propriétés mises en jeu.

29 Définition. *On dit que la condition d'unicité de Dobrushin–Shlosman (DSU) est vérifiée quand il existe un volume Λ et une famille ρ_j tels que :*

$$\forall j \in \Lambda^c, \forall (\mathbf{x}, \mathbf{y}) \in \mathcal{C}_j, \quad d_W(\mu_{\Lambda, \mathbf{x}}, \mu_{\Lambda, \mathbf{y}}) \leq \rho_j |x_j - y_j|, \quad (2.18)$$

$$\sum_{j \in \Lambda^c} \rho_j < 1. \quad (2.19)$$

Dans cette définition, la distance de Wasserstein entre mesures sur \mathbb{R}^Λ provient de la distance L^1 sur \mathbb{R}^Λ :

$$d(\mathbf{x}, \mathbf{y}) = \sum_{i \in \Lambda} |x_i - y_i|.$$

Cette propriété est un analogue « multi-sites » de la condition de Dobrushin. En particulier, elle est *plus faible* que cette dernière (et sa région d'application est donc a priori plus vaste).

Son inconvénient est qu'il n'existe plus de façon simple d'estimer les distances de Wasserstein entre des mesures sur \mathbb{R}^Λ .

30 Définition. *La famille des $\mu_{\Lambda, \mathbf{x}}$ vérifie une propriété de mélange faible exponentiel si il existe une constante C telle que, pour tout volumes $\Delta \subset \Lambda$, tout site j de la frontière (extérieure) $\partial\Delta$, et tout couple de conditions aux bords $(\mathbf{x}, \mathbf{y}) \in \mathcal{C}_j$,*

$$d_W(\mu_{\Lambda, \Delta, \mathbf{x}}, \mu_{\Lambda, \Delta, \mathbf{y}}) \leq C \left(\sum_{k \in \partial\Delta} \exp\left(-\frac{d(\Delta, k)}{C}\right) \right) |x_j - y_j|, \quad (2.20)$$

où $\mu_{\Lambda, \Delta, \mathbf{x}}$ est la marginale de $\mu_{\Lambda, \mathbf{x}}$ sur la boîte Δ .

Ces propriétés garantissent l'unicité de la mesure de Gibbs tempérée.

31 Théorème. *Si la condition de Dobrushin–Shlosman est vérifiée, ou si la propriété de mélange faible exponentiel est vérifiée, la mesure de Gibbs tempérée en volume infini est unique.*

De plus, si les interactions sont attractives et si la variance d'un spin est majorée a priori, la condition de Dobrushin–Shlosman entraîne la propriété de mélange faible exponentiel.

Remarque. *La preuve fera clairement apparaître qu'un mélange faible polynomial suffisamment rapide entraîne lui aussi l'unicité de la mesure de Gibbs.*

Du critère DSU au mélange faible et à l'unicité.

On cherche à utiliser le critère **DSU** sur des translatsés du volume Λ . Pour cela, on notera $\overline{\Lambda}$ la réunion de Λ avec sa frontière $\partial\Lambda$, et pour un (grand) volume L , on notera $\Lambda^{(k)}$ les translatsés de Λ tels que $\overline{\Lambda}^{(k)} \subset L$. On notera $\partial_\Lambda L$ la frontière intérieure épaissie de L :

$$\partial_\Lambda L = \{i \in L, d(i, L^c) \leq \text{diam}(\Lambda)\}.$$

Enfin, pour tout couplage μ de deux mesures sur un volume quelconque, on notera

$$f_i(\mu) \stackrel{\text{def}}{=} \int |x_i - x'_i| d\mu(\mathbf{x}, \mathbf{x}').$$

La définition de la distance de Wasserstein entraîne que, pour tout couplage μ de deux mesures μ_1, μ_2 sur un volume L ,

$$d_W(\mu_1, \mu_2) \leq \sum_{i \in L} f_i(\mu). \quad (2.21)$$

On utilisera deux lemmes.

32 Lemme. *Soit L un volume quelconque, et $\delta > 0$. Si le critère **DSU** est vérifié dans le volume Λ avec une famille (ρ_j) , alors pour toutes configurations \mathbf{y}, \mathbf{z} , il existe $\mu_{L, \mathbf{y}, \mathbf{z}}$ un couplage de $(\mu_{L, \mathbf{y}}, \mu_{L, \mathbf{z}})$ tel que, pour tous les translatsés $\Lambda^{(k)}$,*

$$\sum_{j \in \Lambda^{(k)}} f_j(\mu_{L, \mathbf{y}, \mathbf{z}}) \leq \sum_{j \in \partial\Lambda^{(k)}} \rho_j f_j + \delta. \quad (2.22)$$

En d'autres termes, on peut trouver un couplage tel que chaque sous-volume soit contrôlé par son bord.

33 Lemme. *Il existe une constante C telle que, pour tous volumes $M \subset L$, tout $\delta > 0$ et toutes conditions aux bords \mathbf{y}, \mathbf{z} , il existe un couplage $\mu_{L, \mathbf{y}, \mathbf{z}}$ de $(\mu_{L, \mathbf{y}}, \mu_{L, \mathbf{z}})$ tel que*

$$\sum_{i \in L} \alpha_i f_i(\mu_{L, \mathbf{y}, \mathbf{z}}) \leq C \sum_{i \in \partial_\Lambda L} \alpha_i f_i(\mu_{L, \mathbf{y}, \mathbf{z}}) + \delta |L|,$$

où $\alpha_i = \exp(-\frac{1}{C}d(i, M))$.

Démonstration. Ce sont les lemmes 2.1 et 2.2 de [DS85b]. □

Preuve de l'unicité Soit μ, ν deux mesures de Gibbs tempérées. Pour montrer qu'elles sont égales, il suffit de montrer que pour tout volume M , leurs marginales sur M coïncident. Soit donc M un volume quelconque, on cherche à estimer la distance de Wasserstein entre μ_M et ν_M . Pour tout boîte L contenant M , on peut construire un couplage de μ_M et ν_M de la façon suivante :

- on jette deux configurations \mathbf{y} et \mathbf{z} , respectivement suivant μ et ν ;
- on associe à \mathbf{y} et \mathbf{z} un couplage (mesure sur les couples de configurations dans le volume L) défini par le lemme 33. Ce couplage donne, par marginales, un couplage de μ_M et ν_M (c'est la propriété DLR).

Plus formellement, on définit $\tilde{\mu}$, couplage de (μ_M, ν_M) par :

$$\int f(\mathbf{x}, \mathbf{x}') d\tilde{\mu}(\mathbf{x}, \mathbf{x}') \leq \iiint f(\mathbf{x}, \mathbf{x}') d\mu_{L, \mathbf{y}, \mathbf{z}} d\mu(\mathbf{y}) d\nu(\mathbf{z}),$$

pour toute f à support dans M .

On veut estimer la distance de Wasserstein entre μ_M et ν_M , on utilise (2.21) :

$$\begin{aligned} d_W(\mu_M, \nu_M) &\leq \sum_{i \in M} f_i(\bar{\mu}) \\ &\leq \iint \sum_{i \in M} f_i(\mu_{L,y,z}) d\mu(\mathbf{y}) d\nu(\mathbf{z}). \end{aligned} \quad (2.23)$$

On utilise le lemme 33 pour majorer $f_i(\mu_{L,y,z})$ dans le volume M :

$$\begin{aligned} \sum_{i \in M} f_i(\mu_{L,y,z}) &\leq \exp\left(\frac{1}{C} \text{diam}(M)\right) \sum_{i \in M} \exp\left(-\frac{1}{C} d(i, M)\right) f_i(\mu_{L,y,z}) \\ &\leq C_1 \sum_{i \in L} \exp\left(-\frac{1}{C} d(i, M)\right) f_i(\mu_{L,y,z}) \\ &\leq C_1 C \sum_{i \in \partial_\Lambda L} \exp\left(-\frac{1}{C} d(i, M)\right) f_i(\mu_{L,y,z}) + \delta |L|. \end{aligned} \quad (2.24)$$

On intègre cette relation par rapport à $d\mu(\mathbf{y}) d\nu(\mathbf{z})$; grâce à (2.23), on obtient :

$$d_W(\mu_M, \nu_M) \leq C C_1 \sum_{i \in \partial_\Lambda L} \exp\left(-\frac{1}{C} d(i, M)\right) f_i(\bar{\mu}) + \delta |L|.$$

La quantité $f_i(\bar{\mu})$ est bornée :

$$\begin{aligned} f_i(\bar{\mu}) &= \iiint |x_i - x'_i| d\mu_{L,y,z}(\mathbf{x}, \mathbf{x}') d\mu(\mathbf{y}) d\nu(\mathbf{z}) \\ &\leq \int |x_i| d\mu(\mathbf{x}) + \int |x'_i| d\nu(\mathbf{x}') \\ &\leq c_\mu + c_\nu, \end{aligned}$$

où l'on a utilisé la propriété DLR, puis le caractère tempéré de μ et ν . Quitte à modifier C , on a donc :

$$d_W(\mu_M, \nu_M) \leq C \text{Card}(\partial_\Lambda L) \exp\left(-\frac{1}{C} d(M, L^c)\right) + \delta |L|.$$

Comme les constantes sont indépendantes de δ , on peut faire le faire tendre δ vers 0. Il ne reste plus qu'à prendre pour L la boîte $[-n, n]^d$ et à faire tendre n vers l'infini pour montrer l'égalité de μ_M et ν_M .

Propriété faible de mélange Le même raisonnement que précédemment peut être légèrement modifié pour montrer la propriété de mélange faible (valable sur les volumes finis). Ce paragraphe utilise, en plus de l'article [DS85b] — qui se limite, pour montrer les propriétés de mélange, au cadre des spins bornés — une idée que l'on peut trouver dans [Yos01] (preuve de DS4 \Rightarrow CC).

Considérons deux configurations \mathbf{y}, \mathbf{z} , ne différant qu'en un site i_0 . On veut établir la majoration suivante :

$$d_W(\mu_{L,M,y}, \mu_{L,M,z}) \leq C |y_{i_0} - z_{i_0}| \sum_{k \in \partial L} \exp\left(-\frac{d(k, M)}{C}\right) \quad (2.25)$$

On reprend l'équation (2.24). Il faut majorer les quantités f_i apparaissant dans le terme de droite. Par définition,

$$f_i = \int |x'_i - x_i| d\mu(\mathbf{x}, \mathbf{x}'),$$

où μ est un couplage entre $\mu_{L,\mathbf{y}}$ et $\mu_{L,\mathbf{z}}$. Les preuves des lemmes 32 et 33 montrent que le couplage peut être choisi de façon à ce qu'il préserve l'ordre FKG. En supposant que $y_i \leq z_i$, cela veut dire que :

$$\mu(\{\mathbf{x} \leq \mathbf{x}'\}) = 1.$$

On peut alors majorer f_i :

$$\begin{aligned} f_i &= \int (x'_i - x_i) d\mu(\mathbf{x}, \mathbf{x}') \\ &= \int x'_i d\mu_{L,\mathbf{z}}(\mathbf{x}') - \int x_i d\mu_{L,\mathbf{y}}(\mathbf{x}), \end{aligned}$$

puisque μ a pour marginales $\mu_{L,\mathbf{z}}$ et $\mu_{L,\mathbf{y}}$. On peut alors interpoler entre \mathbf{y} et \mathbf{z} (cet argument sera repris et développé plus bas, cf. l'équation (3.44), page 87). On obtient :

$$f_i \leq |J| |y_{i_0} - z_{i_0}| \sum_{j \sim i_0} \int_0^1 \mu_\theta(x_i; x_j) d\theta.$$

On majore la covariance par $(\mathbf{Var}_{\mu_\theta}(x_i) \mathbf{Var}_{\mu_\theta}(x_j))^{1/2}$, qui est par hypothèse borné uniformément (indépendamment de θ , \mathbf{y} , \mathbf{z}). En reportant cette majoration dans (2.24), on retrouve l'équation (2.25) à établir.

Du mélange faible à l'unicité

Nous montrons ici qu'une propriété faible de mélange exponentiel suffit à entraîner l'unicité de la mesure de Gibbs tempérée.

Soit μ, ν deux mesures tempérées en volume infini. Soit f une fonction cylindrique à support dans la boîte finie Δ . On suppose f Lipschitzienne pour la norme L^1 sur \mathbb{R}^Λ .

Soit Λ une boîte contenant Δ . On utilise les équations DLR pour μ et ν :

$$\begin{aligned} |\mu(f) - \nu(f)| &\leq \left| \int \int f(\mathbf{z}) d\mu_{\Lambda,\mathbf{x}}(\mathbf{z}) d\mu(\mathbf{x}) - \int \int f(\mathbf{z}) d\mu_{\Lambda,\mathbf{y}}(\mathbf{z}) d\nu(\mathbf{y}) \right| \\ &\leq \left| \int \left(\int f(\mathbf{z}) d\mu_{\Lambda,\mathbf{x}}(\mathbf{z}) - \int f(\mathbf{z}) d\mu_{\Lambda,\mathbf{y}}(\mathbf{z}) \right) d\mu \otimes \nu(\mathbf{x}, \mathbf{y}) \right| \\ &\leq \int \left| \left(\int f(\mathbf{z}) d\mu_{\Lambda,\mathbf{x}}(\mathbf{z}) - \int f(\mathbf{z}) d\mu_{\Lambda,\mathbf{y}}(\mathbf{z}) \right) \right| d\mu \otimes \nu(\mathbf{x}, \mathbf{y}) \\ &\leq \int \|f\|_{\text{lip}} d_W(\mu_{\Lambda,\Delta,\mathbf{x}}, \mu_{\Lambda,\Delta,\mathbf{y}}) d\mu \otimes \nu(\mathbf{x}, \mathbf{y}) \\ &\leq C \left(\sum_{j \in \partial\Lambda} \exp\left(-\frac{d(\Delta, j)}{C}\right) \right) \int \|f\|_{\text{lip}} \int \sum_{j \in \partial\Lambda} |x_j - y_j| d\mu \otimes \nu(\mathbf{x}, \mathbf{y}), \end{aligned}$$

où la dernière ligne se déduit de l'hypothèse de mélange.

Comme les mesures μ et ν sont tempérées, il existe c_μ, c_ν tels que pour tout i ,

$$\mu(|x_i|) \leq c_\mu, \quad \nu(|x_i|) \leq c_\nu.$$

On en déduit :

$$|\mu(f) - \nu(f)| \leq C \|f\|_{\text{lip}} (c_\mu + c_\nu) \text{Card}(\partial\Lambda)^2 \exp\left(-\frac{d(\Delta, \Lambda^c)}{C}\right).$$

En choisissant $\Lambda = [-n, n]^d$, et en faisant tendre n vers l'infini, il vient :

$$\mu(f) = \nu(f).$$

Comme ceci est valable pour toute fonction cylindrique lipschitzienne f , les deux mesures μ et ν sont égales (cf. la proposition 66, en annexe). Le passage à la limite fonctionne dès que la fonction de mélange compense le carré du cardinal de $\partial\Lambda$; en particulier, un mélange polynomial assez rapide entraîne également l'unicité de la mesure de Gibbs.

2.6 Construction de la dynamique

On cherche maintenant à définir une évolution temporelle des configurations, qui se comporte bien vis-à-vis des mesures de Gibbs.

Dans un volume fini L , le processus est une simple diffusion de Langevin, définie par l'ÉDS :

$$dX_{i,t}^L = dB_{i,t} - \nabla_i U_{L,z}(X_t^L) dt. \quad (2.26)$$

On peut montrer facilement sa non-explosion grâce à une fonction de Lyapunov (cf. [Roy99], théorème 2.2.19, p. 26 — l'hypothèse (18) étant facilement vérifiée).

Pour montrer que l'on peut définir globalement la dynamique par un passage à la limite, nous aurons besoin de contrôles sur les processus en volume fini : c'est l'objet de la section suivante.

Les contrôles dans $l^2(\alpha)$ de la dynamique

L'objectif ici est d'obtenir des informations sur la dynamique de Langevin. On s'intéressera d'abord à un seul processus dans une boîte finie, puis on comparera les processus dans deux boîtes différentes. On rappellera enfin un résultat de Royer sur la comparaison des processus dans la même boîte, partant de deux configurations différentes.

On suit toujours [Roy99] (lemmes 4.2.9 et 4.2.11), avec quelques adaptations mineures pour tenir compte de la perturbation de l'auto-interaction.

Les contrôles voulus seront obtenus grâce au lemme de Gronwall, qu'on utilisera sous la forme suivante.

34 Lemme (Gronwall). *Soit f et g deux fonctions définies sur \mathbb{R}^+ , continues et positives. Soit k une constante positive. On suppose que g est croissante, et qu'on a l'inégalité :*

$$f(t) \leq g(t) + k \int_0^t f(s) ds.$$

Alors pour tout t positif,

$$f(t) \leq g(t) \exp(kt).$$

Démonstration. Le résultat est connu pour g constante. Ici, on peut fixer un T positif, et par croissance de g , on a :

$$\forall t \leq T, f(t) \leq g(T) + k \int_0^t f(s) ds.$$

On peut alors appliquer la forme classique du lemme sur l'intervalle $[0, T]$, et obtenir :

$$\forall t \leq T, f(t) \leq g(T) \exp(kt).$$

En particulier, l'inégalité est vraie pour $t = T$. Comme T est arbitraire, la preuve est finie. \square

On aura besoin d'utiliser des translatées du poids α ; on notera $\tilde{\alpha}_i$ le poids donné par :

$$\tilde{\alpha}_i(k) = \alpha(k - i).$$

La norme correspondante sur les configurations sera notée $\|\cdot\|_{\tilde{\alpha}_i}$.

35 Lemme (Contrôle du sup du processus). *Soit t un temps, L une boîte finie, et \mathbf{x}, \mathbf{z} deux configurations tempérées.*

Soit $X = (X_{i,t}^{L,\mathbf{z},\mathbf{x}})_{i \in \mathbb{Z}^d}$ le processus dans L , partant de \mathbf{x} avec conditions aux bords \mathbf{z} , et soit Y le processus défini par $Y_i(t) = \sup_{s \in [0,t]} |X_i(s)|$.

Alors il existe des constantes c, k , indépendantes de $t, L, \mathbf{x}, \mathbf{z}$, et telles que

$$\begin{aligned} \mathbf{E} [\|Y(t)\|_{\tilde{\alpha}}^2] &\leq e^{kt} \left(\|(\mathbf{x}^L, \mathbf{z})\|_{\tilde{\alpha}}^2 + c \sum_L (\alpha(i)) \right), \\ \alpha(0) \mathbf{E} [Y_i(t)^2] &\leq e^{kt} \left(\|(\mathbf{x}^L, \mathbf{z})\|_{\tilde{\alpha}_i}^2 + c |\alpha| \right). \end{aligned}$$

Remarque. *La deuxième partie du lemme est plus importante qu'il n'y paraît. En fait, le contrôle global de Y n'est bon que pour les termes proches de l'origine, et on l'utilisera uniquement pour obtenir une majoration de $Y_0(t)$.*

C'est ce qui justifie de prendre, pour la deuxième partie, le poids translaté $\tilde{\alpha}_i$ pour majorer le processus au site i : on se ramène ainsi au contrôle du spin en 0.

Démonstration. On suit la preuve de [Roy99] (lemme 4.2.9). On applique la formule d'Itô pour évaluer $X_i(t)^2$. Comme X_i vérifie l'EDS

$$dX_{i,t} = dB_{i,t} - \partial_i U_{L,\mathbf{z}}(X_{i,s}) ds,$$

on obtient :

$$X_i(t)^2 = x_i^2 + 2 \int_0^t X_i(s) dB_{i,s} - 2 \int_0^t X_i(s) \partial_i U_{L,\mathbf{z}}(X_i(s)) ds + t.$$

Laissant la partie martingale de côté pour l'instant, on peut contrôler l'intégrale classique, grâce à l'expression explicite du hamiltonien : c'est l'inégalité de la proposition 13. Il existe donc a, b deux réels (qu'on peut supposer positifs) tels que :

$$\begin{aligned} -X_i(s) \partial_i U_{L,\mathbf{z}}(X_i(s)) &\leq -a X_i(s)^2 ds + b + \frac{1}{2} \sum_j p(i-j) (X_i(s)^2 + X_j(s)^2) \\ &\leq \left(\frac{\sigma}{2} - a\right) X_i(s)^2 + b + \frac{1}{2} \sum_j p(i-j) X_j(s)^2. \end{aligned} \tag{2.27}$$

où l'on rappelle que σ vaut $\sum_j p(j)$, et que $X_j(\cdot)$ est considéré gelé à la valeur z_j pour tout j extérieur à la boîte L . Comme $\sigma < a$, $(\frac{\sigma}{2} - a) \leq 0$, et on peut oublier le terme correspondant. On intègre entre 0 et t :

$$-2 \int_0^t X_i(s) \partial_i U_{L,\mathbf{z}}(X_i(s)) ds \leq 2bt + \int_0^t \left(\sum_j p(i-j) X_j(s)^2 \right) ds.$$

On s'occupe ensuite de la partie martingale. Notons-la M , et notons N son sup :

$$M_i(t) \stackrel{\text{déf}}{=} \int_0^t X_i(s) dB_{i,s}, \quad N_i(t) = \sup_{[0,t]} M_i(s).$$

Avec cette notation, et en majorant l'intégrale classique, on a donc :

$$X_i(t)^2 \leq x_i^2 + 2N_i(t) + (2b+1)t + \int_0^t \sum_j p(i-j) X_j(s)^2 ds.$$

Comme b et les $p(i-j)$ sont tous positifs, tout le terme de droite est croissant en t . Ce même terme de droite majore donc tous les $X_i(s)^2$, pour $s \leq t$. On en déduit qu'il majore $Y_i(t)^2$, où Y est le processus $\sup_{[0,t]} X$.

On peut alors majorer le terme de droite lui-même en utilisant le processus Y . On obtient finalement :

$$Y_i(t)^2 \leq x_i^2 + 2N_i(t) + (2b+1)t + \int_0^t \sum_j p(i-j) Y_j(s)^2 ds.$$

Le terme provenant de la martingale peut se majorer par l'inégalité de Doob.

On peut ensuite sommer les différents contrôles avec le poids α , puis utiliser le lemme de Gronwall sur la quantité $\|Y_t\|_\alpha$. Pour les détails, cf. la preuve du lemme 4.2.9. de [Roy99]. \square

36 Lemme (Comparaison de processus). *Soit deux boîtes L, M , telles que $L \subset M$, et \mathbf{x} une configuration tempérée. On note X^M, X^L les processus dans les boîtes L, M , partant de \mathbf{x} , à condition aux bords nulles. Enfin, on note Y le processus défini par $Y_i(t) = \sup_{[0,t]} |X_i^M(s) - X_i^L(s)|$.*

Alors on a l'inégalité :

$$\alpha(0)\mathbf{E}[\|Y(s)\|_\alpha^2] \leq e^{k't} \left(\|\mathbf{x}\|_{l^2(\beta_L)}^2 + c|\alpha| \sum_{j \in M \setminus L} \alpha(j) \right)$$

où le poids β_L est défini par :

$$\beta_L(j) = \sum_i \mathbf{1}_{i \notin L} \alpha(i) \alpha(j-i).$$

Remarque. *L'énoncé est légèrement différent de celui du lemme 4.2.11 de [Roy99] (le poids $\alpha * \alpha$ étant remplacé par β_L) ; nous n'avons pu prouver le résultat qu'avec cette modification.*

Démonstration. Comme les processus dans les boîtes L et M sont définies à partir du même mouvement brownien, le processus $X_i^L - X_i^M$ vérifie, pour tout i dans L , l'équation différentielle suivante :

$$d(X_i^L - X_i^M)_t = (-\partial_i U_{L,0}(X^L) + \partial_i U_{M,0}(X^M)) dt.$$

On multiplie des deux côtés par $2(X_i^L - X_i^M)$, et on intègre entre 0 et t pour obtenir :

$$(X_i^L(t) - X_i^M(t))^2 = 2 \int_0^t (X_i^L - X_i^M) (-\partial_i U_{L,0}(X^L) + \partial_i U_{M,0}(X^M)) dt.$$

Pour traiter le terme de droite, on peut remplacer $\partial_i U_{L,0}(X^L)$ par $\partial_i U_{M,0}(X^L)$, puisque X^L est uniformément nulle en dehors de L (donc les termes d'interaction avec les spins de $M \setminus L$ disparaissent). On peut alors appliquer les estimées sur l'hamiltonien obtenues dans la proposition 13 (avec $\mathbf{x} = X^M, \mathbf{y} = X^L$ et $M = L$) :

$$(X_i^L(t) - X_i^M(t))^2 \leq (\sigma - 2c_V) \int_0^t (X_i^L(s) - X_i^M(s))^2 ds + \int_0^t \sum_{j \in M} p(i-j) (X_j^M(s) - X_j^L(s))^2 ds.$$

Comme pour la preuve du lemme précédent, on veut appliquer le lemme de Gronwall au processus Y des sup des différences. On majore donc $|X^L - X^M|_s$ par Y_s dans les termes de droite, puis on passe au sup sur s à gauche :

$$Y_i(t)^2 \leq (\sigma - 2c_V)_+ \int_0^t Y_i(s)^2 ds + \int_0^t \sum_j p(i-j) Y_j(s)^2 ds.$$

Pour les sites $i \in M \setminus L$, on peut appliquer le lemme précédent :

$$Y_i(t)^2 \leq \frac{1}{\alpha(0)} e^{kt} \left(c|\alpha| + \|(\mathbf{x}^M \cdot 0)\|_{\tilde{\alpha}_i}^2 \right).$$

Notons que même si ces estimées sont destinées à être sommées sur tous les sites, on utilise les majorations site par site et pas le contrôle global de Y (cf. la remarque sur le lemme précédent).

On rassemble ces majorations, en écrivant pour tout site i :

$$Y_i(t)^2 \leq \mathbf{1}_{i \notin L} e^{kt} \frac{1}{\alpha(0)} \left(c|\alpha| + \|(\mathbf{x}^L \cdot 0)\|_{\tilde{\alpha}_i}^2 \right) + (\sigma - 2c_V)_+ \int_0^t Y_i(s)^2 ds + \int_0^t \sum_j p(i-j) Y_j(s)^2 ds.$$

On ne perd rien en remplaçant $(\mathbf{x}^M \cdot 0)$ par \mathbf{x} (de toute façon, seules les valeurs de \mathbf{x} dans M importent pour définir les processus étudiés). On fait la somme pondérée sur i , en utilisant pour le dernier terme la majoration $\alpha * p \leq \sigma' \alpha$:

$$\|Y(t)\|_{\alpha}^2 \leq \frac{e^{kt}}{\alpha(0)} \left(c|\alpha| \sum_{i \in M \setminus L} \alpha(i) + \sum_{i \in M \setminus L} \alpha(i) \|\mathbf{x}\|_{\tilde{\alpha}_i}^2 \right) + (\sigma' + (\sigma - 2c_V)_+) \int_0^t \sum_j \alpha(j) Y_j(s)^2 ds$$

Comme pour toute configuration \mathbf{x} , $\sum_{i \in M \setminus L} \alpha(i) \|\mathbf{x}\|_{\tilde{\alpha}_i}^2 \leq \sum_{i,j} \mathbf{1}_{i \notin L} \alpha(i) \alpha(j-i) x_j^2 = \|\mathbf{x}\|_{l^2(\beta_L)}^2$, on a finalement :

$$\|Y(t)\|_{\alpha}^2 \leq \frac{e^{kt}}{\alpha(0)} \left(c|\alpha| \sum_{i \in M \setminus L} \alpha(i) + \|\mathbf{x}\|_{l^2(\beta_L)}^2 \right) + (\sigma' + (\sigma - 2c_V)_+) \int_0^t \|Y(s)\|_{\alpha}^2 ds.$$

Il ne reste plus qu'à appliquer le lemme de Gronwall (lemme 34) à la fonction $f(t) = \|Y(t)\|_{\alpha}^2$. \square

Donnons enfin un dernier résultat de comparaison.

37 Lemme (Comparaison en fonction du point de départ). *Pour tout volume L , toute condition aux bords \mathbf{z} , tout couple (\mathbf{x}, \mathbf{y}) de configurations, si l'on note X et Y les processus dans L partant respectivement de \mathbf{x} et \mathbf{y} , on a :*

$$\|X - Y\|_{\alpha}^2 \leq \exp(k''t) \|\mathbf{x} - \mathbf{y}\|_{\alpha}^2.$$

La preuve suit les mêmes arguments que les deux précédentes (cf. [Roy99], lemme 4.2.11).

La section suivante montre une première application de ces inégalités en boîte finies pour construire la dynamique infinie. Elles seront réutilisées pour comparer les évolutions en boîte finie et infinie.

Le passage en volume infini

Les contrôles précédents permettent de construire une dynamique infinie à partir de n'importe quel point de départ. L'idée est de passer à la limite trajectoire par trajectoire de la façon suivante.

On se place sur un intervalle de temps $[0, T]$. En chaque site, on a un espace de trajectoires $\mathcal{C}([0, T], \mathbb{R})$, que l'on munit de la norme uniforme, ce qui en fait un Banach. Les trajectoires en volume infini appartiennent à l'espace $\mathcal{X} = l^2(\mathcal{C}([0, T]), \alpha)$, qui est lui aussi un Banach. Enfin, en notant Ω l'espace de probabilité, on cherche le processus limite dans $L^2(\Omega, \mathcal{X})$. La norme d'un processus X dans cet espace est donc :

$$\|X\|_T^2 = \mathbf{E} \left[\sum_i \alpha(i) \left(\sup_{[0, T]} |X_i(s)| \right)^2 \right].$$

Remarque. Les espaces l^2 , L^2 considérés ici ne sont pas les espaces habituels (l'espace d'arrivée est un Banach, et non plus \mathbb{R} ou \mathbb{C}). Les preuves du caractère complet de l'espace ne sont cependant pas modifiées (cf. par exemple [Wlo87], théorème 24.5 pour le cas hilbertien).

Soit alors X^n la suite de processus en boîte finie $X^n = X^{L_n, \mathbf{x}, 0}$ (partant de \mathbf{x} avec conditions aux bords nulles). D'après le lemme 36, on peut comparer X^n et X^m , pour $n < m$:

$$\|X^n - X^m\|_T^2 \leq C_T \left(\|\mathbf{x}\|_{l^2(\beta_{L_n})}^2 + c|\alpha| \sum_{L_m \setminus L_n} \alpha(i) \right).$$

Si on montre que $\|\mathbf{x}\|_{l^2(\beta_{L_n})}$ tend vers 0 quand m est grand, on peut assurer que la suite X^n est de Cauchy, et donc qu'elle converge.

38 Lemme. Pour toute configuration tempérée \mathbf{x} , $\|\mathbf{x}\|_{l^2(\beta_{L_n})} \rightarrow 0$.

Démonstration. Rappelons la définition du poids $\beta_L : \beta_L(i) = \sum_{j \notin L} \alpha(j) \alpha(i-j)$. Comme \mathbf{x} est tempérée, il existe $c_{\mathbf{x}}, d_{\mathbf{x}}$ tels que $|x_i|^2 \leq c_{\mathbf{x}}(1+|i|)^{d_{\mathbf{x}}}$. On peut donc écrire :

$$\begin{aligned} \sum_i \beta_L(i) x_i^2 &\leq c_{\mathbf{x}} \sum_i \sum_j \mathbf{1}_{j \notin L} \alpha(j) \alpha(i-j) (1+|i|)^{d_{\mathbf{x}}} \\ &\leq c_{\mathbf{x}} \sum_i \sum_j \mathbf{1}_{j \notin L} \alpha(j) \alpha(i-j) (1+|i-j|+|j|)^{d_{\mathbf{x}}} \\ &\leq c'_{\mathbf{x}} \sum_i \sum_j \mathbf{1}_{j \notin L} \alpha(j) \alpha(i-j) (1+|i-j|)^{d_{\mathbf{x}}} \\ &\quad + c'_{\mathbf{x}} \sum_i \sum_j \mathbf{1}_{j \notin L} \alpha(j) \alpha(i-j) |j|^{d_{\mathbf{x}}}. \end{aligned}$$

On intervertit les doubles sommes, et on fait dans la première le changement d'indice $k = i-j$:

$$\sum_i \beta_L(i) x_i^2 \leq c'_{\mathbf{x}} \sum_j \mathbf{1}_{j \notin L} \alpha(j) \left(\sum_k \alpha(k) (1+|k|)^{d_{\mathbf{x}}} \right) + c'_{\mathbf{x}} |\alpha| \sum_j \mathbf{1}_{j \notin L} \alpha(j) |j|^{d_{\mathbf{x}}}.$$

Comme α est à décroissance exponentielle, la famille $(|k|^{d_{\mathbf{x}}})_k$ est sommable pour le poids α , et les deux sommes tendent donc vers 0 quand $L = L_n$ tend vers \mathbb{Z}^d . On explicite les majorations, car on aura besoin ultérieurement de la vitesse de convergence. On note c_{α}, d_{α} deux constantes telles que $\alpha(i) \leq c_{\alpha} \exp(-d_{\alpha} |i|)$. Il existe une constante $c''_{\mathbf{x}}$ telle que :

$$\sum_i \beta_{L_n}(i) x_i^2 \leq c''_{\mathbf{x}} \sum_{j \notin L_n} \alpha(j) (1+|j|)^{d_{\mathbf{x}}}.$$

Comme le nombre de sites tels que $|j| = n$ croît comme n^{d-1} , on en déduit :

$$\sum_i \beta_{L_n}(i) x_i^2 \leq c_{\mathbf{x}}''' \sum_{l=n+1}^{\infty} (1+l)^{d_{\mathbf{x}}+d-1} \exp(-d_{\alpha} l).$$

Pour tout $d'_{\alpha} < d_{\alpha}$, il existe donc une constante $c(\mathbf{x}, \alpha)$ telle que :

$$\sum_i \beta_{L_n}(i) x_i^2 \leq c(\mathbf{x}, \alpha) e^{-d'_{\alpha} n}, \quad (2.28)$$

et en particulier cette quantité tend vers 0 avec n . □

Toujours d'après Royer ([Roy99], remarque 4.2.14 p. 78, qui reprend les résultats de [DR79]), on peut vérifier que le processus ainsi construit vérifie le système infini d'ÉDS voulu, et que tout mesure de Gibbs est réversible pour ce processus. Par passage à la limite dans la comparaison entre deux boîtes (lemme 36), on montre :

39 Proposition. *Soit X le processus infini-dimensionnel partant de \mathbf{x} , et X^L le processus dans la boîte L , avec conditions aux bords nulles. Alors pour tout t ,*

$$\begin{aligned} \|X - X^L\|_{T=t}^2 &= \mathbf{E} \left[\sum_i \alpha(i) \sup_{[0,t]} (X_{i,s} - X_{i,s}^L)^2 \right] \\ &\leq e^{k't} \left(\|\mathbf{x}\|_{l^2(\beta_L)} + c |\alpha| \sum_{i \notin L} \alpha(i) \right). \end{aligned}$$

Chapitre 3

Inégalités fonctionnelles et mesures de Gibbs

3.1 Unicité de la mesure de Gibbs grâce à Sobolev logarithmique

Maintenant que nous avons construit l'évolution en volume infini, nous allons pouvoir l'utiliser pour montrer que, sous certaines hypothèses sur les constantes dans les inégalités de Sobolev logarithmiques, la mesure de Gibbs tempérée en volume infini est unique.

Comme annoncé dans l'introduction de la thèse, ceci est une généralisation d'un théorème de G. Royer, énoncé pour des constantes de Sobolev logarithmiques uniformément bornées.

40 Théorème. *On fait les hypothèses suivantes.*

- Les mesures de Gibbs sur les cubes de diamètre $2n$, avec conditions aux bords nulles, vérifient une inégalité de Sobolev logarithmique.
- Les constantes correspondantes, que l'on notera c_n^{LS} , n'augmentent pas trop vite avec n :

$$\exists C > 0, \quad c_n^{LS} \leq C \frac{n}{(\log(n))}$$

- La constante C est suffisamment petite (inférieure à une quantité qui ne dépend que du modèle).

Alors, la mesure de Gibbs tempérée est unique.

Remarque. La valeur critique pour C est explicite en fonction du modèle : on verra dans la preuve que l'unicité est garantie pour

$$C < \frac{2d_\alpha}{k'(d + d^{tmp} d_V)},$$

où d_α est la vitesse de décroissance exponentielle du poids α , k' est la constante qui apparaît dans le contrôle dynamique de la proposition 39, d^{tmp} est le degré maximum des configurations chargées par les mesures tempérées (tout réel $> d$ convient), et d_V est le degré d'un polynôme qui domine l'auto-interaction V .

Avant la preuve, on expose un calcul d'entropie dont on aura besoin pour contrôler la dynamique finie.

Entropie relative en boîte finie

On cherche ici à mesurer l'entropie relative du processus en boîte finie par rapport à sa mesure d'équilibre, en fonction de la condition initiale. On fixe donc une configuration tempérée \mathbf{x} , et on regarde la quantité $I(\mathcal{L}(X_t^{\Lambda_n, \mathbf{z}, \mathbf{x}}) | \mu_{\Lambda_n, \mathbf{z}})$, l'entropie relative de la loi du processus (dans la boîte Λ_n , partant de \mathbf{x} et avec une condition aux bords \mathbf{z}) par rapport à sa mesure d'équilibre (la mesure de Gibbs dans la boîte Λ_n avec conditions aux bords \mathbf{z}). Pour simplifier, on prendra $\mathbf{z} = 0$. Les calculs de [Roy99] montrent cependant que les mêmes arguments fonctionnent pour une condition aux bords tempérées.

41 Théorème. *Soit \mathbf{x} une configuration tempérée, et t un temps, tous deux fixés. Il existe une constante $K_{\mathbf{x}, t}$ telle que l'on ait la majoration :*

$$I(\mathcal{L}(X_t^{\Lambda_n, 0, \mathbf{x}}) | \mu_{\Lambda_n, 0}) \leq K_{\mathbf{x}, t} (1+n)^{d+d_{\mathbf{x}}d\nu}$$

On suit l'approche de Royer en montrant ce résultat en deux temps. On donnera d'abord une majoration de l'entropie relative valable pour n'importe quel processus de Langevin. Ensuite, on appliquera cette majoration à notre cas particulier pour montrer le théorème.

Majoration de l'entropie relative pour un processus de Langevin

La première étape est d'explicitier la densité de la loi du processus par rapport à sa mesure d'équilibre. La proposition suivante est valable pour un processus de Langevin quelconque.

42 Proposition (Densités par rapport au brownien, à la mesure d'équilibre). *Soit X un processus de Langevin, c'est-à-dire que X vérifie l'EDS (dans \mathbb{R}^n) :*

$$dX_t = dW_t - \nabla U(X_t) dt,$$

où U est un potentiel suffisamment régulier, et (W_t) un brownien standard dans \mathbb{R}^n , partant d'un point x .

La densité (trajectorielle) de la loi de X par rapport au brownien, sur l'intervalle de temps $[0, t]$, est donnée par :

$$F = \exp\left(U(x) - U(W_t) - \frac{1}{2} \int_0^t (|\nabla U|^2 - \Delta U)(W_s) ds\right). \quad (3.1)$$

La densité en un temps t de X par rapport au brownien est donc :

$$f = \mathbf{E}[F | \mathcal{F}_{\{t\}}].$$

Enfin, la densité en un temps t du processus par rapport à sa mesure d'équilibre μ (de densité $Z^{-1} \exp(-2U)$ par rapport à Lebesgue) est :

$$G = Z \mathbf{E}[\exp(2U(W_t) - 2\nu(W_t)) \times F | \mathcal{F}_{\{t\}}],$$

où $\exp(-2\nu)$ est la densité du brownien au temps t par rapport à la mesure de Lebesgue (i.e. la densité gaussienne de moyenne x et de covariance $t \times Id$).

Démonstration. La première formule se montre en appliquant la formule d'Itô dans la densité donnée par le théorème de Girsanov.

La seconde en est une conséquence immédiate. Pour la dernière, on remarque que les trois mesures que sont la mesure de Lebesgue, la loi du processus à l'instant t et la loi du Brownien

à l'instant t sont toutes absolument continues les unes par rapport aux autres, et on peut donc écrire :

$$\frac{d\mathcal{L}(X_t)}{d\mu} = \frac{d\mathcal{L}(X_t)}{d\mathcal{L}(W_t)} \times \frac{d\mathcal{L}(W_t)}{d\lambda} \times \frac{d\lambda}{d\mu}.$$

Soit :

$$\begin{aligned} G(W_t) &= f(W_t) \times \exp(-2\nu(W_t)) \times Z \exp(2U(W_t)) \\ &= \mathbf{E}[F | \mathcal{F}_{\{t\}}] \times Z \exp(2U(W_t) - 2\nu(W_t)) \\ &= Z \mathbf{E}[F \exp(2U(W_t) - 2\nu(W_t)) | \mathcal{F}_{\{t\}}], \end{aligned}$$

puisque W_t est mesurable par rapport à $\mathcal{F}_{\{t\}}$. \square

Toujours pour un processus de Langevin quelconque, on peut maintenant majorer l'entropie relative de la loi du processus par rapport à sa mesure d'équilibre. Notons cette entropie $I(t)$. Si on note γ la fonction $x \mapsto x \log x$, on a donc :

$$\begin{aligned} I(t) &= \mathbf{E}_\mu[\gamma(G)] = \mathbf{E}\left[\gamma(G(W_t)) \frac{d\mu}{d\mathcal{L}(W_t)}\right] \\ &= \mathbf{E}\left[\gamma(G(W_t)) \frac{\exp(-2U(W_t))}{Z} \exp(2\nu(W_t))\right], \end{aligned} \quad (3.2)$$

où l'espérance est par rapport à la loi du mouvement brownien. Or par l'inégalité de Jensen,

$$\begin{aligned} \gamma(G(W_t)) &= \gamma(\mathbf{E}[ZF \exp(2U(W_t) - 2\nu(W_t)) | \mathcal{F}_{\{t\}}]) \\ &\leq \mathbf{E}[\gamma(ZF \exp(2U(W_t) - 2\nu(W_t))) | \mathcal{F}_{\{t\}}], \end{aligned}$$

et d'après la définition de γ ,

$$\gamma(G(W_t)) \leq \mathbf{E}[ZF \exp(2U(W_t) - 2\nu(W_t)) (\log Z + \log F + 2U(W_t) - 2\nu(W_t)) | \mathcal{F}_{\{t\}}]$$

En multipliant par la quantité (positive) $Z^{-1} \exp(2\nu - 2U)$, que l'on peut faire rentrer dans l'espérance conditionnelle, il vient :

$$\gamma(G(W_t)) Z^{-1} \exp(2\nu - 2U) \leq \mathbf{E}[F \cdot (\log Z + \log F + 2U(W_t) - 2\nu(W_t)) | \mathcal{F}_{\{t\}}].$$

Pour les besoins de la majoration suivante, on a intérêt à faire apparaître $2 \log F$ plutôt que $\log F$. On prend ensuite l'espérance. On reconnaît $I(t)$ à gauche (équation 3.2), et on utilise à droite le fait que $\mathbf{E}[F] = 1$:

$$I(t) \leq \log(Z) + \mathbf{E}[-F \log(F)] + \mathbf{E}[F(2 \log F + 2U(W_t) - 2\nu(W_t))]. \quad (3.3)$$

On développe le $2 \log(F)$:

$$I(t) \leq \log(Z) + \mathbf{E}[-F \log F] + \mathbf{E}\left[F\left(2U(x) - 2\nu(W_t) - \int_0^t (|\nabla U|^2 - \Delta U)(W_s) ds\right)\right]. \quad (3.4)$$

Le terme $2 \log(F)$ a permis de faire disparaître les $U(W_t)$. C'est ce résultat que nous allons maintenant appliquer au système de spins.

Le contrôle de l'entropie relative pour le système de spins

On cherche donc à prouver le théorème 41. Soit donc Λ une boîte finie, et \mathbf{x} une configuration. Soit X le processus de diffusion de Glauber Langevin dans la boîte Λ , partant de \mathbf{x} avec conditions aux bords nulle. Enfin, soit $W^{\Lambda, \mathbf{x}}$ le mouvement brownien dans la boîte Λ , partant de la configuration \mathbf{x} , et $I(t)$ l'entropie relative cherchée.

On peut appliquer l'inégalité (3.4) au processus. On décompose le majorant en plusieurs termes pour obtenir :

$$\begin{aligned} I_t \leq & \log(Z) + \mathbf{E}[-F \log F] + \mathbf{E}[2U(\mathbf{x})F] + \mathbf{E}[-2\nu(W_t)F] \\ & + \mathbf{E} \left[- \left(\int_0^t (|\nabla U|^2 - \Delta U)(W_s) ds \right) F \right], \end{aligned} \quad (3.5)$$

où l'on a noté U le hamiltonien $U_{L,0}$. On s'occupe séparément des différents termes.

Déjà, les encadrements grossiers de la fonction de partition montrent que $\log(Z_{\Lambda_n})$ est majoré par un polynôme en n de degré d . Le deuxième terme est majoré (indépendamment de tous les paramètres) par le maximum de la fonction $x \mapsto -x \log(x)$.

Pour traiter le troisième, $\mathbf{E}[2U(\mathbf{x})F] = 2U(\mathbf{x})$, on utilise le caractère tempéré de la configuration \mathbf{x} , et la croissance polynomiale de l'auto-interaction V à l'infini (V majoré par un polynôme de degré d_V). Soit $c_{\mathbf{x}}, d_{\mathbf{x}}$ des constantes telles que $|x_i| \leq c_{\mathbf{x}}(1 + |i|)^{d_{\mathbf{x}}}$. Alors, pour $L = \Lambda_n$,

$$\begin{aligned} U_{\Lambda_n,0}(\mathbf{x}) & \leq \sum_{i \in \Lambda_n} V(x_i) + c \sum_{i \in \Lambda_n} x_i^2 \\ & \leq \sum_{i \in \Lambda_n} c'_{\mathbf{x}}(1 + |x_i|)^{d_V} + c \sum_{i \in \Lambda_n} (1 + |i|)^{2d_{\mathbf{x}}} \\ & \leq c'_{\mathbf{x}} n^d n^{d_{\mathbf{x}} d_V} + c'' n^{d+2d_{\mathbf{x}}} \\ & \leq c''_{\mathbf{x}} n^{d+d_{\mathbf{x}} d_V} \end{aligned} \quad (3.6)$$

où la première inégalité est une majoration grossière des termes d'interaction, et la dernière vient du fait que d_V est supérieur (ou égal) à 2.

Pour le quatrième terme, $\mathbf{E}[-2\nu(W_t)F]$, il suffit de majorer $-2\nu(W_t)$, puis d'utiliser le fait que $\mathbf{E}(F) = 1$. Ce terme est connu explicitement (c'est le logarithme de la densité gaussienne de W_t), et on peut le majorer par $-\frac{1}{2} \text{Card } \Lambda_n \log(2\pi t)$, et donc par un polynôme en n de degré d .

Le cinquième et dernier terme de (3.5) est le plus délicat à majorer. Commençons par expliciter la quantité $(\Delta U - |\nabla U|^2)(\mathbf{x})$, pour une configuration \mathbf{x} quelconque.

$$\begin{aligned} (\nabla U)_i & = \partial_i U_{L,0} = V'(x_i) + \sum_j J(i-j)(\mathbf{x}^{\cdot L} 0)_j, \\ |\nabla U|^2 & = \sum_{i \in L} \left(V'(x_i) + \sum_j J(i-j)(\mathbf{x}^{\cdot L} 0)_j \right)^2, \\ \Delta U & = \sum_{i \in L} V''(x_i). \end{aligned}$$

On veut utiliser les termes en $V'(x_i)$ venant du gradient pour compenser les $V''(x_i)$ du Laplacien. Pour cela, on utilise l'hypothèse de croissance comparée de V'' et V' (cf. l'hypothèse 1). Pour tout $(x, y) \in \mathbb{R}^2$ et tout $a \in]0, 1[$,

$$(x + y)^2 \geq ax^2 - \frac{a}{1-a} y^2.$$

On applique ceci à chacun des carrés dans $|\nabla U|^2$, avec $a = a''_V$ (qui est tel que $a''_V(V')^2 - V''$ est borné inférieurement). On obtient, en notant $b''_V = \frac{a''_V}{1-a''_V}$:

$$|\nabla U|^2 \geq \sum_{i \in L} \left(a''_V V'(x_i)^2 - b''_V \left(\sum_j J(i-j)(\mathbf{x}^{\cdot 0})_j \right)^2 \right). \quad (3.7)$$

Pour un i quelconque, on majore le terme d'interaction (la deuxième ligne utilise l'inégalité de Cauchy Schwarz sur les familles $\sqrt{p(i-j)}$ et $\sqrt{p(i-j)}|x_j|$) :

$$\begin{aligned} \left(\sum_j J(i-j)(\mathbf{x}^{\cdot 0})_j \right)^2 &\leq \left(\sum_j p(i-j) |x_j| \right)^2 \\ &\leq \left(\sum_j p(i-j) \right) \left(\sum_j p(i-j) |x_j|^2 \right) \\ &\leq \sigma \left(\sum_j p(i-j) |x_j|^2 \right). \end{aligned}$$

En revenant à $\Delta U - |\nabla U|^2$, on a donc montré :

$$\begin{aligned} (\Delta U - |\nabla U|^2)(\mathbf{x}) &\leq \sum_{i \in L} (V''(x_i) - a''_V V'(x_i)^2) + \sum_i b''_V \sigma \sum_j p(i-j) |x_j|^2 \\ &\leq \text{Card}(L) \times \sup_{\mathbb{R}} (V'' - a''_V V'^2) + \sigma^2 b''_V \sum_j |x_j|^2. \end{aligned}$$

On peut alors finir la majoration du dernier terme de (3.5) :

$$\begin{aligned} \mathbf{E} \left[- \left(\int_0^t (|\nabla U|^2 - \Delta U)(W_s) ds \right) F \right] &\leq \text{Card}(L) t \sup_{\mathbb{R}} (V'' - a''_V V'^2) + \sigma^2 b''_V \sum_j \mathbf{E} \left[\int_0^t W_{i,s}^2 ds \times F(W) \right] \\ &\leq K \text{Card}(L) t + \sigma^2 b''_V \sum_i \mathbf{E} \left[\int_0^t X_{i,s}^2 ds \right]. \end{aligned}$$

La dernière espérance est contrôlée grâce au lemme 35 :

$$\begin{aligned} \mathbf{E} \left[\int_0^t X_i(s)^2 ds \right] &\leq \int_0^t \mathbf{E} \left[\sup_{u \leq s} X_i(u)^2 \right] ds \\ &\leq t \mathbf{E} \left[\sup_{[0,t]} X_i(s)^2 \right] \\ &\leq t e^{kt} \left(\|(\mathbf{x}^{\cdot 0})\|_{\tilde{\alpha}_i}^2 + c|\alpha| \right). \end{aligned}$$

La seule chose à vérifier est donc que la quantité $\sum_i \|(\mathbf{x}^{\cdot 0})\|_{\tilde{\alpha}_i}^2$ croît polynomialement en n quand la boîte $L = L_n$ grandit. On la majore de la façon suivante :

$$\begin{aligned} \sum_i \|(\mathbf{x}^{\cdot 0})\|_{\tilde{\alpha}_i}^2 &= \sum_{i \in \mathbb{Z}^d} \sum_{j \in L_n} \alpha(j-i) x_j^2 \\ &\leq |\alpha| \sum_{j \in L_n} x_j^2 \\ &\leq |\alpha| c_{\mathbf{x}} \sum_{j \in L_n} (1+|j|)^{d_{\mathbf{x}}} \\ &\leq |\alpha| c_{\mathbf{x}} \text{Card}(L_n) (1+n)^{d_{\mathbf{x}}}. \end{aligned}$$

On a donc bien majoré tous les termes de l'entropie par un polynôme en n , de degré au plus $d + d_{\mathbf{x}} d_V$: la preuve du théorème 41 est achevée.

La preuve de l'unicité

On peut maintenant prouver le théorème annoncé au début de la section. On veut faire une preuve dynamique, et montrer que pour toute fonction f , la quantité $P_t f$ converge en un certain sens ; si c'est le cas, la limite ne pourra être que μf , pour toute mesure de Gibbs μ , et on aura bien l'unicité.

Comme on maîtrise ce qui se passe en volume fini et avec conditions aux bords nulles, il est naturel de faire la décomposition suivante, valable pour tout n :

$$P_t f = (P_t f - P_t^n f) + (P_t^n f - \mu_n f) + \mu_n f. \quad (3.8)$$

Regardons grossièrement ce qui se passe pour t et n grands. Le premier terme est petit quand n est grand par rapport à t (c'est la propriété de « finitude de la vitesse de propagation », le second quand t est grand par rapport à n (c'est la convergence en boîte finie, donnée par l'inégalité de Sobolev logarithmique). Enfin, le dernier terme converge, au moins si l'on extrait une sous-suite (puisque l'existence d'une mesure de Gibbs est obtenue par compacité d'un ensemble contenant les μ_n , cf. le théorème 20). Pour que $P_t f$ converge, il faut donc faire dépendre n de t d'une façon appropriée. La bonne échelle est en fait de prendre n de l'ordre de t : $n(t) \in [Ct, (C+1)t]$, pour une constante C que nous allons fixer bientôt. Pour cela, on précise les choses en s'intéressant successivement aux trois termes de (3.8).

Premier terme : $P_t f - P_t^n f$. Soit \mathbf{x} une configuration tempérée (le point de départ de la diffusion), et soit f une fonction, que l'on suppose bornée, et lipschitzienne pour la norme $\|\cdot\|_\alpha$. On note $\text{osc } f$ son oscillation, et $\|f\|_{\text{lip},\alpha}$ sa norme de Lipschitz. Soit $L = L_n$ la boîte finie $[-n, n]^d$. En notant $X_t^{L,0,\mathbf{x}}$ (resp. $X_t^\mathbf{x}$) le processus partant de \mathbf{x} , en volume fini avec condition aux bords nulles (resp. en volume infini), on a :

$$\begin{aligned} |P_t f - P_t^n f| &\leq \left| \mathbf{E} \left[f(X_t^{L,0,\mathbf{x}}) - f(X_t^\mathbf{x}) \right] \right| \\ &\leq \mathbf{E} \left[\left| f(X_t^{L,0,\mathbf{x}}) - f(X_t^\mathbf{x}) \right| \right] \\ &\leq \mathbf{E} \left[\|f\|_{\text{lip},\alpha} \|X_t^{L,0,\mathbf{x}} - X_t^\mathbf{x}\|_\alpha \right] \\ &\leq \|f\|_{\text{lip},\alpha} \left(\mathbf{E} \left[\|X_t^{L,0,\mathbf{x}} - X_t^\mathbf{x}\|_\alpha^2 \right] \right)^{\frac{1}{2}}. \end{aligned} \quad (3.9)$$

On se ramène donc à calculer la distance en norme α entre les processus en boîte finie et infinie.

Ce calcul a déjà été fait lors de la construction du processus infini ; on a obtenu (proposition 39) :

$$\mathbf{E} \left[\|X_t^{L,0,\mathbf{x}} - X_t^\mathbf{x}\|_\alpha^2 \right] \leq e^{k't} \left(\|\mathbf{x}\|_{l^2(\beta_L)} + c|\alpha| \sum_{i \in L} \alpha(i) \right).$$

On cherche un majorant plus simple pour cette quantité. Rappelons que α décroît exponentiellement, et qu'il existe donc deux constantes c_α, d_α telles que $\alpha(i) \leq c_\alpha \exp(-d_\alpha |i|)$. La décroissance de $\|\mathbf{x}\|_{l^2(\beta_L)}$ a déjà été étudiée auparavant (équation (2.28), page 59) ; on sait que pour tout $d'_\alpha < d_\alpha$, il existe une constante $C(\alpha, \mathbf{x})$ telle que :

$$\|\mathbf{x}\|_{l^2(\beta_{L_n})} \leq C(\alpha, \mathbf{x}) e^{-d'_\alpha n}.$$

Notons ici que la vitesse exponentielle (le d'_α) est ici indépendant de \mathbf{x} ; en particulier, le même choix vaut pour tous les $\mathbf{x} \in \mathcal{S}'$.

On voit facilement que le terme $c|\alpha|\sum_{i \notin L_n} \alpha(i)$ se majore de la même façon. Quitte à changer $C(\alpha, \mathbf{x})$, on a donc :

$$\mathbf{E} \left[\|X_t^{L,0,\mathbf{x}} - X_t^{\mathbf{x}}\|_{\alpha}^2 \right] \leq C(\alpha, \mathbf{x}) e^{k't} e^{-d'_{\alpha} n}.$$

On fixe alors une constante ρ , qui donnera le rapport entre n et t , telle que $\rho > k'/d'_{\alpha}$. On pose $n(t) = \lfloor \rho t \rfloor + 1$. Par construction, $n(t)$ vérifie :

$$n(t) \in [\rho t, \rho t + 1] \quad (3.10)$$

Ce choix garantit que :

$$k't - d'_{\alpha} n(t) \leq (k' - \rho d'_{\alpha})t,$$

où $k' - \rho d'_{\alpha}$ est une constante strictement négative. La quantité $\mathbf{E} \left[\|X_t^{L,0,\mathbf{x}} - X_t^{\mathbf{x}}\|_{\alpha}^2 \right]$ converge donc vers 0, quand t tend vers ∞ . Il ne reste plus qu'à reporter ce résultat dans la majoration 3.9 pour constater que :

$$\forall \mathbf{x} \in \mathcal{S}', \quad \left| P_t f(\mathbf{x}) - P_t^{L_{n(t)}} f(\mathbf{x}) \right| \xrightarrow[t \rightarrow \infty]{} 0.$$

Deuxième terme : $P_t^n f - \mu_n f$. La majoration de ce terme utilise la convergence en boîte finie assurée par l'inégalité de Sobolev logarithmique.

$$\begin{aligned} \left| P_t^{n(t)} f - \mu_n f \right|^2 &\leq \text{osc}^2(f) \|\mathcal{L}(X_t^{L_{n(t)},0,\mathbf{x}}) - \mu_{n(t)}\|_{vt}^2 \\ &\leq 2 \text{osc}^2(f) I(\mathcal{L}(X_t^{L_{n(t)},0,\mathbf{x}}) | \mu_{n(t)}) \quad (\text{inégalité de Pinsker}) \\ &\leq 2 \text{osc}^2(f) \exp\left(-2 \frac{t-1}{c_{n(t)}^{LS}}\right) I(\mathcal{L}(X_1^{L_{n(t)},0,\mathbf{x}}) | \mu_{n(t)}) \quad (\text{décroissance de l'entropie}) \\ &\leq 2c_{\mathbf{x}} \text{osc}^2(f) \exp\left(-2 \frac{t-1}{c_{n(t)}^{LS}}\right) (1+n(t))^{d_{\mathbf{x}} d_v}. \quad (\text{théorème 41}). \end{aligned} \quad (3.11)$$

Remarque. Si l'on suppose une inégalité de Sobolev logarithmique uniforme en la boîte et les conditions aux bords, on peut ici conclure directement : comme n est de l'ordre de t , la puissance de n est une puissance de t , et le facteur exponentiel assure la convergence vers 0.

On a vu précédemment (théorème 16) que les mesures tempérées ne chargent pas toutes les configurations tempérées ; en particulier, il existe une constante d^{tmp} telle que toute mesure de Gibbs tempérée ne charge que les \mathbf{x} tels que $d_{\mathbf{x}} \leq d^{tmp}$. On peut donc remplacer $d_{\mathbf{x}}$ par d^{tmp} dans la majoration précédente. Comme par hypothèse, c_n^{LS} est majoré par $C \frac{n}{\log(n)}$, et comme $n(t) \leq \rho t + 1$, on a :

$$\begin{aligned} \exp\left(-2 \frac{t-1}{c_{n(t)}^{LS}}\right) &\leq \exp\left(-2 \frac{(t-1) \log(n(t))}{C n(t)}\right) \\ &\leq \exp\left(-2 \frac{(t-1)}{C(\rho t + 1)} \log(n(t))\right) \end{aligned}$$

Pour toute constante $C' > C\rho$, on a pour t assez grand l'inégalité $(t-1)/(C(\rho t + 1)) > 1/C'$, auquel cas :

$$\begin{aligned} \exp\left(-2 \frac{t-1}{c_{n(t)}^{LS}}\right) &\leq \exp\left(-\frac{2}{C'} \log(n(t))\right) \\ &\leq n(t)^{-2/C'} \end{aligned}$$

En reprenant (3.11), on en déduit :

$$\left| P_t^{n(t)} f - \mu_n f \right|^2 \leq c'_x \text{osc}^2(f) n(t)^{-2/C'} (1+n(t))^{d+d^{tmp} d_V}$$

La quantité de droite converge dès que $d + d^{tmp} d_V < 2/C'$, i.e. :

$$C' < \frac{2}{d + d^{tmp} d_V}.$$

Ceci est possible dès que $C < \frac{2}{\rho(d+d^{tmp} d_V)}$.

À cette condition, on a donc garanti la convergence suivante :

$$\forall \mathbf{x} \in \mathcal{S}', \quad \left| P_t^{n(t)} f(\mathbf{x}) - \mu_n f \right| \xrightarrow[t \rightarrow \infty]{} 0.$$

Le dernier terme : $\mu_{n(t)} f$. Pour étudier ce terme, rappelons qu'on a choisi $n(t) = \lfloor \rho t \rfloor + 1$. En particulier, quand t parcourt $[0, \infty[$, $n(t)$ parcourt \mathbb{N} . On a vu dans la partie consacrée au problème d'existence que la suite (μ_n) admettait une sous-suite convergente. Soit donc n_k une suite d'entiers telle que (μ_{n_k}) converge, et soit μ la mesure limite. On pose $t_k = \frac{n_k - 1}{\rho}$; la suite (t_k) est croissante, elle tend vers l'infini, et $n(t_k) = n_k$, pour tout k .

L'unicité. Reprenons maintenant la décomposition (3.8), et écrivons-la aux temps t_k , avec $n = n(t_k) = n_k$. D'après les paragraphes précédents, les deux premiers termes tendent vers 0, et le troisième converge, on peut donc affirmer :

$$\forall \mathbf{x}, \quad P_{t_k} f(\mathbf{x}) = (P_{t_k} f(\mathbf{x}) - P_{t_k}^{n_k} f(\mathbf{x})) + (P_{t_k}^{n_k} f(\mathbf{x}) - \mu_{n_k} f) + \mu_{n_k} f$$

$$\xrightarrow[k \rightarrow \infty]{} \mu f.$$

Soit alors ν une mesure de Gibbs tempérée, différente de μ . Appliquons le théorème de convergence dominée, sous la mesure ν , à la suite de fonctions $g_k : \mathbf{x} \mapsto P_{t_k} f(\mathbf{x})$.

$$\forall \mathbf{x} \in \mathcal{S}', \quad \left. \begin{array}{l} P_{t_k} f(\mathbf{x}) \xrightarrow[k \rightarrow \infty]{} \mu(f) \\ |P_{t_k}(f)| \leq \|f\|_\infty \end{array} \right\} \implies \nu(P_{t_k} f) \xrightarrow[k \rightarrow \infty]{} \nu(\mu(f)) = \mu(f).$$

Il ne reste plus qu'à remarquer que la suite $(\nu(P_{t_k} f))$ est en fait la suite constante $(\nu(f))$: en effet, toute mesure de Gibbs tempérée est invariante par la dynamique infinie P_t .

Finalement $\mu f = \nu f$ pour toute fonction bornée et $l^2(\alpha)$ -lipschitzienne. D'après le résultat de l'annexe A, ceci suffit pour affirmer l'égalité $\mu = \nu$. La mesure de Gibbs tempérée est donc bien unique.

3.2 Vers l'inégalité de Poincaré pour l'unicité de la mesure de Gibbs

La section précédente a montré qu'une inégalité de Sobolev logarithmique valable uniformément sur une famille de boîtes avec une condition aux bords nulle garantissait l'unicité de la mesure de Gibbs. Nous verrons plus loin que, dans le cadre de notre modèle, des phénomènes intéressants apparaissent pour les inégalités fonctionnelles, et qu'en particulier une inégalité de Poincaré uniforme (en un certain sens) implique une inégalité (plus forte!) de Sobolev logarithmique (cf. le théorème 60 ci dessous). On peut alors penser que dans le cadre de ce modèle d'Ising, les inégalités de Poincaré et de Sobolev logarithmique sont d'une certaine manière « équivalentes ». Il est alors naturel d'essayer d'affaiblir les hypothèses de la section précédente :

Question 1. *Si les mesures sur les boîtes L_n avec condition aux bords nulles vérifient, uniformément en n , une inégalité de Poincaré, y a-t-il nécessairement unicité de la mesure de Gibbs ?*

Nous ne répondons malheureusement pas à cette question. L'étude d'un modèle simplifié (introduit dans la section 7 de [BM02]) censé reproduire la dynamique du modèle d'Ising standard dans le régime de transition de phase, laisse penser qu'en toute généralité, ce résultat n'est pas vrai. En effet, il se pourrait que pour certaines valeurs des paramètres, et en dimension supérieure à 3, la constante de Poincaré reste bornée (pour $n \rightarrow \infty$) alors que celle de Sobolev logarithmique explose.

Quoiqu'il en soit, on peut obtenir des résultats en supposant des inégalités plus fortes que Poincaré, mais plus faibles que Sobolev logarithmique.

On considère donc des inégalités de Beckner généralisées (cf. [CGG05, LO00]).

43 Définition. *Soit a un réel de $]0, 1[$. On dit qu'une mesure μ vérifie une inégalité de Beckner généralisée $\mathbf{GBI}(a, C_a)$ si la condition suivante est vérifiée :*

$$\sup_{p \in]1, 2[} \frac{\int f^2 d\mu - (\int f^p d\mu)^{2/p}}{(2-p)^a} \leq C_a \int |\nabla f|^2 d\mu. \quad (3.12)$$

On montre facilement que pour $a = 0$ on retrouve l'inégalité de Poincaré, et pour $a = 1$ l'inégalité de Sobolev logarithmique (cf. le corollaire 1 de [LO00]).

Nous montrerons le résultat suivant.

44 Théorème. *Il existe un réel a_{\min} (défini par l'équation (3.31), ci-dessous) tel que, si les mesures μ_{L_n} (avec condition aux bords nulle) vérifient une inégalité de Beckner uniforme :*

$$\exists a_0, \forall n, \quad \mu_{L_n} \text{ vérifie } \mathbf{GBI}(a_0, C_{a_0}),$$

et si a_0 est suffisamment grand :

$$a_0 > a_{\min}$$

alors la mesure de Gibbs en volume infini est unique.

Remarque. *Nous verrons que l'expression de a_{\min} garantit qu'il appartient à l'intervalle $]0, 1[$. L'inégalité $a_{\min} < 1$ montre donc que l'on a bien obtenu un résultat plus fort que précédemment : une inégalité uniforme strictement plus faible que Sobolev logarithmique suffit à montrer l'unicité de la mesure de Gibbs.*

Remarque. *Pour simplifier l'argument, on ne considère pas ici le cas de constantes qui augmentent avec la taille de la boîte.*

L'idée centrale de la preuve est d'utiliser le lien entre inégalités de Beckner et décroissance de l'entropie, établi par P. Cattiaux, I. Gentil et A. Guillin. Citons un de leurs résultats.

45 Théorème ([CGG05], Th. 5.5). *Soit μ une mesure de probabilité sur \mathbb{R}^n , à densité par rapport à la mesure de Lebesgue, et vérifiant une inégalité $\mathbf{GBI}(a, C_a)$.*

Soit ν (la loi initiale) une mesure vérifiant :

$$\exists t_0, C, \forall t \geq t_0, \forall p \geq 1, \quad \left(\int P_t \nu \log_+^p(P_t \nu) \right)^{1/p} \leq C p. \quad (3.13)$$

Alors l'entropie partant de ν décroît sous-exponentiellement le long du semi-groupe :

$$\forall a' < a, \exists s, t_0 \forall t \geq t_0, \quad \mathbf{Ent}_\mu(P_{s+t} \nu) \leq \exp\left(1 - t^{1/(2-a')}\right). \quad (3.14)$$

Signalons que le paramètre α apparaissant dans [CGG05] est relié à notre a par l'équation $a = (2\alpha - 2)/\alpha$ (cf. l'exemple 4.3 de la même référence), ce qui justifie notre écriture de 3.14.

L'intérêt de ce genre de résultat pour notre étude est immédiat : souvenons nous que le point central de la preuve du théorème 40 d'unicité de la mesure de Gibbs était la compensation de la croissance (polynomiale) de l'entropie au temps 1 par la décroissance (exponentielle) de cette entropie entre le temps 1 et le temps t (cf. l'équation 3.11 et la remarque qui la suit).

Une décroissance plus rapide que n'importe quel polynôme devrait donc suffire à faire converger l'entropie au temps t dans la boîte L_n , quand n est de l'ordre de t .

Malheureusement, le théorème 45 n'est vrai *que pour t suffisamment grand*, et nous voulons l'utiliser pour un t relativement petit (de l'ordre de n). Nous avons donc besoin d'explicitier le temps à partir duquel (3.14) est valable.

Nous procédons en trois temps. Nous rassemblons d'abord des contrôles de moments sur le processus de Langevin, et établissons une hypothèse du type de (3.13) — c'est l'objet du lemme 49. Ensuite, nous suivons l'article [CGG05], en montrant d'abord une décroissance de l'entropie en temps petit, puis en itérant cette estimation pour l'appliquer au temps voulu.

Contrôles de moments et entropies modifiées

Avant d'exposer les contrôles de moments, citons un résultat de décroissance pour certaines quantités le long du semi-groupe, qui sera utilisé plusieurs fois dans la suite.

46 Lemme. *Soit μ une mesure de probabilité à densité $\exp(-2V)$ sur \mathbb{R}^n , et soit P_t le semi-groupe associé. Soit ν une loi initiale, et soit f_t la densité du processus de Langevin partant de ν .*

Alors pour toute fonction convexe ϕ , la fonction

$$t \mapsto \int \phi(f_t) d\mu$$

est décroissante.

Démonstration. Il suffit de montrer que pour toute fonction régulière f ,

$$t \mapsto \int \phi(P_t f) d\mu$$

décroît : en effet, par réversibilité, f_{t_0+t} s'écrit $P_t f_{t_0}$ pour tous t, t_0 , et on obtient bien la décroissance de $\int \phi(P_t f) d\mu$ après tout temps $t_0 > 0$.

Soit donc t, s deux réels, on va montrer la décroissance entre les temps s et $t+s$ (l'argument suivant nous a été communiqué par L. Miclo). On utilise la propriété de semi-groupe, puis l'inégalité de Jensen :

$$\begin{aligned} \phi(P_{t+s} f(x)) &= \phi(P_t P_s f(x)) \\ &= \phi(\mathbf{E}_x [P_s f(X_t)]) \\ &\leq \mathbf{E}_x [\phi(P_s f(X_t))]. \end{aligned}$$

On intègre par rapport à μ , et on utilise l'invariance de μ par rapport à P_t :

$$\begin{aligned} \int \phi(P_{t+s} f(x)) d\mu(x) &\leq \int \mathbf{E}_x [\phi(P_s f(X_t))] d\mu(x) \\ &= \int \phi(P_s f(x)) d\mu(x). \end{aligned}$$

La fonction décroît donc bien entre s et $t+s$. □

47 Lemme (Moments d'ordre supérieur à 2). *Pour toute boîte L , tout site i dans L , toute condition aux bords tempérée \mathbf{z} , et toute configuration de départ \mathbf{x} , on a le contrôle de moments suivant :*

$$\forall p \geq 1, \quad \mathbf{E} \left[|X_i^L(t)|^{2p} \right] \leq \left(\|\mathbf{x}^L \cdot \mathbf{z}\|_{\alpha_i}^{2p} + cp^2 \right) \exp(kp^2 t), \quad (3.15)$$

pour deux constantes c, k ne dépendant que du poids α .

De plus, si l'on note $Y_t = \|X_t^L\|_{\alpha}^2$, on a

$$\mathbf{E} [Y_t^p] \leq \left(\|\mathbf{x}^L \cdot \mathbf{z}\|_{\alpha}^{2p} + cp^2 \right) \exp(kp^2 t) \quad (3.16)$$

Démonstration. La démarche est la même que dans la section 2.6 — on utilise la formule d'Itô, puis l'on cherche à appliquer le lemme de Gronwall. Le principal changement est que l'on ne cherche pas à évaluer de borne supérieure, ce qui permet de simplifier un peu la preuve, en s'affranchissant du contrôle de martingale. Cette version simplifiée s'inspire de [FW86].

On montre d'abord le résultat (3.16) (qui est en fait vérifié sur le processus en dimension infinie), puis on en déduira le contrôle (3.15) sur X .

Comme $dY_t = \sum 2\alpha(i)X^i(t)dX_t^i + dt$ et que par définition de X , $dX_t^i = dB_t^i - \partial_i U^L(X_t)dt$ pour les spins dans L (et 0 sinon), on obtient successivement :

$$\begin{aligned} (X_t^i)^2 &= (X_0^i)^2 + 2 \int_0^t X_s^i dX_s^i + t, \\ Y_t &= Y_0 + \sum \alpha(i) \left(2 \int_0^t X_s^i dX_s^i + t \right), \end{aligned}$$

et la variation quadratique :

$$\langle Y \rangle_t = 4 \sum \alpha(i)^2 \int_0^t (X_s^i)^2 ds.$$

On cherche des moments d'ordre p , on va donc étudier Y_t^p . On applique la formule d'Itô :

$$\begin{aligned} Y_t^p &= Y_0^p + p \int_0^t Y_s^{p-1} dY_s + \frac{1}{2} p(p-1) \int_0^t Y_s^{p-2} d\langle Y \rangle_s \\ &= Y_0^p + p \int_0^t \sum_{i \in L} \alpha(i) \left(Y_s^{p-1} (2X_s^i dX_s^i + ds) \right) \\ &\quad + \frac{1}{2} p(p-1) \int_0^t \sum_{i \in L} \alpha(i)^2 \times 4Y_s^{p-2} (X_s^i)^2 ds. \end{aligned} \quad (3.17)$$

Le dernier terme se traite facilement, en majorant $\alpha(i)^2$ par $\alpha(0)\alpha(i)$ (on peut en effet toujours supposer que $\alpha(0)$ est le maximum des $\alpha(i)$, quitte à translater le poids α) :

$$\begin{aligned} \int_0^t \sum_{i \in L} \alpha(i)^2 Y_s^{p-2} (X_s^i)^2 ds &\leq \alpha(0) \int_0^t Y_s^{p-2} \left(\sum_{i \in \mathbb{Z}^d} \alpha(i) (X_s^i)^2 \right) ds \\ &\leq \alpha(0) \int_0^t Y_s^{p-1} ds. \end{aligned}$$

On regroupe le terme obtenu avec l'intégrale en ds du deuxième terme de (3.17) :

$$Y_t^p \leq Y_0^p + p \sum \alpha(i) \int_0^t 2Y_s^{p-1} X_s^i dX_s^i + (p|\alpha| + 2p(p-1)\alpha(0)) \int_0^t Y_s^{p-1} ds, \quad (3.18)$$

où l'on rappelle que $|\alpha| = \sum_{i \in \mathbb{Z}^d} \alpha(i)$.

L'intégrale stochastique par rapport à X se majore maintenant comme dans le lemme 35 (équation (2.27), et ce qui la suit) :

$$\begin{aligned} 2 \int_0^t Y_s^{p-1} X_s^i dX_s^i &= 2 \int_0^t Y_s^{p-1} X_s^i dB_s^i - 2 \int_0^t Y_s^{p-1} X_s^i \partial_i U_{L,\mathbf{z}}(X_s) \\ &\leq M_t + 2b \int_0^t Y_s^{p-1} ds + 2 \int_0^t \left(\sum_{j \in \mathbb{Z}^d} p(i-j)(X_s^j)^2 \right) Y_s^{p-1} ds. \end{aligned} \quad (3.19)$$

Sommons maintenant (3.19) sur i (avec les poids $\alpha(i)$) et prenons l'espérance. La partie martingale disparaît, et la double somme sur les sites se simplifie grâce à la propriété fondamentale de α (cf. la proposition 17) :

$$\begin{aligned} 2\mathbf{E} \left[\sum \alpha(i) \int_0^t Y_s^{p-1} X_s^i dX_s^i \right] &\leq 2b|\alpha| \int_0^t \mathbf{E} \left[Y_s^{p-1} \right] + 2\mathbf{E} \left[\int_0^t \sum_{i,j} \alpha(i)p(i-j)(X_s^j)^2 Y_s^{p-1} ds \right] \\ &\leq 2b|\alpha| \int_0^t \mathbf{E} \left[Y_s^{p-1} \right] + 2\sigma' \mathbf{E} \left[\int_0^t \sum_j \alpha(j)(X_s^j)^2 Y_s^{p-1} ds \right] \\ &\leq 2\sigma' \int_0^t \mathbf{E} \left[Y_s^p \right] ds + 2b|\alpha| \int_0^t \mathbf{E} \left[Y_s^{p-1} \right] ds. \end{aligned}$$

On peut maintenant prendre l'espérance dans (3.18), et y reporter la dernière expression.

$$\begin{aligned} \mathbf{E}[Y_t^p] &\leq Y_0^p + 2p\sigma' \int_0^t \mathbf{E} \left[Y_s^p \right] ds + 2pb|\alpha| \int_0^t \mathbf{E} \left[Y_s^{p-1} \right] ds \\ &\quad + (p|\alpha| + 2p(p-1)\alpha(0)) \int_0^t \mathbf{E} \left[Y_s^{p-1} \right] ds \\ &\leq Y_0^p + 2p\sigma' \int_0^t \mathbf{E} \left[Y_s^p \right] ds + ((2b+1)p|\alpha| + 2p(p-1)\alpha(0)) \int_0^t \mathbf{E} \left[Y_s^{p-1} \right] ds. \end{aligned}$$

Comme pour tout réel positif x , $x^{p-1} \leq 1 + x^p$, on en déduit une inégalité différentielle sur la fonction $f(t) = \mathbf{E}[Y_t^p]$:

$$f(t) \leq f(0) + c_p t + d_p \int_0^t f(s) ds,$$

avec $c_p = (2b+1)p|\alpha| + 2p(p-1)\alpha(0)$, $d_p = c_p + 2p\sigma'$. Il ne reste plus qu'à appliquer le lemme de Gronwall (lemme 34) pour obtenir :

$$f(t) \leq (f(0) + c_p t) \exp(d_p t).$$

En majorant t par e^t , on a :

$$f(t) \leq (f(0) + c_p) \exp((1 + d_p) t).$$

On peut maintenant choisir c et d tels que $c_p \leq cp^2$, $1 + d_p \leq dp^2$, et la preuve de (3.16) est achevée (car $f(0) = \|(\mathbf{x}^t; \mathbf{z})\|_\alpha^{2p}$).

Le contrôle sur X_t^i s'en déduit facilement. Quitte à considérer des poids $\tilde{\alpha}$ translétés, on peut supposer $i = 0$, et il suffit alors de remarquer :

$$\begin{aligned} \alpha(0) (X_t^0)^2 &\leq \sum \alpha(i) (X_t^i)^2 = Y_t, \\ \alpha(0)^p (X_t^0)^{2p} &\leq Y_t^p, \end{aligned}$$

puis de prendre l'espérance, et d'utiliser le contrôle (3.16) de Y_t pour conclure. \square

Utilisons maintenant ce résultat pour établir une hypothèse de type (3.13).
Pour une fonction positive f , appelons p -entropie la quantité

$$\left(\int f \log_+(f)^p d\mu \right)^{1/p}. \quad (3.20)$$

L'objectif est de la contrôler cette p -entropie en temps petit de la même façon que l'entropie (cf. le théorème 41 de la section précédente, page 62), de manière à pouvoir reproduire la preuve de Cattiaux, Gentil et Guillin du théorème 45.

48 Définition. Dans toute la suite, on fixera un point de départ \mathbf{x} et des conditions aux bords nulles.

On notera :

- h_t^n la densité du processus partant de \mathbf{x} par rapport à sa mesure d'équilibre dans la boîte L_n ;
- H_t^n l'entropie de la loi du processus par rapport à sa mesure d'équilibre ($H_t^n = \mathbf{Ent}_{\mu_{L_n}}(h_t^n)$);
- $H_{p,t}^n$ la p -entropie de h_t^n par rapport à μ_{L_n} .

Le résultat principal de cette section est le contrôle suivant.

49 Lemme. Pour tout point de départ \mathbf{x} , il existe un polynôme Q tel que pour tout $p \geq 1$, pour tout $t \geq 1$,

$$H_{p,t}(n) \leq Q(n)p.$$

Démonstration. Le calcul est très similaire à celui déjà effectué pour l'entropie (théorème 41, page 41). Reprenons l'équation (3.2) (page 63), mais avec la fonction $\gamma : x \mapsto x \log_+^p(x)$, en notant $I_p(t) = (H_{p,t}(n))^p$:

$$\begin{aligned} I_p(t) &= \mathbf{E}_\mu [\gamma(G)] = \mathbf{E} \left[\gamma(G(W_t)) \frac{d\mu}{d\mathcal{L}(W_t)} \right] \\ &= \mathbf{E} \left[\gamma(G(W_t)) \frac{\exp(-2U(W_t))}{Z} \exp(2\nu(W_t)) \right], \end{aligned}$$

où l'espérance est par rapport à la loi du mouvement brownien. La nouvelle fonction γ est toujours convexe, on peut donc appliquer l'inégalité de Jensen :

$$\begin{aligned} \gamma(G(W_t)) &= \gamma(\mathbf{E}[ZF \exp(2U(W_t) - 2\nu(W_t)) | \mathcal{F}_{\{t\}}]) \\ &\leq \mathbf{E}[\gamma(ZF \exp(2U(W_t) - 2\nu(W_t))) | \mathcal{F}_{\{t\}}], \end{aligned}$$

et d'après la définition de γ ,

$$\gamma(G(W_t)) \leq \mathbf{E} \left[ZF \exp(2U(W_t) - 2\nu(W_t)) \log_+^p(ZF \exp(2U(W_t) - 2\nu(W_t))) | \mathcal{F}_{\{t\}} \right]$$

Toujours par le même calcul que pour l'entropie simple, on multiplie par $Z^{-1} \exp(2\nu - 2U)$:

$$\gamma(G(W_t)) Z^{-1} \exp(2\nu - 2U) \leq \mathbf{E} \left[F \log_+^p(ZF \exp(2U(W_t) - 2\nu(W_t))) | \mathcal{F}_{\{t\}} \right].$$

On prend ensuite l'espérance, pour obtenir :

$$I_p(t) \leq \mathbf{E} \left[F \log_+^p(ZF \exp(2U(W_t) - 2\nu(W_t))) \right].$$

Comme la fonction \log_+^p est croissante, on peut se ramener à des quantités plus simples. En effet, la définition de F (équation (3.1)) et les calculs faits précédemment (ceux qui suivent l'équation (3.5)) entraînent :

$$F \leq \exp \left(U(\mathbf{x}) - U(W_t) + c \text{Card}(L)t + b_V'' \sigma^2 \sum_i \int_0^t W_{i,s}^2 ds \right),$$

où b_V'' est la constante définie juste avant l'équation (3.7), page 65. On a donc pour $I_p(t)$:

$$I_p(t) \leq \mathbf{E} \left[F \log_+^p \left(Z \exp \left(U(\mathbf{x}) + U(W_t) - 2\nu(W_t) + c \text{Card}(L)t + b_V'' \sigma^2 \sum_i \int_0^t W_{i,s}^2 ds \right) \right) \right]. \quad (3.21)$$

On peut encore majorer l'argument du logarithme. En effet, $U_L(\mathbf{x})$ est majoré par un polynôme en $\text{Card}(L)$ (cf. l'équation (3.6), dans la preuve du théorème 41). De plus, on peut trouver des constantes telles que :

$$U(W_t) - 2\nu(W_t) \leq c \left(1 + \log_+ \left(\frac{1}{2\pi t} \right) \right) \text{Card}(L) + \sum W_{i,t}^q,$$

où q est une puissance suffisamment grande pour absorber les termes d'auto-interaction.

Majorons alors le terme en \log_+ de (3.21) :

$$\begin{aligned} \log_+(\cdot) &\leq |\log(\cdot)| \\ &\leq |\log(Z_L)| + P(\text{Card}(L)) + c(1 + \log_+(1/(2\pi t) + t)) \text{Card}(L) \\ &\quad + \sum_{i \in L} W_{i,t}^q + b_V'' \sigma^2 \int_0^t \sum_{i \in L} W_{i,s}^2 ds. \end{aligned}$$

On a déjà vu que $|\log(Z_L)|$ admettait une majoration polynomiale en $\text{Card}(L)$: on peut donc modifier le polynôme P pour en tenir compte. Pour séparer les termes après avoir pris la puissance p -ième, on utilise la propriété facile valable pour une famille quelconque de réels :

$$\left| \sum_{i \in I} x_i \right|^p \leq \left(\sum_{i \in I} |x_i| \right)^p \leq |I|^{p-1} \sum |x_i|^p.$$

En reprenant la majoration (3.21), et en se rappelant que F est d'espérance 1, on obtient donc :

$$\begin{aligned} I_p(t) &\leq 4^{p-1} (|P(\text{Card}(L))|^p + c^p (1 + \log_+(1/(2\pi t) + t))^p \text{Card}(L)^p) \\ &\quad + 4^{p-1} \left(\mathbf{E} \left[F \left(\sum_{i \in L} W_{i,t}^q \right)^p \right] + \mathbf{E} \left[F \left(b_V'' \sigma^2 \int_0^t \sum_{i \in L} W_{i,s}^2 ds \right)^p \right] \right). \end{aligned}$$

On élève ensuite à la puissance $1/p$, et on remajore :

$$\begin{aligned} I_p(t)^{1/p} &\leq 4 (P(\text{Card}(L)) + c(1 + \log_+(1/(2\pi t) + t)) \text{Card}(L)) \\ &\quad + 4 \left(\mathbf{E} \left[F \left(\sum_{i \in L} W_{i,t}^q \right)^p \right]^{1/p} + \mathbf{E} \left[F \left(b_V'' \sigma^2 \int_0^t \sum_{i \in L} W_{i,s}^2 ds \right)^p \right]^{1/p} \right). \quad (3.22) \end{aligned}$$

Notons A le troisième terme de cette somme. Pour le majorer, on fait rentrer la puissance p dans la somme sur les sites, quitte à perdre un facteur $\text{Card}(L)^{p-1}$:

$$A = \mathbf{E} \left[F \left(\sum_{i \in L} W_{i,t}^q \right)^p \right]^{1/p} \leq \text{Card}(L)^{(p-1)/p} \mathbf{E} \left[F \sum_{i \in L} W_{i,t}^{pq} \right]^{1/p}.$$

Comme F est par définition la densité du processus de Langevin par rapport au Brownien, on est ramené à :

$$A \leq \text{Card}(L) \mathbf{E} \left[\sum_{i \in L} X_{i,t}^{pq} \right]^{1/p}.$$

On peut appliquer le contrôle de moments (3.15) :

$$\begin{aligned} A &\leq \text{Card}(L) \left(\sum_{i \in L} \left(\|(\mathbf{x}^L; \mathbf{z})\|_{\tilde{\alpha}_i}^{2pq} + cp^2q^2 \right) \exp(kp^2q^2t) \right)^{1/p} \\ &\leq \text{Card}(L) \exp(kpq^2t) \left(\sum_{i \in L} \left(\|(\mathbf{x}^L; \mathbf{z})\|_{\tilde{\alpha}_i}^{2q} \right) + \text{Card}(L)(cp^2q^2)^{1/p} \right). \end{aligned}$$

Rappelons que l'on travaille dans la boîte $L = L_n$. Comme la configuration \mathbf{x} est tempérée, on peut majorer $\|(\mathbf{x}^L; \mathbf{z})\|_{\tilde{\alpha}_i}$ par un polynôme en n . Le terme $(cp^2q^2)^{1/p}$ est borné indépendamment de p .

On peut maintenant choisir $t = 1/p$ pour obtenir une majoration de A par un polynôme en n , indépendant de p .

Pour ce choix de t , revenons à l'équation (3.22). Le quatrième terme se majore de la même façon que celui que nous venons de traiter, et l'on obtient l'existence d'un polynôme Q_1 , ne dépendant que de \mathbf{x} et du potentiel V (à travers b_V'' et la valeur de q), tel que :

$$I_p(1/p)^{1/p} \leq 4P(\text{Card}(L_n)) + 4c(1 + \log_+(p/(2\pi)) + 1/p) \text{Card}(L_n) + Q_1(n).$$

Comme $\text{Card}(L_n)$ est un polynôme en n , et comme $\log_+(p/(2\pi))$ est majoré par p , on aboutit à :

$$I_p(1/p)^{1/p} \leq p \times Q(n),$$

pour un certain polynôme Q .

Pour conclure, on note que $I_p(\cdot)$ est une fonction décroissante du temps, car la fonction $x \mapsto x \log_+^p$ est convexe (cf. le lemme 46), donc pour tout $t \geq 1/p$,

$$I_p(t)^{1/p} \leq p \times Q(n).$$

Comme p est supérieur à 1, ce résultat est vérifié pour tout $t \geq 1$, ce qui achève la preuve. \square

Première estimation de l'entropie

On cherche ici à montrer un premier résultat de décroissance de l'entropie le long du semi-groupe. On reprend l'idée de la section 5 de [CGG05] (*Convergence to equilibrium for diffusion processes*), idée que l'on peut résumer de la manière suivante.

Rappelons qu'une inégalité de Sobolev logarithmique assure une décroissance exponentielle de l'entropie le long du semi-groupe :

$$\exists c, \forall f, \forall t, \quad \mathbf{Ent}_\mu(P_t f) \leq \exp\left(-\frac{t}{c}\right) \mathbf{Ent}_\mu(f).$$

Si l'on ne considère que des inégalités plus faibles, comme les inégalités de Beckner (3.12), on peut toujours obtenir des décroissances exponentielles, mais uniquement pour des fonctions bornées (et la vitesse dépend de la borne). On a donc des résultats du type :

50 Lemme ([CGG05], exemple 4.3). *Si μ vérifie une inégalité de Beckner $\mathbf{GBI}(a, C_a)$, il existe C'_a telle que, pour toute densité de probabilité bornée h , on ait la décroissance :*

$$\mathbf{Ent}_\mu(P_t h) \leq \mathbf{Ent}_\mu(h) \times \exp\left(-\frac{t}{C'_a(1 + \log^{1-a}(\|h\|_\infty))}\right). \quad (3.23)$$

Cattiaux, Gentil et Guillin ont montré que l'on peut en déduire une décroissance de l'entropie pour toutes les fonctions, mais que la vitesse devient sous-exponentielle : c'est le résultat du théorème 45, cité précédemment.

La preuve utilise un argument de troncature : en décomposant une fonction h en $h\mathbf{1}_{h \leq K} + h\mathbf{1}_{h > K}$, on peut appliquer (3.23) pour contrôler le premier terme, et estimer directement le second, puis choisir un bon niveau K .

C'est ce que nous allons faire ici, pour obtenir le résultat suivant. Rappelons que H_t^n est l'entropie relative du processus par rapport à sa mesure d'équilibre (cf. la définition 48).

51 Proposition. *Si les mesures μ_n vérifient, uniformément en n , une inégalité de Beckner (hypothèse du théorème 44), alors pour tout $a < a_0$, il existe un polynôme $Q = Q_{a,\mathbf{x}}$, dont le degré ne dépend que de l'interaction V et de la dimension du réseau, et un réel $t_0(a)$, tels que :*

$$\forall s \geq 1, \forall t \geq t_0(a), \quad H_{s+t}^n \leq \frac{1}{c_{t,n}} \phi(H_s), \quad (3.24)$$

où $\phi(x) = x(1 + \log_+(1/x))$, et $c_{t,n} = t^{1/(1-a)}/Q(n)$.

La preuve nécessite deux lemmes.

52 Lemme ([CGG05], lemme 5.3). *Soit h une densité de probabilité par rapport à μ . Si il existe $c > 0$ tel que la p -entropie est majorée :*

$$\forall p > 1, \quad H_{p,t} \leq cp,$$

et si K est tel que :

$$K \geq e^2, \quad \log(K) \geq 2e \times \mathbf{Ent}_\mu(h),$$

alors

$$\mathbf{Ent}_\mu(h\mathbf{1}_{h > K}) \leq (ec + 2) \frac{\mathbf{Ent}_\mu(h)}{\log(K)} \log\left(\frac{\log(K)}{\mathbf{Ent}_\mu(h)}\right).$$

Nous renvoyons à [CGG05] pour une preuve.

Le second résultat donne un contrôle de l'entropie des fonctions positives bornées.

53 Lemme. *Soit μ une mesure vérifiant une inégalité de Beckner $\mathbf{GBI}(a, C_a)$. Il existe C'_a telle que si h est une densité de probabilité, H l'entropie de h , et si :*

$$K \geq e^2, \quad \log(K) \geq 4H,$$

alors

$$\mathbf{Ent}_\mu(P_t(h\mathbf{1}_{h \leq K})) \leq H \times \exp\left(\frac{t}{C'_a \log^{1-a}(K)}\right).$$

Démonstration. Ce lemme est un corollaire de la majoration (3.23). Pour le voir, on voudrait renormaliser la fonction $h\mathbf{1}_{h \leq K}$ pour en faire une densité de probabilité, et pouvoir lui appliquer le résultat précédent. Il faut donc s'assurer que son intégrale est non-nulle. Or, par le lemme 3.4 de [CG06], on a pour $K \geq e^2$:

$$\int h\mathbf{1}_{h > K} \leq \frac{2H}{\log K}.$$

Comme par hypothèse, $\log(K) \geq 4H$, il vient :

$$\int h\mathbf{1}_{h > K} \leq \frac{1}{2},$$

et comme $\int h = 1$, on a finalement

$$\int h\mathbf{1}_{h \leq K} = 1 - \int h\mathbf{1}_{h > K} \geq 1/2. \quad (3.25)$$

On peut donc renormaliser $h\mathbf{1}_{h \leq K}$: notons \tilde{h} la fonction renormalisée. C'est une densité de probabilité bornée, et on peut lui appliquer la majoration (3.23) :

$$\mathbf{Ent}_\mu(P_t \tilde{h}) \leq \exp\left(-\frac{t}{C'_a(1 + \log^{1-a}(\|\tilde{h}\|_\infty))}\right) \mathbf{Ent}_\mu(\tilde{h}).$$

Multiplions cette inégalité par $\int h\mathbf{1}_{h \leq K} d\mu$, que l'on peut faire rentrer dans les entropies (par homogénéité).

$$\mathbf{Ent}_\mu(P_t(h\mathbf{1}_{h \leq K})) \leq \exp\left(-\frac{t}{C'_a(1 + \log^{1-a}(\|\tilde{h}\|_\infty))}\right) \mathbf{Ent}_\mu(h\mathbf{1}_{h \leq K}). \quad (3.26)$$

Enfin, la norme infinie de \tilde{h} se contrôle :

$$\begin{aligned} \tilde{h} &= \frac{h\mathbf{1}_{h \leq K}}{\int h\mathbf{1}_{h \leq K} d\mu} \\ &\leq \frac{K}{1/2}, \end{aligned}$$

en réutilisant la minoration (3.25) de l'intégrale. Comme $K \geq e$, on peut en déduire la majoration suivante du dénominateur de (3.26) :

$$\begin{aligned} C'_a(1 + \log^{1-a}(\|\tilde{h}\|_\infty)) &\leq C'_a(\log^{1-a}(K) + \log^{1-a}(2K)) \\ &\leq C''_a \log^{1-a}(K). \end{aligned}$$

Le lemme est démontré. □

Avec ces deux résultats, nous pouvons maintenant montrer la proposition 51.

Démonstration. Par définition, $H_t = \mathbf{Ent}_{\mu_{L_n}}(h_t)$. On se place au temps $s + t$, et comme annoncé précédemment, on tronque h_s : pour tout K ,

$$h_s = h_s\mathbf{1}_{h_s \leq K} + h_s\mathbf{1}_{h_s > K}.$$

Pour tout couple (f, g) de fonctions positives, $\mathbf{Ent}(f + g) \leq \mathbf{Ent}(f) + \mathbf{Ent}(g)$ (ceci se voit facilement à partir de la formule variationnelle de l'entropie : $\mathbf{Ent}_\mu(f) = \sup \{ \int f h, \int e^h d\mu = 1 \}$).

$$\begin{aligned} \forall s, t, \forall K, \quad H_{t+s} &= \mathbf{Ent}(P_t h_s) \leq \mathbf{Ent}(P_t(h_s \mathbf{1}_{h_s \leq K})) + \mathbf{Ent}(P_t(h_s \mathbf{1}_{h_s > K})) \\ &\leq \mathbf{Ent}(P_t(h_s \mathbf{1}_{h_s \leq K})) + \mathbf{Ent}(h_s \mathbf{1}_{h_s > K}), \end{aligned}$$

car l'entropie décroît le long de P_t . Supposons que K vérifie :

$$\begin{cases} K \geq e^e, \\ \log(K) \geq 2eH_s. \end{cases} \quad (3.27)$$

On peut alors appliquer le lemme 53 au premier terme. Le contrôle de p -entropie de la section précédente (lemme 49) garantit que l'on peut appliquer le lemme 52 au second terme. Pour tout K vérifiant les deux hypothèses, on obtient donc :

$$H_{t+s} \leq \exp\left(-\frac{t}{C_{a_0} \log(K)^{1-a_0}}\right) H_s + Q(n) \frac{H_s}{\log(K)} \log\left(\frac{\log(K)}{H_s}\right). \quad (3.28)$$

On choisit maintenant K en fonction de t , comme unique solution sur $]e^e, \infty[$ de l'équation suivante :

$$\log(K) = \left(\frac{t}{C_{a_0} \log \log K}\right)^{1/(1-a_0)}, \quad (3.29)$$

Ce $K(t)$ est bien défini car $K \mapsto \log(K) \log \log(K)^{1/(1-a_0)}$ est bijective de $]e^e, \infty[$ sur $]0, \infty[$. Supposons pour le moment que K vérifie la seconde condition de (3.27). L'inégalité (3.28) devient :

$$H_{t+s} \leq \frac{1}{\log(K)} H_s + Q(n) \frac{H_s}{\log(K)} \log\left(\frac{\log(K)}{H_s}\right).$$

On cherche un majorant plus simple à manipuler. Comme $K \geq e^e$, $\log \log K \geq 1$, et on peut toujours supposer $Q(n) \geq 1$. On obtient alors :

$$\begin{aligned} H_{t+s} &\leq \frac{\log \log(K)}{\log(K)} H_s + \frac{Q(n) H_s}{\log(K)} (\log \log(K) + \log_+(1/H_s)) \\ &\leq \frac{\log \log(K)}{\log(K)} H_s + \frac{Q(n) H_s}{\log(K)} \log \log(K) (1 + \log_+(1/H_s)) \\ &\leq \left(\frac{\log \log(K)}{\log(K)}\right) Q(n) H_s (2 + \log_+(1/H_s)). \end{aligned} \quad (3.30)$$

Le choix de K assure l'existence d'une constante c_a telle que :

$$\log \log(K) / \log(K) \leq \frac{c_a}{t^{1/(1-a)}},$$

pour t plus grand qu'un certain $t_0(a)$.

En reportant ceci dans (3.30), et en posant $Q_a(n) = 2c_a Q(n)$, pour faire disparaître le facteur 2, on obtient finalement :

$$H_{t+s} \leq \frac{Q_a(n)}{t^{1/(1-a)}} H_s (1 + \log_+(1/H_s)),$$

qui est l'inégalité (3.24) annoncée.

Supposons maintenant que la quantité K (solution de (3.29)) ne vérifie pas l'hypothèse (3.27). Comme par construction, $K \geq e^e$, le seul cas qui pose problème est celui où $\log(K) \leq 2eH_s$. On sait qu'il existe un polynôme $Q'_{a,x}$ tel que $H_s^n \leq Q'_{a,x}(n)$, pour tout $s \geq 1$. Dans ce cas,

$$\log(K) \leq Q''_{a,x}(n).$$

Autrement dit,

$$1 \leq \frac{Q''_{a,x}(n)}{\log(K)}.$$

Comme $\log \log(K) \geq 1$, on en déduit :

$$1 \leq \frac{\log \log K}{\log K} Q''_{a,x}(n).$$

Enfin, l'entropie H décroît le long du semi-groupe. Pour tout $s \geq 1$ et $t \geq t_0(a)$, on en déduit :

$$H_{t+s} \leq H_s \leq \left(\frac{\log \log K}{\log K} \right) Q''_{a,x}(n) H_s (2 + \log_+(1/H_s)).$$

L'inégalité (3.30) est donc toujours valable dans ce cas, et on peut répéter la fin de la preuve. \square

Itération de l'estimation et conclusion

Nous sommes maintenant en mesure de prouver le théorème 44.

L'estimation précédente (3.24) est intéressante dès que $c_{t,n}$ est strictement supérieur à 1. Notons D le degré de Q (il est indépendant de x et de a). On supposera :

$$a_0 > a_{min} = \frac{D-1}{D}. \quad (3.31)$$

Remarquons que l'on peut alors choisir, dans le lemme 51, $a > a_{min}$.

54 Lemme. *Les propriétés suivantes sont vérifiées.*

- Il existe $t_0(n)$ tel que, pour tout $t > t_0(n)$, $c_{t,n} > 1$;
- Il existe $u_0(n)$ tel que, pour tout $u > u_0$, u s'écrive $t(c_{t,n})^2$, avec un $t > t_0(n)$;
- La quantité $u_0(n)$ est relativement petite :

$$u_0(n) = o(n). \quad (3.32)$$

De plus, pour tout $u > u_0(n)$, pour tout $s \geq 1$,

$$H_{s+u}^n \leq (e + H_s) \exp \left(- \frac{u^{1/(3-a)}}{Q(n)^{(1-a)/(3-a)}} \right). \quad (3.33)$$

Ce lemme entraîne le théorème 44.

Pour le voir, on reprend point par point la preuve du théorème 40, à la fin de la section 3.1. La comparaison entre boîte finie et infinie, et l'argument de convergence faible des mesures (qui servent à traiter le premier et le dernier terme de (3.8), page 66) sont identiques — en particulier, on choisit n en fonction de t comme dans (3.10). On adapte uniquement l'argument de convergence en volume fini.

On utilise le lemme qui vient d'être prouvé. D'après (3.32), on peut l'appliquer à $u = t$, $n = n(t)$ (c'est-à-dire que l'on peut contrôler l'entropie dans une boîte pour des temps de l'ordre du diamètre de la boîte).

On fixe $s = 1$. Comme $Q(n)$ est (par définition) de degré D , il est majoré par n^D à une constante multiplicative près, et il existe c tel que :

$$H_{1+t}^{n(t)} \leq \left(e + H_1^{n(t)} \right) \exp \left(-c \left(\frac{n(t)}{n(t)^{D(1-a)}} \right)^{1/(3-a)} \right).$$

Comme $H_1^{n(t)}$ est à croissance polynomiale en n , donc en t (on réutilise le théorème 41 de la section précédente), il suffit que la puissance de n dans l'exponentielle soit strictement positive pour que toute la quantité tende vers 0. Or cette puissance vaut :

$$\frac{1}{3-a} (1-D+aD),$$

qui est bien strictement positive, car $a > a_{min}$ (défini par (3.31)).

L'entropie dans la boîte $L_{n(t)}$ au temps t tend bien vers 0, comme lorsque l'on supposait une inégalité de Sobolev logarithmique (cf. les arguments sur la convergence en boîte finie dans la section précédente, page 67).

Le reste de la preuve est identique. Le théorème 44 est donc prouvé.

Il ne reste plus qu'à montrer le lemme.

Démonstration du lemme 54. Montrons d'abord l'existence des réels t_0, u_0 .

Rappelons que $c_{t,n} = t^{1/(1-a)}/Q(n)$, et que Q est de degré D . On choisit un a' tel que :

$$a_{min} < a' < a < a_0,$$

En posant alors $t_0 = cn^{D(1-a')}$ pour une certaine constante c , on peut bien garantir

$$c_{t,n} \geq \frac{t^{1/(1-a)}}{n^D} > 1,$$

pour $t \geq t_0$.

Ensuite, si l'on pose $u = u(t, n) = t(c_{t,n})^2$, u est croissant en t et on peut choisir

$$u_0(n) = u(t_0(n), n) = cn^{D(1-a')(3-a)/(1-a)}/Q(n)^2.$$

On veut que u_0 soit petit devant n (on ne veut pas attendre un temps supérieur au diamètre de la boîte). Le choix précédent garantit :

$$u_0(n) \sim cn^{D(1-a')(3-a)/(1-a)-2D}.$$

Ce u_0 est négligeable devant n si

$$D \frac{(1-a')(3-a)}{1-a} - 2D < 1.$$

Or ceci est vérifié pour $a' = a$ (puisque $D(3-a)-2D = D(1-a) < 1$, car $a > a_{min}$). Par continuité, on peut donc choisir a' tel que l'inégalité reste vraie.

Venons en maintenant au principal argument, et montrons (3.33). L'idée est d'itérer l'estimation donnée par le lemme 51. Pour cela, fixons un t supérieur à $t_0(n)$, et définissons la suite (u_k) par $u_k = H_{s+kt}$. Pour contrôler u_k , on la compare à la suite v_k définie par récurrence :

$$\begin{cases} v_0 = u_0, \\ v_{k+1} = f(v_k), \end{cases}$$

où $f(x) = \frac{1}{c_{t,n}}\phi(x)$ (cf. le lemme 51). Comme f est croissante, et comme

$$u_{k+1} \leq f(u_k)$$

(par l'équation (3.24) du lemme 51, appliqué à $s = s + tk$ et $t = t$), on montre facilement par récurrence que $u_k \leq v_k$.

Or la suite v_k est définie simplement par itération, et la condition $c_{t,n} > 1$ entraîne que f a un seul point stationnaire stable, $x_e = \exp(1 - c_{t,n})$, et que v_k converge vers ce point. Si l'on part d'un point à gauche de x_e , v_k reste inférieur à x_e . À droite de x_e , l'application f est $\left(1 - \frac{1}{c_{t,n}}\right)$ -contractante, et l'on a donc, pour tout k :

$$v_k \leq x_e + \left(1 - \frac{1}{c_{t,n}}\right)^k (v_0 - x_e)_+. \quad (3.34)$$

La valeur explicite de x_e , et la majoration $(1 - 1/c)^k \leq \exp(-k/c)$ montre :

$$\forall k, \quad v_k \leq \exp(1 - c_{t,n}) + v_0 \exp(-k/c_{t,n}).$$

Revenons à l'entropie H , évaluée au temps $s + u$. Si u peut s'écrire kt avec un $t > t_0(n)$, l'itération précédente fournit :

$$H_{s+u} \leq \exp(1 - c_{t,n}) + v_0 \exp(-k/c_{t,n}).$$

Pour u supérieur à $u_0(n)$, on choisit t tel que $t(c_{t,n})^2 = u$, et $k = c_{t,n}^2$ (plus précisément, on choisit k l'entier le plus proche de cette quantité). Par définition de u_0 , t est bien supérieur à t_0 . Les deux quantités t et $c_{t,n}$ se réexpriment en fonction de u :

$$tc_{t,n}^2 = u, \text{ donc } t = (uQ(n)^2)^{(1-a)/(3-a)}, \\ c_{t,n} = \frac{1}{Q(n)}(uQ(n)^2)^{1/(3-a)}.$$

Comme $u_0 = H_s$, on en déduit :

$$H_{s+u} \leq (e + H_s) \exp(-c_{t,n}) \\ \leq (1 + H_s) \exp\left(-\frac{u^{1/(3-a)}}{Q(n)^{(1-a)/(3-a)}}\right),$$

ce qui était annoncé. □

3.3 Des inégalités fonctionnelles aux conditions de mélange

La section 3.1 a montré que, lorsque les mesures de Gibbs en volume fini vérifient une inégalité de Sobolev logarithmique avec des constantes n'augmentant que lentement avec le diamètre de la boîte, la mesure de Gibbs en volume infini est unique. On peut se demander si des conditions du même genre, sur les constantes de Poincaré ou de Sobolev logarithmique, donnent d'autres informations sur les mesures en volume fini et infini. On sait en effet que si les constantes dans l'une ou l'autre inégalité sont uniformes (en la boîte et la condition aux bords), les mesures vérifient des conditions dites de mélange exponentiel et de décroissance

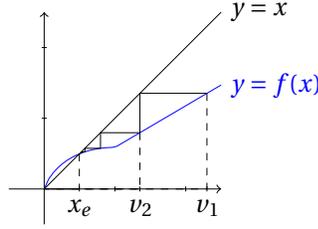


FIG. 3.1: Étude de la suite récurrente v_k . La fonction $f : x \mapsto \frac{1}{c}x(1 + \log_+(1/x))$ admet, pour $c > 1$, un unique point stationnaire stable en $x_e = \exp(1 - c)$.

des corrélations exponentielles (*cf.* l'introduction de la thèse). On sait aussi que sous certaines conditions, on a l'implication inverse : des conditions de mélange sur les mesures en boîtes finies impliquent la validité d'inégalités de Sobolev logarithmiques (et *a fortiori* de Poincaré) avec des constantes uniformes ([Yos01]).

L'objet de cette section d'adapter l'article de N. Yoshida ([Yos01]) pour montrer qu'on peut relâcher l'hypothèse, quitte à affaiblir la conclusion : il suffit que les constantes de Poincaré n'augmentent pas trop vite avec la taille de la boîte pour que l'on obtienne un mélange polynomial. Nous verrons dans la section suivante que, sous certaines conditions, ce mélange polynomial entraîne en fait des inégalités fonctionnelles uniformes : dans ce cas, on sait donc que si les constantes n'augmentent pas trop vite, elles sont en réalité uniformément bornées.

Pour toute fonction cylindrique f , on notera S_f le support de f , *i.e.* le plus petit ensemble L de sites tel que $f(\mathbf{x})$ ne dépende que des $(x_i, i \in L)$.

Le résultat précis pour passer du trou spectral au mélange est le suivant :

55 Théorème. *Supposons qu'il existe $C_{3.35}$ telle que les mesures de Gibbs dans les boîtes finies Λ vérifient :*

$$\forall \mathbf{z}, \forall \Lambda, \quad \gamma_{SG}(\Lambda, \mathbf{z}) \leq C_{3.35} \psi(\text{diam } \Lambda) \quad (3.35)$$

où $\psi(t) = t / \log t$. Alors il existe une constante C telle que la condition de mélange suivante soit vérifiée : pour toute boîte Λ , toute $f \in \mathcal{C}_\Lambda$, tout site frontière $i \in \partial\Lambda$, et toutes conditions aux bords $\mathbf{z}, \bar{\mathbf{z}}$ ne différant qu'en i ,

$$|\mathbf{E}^{\Lambda, \mathbf{z}}(f) - \mathbf{E}^{\Lambda, \bar{\mathbf{z}}}(f)| \leq |z_i - \bar{z}_i| \cdot \|f\| \phi(d(i, S_f)), \quad (3.36)$$

où $\phi(\rho) = \rho^{d-1-C/C_{3.35}}$, S_f est le support de la fonction cylindrique f et $d(\cdot, \cdot)$ la distance sur \mathbb{Z}^d donnée par la norme infinie.

On peut faire quelques remarques sur ce résultat. Le choix de la distance est arbitraire, puisque l'on passe d'un choix à l'autre quitte à perdre des constantes multiplicatives.

Notons ensuite que l'hypothèse peut être légèrement affaiblie, et que l'on peut par exemple supposer :

$$\limsup_{D \rightarrow \infty} \sup_{\Lambda, \mathbf{z}} \left\{ \frac{\gamma_{SG}(\Lambda, \mathbf{z})}{\psi(\text{diam } \Lambda)}; \text{diam } \Lambda \geq D \right\} < C_{3.35} \quad (3.37)$$

C'est sous cette forme que l'hypothèse apparaît dans la preuve pour les spins compacts dans [Yos97]. Comme la décroissance finale (la valeur de la fonction ϕ) dépend de la constante $C_{3.35}$, cet affaiblissement de l'hypothèse peut, en permettant de se limiter à des grandes boîtes, autoriser un meilleur choix de $C_{3.35}$ et donc un meilleur résultat. Les modifications à faire dans la preuve sont minimes.

Ensuite, on peut comparer l'hypothèse avec celle utilisée pour montrer l'unicité de la mesure de Gibbs (théorème 40). La différence fondamentale est qu'ici, on demande une majoration de la constante pour *toutes* les boîtes, et surtout *toutes les conditions aux bords*, alors que la preuve de l'unicité fonctionne dès qu'on contrôle les constantes pour une famille de boîtes et une condition aux bords (la condition nulle). En contrepartie, on ne demande ici qu'une hypothèse sur le trou spectral, qui est *a priori* plus facile à obtenir que les constantes de Sobolev logarithmique ou celles des inégalités de Beckner.

Notons aussi que la constante C , qui apparaît dans le théorème comme une valeur critique (la fonction ϕ ne devenant décroissante que si C est suffisamment grande vis-à-vis de $C_{3.35}$), peut être explicitée en fonction de la dimension et du potentiel. On verra dans la preuve que l'on peut choisir pour C tout réel strictement inférieur à $d_\beta/(6k)$, où d_β et k sont les constantes qui apparaissent dans le théorème 58.

Enfin, nous devons signaler qu'un résultat très proche du nôtre apparaît dans l'article de synthèse de T. Bodineau et B. Helffer ([BH99a], théorème 3.7) On peut reformuler ce résultat ainsi :

56 Théorème. *Soit C_N une suite telle que, pour toute boîte Λ dont les côtés sont entre N et $2N$,*

$$\gamma_{SG}(\Lambda, \mathbf{z}) \leq C_N,$$

et supposons que $C_N/N \rightarrow 0$.

Alors il existe N_0 et C tel que, pour toute boîte Λ dont les côtés sont plus grands que N_0 , et pour toute condition aux bords \mathbf{z} ,

$$\gamma_{SG}(\Lambda, \mathbf{z}) \leq C.$$

En particulier, $(C_N)_{N \geq N_0}$ est bornée par C .

Ce dernier théorème admet une preuve plus simple que la nôtre. L'hypothèse est plus faible (on demande des contrôles sur les boîtes suffisamment « épaisses ») et la conclusion également (on ne récupère des contrôles que sur ces boîtes épaisses). Notre hypothèse sur la vitesse de croissance des constantes est également légèrement plus restrictive (on impose une croissance en $\text{diam } \Lambda / \log(\text{diam } \Lambda)$, alors que n'importe quelle croissance strictement sous-linéaire convient dans le théorème 56).

Résultats préliminaires

On prouve le théorème 55 en trois temps : après avoir donné deux estimations préliminaires, on prouvera une forme restreinte du théorème, puis on passera de cette forme restreinte à la forme générale.

Fixons d'abord quelques notations. Pour une boîte Λ et une configuration \mathbf{z} , on note $\mathbf{E}^{\Lambda, \mathbf{z}}$ l'intégration par rapport à la mesure de Gibbs $\mu_{\Lambda, \mathbf{z}}$. Quand \mathbf{z} est omis, \mathbf{E}^Λ est le noyau $\mathbf{z} \mapsto \mathbf{E}^{\Lambda, \mathbf{z}}$. On introduit également de nouvelles notations pour la dynamique, en notant :

$$\mathbf{P}_t(f)(\mathbf{x}) : \mathbf{z} \mapsto \mathbf{E} \left[f(X_t^{\Lambda, \mathbf{z}, \mathbf{x}}) \right].$$

Enfin, pour toute mesure μ , on note $\mu(f; g)$ la covariance sous μ :

$$\mu(f; g) = \int (f - \mu f)(g - \mu g) d\mu.$$

On fera l'hypothèse suivante.

Hypothèse 4. *La variance du spin à l'origine est majorée uniformément en la boîte et la condition aux bords :*

$$\exists C_{3.38}, \forall \Lambda, \forall \mathbf{z}, \quad \mathbf{E}^{\Lambda, \mathbf{z}}(x_0; x_0) \leq C_{3.38}. \quad (3.38)$$

L'article de Yoshida [Yos00] indique que cette propriété est toujours vérifiée quand l'auto-interaction est somme d'une fonction convexe et d'une fonction lipschitzienne bornée, mais la preuve semble faire appel à des hypothèses supplémentaires de convexité ([Yos99], hypothèse U2).

Notons que dans les théorèmes d'origine, ce point ne pose pas de problème, car toutes les propriétés dont on veut montrer l'équivalence (trou spectral uniforme, décroissance des corrélations, etc.) impliquent facilement l'hypothèse.

Sous cette hypothèse, on a la proposition :

57 Proposition. *La variance d'une fonction est contrôlée par sa norme triple :*

$$\forall \Lambda, \mathbf{z}, f \in \mathcal{C}_\Lambda, \quad \mathbf{E}^{\Lambda, \mathbf{z}}(f; f) \leq C_{3.38} \|f\|^2,$$

où l'on rappelle que $\|f\| = \sum_i \|\nabla_i f\|_\infty$.

Démonstration. Notons μ la mesure $\mu_{\Lambda, \mathbf{z}}$. La variance s'écrit :

$$\mathbf{Var}_\mu(f) = \frac{1}{2} \iint (f(\mathbf{x}) - f(\mathbf{y}))^2 d\mu(\mathbf{x}) d\mu(\mathbf{y}). \quad (3.39)$$

Notons i_1, i_2, \dots, i_n une énumération de la boîte Λ . Pour deux configurations \mathbf{x} et \mathbf{y} , on peut construire une suite de configurations $(\mathbf{x}^l)_{l=0..n}$ pour passer de \mathbf{x} à \mathbf{y} , en posant :

$$x_{i_k}^l = \begin{cases} x_{i_k} & \text{si } k > l, \\ y_{i_k} & \text{si } k \leq l. \end{cases}$$

On fait donc apparaître une somme télescopique :

$$|f(\mathbf{x}) - f(\mathbf{y})| \leq \sum_{l=1}^n |f(\mathbf{x}^{l-1}) - f(\mathbf{x}^l)|.$$

Comme \mathbf{x}^{l-1} et \mathbf{x}^l ne diffèrent qu'au site i_l , on a grâce aux accroissements finis :

$$\begin{aligned} |f(\mathbf{x}) - f(\mathbf{y})| &\leq \sum_{l=1}^n \|\nabla_{i_l} f\|_\infty |x_{i_l} - y_{i_l}| \\ &\leq \|f\| \sum_{l=1}^n \frac{\|\nabla_{i_l} f\|_\infty}{\|f\|} |x_{i_l} - y_{i_l}|. \end{aligned}$$

On élève au carré, puis on applique l'inégalité de Hölder, en se souvenant du fait que $\|f\| = \sum \|\nabla_i f\|_\infty$:

$$|f(\mathbf{x}) - f(\mathbf{y})|^2 \leq \|f\|^2 \sum_{l=1}^n \frac{\|\nabla_{i_l} f\|_\infty}{\|f\|} |x_{i_l} - y_{i_l}|^2.$$

En revenant à l'expression (3.39) de la variance, il vient :

$$\begin{aligned} \mathbf{Var}_\mu(f) &\leq \frac{1}{2} \iint \|f\| \sum_{l=1}^n \|\nabla_{i_l} f\|_\infty (x_{i_l} - y_{i_l})^2 d\mu(\mathbf{x}) d\mu(\mathbf{y}) \\ &\leq \|f\| \sum_{l=1}^n \|\nabla_{i_l} f\|_\infty \times \frac{1}{2} \iint (x_{i_l} - y_{i_l})^2 d\mu(\mathbf{x}) d\mu(\mathbf{y}). \end{aligned}$$

On reconnaît dans la somme la variance de chaque spin, qui par hypothèse est uniformément majorée par une constante $C_{3.38}$. Finalement, on a donc comme annoncé :

$$\mathbf{Var}_\mu(f) \leq C_{3.38} \|f\|^2. \quad \square$$

Notons maintenant $\mathbf{P}_t^{\Lambda, \mathbf{z}}$ le semi-groupe dans le volume Λ avec condition aux bords \mathbf{z} — on sous-entendra parfois Λ et \mathbf{z} lorsqu'ils seront fixés. Le deuxième résultat préliminaire est le suivant.

58 Théorème (Finitude de la vitesse de propagation). *Il existe une constante k et un poids β , à décroissance exponentielle, tels que pour tout Λ , tout \mathbf{z} et toutes fonctions f et g de \mathcal{C}_Λ ,*

$$\mathbf{P}_t(f; g) \leq C_{3.40} e^{kt} \sum_{i, j \in \mathbb{Z}^d} \beta(i-j) \|\nabla_i f\|_\infty \|\nabla_j g\|_\infty \quad (3.40)$$

Remarque. *Cette formulation est légèrement différente de celle de l'article d'origine ([Yos01], lemme 3.2) qui précise que l'on peut choisir β sous la forme $\beta(i) = \exp(-\epsilon|i|)$, pour tout ϵ , quitte à changer la constante devant le t . En fait, ceci reste possible avec la version ci-dessus. Pour le voir, il suffit de remarquer que l'on peut fixer la vitesse de décroissance du poids α (défini juste avant la proposition 17, page 39) en changeant σ' . Vu la façon dont β est défini (cf. la preuve ci-dessous), la vitesse de décroissance de β peut donc elle aussi être fixée arbitrairement, quitte à changer σ' . Ce choix de σ' modifie alors les constantes dans les contrôles dynamiques présentés dans la section 2.6, ce qui se traduit par un changement de la constante k .*

Remarque. *L'appellation « finitude de la vitesse de propagation » peut paraître trompeuse. Dans les modèles discrets analogues (comme la dynamique de Glauber pour le modèle d'Ising), elle signifie qu'une modification de la condition initiale n'a d'impact en temps fini que sur un ensemble de sites fini. L'information (qu'un site a été modifié) se propage donc à vitesse finie dans \mathbb{Z}^d .*

Ici, le caractère continu du modèle fait qu'une modification quelconque de la condition initiale modifie immédiatement toute la dynamique. Cependant, l'effet de cette modification décroît exponentiellement avec la distance, et c'est ce qu'indique ce lemme. À titre d'exemple, si l'on prend pour f et g les valeurs des spins en deux sites distants i et j , le théorème nous donne :

$$\mathbf{P}_t(x_i, x_j) \leq C e^{kt} \exp(-|i-j|),$$

et la corrélation entre les deux spins est faible quand le temps t est petit par rapport à la distance entre les sites.

Démonstration. Pour montrer le théorème, on s'inspire de [Yos01], en utilisant le poids α introduit précédemment. On fixe t , la boîte Λ et la configuration \mathbf{z} , et on note \mathbf{P}_t pour $\mathbf{P}_t^{\Lambda, \mathbf{z}}$. L'idée est de passer par le semi-groupe en introduisant la fonction suivante :

$$\phi(s) = \mathbf{P}_s(\mathbf{P}_{t-s} f \mathbf{P}_{t-s} g).$$

La covariance cherchée s'exprime alors en fonction de ϕ :

$$\mathbf{P}_t(f; g) = \mathbf{P}_t(fg) - \mathbf{P}_t f \mathbf{P}_t g = \phi(t) - \phi(0).$$

La fonction ϕ est dérivable, et on a, en notant A le générateur du semi-groupe,

$$\begin{aligned} \phi'(s) &= \mathbf{P}_s A(\mathbf{P}_{t-s} f \mathbf{P}_{t-s} g) + \mathbf{P}_s (-(A \mathbf{P}_{t-s} f) \mathbf{P}_{t-s} g - \mathbf{P}_{t-s} f (A \mathbf{P}_{t-s} g)) \\ &= \mathbf{P}_s (\Gamma(\mathbf{P}_{t-s} f, \mathbf{P}_{t-s} g)) \\ &= \mathbf{P}_s \left(\sum_i \nabla_i(\mathbf{P}_{t-s} f) \cdot \nabla_i(\mathbf{P}_{t-s} g) \right) \end{aligned}$$

On cherche alors à étudier $\nabla_i \mathbf{P}_s f$ (ce qui revient à évaluer l'influence de la condition initiale au site i sur le processus pris au temps s). Pour cela, on prend \mathbf{x} une configuration quelconque, et on pose $\mathbf{y} = \mathbf{x} + h e_i$, où e_i est la configuration nulle partout sauf au site i où elle vaut 1. On regarde alors :

$$\begin{aligned} |\mathbf{P}_s f(\mathbf{x}) - \mathbf{P}_s f(\mathbf{y})| &= \mathbf{E} [|f(X_s^{\mathbf{x}}) - f(X_s^{\mathbf{y}})|] \\ &\leq \sum_j \|\nabla_j f\|_{\infty} \mathbf{E} \left[|X_{s,j}^{\mathbf{x}} - X_{s,j}^{\mathbf{y}}| \right], \end{aligned} \quad (3.41)$$

où la dernière ligne est obtenue en écrivant $X^{\mathbf{x}} - X^{\mathbf{y}}$ comme somme télescopique. On utilise le lemme 37 pour comparer les processus partant de \mathbf{x} et \mathbf{y} .

$$\begin{aligned} \alpha(0) \mathbf{E} \left[\left(X_{s,j}^{L,\mathbf{z},\mathbf{x}} - X_{s,j}^{L,\mathbf{z},\mathbf{y}} \right)^2 \right] &\leq e^{k'''s} \sum_l \alpha(j-l) (x_l - y_l)^2 \\ &\leq e^{k'''s} \alpha(j-i) h^2, \end{aligned}$$

puisque \mathbf{x} et \mathbf{y} ne diffèrent qu'au site i , où leur différence vaut h . L'inégalité de Hölder donne alors :

$$\mathbf{E} \left[|X_j^{L,\mathbf{z},\mathbf{x}} - X_j^{L,\mathbf{z},\mathbf{y}}| \right] \leq e^{k'''s/2} \left(\frac{\alpha(j-i)}{\alpha(0)} \right)^{1/2} |h|.$$

On reporte cette inégalité dans (3.41), puis on fait tendre h vers 0.

$$\begin{aligned} |\mathbf{P}_s f(\mathbf{x}) - \mathbf{P}_s f(\mathbf{y})| &\leq |h| e^{k'''s/2} \frac{1}{\sqrt{\alpha(0)}} \sum_j \|\nabla_j f\|_{\infty} \sqrt{\alpha(i-j)}; \\ |\nabla_i \mathbf{P}_s f| &\leq e^{k'''s/2} \frac{1}{\sqrt{\alpha(0)}} \sum_j \|\nabla_j f\|_{\infty} \sqrt{\alpha(i-j)}. \end{aligned}$$

On peut maintenant revenir au calcul de $\phi'(s)$ commencé plus haut :

$$\begin{aligned} |\phi'(s)| &\leq \mathbf{P}_s \left(\sum_i |\nabla_i(\mathbf{P}_{t-s} f) \cdot \nabla_i(\mathbf{P}_{t-s} g)| \right) \\ &\leq e^{k'''(t-s)} \sum_{i,j,l} \|\nabla_j f\|_{\infty} \|\nabla_l g\|_{\infty} \sqrt{\alpha(i-j)\alpha(i-l)}. \end{aligned}$$

On passe la somme sur i à l'intérieur, et on note β le poids $\sqrt{\alpha} * \sqrt{\alpha}$:

$$\phi'(s) \leq e^{k'''(t-s)} \sum_{j,l} \beta(j-l) \|\nabla_j f\|_{\infty} \|\nabla_l g\|_{\infty}$$

En intégrant entre 0 et t , on obtient :

$$\mathbf{P}_t(f; g) \leq \frac{1}{k'''} e^{k'''t} \sum_{j,l} \beta(j-l) \|\nabla_j f\|_{\infty} \|\nabla_l g\|_{\infty},$$

ce qui est le résultat annoncé (équation (3.40)).

Il reste à vérifier que le poids β décroît bien exponentiellement. Comme α est à décroissance exponentielle, il en est de même de $\sqrt{\alpha}$. Si c_{α}, d_{α} sont deux constantes telles que $\alpha(i) \leq c_{\alpha} \exp(-d_{\alpha} |i|)$, on a :

$$\beta(i) = \sum_j \sqrt{\alpha(j)\alpha(i-j)} \leq c_{\alpha} \sum_j \exp\left(-\frac{d_{\alpha}}{2} (|i-j| + |j|)\right).$$

Pour tout ϵ , on peut écrire :

$$\begin{aligned} |i-j| + |j| &\geq (1-\epsilon)(|i-j| + |j|) + \epsilon|j| \\ &\geq (1-\epsilon)|i| + \epsilon|j| \end{aligned}$$

En sommant sur j , on a donc :

$$\begin{aligned} \beta(i) &\leq \exp\left(-\frac{d_\alpha(1-\epsilon)}{2}|i|\right) \left(\sum_j e^{-\frac{\epsilon d_\alpha}{2}|j|}\right) \\ &\leq C(\alpha, \epsilon) \exp\left(-\frac{d_\alpha(1-\epsilon)}{2}|i|\right). \end{aligned}$$

Le poids β décroît donc exponentiellement. \square

Preuve du mélange pour les sites distants

Pour montrer le théorème 55, on procède en deux temps. La première étape consiste à montrer la propriété de mélange dans le cas où le site frontière dont on regarde l'influence est relativement loin du support de la fonction.

59 Théorème. *On suppose qu'il existe $C_{3.35}$ telle que :*

$$\forall \mathbf{z}, \forall \Lambda, \quad \gamma_{SG}(\Lambda, \mathbf{z}) \leq C_{3.35} \psi(\text{diam } \Lambda) \quad (3.42)$$

où $\psi(t) = t / \log t$. Alors il existe $C_{3.43}$ et une constante C telles que, pour tout Λ , $f \in \mathcal{C}_\Lambda$, $i \in \partial\Lambda$ vérifiant $\text{diam } \Lambda \leq 3d(i, S_f)$, et pour toutes conditions \mathbf{z} et $\tilde{\mathbf{z}}$ ne différant qu'en i ,

$$|\mathbf{E}^{\Lambda, \mathbf{z}}(f) - \mathbf{E}^{\Lambda, \tilde{\mathbf{z}}}(f)| \leq C_{3.43} |z_i - \tilde{z}_i| \cdot \|f\| d(i, S_f)^{-\frac{C}{C_{3.35}}}. \quad (3.43)$$

Remarque. *La constante C est explicite en fonction de l'interaction (cf. la discussion qui suit le théorème 55).*

Démonstration. Soit donc une boîte Λ , une fonction f à support dans Λ , et i un site frontière. On note ρ la distance $d(i, S_f)$, et on suppose $3\rho > \text{diam } \Lambda$.

En interpolant entre les conditions aux bords \mathbf{z} et $\tilde{\mathbf{z}}$ par $\mathbf{z}_\theta = \theta\mathbf{z} + (1-\theta)\tilde{\mathbf{z}}$, on prouve :

$$|\mathbf{E}^{\Lambda, \mathbf{z}} f - \mathbf{E}^{\Lambda, \tilde{\mathbf{z}}} f| \leq |z_i - \tilde{z}_i| \int_0^1 d\theta \sum_{j \in \Lambda} |J_{i,j}| \cdot |\mathbf{E}^{\Lambda, \mathbf{z}_\theta}(f; x_j)|. \quad (3.44)$$

On se ramène donc à étudier la covariance de deux fonctions. On procède alors comme dans [Yos01] (section 3.2). Pour toute fonction f , notons \tilde{f} la fonction recentrée $f - \mathbf{E}^{\Lambda, \mathbf{z}} f$. On introduit le semi-groupe, pour lequel $\mathbf{E}^{\Lambda, \mathbf{z}}$ est invariante, puis on fait apparaître une covariance :

$$\begin{aligned} \mathbf{E}^{\Lambda, \mathbf{z}}(f; g) &= \mathbf{E}^{\Lambda, \mathbf{z}}(\tilde{f} \tilde{g}) \\ &= \mathbf{E}^{\Lambda, \mathbf{z}}(\mathbf{P}_t(\tilde{f} \tilde{g})) && \text{(invariance)} \\ &= \underbrace{\mathbf{E}^{\Lambda, \mathbf{z}} \mathbf{P}_t(\tilde{f} \tilde{g}) - \mathbf{E}^{\Lambda, \mathbf{z}}(\mathbf{P}_t \tilde{f} \cdot \mathbf{P}_t \tilde{g})}_{J_1} + \underbrace{\mathbf{E}^{\Lambda, \mathbf{z}}(\mathbf{P}_t \tilde{f} \cdot \mathbf{P}_t \tilde{g})}_{J_2}. \end{aligned}$$

Le terme J_1 se majore grâce à la finitude de la vitesse de propagation. On majore d'abord ce qui apparaît dans l'espérance :

$$\begin{aligned} \mathbf{P}_t(\bar{f}\bar{g}) - \mathbf{P}_t\bar{f}\mathbf{P}_t\bar{g} &= \mathbf{P}_t(\bar{f}; \bar{g}) \\ &\leq C_{3.40} e^{kt} \sum_{i,j} \beta(i-j) \|\nabla_i \bar{f}\|_\infty \cdot \|\nabla_j \bar{g}\|_\infty \\ &\leq C_{3.40} e^{kt} \beta(d(S_f, S_g)) \|\bar{f}\| \cdot \|\bar{g}\| \end{aligned}$$

où dans la dernière ligne, $\tilde{\beta}(n) = \sup\{\beta(i), |i| \geq n\}$. Le terme de droite ne dépendant plus du point de départ, la même majoration vaut pour J_1 .

Pour majorer J_2 , on fait appel à l'inégalité de trou spectral donnée par les hypothèses. Pour cela, on remarque que :

$$\mathbf{P}_t\bar{f} = \mathbf{P}_t f - \mathbf{E}^{\Lambda, \mathbf{z}}(f)$$

et que l'on a donc :

$$\begin{aligned} J_2 &= \mathbf{E}^{\Lambda, \mathbf{z}}(\mathbf{P}_t\bar{f} \cdot \mathbf{P}_t\bar{g}) \\ &= \mathbf{E}^{\Lambda, \mathbf{z}}(\mathbf{P}_t f; \mathbf{P}_t g) \\ &\leq (\mathbf{E}^{\Lambda, \mathbf{z}}(\mathbf{P}_t f; \mathbf{P}_t f) \times \mathbf{E}^{\Lambda, \mathbf{z}}(\mathbf{P}_t g; \mathbf{P}_t g))^{\frac{1}{2}} \end{aligned}$$

Les variances qui apparaissent sont majorées par l'inégalité de trou spectral :

$$\mathbf{E}^{\Lambda, \mathbf{z}}(\mathbf{P}_t f; \mathbf{P}_t f) \leq \exp\left(-\frac{t}{2\gamma_{SG}^{\Lambda, \mathbf{z}}}\right) \mathbf{E}^{\Lambda, \mathbf{z}}(f; f)$$

En majorant les variances de f et g par les normes triples (grâce à la proposition 57), on obtient :

$$J_2 \leq C_{3.38} \exp\left(-\frac{t}{2\gamma_{SG}^{\Lambda, \mathbf{z}}}\right) \|f\| \|g\|.$$

On choisit maintenant le temps t pour minimiser $J_1 + J_2$, en fixant :

$$t = \frac{1}{k + 1/(2\gamma)} \log\left(\frac{C_{3.40}}{2C_{3.38} k \tilde{\beta}(d(S_f, S_g))}\right),$$

où k est la constante apparaissant dans le contrôle de la vitesse de propagation. Cette valeur donne la majoration suivante :

$$J_1 + J_2 \leq 2 \|f\| \|g\| (C_{3.38} \tilde{\beta}(d(S_f, S_g)))^{\frac{1}{1+2\gamma k}} \left(\frac{C_{3.40}}{2k\gamma}\right)^{\frac{2\gamma k}{1+2\gamma k}}.$$

On peut toujours supposer que les constantes γ_{SG} sont supérieures à 1 (quitte à les remplacer par $\gamma \vee 1$). Toute la deuxième parenthèse peut alors être majorée par une constante (indépendante de Λ et \mathbf{z}). On obtient donc, pour une certaine constante $C_{3.45}$:

$$\mathbf{E}^{\Lambda, \mathbf{z}}(f; g) \leq C_{3.45} \|f\| \|g\| (\tilde{\beta}(d(S_f, S_g)))^{\frac{1}{1+2k\gamma(\Lambda, \mathbf{z})}}. \quad (3.45)$$

Simplifions encore un peu l'exposant. Comme le poids β (qui décroît exponentiellement) peut être supposé inférieur à 1, il en est de même de $\tilde{\beta}$, et il faut donc minorer l'exposant, ce qui

revient à majorer $1 + 2k\gamma$. Quitte à remplacer k par $k' > k$, et à supposer γ suffisamment grand, on peut supposer : $\frac{1}{1+2k\gamma} \leq \frac{1}{2k'\gamma}$, ce qui donne :

$$\mathbf{E}^{\Lambda, \mathbf{z}}(f; g) \leq C_{3.45} \|f\| \|g\| \left(\tilde{\beta}(d(S_f, S_g)) \right)^{\frac{1}{2k'\gamma(\Lambda, \mathbf{z})}}.$$

Revenons alors à (3.44) : on va utiliser le dernier résultat avec $\mathbf{z} = \mathbf{z}_\theta$, et $g = x_j$ pour les sites j tels que $J(i - j)$ est non-nul. La distance $d(S_f, S_{x_j})$ est donc supérieure à $d(i, S_f) - R$, où R est la portée de l'interaction. D'après les hypothèses, on a un encadrement de $\text{diam } \Lambda$ en fonction de $\rho = d(S_f, i)$, et on peut en déduire une majoration de $\gamma(\Lambda, \mathbf{z})$ en fonction de ρ :

$$\begin{aligned} \rho &\leq \text{diam } \Lambda \leq 3\rho \\ \gamma &\leq C_{3.35} \frac{\text{diam } \Lambda}{\log \text{diam } \Lambda} \leq 3C_{3.35} \frac{\rho}{\log(\rho)}. \end{aligned} \quad (3.46)$$

On obtient donc :

$$|\mathbf{E}^{\Lambda, \mathbf{z}_\theta}(f; x_j)| \leq \|f\| \|g\| \left(\tilde{\beta}(\rho - R) \right)^{\frac{1}{2k'\gamma}} \quad (3.47)$$

Comme précédemment, on continue la majoration en minorant l'exposant, grâce à 3.46 :

$$|\mathbf{E}^{\Lambda, \mathbf{z}_\theta}(f; x_j)| \leq \|f\| \|g\| \left(\tilde{\beta}(\rho - R) \right)^{\frac{\log(\rho)}{6k'C_{3.35}\rho}}.$$

Enfin, on note c_β, d_β deux constantes telles que $\tilde{\beta}(n) \leq c_\beta \exp(-d_\beta n)$, et on obtient :

$$\begin{aligned} |\mathbf{E}^{\Lambda, \mathbf{z}_\theta}(f; x_j)| &\leq c_\beta \|f\| \|g\| \exp\left(-d_\beta \frac{(\rho - R) \log(\rho)}{6k'C_{3.35}\rho}\right) \\ &\leq c_\beta \|f\| \|g\| \exp\left(-d_\beta \frac{\log(\rho)}{6k'C_{3.35}} + \frac{d_\beta R}{6k'C_{3.35}} \times \frac{\log(\rho)}{\rho}\right). \end{aligned}$$

Comme $\log(\rho)/\rho$ est borné, on peut, quitte à intégrer le dernier terme dans une constante $C_{3.48}$, obtenir :

$$\begin{aligned} |\mathbf{E}^{\Lambda, \mathbf{z}_\theta}(f; x_j)| &\leq C_{3.48} \|f\| \|g\| \exp\left(-\frac{d_\beta \log(\rho)}{6k'C_{3.35}}\right) \\ &= C_{3.48} \|f\| \|g\| \rho^{-\frac{d_\beta}{6k'C_{3.35}}} \end{aligned} \quad (3.48)$$

Il ne reste plus qu'à reporter cette majoration dans l'interpolation (3.44) pour conclure. \square

Preuve générale du mélange

On vient de montrer que si les constantes de trou spectral n'augmentent pas trop vite, l'influence du bord sur les sites distants (approximativement, éloignés de plus du tiers de la boîte) décroît polynomialement. L'objectif est maintenant de généraliser ceci à tous les sites, y compris ceux relativement proches : on montre donc ici le théorème 55.

On considère donc Λ , f vérifiant les conditions du théorème, et i un site frontière. L'idée est d'appliquer la forme restreinte montrée dans le paragraphe précédent à une boîte et une fonction auxiliaires. Pour cela notons Γ le plus grand cube centré en i et ne rencontrant pas S_f (voir la figure 3.2). Ce cube Γ a donc pour rayon $\rho = d(i, S_f)$ (on a pris pour $d(\cdot, \cdot)$ la distance donnée par la norme sup).

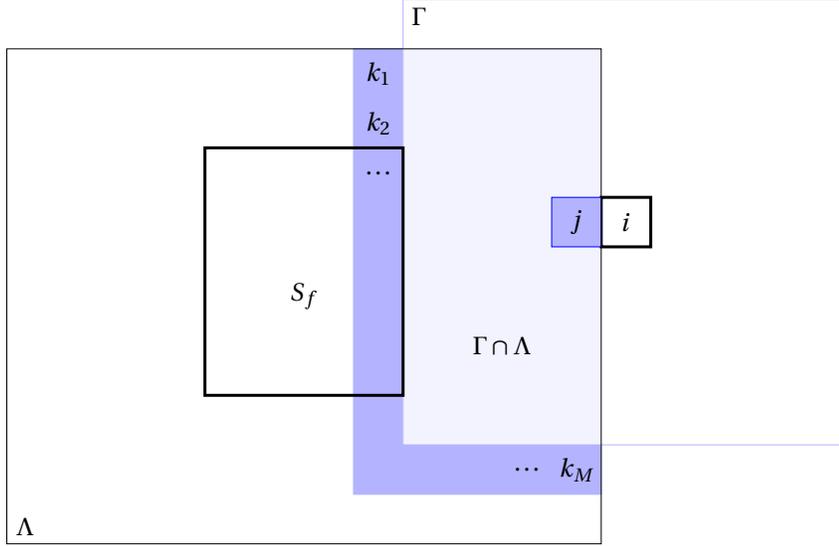


FIG. 3.2: Les différentes boîtes considérées. Plutôt que de regarder l'influence du spin en i sur la moyenne de f dans la boîte Λ , on regarde l'influence des spins aux points $k = k_1, k_2, \dots, k_M$ sur les voisins de i (ici, j), dans la boîte $\Lambda \cap \Gamma$.

Considérons maintenant un point k situé à la frontière extérieure $\partial\Gamma$, et dans la boîte Λ , et un point j voisin de i , appartenant également à Λ . La distance entre j et k est donc supérieure à $\rho - R$, et on a :

$$\begin{aligned} \text{diam}(\Lambda \cap \Gamma) &\leq \text{diam}(\Gamma) \\ &\leq 2\rho + 1 \\ &\leq 2R + 1 + 2d(j, k). \end{aligned}$$

Remarquons que l'on peut toujours supposer que $\rho = d(i, k)$ est suffisamment grand, en particulier supérieur à $3R + 1$ — en effet, la propriété de mélange peut être montrée facilement pour les ρ inférieurs à une borne fixée (il suffit d'utiliser l'interpolation (3.44), l'inégalité de Cauchy-Schwarz, puis la majoration uniforme de la variance d'un spin pour l'obtenir). Si ρ vérifie cette inégalité, alors $d(j, k) \geq 2R + 1$, et par conséquent :

$$\text{diam}(\Lambda \cap \Gamma) \leq 3d(j, k).$$

On peut donc appliquer le théorème 59 aux sites j et k dans la boîte $\Lambda \cap \Gamma$.

Passons au calcul proprement dit. En reprenant l'argument d'interpolation précédent, on écrit :

$$|\mathbf{E}^{\Lambda, z} f - \mathbf{E}^{\Lambda, \tilde{z}} f| \leq |z_i - \tilde{z}_i| \int_0^1 d\theta \sum_{j \in \Lambda} |J_{i,j}| \cdot |\mathbf{E}^{\Lambda, z_\theta}(f; x_j)|. \quad (3.49)$$

Notons g_j la fonction qui à une configuration associe sa valeur en j . On réécrit la covariance avec des intégrales :

$$\mathbf{E}^{\Lambda, z_\theta}(f; g_j) = \int (f(\mathbf{x}) - \mathbf{E}^{\Lambda, z_\theta}(f)) (g_j(\mathbf{x}) - \mathbf{E}^{\Lambda, z_\theta}(g_j)) d\mu_{\Lambda, z_\theta}(\mathbf{x}).$$

On utilise maintenant la propriété de Gibbs sur la sous-boîte $\Lambda \cap \Gamma$ de Λ , en notant $(\mathbf{x}'^{\Lambda \cap \Gamma}, \mathbf{x})$ la configuration qui coïncide avec \mathbf{x}' sur $\Lambda \cap \Gamma$, et avec \mathbf{x} sur le reste de la boîte Λ .

$$\begin{aligned} \mathbf{E}^{\Lambda, \mathbf{z}_0}(f; g_j) &= \iint \left(f((\mathbf{x}'^{\Lambda \cap \Gamma}, \mathbf{x})) - \mathbf{E}^{\Lambda, \mathbf{z}_0}(f) \right) \left(g_j((\mathbf{x}'^{\Lambda \cap \Gamma}, \mathbf{x})) - \mathbf{E}^{\Lambda, \mathbf{z}_0}(g_j) \right) d\mu_{\Lambda \cap \Gamma, \mathbf{x}}(\mathbf{x}') d\mu_{\Lambda, \mathbf{z}_0}(\mathbf{x}) \\ &= \int (f(\mathbf{x}) - \mathbf{E}^{\Lambda, \mathbf{z}_0}(f)) \left(\int g_j(\mathbf{x}') d\mu_{\Lambda \cap \Gamma, \mathbf{x}}(\mathbf{x}') - \mathbf{E}^{\Lambda, \mathbf{z}_0}(g_j) \right) d\mu_{\Lambda, \mathbf{z}_0}(\mathbf{x}), \end{aligned} \quad (3.50)$$

où l'on a utilisé le fait que g_j ne dépendait que de \mathbf{x}' , et f uniquement de \mathbf{x} .

Dans le terme concernant g_j , on réexprime la moyenne de g_j en réutilisant la propriété de Gibbs :

$$\begin{aligned} \int g_j(\mathbf{x}') d\mu_{\Lambda \cap \Gamma, \mathbf{x}}(\mathbf{x}') - \mathbf{E}^{\Lambda, \mathbf{z}_0}(g_j) &= \int g_j(\mathbf{x}') d\mu_{\Lambda \cap \Gamma, \mathbf{x}}(\mathbf{x}') - \iint g_j(\mathbf{x}') d\mu_{\Lambda \cap \Gamma, \mathbf{x}''}(\mathbf{x}') d\mu_{\Lambda, \mathbf{z}_0}(\mathbf{x}'') \\ &= \int \left(\mathbf{E}^{\Lambda \cap \Gamma, \mathbf{x}}(g_j) - \mathbf{E}^{\Lambda \cap \Gamma, \mathbf{x}''}(g_j) \right) d\mu_{\Lambda, \mathbf{z}_0}(\mathbf{x}''). \end{aligned}$$

Pour majorer la différence d'espérances qui apparaît, on voudrait maintenant appliquer le résultat précédent, en prenant pour fonction g_j et pour boîte $\Lambda \cap \Gamma$. Comme ce résultat demande que les conditions aux bords ne diffèrent qu'en un site, on construit une suite de configurations pour passer de \mathbf{x} à \mathbf{x}'' , de la même façon que dans la preuve de la proposition 57. Les seuls sites à prendre en considération sont ceux situés sur $\partial \Gamma \cap \Lambda$, puisque les autres n'ont pas d'influence sur ce qui se passe dans $\Gamma \cap \Lambda$.

Plus précisément, on note k_1, k_2, \dots, k_M une énumération de $\partial \Gamma \cap \Lambda$, et on construit une suite \mathbf{x}^l de configurations en posant : $\mathbf{x}^0 = \mathbf{x}$, et pour l de 1 à M ,

$$x_k^l = \begin{cases} x_k'' & \text{si } k = k_{l'}, \text{ avec } l' \leq l \\ x_k & \text{sinon.} \end{cases}$$

On peut commencer les majorations. Notons $\tilde{\phi}(\cdot) = (\cdot)^{-\frac{C}{c_{3.35}}}$ la fonction puissance apparaissant dans le théorème 59.

$$\begin{aligned} \left| \mathbf{E}^{\Lambda \cap \Gamma, \mathbf{x}}(g_j) - \mathbf{E}^{\Lambda \cap \Gamma, \mathbf{x}''}(g_j) \right| &\leq \sum_{l=1}^M \left| \mathbf{E}^{\Lambda \cap \Gamma, \mathbf{x}^l}(g_j) - \mathbf{E}^{\Lambda \cap \Gamma, \mathbf{x}^{l-1}}(g_j) \right| \\ &\leq \sum_{l=1}^M \left| x_{k_l}^l - x_{k_l}^{l-1} \right| \tilde{\phi}(d(j, k_l)) \quad (\text{Th. 59, avec } \|g_j\| = 1) \\ &\leq \sum_{l=1}^M \left| x_{k_l} - x_{k_l}'' \right| \tilde{\phi}(\rho - R) \end{aligned}$$

Finalement,

$$\left| \int g_j(\mathbf{x}') d\mu_{\Lambda \cap \Gamma, \mathbf{x}}(\mathbf{x}') - \mathbf{E}^{\Lambda, \mathbf{z}_0}(g_j) \right| \leq \tilde{\phi}(\rho - R) \int \sum_{l=1}^M \left| x_{k_l} - x_{k_l}'' \right| d\mu_{\Lambda, \mathbf{z}_0}(\mathbf{x}'')$$

On reporte ceci dans la majoration (3.50) de la covariance :

$$\begin{aligned} \left| \mathbf{E}^{\Lambda, \mathbf{z}_0}(f; g_j) \right| &\leq \int |f(\mathbf{x}) - \mathbf{E}^{\Lambda, \mathbf{z}_0}(f)| \times \left(\int |g_j(\mathbf{x}') d\mu_{\Lambda \cap \Gamma, \mathbf{x}}(\mathbf{x}') - \mathbf{E}^{\Lambda, \mathbf{z}_0}(g_j)| \right) d\mu_{\Lambda, \mathbf{z}_0}(\mathbf{x}) \\ &\leq \sum_{l=1}^M \tilde{\phi}(\rho - R) \int |f(\mathbf{x}) - \mathbf{E}^{\Lambda, \mathbf{z}_0}(f)| \times \int \left| x_{k_l} - x_{k_l}'' \right| d\mu_{\Lambda, \mathbf{z}_0}(\mathbf{x}'') d\mu_{\Lambda, \mathbf{z}_0}(\mathbf{x}) \end{aligned}$$

Il ne reste alors plus qu'à appliquer l'inégalité de Cauchy-Schwarz à la mesure produit $\mu_{\Lambda, \mathbf{z}_0} \otimes \mu_{\Lambda, \mathbf{z}_0}$, en se souvenant de la proposition 57 pour majorer les variances par les normes triples des fonctions :

$$\begin{aligned} |\mathbf{E}^{\Lambda, \mathbf{z}_0}(f; g_j)| &\leq 2\tilde{\phi}(\rho - R) \sum_{i=1}^M \|f\| \|\mathbf{E}^{\Lambda, \mathbf{z}_0}(g_{k_i}; g_{k_i})\|^{1/2} \\ &\leq C_{3.51} M \tilde{\phi}(\rho - R) \|f\| \end{aligned} \quad (3.51)$$

Comme M est inférieur au cardinal de $\partial\Gamma \cap \Lambda$, on peut le majorer par $C_d \rho^{(d-1)}$ (puisque ρ est le rayon de Γ). On reporte la dernière majoration dans l'expression (3.49), et on utilise l'expression explicite de $\tilde{\phi}$ pour conclure la preuve du théorème 55.

3.4 Du mélange polynomial au trou spectral uniforme.

Le résultat

Dans les sections précédentes, on a vu comment passer d'inégalités fonctionnelles à des propriétés de mélange. Les résultats présents dans la littérature (et en particulier [Yos01]) nous indiquent que, sous certaines hypothèses, un mélange *exponentiel* est en fait équivalent à une inégalité fonctionnelle uniforme. Il est alors naturel de se demander si des propriétés de mélange polynomiales (comme celles obtenues dans la section précédente) entraînent à leur tour un contrôle dans les inégalités fonctionnelles.

On cherche donc à adapter la preuve du passage du mélange à une inégalité de Poincaré dans [Yos01] (qui s'appuie également sur [Yos99]) en relâchant la condition de mélange à une décroissance polynomiale.

On aura besoin de rajouter une hypothèse sur le potentiel.

Hypothèse 5. *L'auto-interaction V se décompose, pour tout m , en la somme d'une fonction convexe U telle que $U'' \geq m$, et d'une fonction lipschitzienne bornée.*

On a alors le résultat suivant :

60 Théorème. *Supposons que la condition suivante (condition de mélange) soit vérifiée : il existe $C_{3.52}$ et une fonction décroissante ϕ telle que, pour tout Λ , tout $i \in \Lambda$, $j \notin \Lambda$ et pour toutes conditions aux limites $\mathbf{z}, \bar{\mathbf{z}}$ ne différant qu'en j ,*

$$|\mathbf{E}^{\Lambda, \bar{\mathbf{z}}}(x_i) - \mathbf{E}^{\Lambda, \mathbf{z}}(x_i)| \leq C_{3.52} |z_j - \bar{z}_j| \phi(d(i, j)). \quad (3.52)$$

Alors, si ϕ décroît suffisamment vite (précisément, $\phi(k)$ doit être $O(k^{-\alpha})$, avec $\alpha > 3d/2$), les mesures de Gibbs en volume fini vérifient une inégalité de trou spectral avec une constante uniforme en Λ, \mathbf{z} .

Dans ce cas, grâce aux résultats de N. Yoshida ([Yos01]), le mélange est en fait exponentiel (on peut choisir ϕ à décroissance exponentielle), et les mesures de Gibbs vérifient également une inégalité de Sobolev logarithmique uniforme.

L'idée de la preuve, développée ci-dessous, est la suivante ([Yos99]) : on note γ_n la borne supérieure des trous spectraux inverses sur les boîtes de taille (volume) inférieure à n . La quantité γ_n croît avec n , on veut montrer qu'elle est bornée. Pour cela, on essaie de contrôler γ_{2n} à partir de γ_n : on obtiendra une inégalité du type

$$\gamma_{2n} \leq A\gamma_n + B, \quad (3.53)$$

avec $A < 1$. Ceci implique que γ_{2^n} est borné, et comme γ_n est croissante, cela suffit pour conclure.

Quelques notations, et une inégalité de Sobolev logarithmique

La preuve repose donc sur l'établissement de (3.53). Soit donc Λ une boîte quelconque, de taille inférieure à $2n$, et Λ_0 une sous boîte, telle que $\max\{|\Lambda_0|, |\Lambda \setminus \Lambda_0|\} \leq n$. On note i_1, \dots, i_m une énumération de $\Lambda \setminus \Lambda_0$, et on définit une suite de boîtes (Λ_l) pour passer de Λ_0 à Λ :

$$\Lambda_l = \Lambda_0 \cup \{i_1, \dots, i_l\}.$$

On aura également besoin de découper chacun de ces boîtes en plusieurs « couches » ; pour $l \leq m-1$, on définit donc les boîtes $\Lambda_{l,k}$, sous-boîtes de Λ_l , par :

$$\Lambda_{l,k} = \left\{ i \in \Lambda_l, d(i, i_{l+1}) < \frac{k}{2} \right\}.$$

Ainsi $\Lambda_{l,0} = \emptyset$, puis on gagne les sites de Λ_l voisins de i_{l+1} , puis leurs voisins, jusqu'à obtenir tout Λ_l . Le choix de $k/2$ assure que le diamètre de $\Lambda_{l,k}$ est au plus k . Enfin, on note $f_{l,k}$ l'espérance conditionnelle de f par rapport au complémentaire de la boîte $\Lambda_{l,k}$:

$$f_{l,k}(\mathbf{x}) = \int f(\mathbf{y}^{\Lambda_{l,k}} \mathbf{x}) d\mu_{\Lambda_{l,k}, \mathbf{x}}(\mathbf{y})$$

Ces notations étant fixées, notons ici, sans preuve, deux résultats tirés directement de [Yos99]. Le premier est une inégalité de Sobolev logarithmique, où la constante ne dépend que de la taille du support de la fonction.

61 Lemme. *Supposons que l'interaction vérifie l'hypothèse de convexité 5 (page 92).*

Alors il existe des constantes $C_{3.54}(n)$ telles que, pour toute boîte, toute condition aux bords, et toute fonction de \mathcal{C}_Λ ,

$$\mathbf{E}^{\Lambda, \mathbf{z}} \left(f^2 \log \frac{f^2}{\mathbf{E}^{\Lambda, \mathbf{z}}(f^2)} \right) \leq C_{3.54}(|S_f|) \mathbf{E}^{\Lambda, \mathbf{z}} |\nabla_\Lambda f|^2. \quad (3.54)$$

A fortiori, on a également une inégalité de Poincaré :

$$\mathbf{E}^{\Lambda, \mathbf{z}}(f; f) \leq C_{3.55}(|S_f|) \mathbf{E}^{\Lambda, \mathbf{z}} |\nabla_\Lambda f|^2. \quad (3.55)$$

Remarque. *L'hypothèse de convexité 5 paraît assez forte, et n'apparaît pas dans les articles de T. Bodineau et B. Helffer consacrés à ce modèle. Ceci s'explique relativement facilement. En effet, la preuve du lemme fait intervenir une comparaison entre l'interaction J et la borne de convexité de l'auto-interaction. Si l'on se place en régime perturbatif comme le font Bodineau et Helffer, on peut rendre J aussi petit que l'on veut devant cette borne, et obtenir le résultat. N. Yoshida s'intéresse en revanche au régime non-perturbatif : comme il ne peut pas faire d'hypothèse sur J , c'est sur la borne de convexité qu'il doit en faire une.*

On pourra consulter à ce sujet le lemme 9.2.1 et la remarque qui le suit dans le livre d'Helffer [Hel02].

Grâce à cela, on peut montrer le lemme suivant (c'est le lemme 4.1 de [Yos99]), pour commencer la majoration de la variance d'une fonction.

62 Lemme. *Il existe $C_{3.56}$ telle que pour toute f ,*

$$\mathbf{E}^\Lambda(f; f) \leq \gamma_n \mathbf{E}^\Lambda(|\nabla_{\Lambda_0} f|^2) + C_{3.56} \mathbf{E}^\Lambda(|\nabla_{\Lambda \setminus \Lambda_0} f|^2) + C_{3.56} \sum_0^{m-1} \mathbf{E}^\Lambda \mathcal{D}_l(f), \quad (3.56)$$

où $\mathcal{D}_l(f) \stackrel{\text{déf}}{=} \sup \{(\mathbf{E}^{\Lambda_l}(f; x_i))^2, i \in \Lambda_l, d(i, i_{l+1}) \leq R\}$.

Pour étudier les différents termes de cette décomposition, et en particulier les termes \mathcal{D}_l , à ce stade, on va faire intervenir les corrélations, ce qui nécessite donc quelques adaptations.

Fin de la preuve

La condition de mélange (l'hypothèse du théorème 60) sert à montrer le résultat suivant : pour un site i éloigné du support de la fonction f , la covariance entre f et x_i est contrôlée par la variance de f .

63 Lemme. *Sous la condition de mélange (3.52), il existe une constante $C_{3,57}$ telle que pour toute f cylindrique, toute boîte $S_f \cap \Lambda \subset \Delta \subset \Lambda$, et toute condition aux bords \mathbf{z} ,*

$$|\mathbf{E}^{\Lambda, \mathbf{z}}(f; x_i)| \leq C_{3,57} |\Delta| \phi(d(i, \Delta)) \sqrt{\mathbf{E}^{\Lambda, \mathbf{z}}(f; f)}. \quad (3.57)$$

Démonstration. Le passage correspondant dans [Yos99] est la preuve de l'inégalité 3.19. Si l'on observe cette preuve, on se rend compte que la vitesse de mélange est transférée directement en vitesse de décroissance pour les corrélations. On peut donc adapter la preuve en remplaçant la vitesse exponentielle par une vitesse quelconque ϕ . \square

L'étape suivante est le contrôle des termes \mathcal{D}_l apparaissant dans (3.56).

64 Lemme. *Il existe une constante $C_{3,58}$ telle que :*

$$\mathcal{D}_l(f) \leq C_{3,58} \sum_{k=0}^{\infty} (k+1)^{2d+\varepsilon-1} \phi^2(k/2 - R) \mathbf{E}^{\Lambda_l} (\mathbf{E}^{\Lambda_{l,k+1}}(f; f)^{1/2})^2. \quad (3.58)$$

Démonstration. On veut majorer $\mathbf{E}^{\Lambda_l}(f; x_i)^2$, où i est un site de Λ vérifiant $d(i, l_{l+1}) \leq R$. On a :

$$\begin{aligned} \mathbf{E}^{\Lambda_l}(f; x_i)^2 &= (\mathbf{E}^{\Lambda_l}((f - f_l)x_i))^2 \\ &= \left(\sum_{k=0}^{\infty} \mathbf{E}^{\Lambda_l}((f_{l,k} - f_{l,k+1})x_i) \right)^2, \end{aligned} \quad (3.59)$$

en écrivant $f - f_l$ comme somme télescopique (pour $k=0$, $f_{l,k} = f$, et pour k grand, $f_{l,k} = f_l$). Ensuite, comme par définition, $f_{l,k} = \mathbf{E}^{\Lambda_{l,k}} f$, et comme $\Lambda_{l,k} \subset \Lambda_{l,k+1}$, on a $\mathbf{E}^{\Lambda_{l,k+1}}(f_{l,k}) = f_{l,k+1}$, donc :

$$\begin{aligned} \mathbf{E}^{\Lambda_l}((f_{l,k} - f_{l,k+1})x_i) &= \mathbf{E}^{\Lambda_l}((f_{l,k} - \mathbf{E}^{\Lambda_{l,k+1}}(f_{l,k}))x_i) \\ &= \mathbf{E}^{\Lambda_l}(\mathbf{E}^{\Lambda_{l,k+1}}((f_{l,k} - \mathbf{E}^{\Lambda_{l,k+1}}(f_{l,k}))x_i)) \\ &= \mathbf{E}^{\Lambda_l}(\mathbf{E}^{\Lambda_{l,k+1}}(f_{l,k}; x_i)). \end{aligned}$$

En reportant cette expression dans (3.59), puis en utilisant l'inégalité de Hölder, on a pour tout ε l'existence d'une constante C_ε telle que :

$$\begin{aligned} \mathbf{E}^{\Lambda_l}(f; x_i)^2 &= \left(\sum_0^{\infty} \left(\frac{1+k}{1+k} \right)^{\frac{1+\varepsilon}{2}} \mathbf{E}^{\Lambda_l} \mathbf{E}^{\Lambda_{l,k+1}}(f_{l,k}; x_i) \right)^2 \\ &\leq C_\varepsilon \sum_0^{\infty} (k+1)^{1+\varepsilon} (\mathbf{E}^{\Lambda_l} \mathbf{E}^{\Lambda_{l,k+1}}(f_{l,k}; x_i))^2 \end{aligned} \quad (3.60)$$

On va ensuite majorer les covariances $\mathbf{E}^{\Lambda_{l,k+1}}(f_{l,k}; x_i)$ par des variances, en appliquant (3.57), avec la boîte $\Lambda_{l,k+1}$, la fonction $f_{l,k}$ et la boîte :

$$\Delta = \Lambda_{l,k+1} \cap S_{f_{l,k}} \subset \Lambda_{l,k+1} \setminus \Lambda_{l,k}$$

On a donc $|\Delta| \leq C_d(k+1)^{d-1}$. De plus, la distance de i à Δ est au moins $k/2 - R$ (par définition de Δ et car $d(i, i_{l+1}) \leq R$). On obtient finalement :

$$|\mathbf{E}^{\Lambda_{l,k+1}}(f_{l,k}; x_i)| \leq C_{3.61}(k+1)^{d-1} \phi(k/2 - R) \mathbf{E}^{\Lambda_{l,k+1}}(f_{l,k}; f_{l,k})^{1/2}. \quad (3.61)$$

On peut majorer $\mathbf{E}^{\Lambda_{l,k+1}}(f_{l,k}; f_{l,k})$ par $\mathbf{E}^{\Lambda_{l,k+1}}(f; f)$, grâce à l'inégalité de Jensen :

$$\mathbf{E}^{\Lambda_{l,k+1}}(f_{l,k}^2) = \mathbf{E}^{\Lambda_{l,k+1}}(\mathbf{E}^{\Lambda_{l,k}}(f^2)) \leq \mathbf{E}^{\Lambda_{l,k+1}} \mathbf{E}^{\Lambda_{l,k}}(f^2) = \mathbf{E}^{\Lambda_{l,k+1}}(f^2)$$

On reporte ceci dans l'inégalité de départ :

$$\mathbf{E}^{\Lambda_l}(f; x_i)^2 \leq C_{d,\varepsilon} \sum_{k=0}^{\infty} (k+1)^{\varepsilon+2d-1} \phi(k/2 - R)^2 (\mathbf{E}^{\Lambda_l}(\mathbf{E}^{\Lambda_{l,k+1}}(f; f)^{1/2}))^2.$$

Comme \mathcal{D}_l est le maximum de ces quantités sur les sites i , et que le membre de droite ne dépend pas de i , la preuve est finie. \square

Revenons au problème de départ : il s'agit de majorer $\mathbf{E}^{\Lambda}(f; f)$. On va donc maintenant reporter la majoration précédente de \mathcal{D}_l dans l'inégalité (3.56).

65 Lemme. *Posons $k_1 = \lfloor n^{\frac{1}{d}} \rfloor$. Il existe une constante $C_{3.62}$, une fonction Ψ et, pour tout $k_0 \leq k_1$, une constante $D_{3.62}(k_0)$ telles que :*

$$\begin{aligned} \mathbf{E}^{\Lambda}(f; f) &\leq \gamma_n \mathbf{E}^{\Lambda}(|\nabla_{\Lambda_0} f|^2) + \gamma_n C_{3.62} \Psi(k_0) \mathbf{E}^{\Lambda}(|\nabla_{\Lambda} f|^2) \\ &\quad + D_{3.62}(k_0) \mathbf{E}^{\Lambda}(|\nabla_{\Lambda} f|^2) + C_{3.62} \Psi(k_1) \mathbf{E}^{\Lambda}(f; f). \end{aligned} \quad (3.62)$$

La fonction Ψ est définie par :

$$\Psi(k) = \sum_{l=k}^{\infty} \phi\left(\frac{l}{2} - R\right)^2 (1+l)^{3d+\varepsilon-1}.$$

Démonstration. On suit la preuve du lemme 4.3 de [Yos99]. Rappelons tout d'abord l'inégalité (3.56) :

$$\mathbf{E}^{\Lambda}(f; f) \leq \gamma_n \mathbf{E}^{\Lambda}(|\nabla_{\Lambda_0} f|^2) + C_{3.56} \mathbf{E}^{\Lambda}(|\nabla_{\Lambda \setminus \Lambda_0} f|^2) + C_{3.56} \sum_{l=0}^{m-1} \mathbf{E}^{\Lambda} \mathcal{D}_l(f),$$

Vu ce que l'on cherche à établir, il nous faut majorer les deux derniers termes. Le second est contrôlé, à une constante près, par $\mathbf{E}^{\Lambda}(|\nabla_{\Lambda} f|^2)$: il sera intégré au terme contenant $D_{3.62}(k_0)$ dans l'inégalité (3.62).

Il suffit donc de majorer le troisième terme. Comme le lemme 64 garantit

$$\sum_{l=0}^{m-1} \mathbf{E}^{\Lambda} \mathcal{D}_l(f) \leq C_{3.58} \sum_{k=0}^{\infty} \psi(k) \sum_{l=0}^{m-1} \mathbf{E}^{\Lambda} \mathbf{E}^{\Lambda_{l,k+1}}(f; f),$$

où $\psi(k) = \phi(k/2 - R)^2 (k+1)^{2d+\varepsilon-1}$, il suffit pour conclure de montrer :

$$\sum_{k=0}^{\infty} \psi(k) \sum_{l=0}^{m-1} \mathbf{E}^{\Lambda} \mathbf{E}^{\Lambda_{l,k+1}}(f; f) \leq D(k_0) \mathbf{E}^{\Lambda}(|\nabla f|^2) + \gamma_n C \Psi(k_0) \mathbf{E}^{\Lambda}(|\nabla f|^2) + C \Psi(k_1) \mathbf{E}^{\Lambda}(f; f). \quad (3.63)$$

En effet, on pourra alors poser $C_{3.62} = C_{3.56} C_{3.58} C$, et $D_{3.62}(k_0) = C_{3.56}(1 + C_{3.58} D(k_0))$, pour obtenir le résultat voulu.

On s'intéresse donc au membre de gauche de (3.63), et on veut majorer $\mathbf{E}^{\Lambda_{l,k+1}}(f; f)$, qui est une variance dans une boîte de diamètre proche de k . On sépare la somme en trois, en coupant en deux indices k_0 et k_1 , avec k_1 comme dans l'énoncé, et $k_0 \leq k_1$ que nous fixerons plus tard.

Pour la première partie de la somme, la variance est prise dans une « petite » boîte. On la majore par une inégalité de Poincaré : d'après le lemme 61, il existe une constante $C(k_0)$ telle que, pour toutes les boîtes $\Lambda_{l,k+1}$ ($k \leq k_0$) et toutes les conditions aux bords, $\mathbf{E}^{\Lambda_{l,k+1}}(f; f) \leq C(k_0)\mathbf{E}^{\Lambda_{l,k+1}}(|\nabla f|^2)$.

$$\begin{aligned} \sum_{k=0}^{k_0-1} \psi(k) \sum_{l=0}^{m-1} \mathbf{E}^{\Lambda} \mathbf{E}^{\Lambda_{l,k+1}}(f; f) &\leq \sum_{k=0}^{k_0-1} \psi(k) \sum_{j=0}^{m-1} C(k_0) \mathbf{E}^{\Lambda} \left(|\nabla_{\Lambda_{l,k+1}} f|^2 \right) \\ &= C(k_0) \sum_{k=0}^{k_0-1} \psi(k) \sum_{l=0}^{m-1} \sum_{i \in \Lambda} \mathbf{1}_{i \in \Lambda_{l,k+1}} \mathbf{E}^{\Lambda} \left(|\nabla_i f|^2 \right) \\ &\leq C(k_0) \sum_{i \in \Lambda} \mathbf{E}^{\Lambda} \left(|\nabla_i f|^2 \right) \sum_{k=0}^{k_0-1} \psi(k) \left(\sum_{l=0}^{m-1} \mathbf{1}_{i \in \Lambda_{l,k+1}} \right) \end{aligned}$$

Or dans la somme indiquée par l , $\mathbf{1}_{i \in \Lambda_{l,k+1}}$ ne vaut 1 que quand i est à une distance inférieure à $(k/2)$ d'un certain site (i_{l+2}). Autrement dit, à k donné, il n'y a pas plus de $(k+1)^d$ indices l tels que i puisse appartenir à $\Lambda_{l,k+1}$.

On en déduit :

$$\begin{aligned} \sum_{k=0}^{k_0-1} \psi(k) \sum_{l=0}^{m-1} \mathbf{E}^{\Lambda} \mathbf{E}^{\Lambda_{l,k+1}}(f; f) &\leq \left(\sum_{k=0}^{k_0-1} \psi(k) C(k_0) (k+1)^d \right) \times \sum_{\Lambda} \mathbf{E}^{\Lambda} \left(|\nabla_i f|^2 \right) \\ &\leq C_{3.64}(k_0) \mathbf{E}^{\Lambda} \left(|\nabla_{\Lambda} f|^2 \right), \end{aligned} \quad (3.64)$$

où $C_{3.64}(k_0) \stackrel{\text{def}}{=} C(k_0) \sum_{k=0}^{k_0-1} \psi(k) (k+1)^d$.

La deuxième partie de la somme ($k_0 \leq k \leq k_1 - 1$) concerne des boîtes de taille « moyenne ». On la traite de la même manière que la première somme, à deux détails près : la constante de Poincaré est maintenant majorée par γ_n (ce qui est légitime car $k+1 \leq k_1$, donc $(k+1)^d \leq n$), et la somme $\sum_{k=k_0}^{k_1} \psi(k) (k+1)^d$ par le reste de la série, qui est par définition $\Psi(k_0)$.

$$\sum_{k=k_0}^{k_1-1} \psi(k) \sum_{l=0}^{m-1} \mathbf{E}^{\Lambda} \mathbf{E}^{\Lambda_{l,k+1}}(f; f) \leq \gamma_n C_{3.65} \Psi(k_0) \mathbf{E}^{\Lambda} \left(|\nabla_{\Lambda} f|^2 \right). \quad (3.65)$$

Enfin, pour la troisième partie de la somme (sur les « grandes » boîtes), on majore directement $\mathbf{E}^{\Lambda} \mathbf{E}^{\Lambda_{l,k+1}}(f; f)$ par $\mathbf{E}^{\Lambda}(f; f)$, pour obtenir :

$$\sum_{k=k_1}^{\infty} \psi(k) \sum_{l=0}^{m-1} \mathbf{E}^{\Lambda} \mathbf{E}^{\Lambda_{l,k+1}}(f; f) \leq \Psi(k_1) \mathbf{E}^{\Lambda}(f; f).$$

En reportant les majorations des trois sommes dans (3.56), on obtient (3.63), ce qui achève la preuve. \square

Grâce à ce dernier lemme, nous allons pouvoir montrer la majoration voulue de γ_{2n} . On peut en effet l'appliquer à deux reprises, d'abord directement, ensuite en échangeant Λ_0 et $\Lambda \setminus \Lambda_0$. Si l'on ajoute les résultats, on obtient :

$$\begin{aligned} 2\mathbf{E}^{\Lambda}(f; f) &\leq \gamma_n \mathbf{E}^{\Lambda} \left(|\nabla_{\Lambda} f|^2 \right) + 2\gamma_n C_{3.62} \Psi(k_0) \mathbf{E}^{\Lambda} \left(|\nabla_{\Lambda} f|^2 \right) \\ &\quad + 2D_{3.62}(k_0) \mathbf{E}^{\Lambda} \left(|\nabla_{\Lambda} f|^2 \right) + 2C_{3.62} \Psi(k_1) \mathbf{E}^{\Lambda}(f; f). \end{aligned}$$

En passant $\mathbf{E}^\Lambda(f; f)$ à gauche et en divisant par 2, il vient :

$$\mathbf{E}^\Lambda(f; f) \leq \frac{\gamma_n/2 + 2\gamma_n C_{3.62} \Psi(k_0) + D_{3.62}(k_0)}{1 - C_{3.62} \Psi(k_1)} \mathbf{E}^\Lambda(|\nabla_\Lambda f|^2)$$

Remarquons maintenant que l'on peut choisir ϵ pour que $\Psi(k)$ tende vers 0 (par hypothèse, $\phi = O(k^{-\alpha})$, avec $\alpha > 3d/2$, donc pour ϵ petit, Ψ est le reste d'une série convergente). Prenons alors k_0 assez grand pour garantir : $(\frac{1}{2} + 2C_{3.62} \Psi(k_0)) \leq 3/5$. Rappelons que l'on a pris $k_1 = \lfloor n^{1/d} \rfloor$: pour n suffisamment grand, k_1 est supérieur au k_0 fixé, et on peut assurer :

$$1 - C_{3.62} \Psi(k_1) \geq \frac{4}{5}.$$

On en déduit l'existence d'une constante B telle que, pour n suffisamment grand,

$$\gamma_{2n} \leq \frac{3}{4} \gamma_n + B,$$

ce qui est l'inégalité (3.53) cherchée : on a donc bien un trou spectral uniforme.

Annexe A

Topologie et mesures sur l'espace des configurations

On rassemble ici les détails sur la topologie de Ω et sur les questions de convergence de mesures.

L'espace Ω est muni de la topologie produit ; celle-ci est métrisable, par exemple par :

$$d(\mathbf{x}, \mathbf{y}) = \sum \frac{1}{2^l} \frac{|x_l|}{1 + |x_l|},$$

où on a énuméré les sites de \mathbb{Z}^d dans un ordre quelconque. Cet espace topologique est de plus séparable et complet (c'est donc un espace Polonais).

Pour étudier les mesures sur cet espace, on le munit de la tribu borélienne $\mathcal{B}(\Omega)$. Cette tribu coïncide avec la tribu produit $\otimes \mathcal{B}(\mathbb{R})$ ([Kal02], lemme 1.2). Une mesure est donc déterminée par ses valeurs sur les cylindres élémentaires (les ensembles du type $\prod A_l$, où A_l est un intervalle ouvert, et $A_l = \mathbb{R}$ sauf sur un ensemble fini de sites).

Soit \mathcal{C} une classe de fonctions de \mathcal{S} dans \mathbb{R} . On dira que \mathcal{C} *détermine les mesures* si on a l'implication :

$$\forall f \in \mathcal{C}, \mu f = \nu f \implies \mu = \nu.$$

66 Proposition. *Soit α un poids sur les sites de \mathbb{Z}^d , partout positif, et $\|\cdot\|_\alpha$ la norme l^2 correspondante.*

La classe des fonctions cylindriques, bornées, et α -Lipschitziennes détermine les mesures.

Le même résultat est vrai si l'on remplace α -lipschitziennes par « lipschitziennes pour la norme l^1 sur leur support », ou par « de norme triple finie ».

Démonstration. Prouvons d'abord le premier résultat. Supposons que μ et ν coïncident sur la classe indiquée. Il suffit de montrer que pour tout cylindre élémentaire C , $\mu \mathbf{1}_C = \nu \mathbf{1}_C$. Soit i_1, \dots, i_m des sites de \mathcal{S} , $a_1, b_1, \dots, a_m, b_m$ des réels tels que :

$$\mathbf{x} \in C \iff \forall l \leq m, x_{i_l} \in]a_l, b_l[.$$

Pour tout $l \leq m$, on pose $g_l^1, g_l^2, \dots, g_l^n, \dots$ une suite de fonctions de $\mathcal{C}^\infty(\mathbb{R}, \mathbb{R})$ qui croît vers $\mathbf{1}_{]a_l, b_l[}$. On définit alors les fonctions g^n par :

$$g_n(\mathbf{x}) = \prod_{l=1}^m g_l^n(x_{i_l}). \tag{A.1}$$

La suite g_n croît vers $\mathbf{1}_C$, et g_n est cylindrique et bornée. Montrons qu'elle est α -Lipschitzienne. Comme les g_l^n sont \mathcal{C}^∞ à support compact, leur dérivée est bornée, disons par M_l . En passant de \mathbf{x} à \mathbf{y} en changeant un site après l'autre, et en utilisant le fait que $|g_l^n| \leq 1$, on obtient :

$$\begin{aligned} |g^n(\mathbf{x}) - g^n(\mathbf{y})| &\leq \sum_1^m M_l |x_{i_l} - y_{i_l}| \\ &\leq \sqrt{(\sum M_l^2)} \sqrt{\sum (x_{i_l} - y_{i_l})^2}. \end{aligned} \quad (\text{A.2})$$

Comme le poids α est partout (strictement) positif, $c \stackrel{\text{d\'ef}}{=} \inf_l \alpha_{i_l}$ est strictement positif. On peut donc écrire :

$$\begin{aligned} |g^n(\mathbf{x}) - g^n(\mathbf{y})| &\leq \frac{1}{\sqrt{c}} \sqrt{(\sum M_l^2)} \sqrt{\sum \alpha(i_l) (x_{i_l} - y_{i_l})^2} \\ &\leq c' \|\mathbf{x} - \mathbf{y}\|_\alpha. \end{aligned}$$

Les g^n sont donc α -Lipschitziennes, cylindriques et bornées. Par hypothèse, on a pour tout n :

$$\mu g^n = \nu g^n.$$

Par convergence monotone, $\mu(C) = \nu(C)$ pour tout cylindre élémentaire C , donc $\mu = \nu$.

La preuve se modifie aisément pour traiter les deux autres classes de fonctions considérées. En effet, il suffit de montrer que les fonctions g_n définies par (A.1) sont dans ces classes pour pouvoir conclure. Comme les g_l^n sont bornées à dérivées bornées, la norme triple de g_n est bien finie. De plus, l'inégalité (A.2) entraîne immédiatement que g_n est $(\max(M_l))$ -lipschitzienne, pour la norme $|x_\Delta| = \sum_{i \in \Delta} |x_i|$. La fin du raisonnement est identique. \square

On utilise dans la section 2.3 le résultat de continuité suivant.

67 Proposition. *Soit f une fonction cylindrique bornée sur \mathcal{S} , lipschitzienne pour la norme l^1 sur son support, et soit Λ une boîte finie. Alors la fonction $\mu_\Lambda f$ définie par :*

$$\mu_\Lambda f : \mathbf{z} \mapsto \int f((\mathbf{x}^\wedge \mathbf{z})) d\mu_{\Lambda, \mathbf{z}}(d\mathbf{x})$$

est continue bornée.

Démonstration. Le caractère borné est évident.

Notons $f_{\mathbf{z}}$ la fonction $\mathbf{x} \mapsto f((\mathbf{x}^\wedge \mathbf{z}))$, et S_f le support de f . Soit \mathbf{z}_n une suite de configurations tendant vers \mathbf{z} pour la topologie produit. La suite $f_{\mathbf{z}_n}$ converge uniformément vers f : en effet,

$$\begin{aligned} |f(\mathbf{x}^\wedge \mathbf{z}_n) - f(\mathbf{x}^\wedge \mathbf{z})| &\leq C_f \sum_{i \in S_f} C_f |(\mathbf{x}^\wedge \mathbf{z}_n)_i - (\mathbf{x}^\wedge \mathbf{z})_i| \\ &\leq C_f \sum_{i \in S_f \cap \Lambda^c} |z_{n,i} - z_i|. \end{aligned}$$

D'autre part, $\mu_{\Lambda, \mathbf{z}_n}$ converge faiblement vers $\mu_{\Lambda, \mathbf{z}}$ (par convergence dominée sur l'expression explicite des mesures de Gibbs). Donc :

$$\begin{aligned} |\mu_\Lambda f(\mathbf{z}_n) - \mu_\Lambda f(\mathbf{z})| &\leq |\mu_{\Lambda, \mathbf{z}_n} f_{\mathbf{z}_n} - \mu_{\Lambda, \mathbf{z}_n} f_{\mathbf{z}}| + |\mu_{\Lambda, \mathbf{z}_n} f_{\mathbf{z}} - \mu_{\Lambda, \mathbf{z}} f_{\mathbf{z}}| \\ &\leq \|f_{\mathbf{z}_n} - f_{\mathbf{z}}\|_\infty + |\mu_{\Lambda, \mathbf{z}_n} f_{\mathbf{z}} - \mu_{\Lambda, \mathbf{z}} f_{\mathbf{z}}| \end{aligned}$$

D'après les remarques précédentes, les deux termes tendent vers 0, donc $\mu_\Lambda f$ est bien continue. \square

Annexe B

Décomposition d'une fonction strictement convexe à l'infini

68 Proposition. Soit V une fonction réelle, de class \mathcal{C}^∞ . Les deux propriétés suivantes sont équivalentes :

- V est strictement convexe à l'infini : il existe une constante $a > 0$ et un compact K tel que $V'' \geq a$ en dehors de K .
- V se décompose en une somme $V_1 + V_2$, où V_1 est strictement convexe ($V_1 \geq a' > 0$) et V_2 est à support compact.

Démonstration. On voit immédiatement que la deuxième propriété implique la première.

Pour montrer la réciproque, on suppose donc que V'' est supérieur à a en dehors d'un compact. Il existe alors un intervalle réel $[x_1, x_2]$ tel que V est décroissante sur $] -\infty, x_1]$, croissante sur $[x_2, \infty[$, et $V'' \geq a$ en dehors de $[x_1, x_2]$. Quitte à agrandir l'intervalle, on peut supposer $V(x_1) = V(x_2)$, $V'(x_1) < 0$, et $V'(x_2) > 0$. On peut encore supposer $x_1 = -1, x_2 = 1$, $V(-1) = V(1) = 0$.

Plutôt que de recoller les deux bouts de V , on va recoller ceux de V' . En appelant a_i, b_i les dérivées successives de V en -1 et 1 , on se ramène au problème suivant.

Soit (a_i, b_i) deux familles de réels tels que $a_0 < 0, b_0 > 0, a_1 > 0$ et $b_1 > 0$. Alors il existe une fonction g , définie sur $[0, 1]$, telle que :

- $\inf g' > 0$,
- $g^{(k)}(-1) = a_k, g^{(k)}(1) = b_k$,
- $\int g = 0$.

Soit ϵ un réel inférieur à $b_0 - a_0$. Cherchons g sous la forme $g = h + \epsilon x$. On a résolu le problème si on trouve une fonction h vérifiant :

- h croissante,
- $h^{(k)}(-1) = a'_k, h^{(k)}(1) = b'_k$,
- $\int h = 0$.

où a'_0, b'_0 sont de signe opposé et tels que $b_0 - a_0 = b'_0 - a'_0 + 2\epsilon$, et $(a'_i, b'_i) = (a_i, b_i)$ pour $i \geq 1$. On peut maintenant construire h . Sur $[1/2, 1]$, on peut la définir telle que $h = 1/2$ sur $[1/2, 3/4]$, h croissante et de façon qu'elle ait les bonnes dérivées en 1. On la construit ensuite sur $[0, 1]$ de manière à ce qu'elle soit nulle au voisinage de 0, et que $\int_0^1 h = 1/2$. On fait une construction symétrique sur $[-1, 0]$. La fonction h vérifie alors bien les conditions voulues.

Pour reconstruire V , il ne reste plus qu'à choisir la bonne primitive de $g = h + \epsilon x$. \square

Deuxième partie

Recuit simulé

Cette partie reprend le texte de la prépublication « Annealing diffusions in a slowly growing potential », soumise à *Stochastic Processes and their Applications*.

Chapitre 4

Annealing diffusions in a slowly growing potential

Abstract

We consider a continuous analogue of the simulated annealing algorithm in \mathbb{R}^d , namely the solution of the SDE $dX_t = \sigma(t)dB_t - \nabla V(X_t)dt$, where V is a function called potential. We prove a convergence result, similar to the one in [Mic91], under weaker hypotheses on the potential function. In particular, we cover cases where the gradient of the potential goes to zero at infinity. The main idea is to replace the Poincaré and log-Sobolev inequalities used in [Mic91, HCS87] by weak Poincaré inequalities (introduced in [RW01]), and to estimate constants with measure-capacity criteria. We show that the convergence still holds for the “classical” schedule $\sigma(t) = c/\ln(t)$, where c is bigger than a constant related to V .

Keywords: simulated annealing, weak Poincaré inequality
MSC 2000: 90C59, 60J60, 60F99

Introduction

The goal of this article is to study a continuous analogue of a discrete optimization algorithm called *simulated annealing*. This algorithm was introduced in 1983 by Kirkpatrick, Gelatt and Vecchi, and aims at finding “good” (if not perfect) solutions to complex problems. The crucial idea is to perturb the standard gradient descent by a random noise; hopefully this noise will get the process out of traps (local minimas), and help it reach the global minimum. The noise is taken relatively big at the beginning, so that the process explores the space, and is gradually reduced thereafter.

The standard case is the discrete case (in time and space); here we consider a process on \mathbb{R}^d in continuous time. Note that more complicated state spaces have been studied, see for example [Jac94, JR95, Jac96]; here we will stick to \mathbb{R}^d . This “annealing diffusion” process has already been studied by several authors. Hwang, Chiang and Sheu ([HCS87]) proved the convergence under quite strong assumptions, using comparisons with the associated (ordinary) differential equation and results on the trajectories (estimates of exit times from domains, etc.). The result was enhanced by Royer ([Roy89]). The approach we follow was developed by L. Miclo in [Mic92] (and in his doctoral dissertation [Mic91]), and reduces the problem to the convergence of a single quantity, the free energy. Since then, other questions have been asked: speed of convergence, choice of a better algorithm etc. (see *e.g.* the survey [Loc00]). Let us also note that the “functional inequalities” approach, which may be dated back to the works of Holley, Kusuoka and Stroock ([HS88, HKS89]) has also been used extensively for other (possibly

discrete) models, and other closely related algorithms (see *e.g.* [DMM99] for a study of a generalized simulated annealing process).

A common feature of these works on global optimization on \mathbb{R}^d is the quite strong assumptions they require on the growth of the potential. In particular, the norm of the gradient is supposed to go to infinity at infinity. These hypotheses are technically useful: they guarantee that, at any fixed temperature, the generator has a spectral gap, which in turn gives estimates on the rate of convergence. Let us note that the “cooling schedule” (*i.e.* the choice of the temperature as a function of time) for which the process converges is linked with the speed of explosion of the spectral gap, but that it can be read directly on the potential (see below the remarks on the constant d^*).

A natural question arises: what happens when the gradient of the potential does not go to infinity, and when there is no spectral gap? Do we need to change the cooling schedule to reflect the slow-down of the diffusions at fixed temperature, or does the local structure of the potential dictate the optimal schedule?

Before we answer this question, let us be more precise and give our hypotheses.

We study the following optimization problem: how to find the minimum of a function V on the space \mathbb{R}^d . To solve this problem, we introduce the following stochastic differential equation:

$$\begin{cases} dX_t &= \sqrt{\sigma(t)}dB_t - \frac{1}{2}\nabla V(X_t)dt, \\ X_0 &\sim m_0. \end{cases}$$

The function σ will be called *temperature*, and will be a (deterministic) function of time, decreasing to zero.

Intuitively, this process is similar to simulated annealing: we perturb a gradient descent by a stochastic term whose intensity decreases over time.

We would like to know if the process finds a point where the global minimum is reached; we will show that it does, in a weak sense.

69 Definition. *The annealing process starting from a law m_0 is said to converge if its law m_t at time t converges weakly to a measure supported by $\operatorname{argmin} V$. In particular, if the global minimum of V is reached in a single point x_0 , the process converges if m_t goes to a Dirac mass at x_0 .*

Let us now recall the result we would like to generalize: this is the main result of [HCS87, Roy89, Mic92], as it appears in [Mic92].

70 Theorem (L. Miclo). *If V satisfies some regularity assumptions, and the following conditions:*

- $V \xrightarrow{x \rightarrow \infty} \infty$,
- $|\nabla V| \xrightarrow{x \rightarrow \infty} \infty$,
- $|\nabla V| - \Delta V$ is bounded from below,

then there exists a constant d^ such that, for any $c > d^*$, and for $\sigma(t) = c/\ln(t)$, the annealing process converges.*

This d^* will be explicitly defined below, in terms of the structure of the wells of V . It is known to be optimal, at least in the discrete space case : if the process is cooled faster, then its limiting law charges local minima (*cf.* the last remark of section 4.1, p. 116).

To understand the direction in which we generalize this result, let us note that this theorem applies for any potential V which is equal to $|x|^\alpha$ outside a compact set, whenever α is strictly bigger than 1. It is then a quite natural question to ask whether this still holds when α is strictly less than 1. Our hypotheses, which we now state, allow us to treat this case.

Hypothesis 1 (Global minimum). *The potential has a unique global minimum, located at the origin and $V(0) = 0$. Moreover, this minimum is non degenerate: $\text{Hess}V(0)$ is positive definite.*

Hypothesis 2 (Growth at infinity). *The potential V goes to infinity at infinity faster than a logarithm:*

$$\exists m_V > 1, \exists C, \quad V(x) \geq \ln(|x|)^{m_V} - C.$$

Hypothesis 3 (Bounded gradient). *The potential V is continuously differentiable, and its gradient is bounded:*

$$\|\nabla V\|_\infty < \infty.$$

Hypothesis 4 (Concavity). *The Laplacian of V is negative at infinity: there exists a compact set K compact such that*

$$\forall x \notin K, \quad \Delta V(x) \leq 0.$$

One last hypothesis will be added in section 4.3, regarding the structure of local minima of V .

These hypotheses call for a few remarks.

The first one simplifies the problem at hand: there is only one goal to go after. If the weak limit of the equilibrium measures μ_σ (*cf.* *infra*) is known (some results in this direction may be found in [Mic92, Hwa80]), the arguments given here should work in the same way. The non-degeneracy hypothesis may be weakened too (see *e.g.* section 4.2 for a slight generalization in $d = 1$) However, this restriction allows for two simplifications: it gives an estimate of the partition function Z_σ , and avoids more intricate reasonings in the computation of the weak inequalities (section 4.3).

The growth hypothesis is not very restrictive. In particular, V may grow like $|x|^\alpha$ with $\alpha < 1$ (or even slower). These cases were not covered in the literature. Let us note that we do not know what happens in the limit case (when $m_V = 1$, *i.e.* the tails of the equilibrium measures are polynomial).

In the light of the previously known results, the bounded gradient assumption seems less stringent: in some sense, we already know what happens when the gradient is big. The hypothesis could probably be lifted if we allowed a polynomial growth, or a control by V , but we keep it for the sake of clarity.

Finally, the condition on ΔV seems more restrictive. It will only be used in the proof of the moment bound (section 4.B). It could probably be replaced by a condition like $\Delta V \leq C|\nabla V|^2$. However, in the “natural example” where $V(x) = |x|^\alpha$ at infinity, the Laplacian is indeed negative if $\alpha < 1$, and this example was one motivation for investigating the problem. Moreover, even this weakened hypothesis would not allow the existence of traps at infinity, however shallow they may be. It would be interesting to know what could happen if there were such traps: either they have no effect (in the sense that the same cooling schedule may be chosen), or they slow down the process too much and destroy the convergence.

Our principal result is the following.

71 Theorem. *If the potential V satisfies the hypotheses above, there exists a constant d^* such that, if we choose*

$$\sigma(t) = \frac{c}{\ln(t)},$$

with $c > d^$, the annealing process converges.*

This result generalizes theorem 70 by allowing more general choices for the potential function. In particular, as we will see in the sequel, the equilibrium measures need not satisfy a Poincaré inequality. Nonetheless, the critical cooling schedule is the same, which contradicts the intuition that the speed was given by the Poincaré constants. In fact, what seems to prevail is the behavior of V in a compact set, and from a certain point of view, that is precisely what the weak inequalities capture.

The remainder of the paper is organized in the following way. Firstly, we explain the analytic approach of L. Miclo and give the main line of the proof.

This proof, under our weakened hypotheses, uses *weak Poincaré inequalities*. We will need controls over their dependence on temperature: these are established in sections 4.2 and 4.3, respectively in the one- and multi-dimensional case. These three sections are the core of the proof of the convergence result.

The quite technical 4.4th section gathers definitions and results about Orlicz norms and weak inequalities. Finally, we postpone to the annexes a comparison between functions centered by their mean or by their median, a moment bound for the annealing process, and a brief proof of the estimation of the partition function.

4.1 The convergence of the process (main line of the proof)

A differential inequality for the free energy

Before we describe the main idea, we introduce some notation. Consider the SDE defining the annealing diffusion, but with a constant temperature σ . The process is then a classical diffusion with a gradient drift. The corresponding generator is given by:

$$L_\sigma : f \mapsto \frac{\sigma}{2} \Delta f - \frac{1}{2} \nabla V \nabla f.$$

The measure μ_σ defined by

$$d\mu_\sigma = \frac{1}{Z_\sigma} \exp\left(-\frac{V}{\sigma}\right) d\lambda,$$

is reversible for this process (Z_σ is a normalization constant). We will call μ_σ the *instantaneous equilibrium measure*.

It's easy to see that, as σ goes to zero, the measures μ_σ concentrate around the global minimum of the potential (which is found at the origin by hypothesis). In fact, we even have the following convergence.

72 Proposition. *The measures μ_σ converge weakly:*

$$\mu_\sigma \xrightarrow{\sigma \rightarrow 0} \delta_0.$$

Moreover, the normalization constant Z_σ behaves like $\sigma^{d/2}$.

The asymptotic behavior of Z_σ is proved in annex 4.C.

In order to prove that the process converges, we follow the approach of L. Miclo ([Mic92]) and show that the relative entropy of the law of the process with respect to its instantaneous equilibrium measure goes to zero.

More precisely, let f_t be the density of $m_t = \mathcal{L}(X_t)$ with respect to the equilibrium measure μ_t (we slightly abuse notations and define $\mu_t = \mu_{\sigma(t)}$). The relative entropy (also called free energy) is $I_t = \int f_t \log f_t d\mu_t$, which can be rewritten as $I_t = \mathbf{Ent}_{\mu_t}(\sqrt{f_t}^2)$. The finiteness of I_t is established in annex 4.C. We would like to study the evolution of I_t ; the natural idea is to differentiate it. One can justify the following formal computation:

73 Proposition (Differentiation of the free energy). *The derivative of the free energy is given by:*

$$\begin{aligned} \frac{dI_t}{dt} &= \frac{1}{\sigma(t)^2} \sigma'(t) \int V \times (1 - f_t) d\mu_t - 2\sigma(t) \int |\nabla \sqrt{f_t}|^2 \\ &= \frac{1}{\sigma(t)^2} \sigma'(t) \int V \times (1 - f_t) d\mu_t - 2\sigma(t) \mathcal{E}_{\mu_t}(\sqrt{f_t}). \end{aligned}$$

Remark. By $\mathcal{E}_{\mu_t}(f)$ we denote $\int |\nabla f|^2 d\mu_t$. This is somewhat improper — strictly speaking, this is the energy associated with the generator $\Delta - (1/\sigma_t)\nabla V \nabla$ (so we should multiply our energy by $(\sigma_t/2)$ to get the “real” one). However, the classical criteria for functional inequalities are written for this form of the energy.

The first term is set aside for the time being, we shall bound it later directly by a function of t .

Following the classical path leading from functional inequalities to semigroup estimates, we now try to control the energy term on the right hand side.

If the measures μ_t satisfied logarithmic Sobolev inequalities, everything would be fine: the energy of $\sqrt{f_t}$ could be controlled by its entropy with respect to μ_t , and we would get I_t back on the right hand side of the inequality. We would still have to know how the constants in the logarithmic Sobolev inequality depend on the small parameter σ , and get an upper bound for the first term, but we could get the convergence of I_t to zero.

Unfortunately, the scaling behavior of the constants in the logarithmic Sobolev inequality (*i.e.* the way they behave when σ goes to zero) is not clear. Moreover, this inequality need not hold, and in fact it won't under our hypotheses.

In Miclo's paper, the first difficulty is overcome thanks to a Poincaré inequality, weaker than the logarithmic Sobolev inequality, but for which the constants are well known. However, even this inequality won't be satisfied in our case, and we have to find another way.

Our idea is to consider a still weaker functional inequality, namely a *weak Poincaré inequality*, written with an Orlicz norm. Weak Poincaré inequalities were introduced by M. Röckner and F.-Y. Wang in [RW01], originally with an L^∞ norm and the mean of f instead of a median on the right hand side. We will give a brief account on weak inequalities and Orlicz norms in section 4.4, and explain the link between the original inequality and the one we use.

For now, let us just state this inequality. It reads:

$$\forall f, \forall r, \quad \mathbf{Var}_{\mu_t}(f) \leq \alpha_t(r) \mathcal{E}_{\mu_t}(f) + r \|f - m_f\|_\phi^2, \quad (4.1)$$

where m_f is a median of f under μ , $\|\cdot\|_\phi$ is the Orlicz norm associated to a function ϕ (to be precised later), and α_t , a decreasing function of r , is the *compensating function*. The Orlicz norm

is not easily tractable, but we will see (cf. lemma 98) that it can be bounded by the entropy: there exists a C such that, for all positive f ,

$$\|f - m_f\|_{\phi}^2 \leq C(\mu(f^2) + \mathbf{Ent}(f^2)).$$

At this point, the energy is bounded above by three terms: $\mu(f^2)$, the entropy of f and its variance. To get rid of the variance term, we would like to bound it by entropy-like quantities. To this end we introduce the following definition.

74 Definition. For any probability measure μ and any positive f , we will call pseudo-entropy the quantity:

$$\mathbf{Ps-Ent}(f) = \int f \log^2 \left(e + \frac{f}{\|f\|_{1,\mu}} \right) d\mu.$$

With this definition in hand, we can state ([Mic92], lemma 4):

75 Lemma. There exists a δ_0 such that, for all probability measure μ and all positive f with $\mu(f^2) = 1$,

$$\forall \delta < \delta_0, \frac{1}{\delta} \mathbf{Var}_{\mu}(f) + 4\delta \mathbf{Ps-Ent}_{\mu}(f^2) \geq \mathbf{Ent}_{\mu}(f^2).$$

Let us glue all these inequalities together: we get that for all probability measure μ , if μ satisfies the weak Poincaré inequality 4.1, then for all positive f with $\int f^2 d\mu = 1$,

$$\delta \mathbf{Ent}_{\mu}(f^2) - 4\delta^2 \mathbf{Ps-Ent}(f^2) \leq \mathbf{Var}_{\mu}(f) \leq \alpha(r) \mathcal{E}_{\mu}(f, f) + Cr \mathbf{Ent}_{\mu}(f^2) + Cr.$$

This entails a lower bound on the energy:

$$\mathcal{E}_{\mu}(f, f) \geq -\frac{4}{\alpha(r)} \delta^2 \mathbf{Ps-Ent}(f^2) - C \frac{r}{\alpha(r)} + \frac{1}{\alpha(r)} (\delta - Cr) \mathbf{Ent}_{\mu}(f^2).$$

Let us get back into our special case, and take $\mu = \mu_t$, $f = \sqrt{f_t}$. The entropy $\mathbf{Ent}(f^2)$ just becomes I_t , and we can plug the inequality back in the differential equation for I_t :

$$\frac{dI_t}{dt} \leq \frac{1}{\sigma(t)^2} \sigma'(t) \int V \times (1 - f_t) d\mu_t + 8\delta^2 \frac{\sigma(t)}{\alpha_t(r)} \mathbf{Ps-Ent}(f_t) + 2C\sigma(t) \frac{r}{\alpha_t(r)} - 2(\delta - Cr) \frac{\sigma(t)}{\alpha_t(r)} I_t$$

Since σ is non-increasing in time, we may omit the 1 in $(1 - f_t)$ in the first term, and since $f_t d\mu_t = dm_t$,

$$\frac{dI_t}{dt} \leq \frac{d}{dt} \left(\frac{1}{\sigma(t)} \right) \int V dm_t + 8\delta^2 \frac{\sigma(t)}{\alpha_t(r)} \mathbf{Ps-Ent}(f_t) + 2C\sigma(t) \frac{r}{\alpha_t(r)} - 2(\delta - Cr) \frac{\sigma(t)}{\alpha_t(r)} I_t \quad (4.2)$$

Our goal is to obtain a differential inequality involving only I_t and explicit functions of t , so that we may deduce information on the evolution of I_t . Since σ is known, this leaves us with three questions. First, we have to obtain controls on $\int V dm_t$ and on the pseudo-entropy — we will get explicit bounds in t . Once this is done, we have to estimate the compensating function α_t . Finally we must choose r and δ depending on t in a suitable way, so that the inequality on I_t is good enough to prove the convergence to zero.

We now deal with the first problem.

Moment bounds and pseudo-entropy

The first inequality is a moment bound on the value of the potential at time t . The proof is postponed to the annexes.

76 Lemma. *Suppose that hypotheses 3 and 4 hold, and that the initial law m_0 satisfies:*

$$\int V^p m_0(dx) < \infty,$$

for some $p \geq 1$. Then there exists an M such that:

$$\int V^p(x) m_t(dx) \leq M \sigma(t)^p \ln(t)^p (\ln \ln t)^{3p}.$$

The last result will be used directly, but it also helps us prove the following bound.

77 Lemma. *Suppose that $\int V^2 dm_0$ is finite, and that the cooling schedule has the form: $\sigma(t) = c/\ln(t)$, for a positive constant c . Then there exists an A such that, for all big enough t ,*

$$\mathbf{Ps-Ent}(f_t) \leq A \ln(t)^2 (\ln \ln(t))^6.$$

Proof. We differentiate the quantity under scrutiny, namely $J_t = \mathbf{Ps-Ent}_{\mu_t}(f_t)$. The following formal computation can be justified (cf. [Mic92]):

$$\frac{dJ_t}{dt} = -\frac{\sigma(t)}{2} \int F'(f_t) |\nabla f_t|^2 d\mu_t + 2 \frac{d}{dt} \left(\frac{1}{\sigma(t)} \right) \int \log(e + f_t) \frac{f_t}{e + f_t} \left(V - \int V(x) d\mu_t \right) dm_t,$$

where $F(x) = \frac{2x}{x+e} \log(x+e) + \log^2(x+e)$. Since F is non decreasing (in x), and σ is positive, the first term is bounded above by 0. Moreover, since V is positive and $1/\sigma$ increases, we may also forget the $\int V(x) d\mu_t$ in the second term. We get:

$$\begin{aligned} \frac{dJ_t}{dt} &\leq 2 \frac{d}{dt} \left(\frac{1}{\sigma(t)} \right) \int \frac{f_t}{e + f_t} \log(e + f_t) V dm_t \\ &\leq 2 \frac{d}{dt} \left(\frac{1}{\sigma(t)} \right) \int \log(e + f_t) V dm_t \\ &\leq 2 \frac{d}{dt} \left(\frac{1}{\sigma(t)} \right) \left(\int \log^2(e + f_t) dm_t \right)^{\frac{1}{2}} \left(\int V^2 dm_t \right)^{\frac{1}{2}} \\ &= 2 \frac{d}{dt} \left(\frac{1}{\sigma(t)} \right) J_t^{\frac{1}{2}} \left(\int V^2 dm_t \right)^{\frac{1}{2}}. \end{aligned}$$

After dividing by $2J_t^{\frac{1}{2}}$, the left hand side becomes the derivative of $\sqrt{J_t}$. The right hand side may then be bounded (cf. previous lemma):

$$\begin{aligned} \frac{d\sqrt{J_t}}{dt} &\leq \frac{d}{dt} \left(\frac{1}{\sigma(t)} \right) \left(\int V^2 dm_t \right)^{\frac{1}{2}} \\ &\leq \frac{d}{dt} \left(\frac{1}{\sigma(t)} \right) \sqrt{M} \sigma(t) \ln(t) (\ln \ln(t))^3. \end{aligned}$$

The explicit value of σ allows us to simplify (the multiplicative constant c cancels):

$$\frac{d\sqrt{J_t}}{dt} \leq \sqrt{M} \frac{1}{t} (\ln \ln(t))^3.$$

An easy computation shows that the right hand side may be bounded by:

$$\sqrt{M} \frac{d}{dt} (\ln(t) (\ln \ln(t))^3).$$

To conclude the proof, we integrate this inequality between a (fixed and big enough) t_0 and the current time t . The constant A naturally depends on the initial law m_0 (through the value of M and through the pseudo-entropy at time t_0). \square

From the differential inequality to the convergence of the entropy

It is now time to get back to our differential inequality and apply the bounds we just derived. We fix a logarithmic cooling schedule:

$$\sigma(t) = \frac{c}{\ln(t)}.$$

Recall that we showed (inequality 4.2):

$$\frac{dI_t}{dt} \leq \frac{d}{dt} \left(\frac{1}{\sigma(t)} \right) \int V dm_t + 8\delta^2 \frac{\sigma(t)}{\alpha_t(r)} \mathbf{Ps-Ent}(f_t) + 2C\sigma(t) \frac{r}{\alpha_t(r)} - 2(\delta - Cr) \frac{\sigma(t)}{\alpha_t(r)} I_t$$

We use the moment bound (lemma 76) to deal with the first term, and lemma 77 to bound the second one.

$$\frac{dI_t}{dt} \leq \frac{d}{dt} \left(\frac{1}{\sigma(t)} \right) M \ln \ln(t)^3 + 8M\delta^2 \frac{\sigma(t)}{\alpha_t(r)} (\ln(t))^2 (\ln \ln(t))^6 + 2C\sigma(t) \frac{r}{\alpha_t(r)} - 2(\delta - Cr) \frac{\sigma(t)}{\alpha_t(r)} I_t$$

We number our four terms and define:

$$\begin{aligned} \textcircled{1} &= \frac{d}{dt} \left(\frac{1}{\sigma(t)} \right) M \ln \ln(t)^3 & \textcircled{3} &= 2C\sigma(t) \frac{r}{\alpha_t(r)} \\ \textcircled{2} &= 8M\delta^2 \frac{\sigma(t)}{\alpha_t(r)} (\ln(t))^2 (\ln \ln(t))^6 & \textcircled{4} &= 2(\delta - Cr) \frac{\sigma(t)}{\alpha_t(r)} \end{aligned}$$

The inequality becomes:

$$\frac{dI_t}{dt} \leq \textcircled{1} + \textcircled{2} + \textcircled{3} - \textcircled{4} I_t \tag{4.3}$$

This last inequality will allow us to prove that the free energy goes to zero. To this end, we use the same lemma as L. Miclo:

78 Lemma. *Let I be a positive function, and suppose:*

$$\frac{dI_t}{dt} \leq a(t) - b(t)I(t),$$

where a, b are positive functions and satisfy:

1. $\int^\infty b(t) = \infty$,
2. $\frac{a(t)}{b(t)} \xrightarrow{t \rightarrow \infty} 0$.

Then I goes to zero when t goes to infinity.

Our goal is now to use the inequality 4.3 to check the hypotheses of this lemma. We choose δ and r as follows.

$$\begin{cases} \delta_t &= \frac{1}{\ln(t)^2 (\ln \ln(t))^7} \\ r_t &= \frac{1}{C \ln(t)^2 (\ln \ln(t))^8}, \end{cases} \quad (4.4)$$

where the factor C in r_t appears only for cosmetic reasons. This choice ensures:

$$\begin{aligned} \frac{\textcircled{3}}{\textcircled{4}} &= \frac{Cr_t}{\delta_t - Cr_t} \sim \frac{C}{\ln \ln(t)} \rightarrow 0, \\ \frac{\textcircled{2}}{\textcircled{4}} &= \frac{4M\delta_t^2}{\delta_t - Cr_t} \ln^2(t) (\ln \ln(t))^6 \sim 4M\delta_t \ln^2(t) (\ln \ln(t))^6 \rightarrow 0. \end{aligned}$$

Two things remain to check:

$$\frac{\textcircled{1}}{\textcircled{4}} \rightarrow 0 \quad \text{and} \quad \int^{\infty} \textcircled{4} = \infty.$$

This is where we need bounds on the weak Poincaré inequalities: we have to know how α_t behaves for our particular choice of r . This is the aim of the following sections, in one or many dimensions.

In both cases, we will get:

79 Lemma. *There exists a constant d^* such that, for all $D^* > d^*$,*

$$\exists C_\alpha \quad \alpha_t(r_t) \leq C_\alpha \exp\left(\frac{D^*}{\sigma(t)}\right).$$

For the cooling schedule $\sigma(t) = c/\ln(t)$, we get:

$$\alpha_t(r_t) \leq C_\alpha t^{D^*/c}.$$

In the one-dimensional case, this follows from theorem 80 below, and the choice of r_t . The multi-dimensional case is proved in theorem 85 and the discussion that follows it.

Remark. *The approach in the one- and multi-dimensional case will differ slightly. In the former, we prove a (full) weak Poincaré inequality, i.e. we estimate the whole function α_t , and then use this estimate at the point r_t . In the latter, we will only prove a bound on α_t at r_t and disregard the other points.*

We may now get back to our proof. Recall that we have assumed:

$$\sigma(t) = \frac{c}{\ln(t)}, \quad c > d^*,$$

so that we may always pick a D^* strictly less than c .

Let us check the two remaining points. First we must prove that $\textcircled{1}/\textcircled{4}$ converges to zero. Since $\sigma(t)$ is explicit and we know a bound on $\alpha(r)$, we see that:

$$\begin{aligned} \frac{\textcircled{1}}{\textcircled{4}} &= \frac{d}{dt} \left(\frac{1}{\sigma(t)} \right) M (\ln \ln(t))^3 \times \frac{\alpha_t(r_t)}{2(\delta_t - Cr_t)\sigma(t)} \\ &\leq M' \frac{1}{t} \ln(t)^3 \ln \ln(t)^{10} \alpha_t(r_t). \end{aligned}$$

where M, M' are constants.

Using the bound on α we just recalled (lemma 79), we get:

$$\frac{\textcircled{1}}{\textcircled{4}} \leq M'' \frac{t^{D^*/c}}{t} ((\ln t)^3 (\ln \ln t)^{10})$$

Since $c > D^*$, $\textcircled{1}/\textcircled{4}$ goes to zero, as was claimed.

Just in the same way, we have, for t big enough:

$$\begin{aligned} \textcircled{4} &= 2(\delta_t - Cr_t) \frac{\sigma(t)}{\alpha_t(r_t)} \\ &\sim M'' (\ln t)^{-3} (\ln \ln t)^{-7} \frac{1}{t^{\frac{D^*}{c}}}. \end{aligned}$$

Once more, the condition $c > D^*$ guarantees that the integral of this quantity diverges, which was expected.

This allows us to apply lemma 78, and prove that I_t converges to 0. Thanks to Pinsker's inequality, the total variation between m_t (law of the process) and μ_t (the instantaneous equilibrium) converges too. Since we already know that μ_t converges weakly to the Dirac mass δ_0 , this concludes the proof.

Some remarks

Our theorem immediately raises a few questions. Some of these have already been asked when we discussed the hypotheses — equilibrium measures with polynomial tails are not covered, and we do not know what happens when there are traps at infinity.

It would also be interesting to know what happens if we cool faster than the “good” schedule (*i.e.* $\sigma(t) = c/\ln(t), c < d^*$). In a discrete setting (or on a compact manifold), this has been studied *e.g.* in [Mic95]: a fast cooling may freeze the system in a local trap, and the possible traps may be identified precisely for (almost) any value of c . It may be possible to derive the same results in our setting.

In the same discrete or compact setting, if V has several global minima, another critical constant $\tilde{d} < d^*$ appears, and for $\tilde{d} < c < d^*$, the process may approach some, but not all, global minima, depending on its starting point (see [Mic95]). Once more, the fact that local properties seem to dictate the behaviour of the process leads us to expect similar results on \mathbb{R}^d .

4.2 The one-dimensional case

In this section we treat the case of a one-dimensional potential, for which we derive a weak Poincaré inequality (more precisely we prove lemma 79).

The major advantage of this case is that, in one dimension, explicit (Hardy-like) criteria are known for weak inequalities. Thus we are able to prove a quite general result (the de-coupling of the parameters s and σ in the weak inequality). This has a small price: we restrict ourselves to potentials that grow like a power of x , and do not cover the case $V(x) = \log(|x|)^\alpha$ at infinity (for some $\alpha > 1$). It should be noted that the multidimensional argument (*cf.* next section) may still be used in this logarithmic case.

Let us write down a few notations. The potential V is a real \mathcal{C}^2 function, which satisfies the hypothesis 2. For any (small) σ , we denote by V_σ the function $\frac{1}{\sigma}V$, and by $Z_\sigma = \int e^{-V_\sigma(x)} dx$

the partition function . We normalize V_σ by defining Φ_σ : $\Phi_\sigma = V_\sigma + \log Z_\sigma$. The equilibrium measure μ_σ reads:

$$d\mu_\sigma = \frac{1}{Z_\sigma} \exp(-V_\sigma) d\lambda = \exp(-\Phi_\sigma) d\lambda.$$

We now state our hypotheses on V . We suppose there exists a compact set $[K_1, K_2]$ such that the following holds.

Hypothesis U 1 (Behavior near the minimum). *In $[K_1, K_2]$, V is bounded below by 0 and above $V(K_1) = V(K_2)$. It reaches its minimum only once, at x_1 . Near this point, V behaves like:*

$$V(x) \sim (x - x_1)^b,$$

with $b > 1$. Finally, there exists δ such that V is bijective from $[x_1, x_1 + \delta]$ onto its image, and from $[x_1 - \delta, x_1]$ onto its image.

This generalizes a little the general assumptions on the minimum: if $\text{Hess}V$ is positive definite at x_1 , it satisfies this hypothesis with $b = 2$.

Hypothesis U 2 (Behavior outside the compact). *Outside the compact, V' and $|V''|/(V')^2$ are bounded:*

$$\exists C_V \forall x \notin [K_1, K_2], \quad \frac{|V''|}{V'^2} \leq C_V. \quad (4.5)$$

Moreover, $V'(K_1) \neq 0$ and $V'(K_2) \neq 0$.

In particular, V' has no zero, V decreases before K_1 and increases after K_2 .

Hypothesis U 3 (The function β). *There exists a function β such that, for all x outside the compact,*

$$\beta \left(\frac{\exp(-V(x))}{V'(x)} \right) \geq \frac{1}{V'(x)^2}. \quad (4.6)$$

To apply the result to the annealing diffusion, we need an additional growth condition on β . Since V satisfies hypothesis 2 (p. 109) and U2, it is easily seen that $\exp(-V)/V'$ must go to zero at infinity (just integrate $V''/(V')^2$ between K_2 and x and use (4.5) to see that $1/V'(x)$ is bounded above by $cst. + C_V x$, then use the growth hypothesis on V), so we need bounds on β near 0.

Hypothesis U 4 (Behavior of β near the origin). *There exist constants A, C such that, near 0, the following holds:*

$$\beta(s) \leq C \left(\log \frac{1}{s} \right)^A.$$

Remark. *We shall note here that the last two hypotheses hold if $V(x) = |x|^\alpha$ outside a compact, with $\alpha \in (0, 1]$, if we choose $\beta = C (\log(1/s))^{\frac{2}{\alpha}-2}$ for $s \in (0, 1)$ (cf. [RW01],[BCR05] cor. 4). If V grows like $|\log(x)|^m$, this is not true (β behaves like $\exp(c \log(1/s)^{1/m})$ for $m > 1$, like a power of s if $m = 1$). This explains the small loss of generality we spoke about above.*

We define, for all $x \geq x_1$, $i(x) = \inf\{V(y), y > x\}$ and $s(x) = \sup\{V(y), y \in [x_1, x]\}$. In the same way, $i(x) = \inf\{V(y), y < x\}$ and $s(x) = \sup\{V(y), y \in [x, x_1]\}$ for x less than x_1 .

Outside $[K_1, K_2]$, we have $i = V = s$, so $s - i$ is continuous with compact support. We call d^* its maximum value.

The main result of this section may now be stated as follows (see definition 92, page 131 for a rigorous definition of a weak Poincaré inequality and a compensating function).

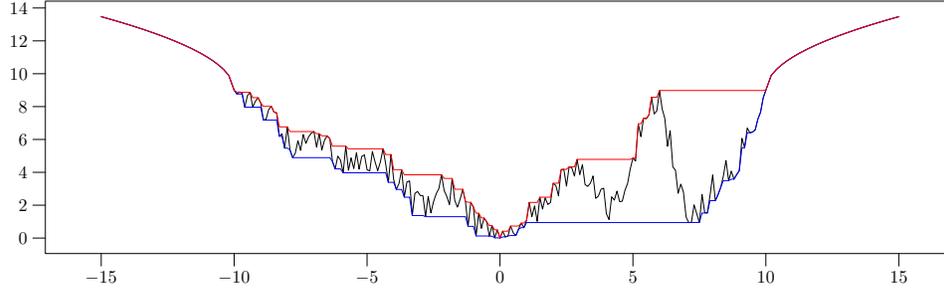


Figure 4.1: The potential V and the associated functions i and s . Here, $x_1 = 0$, $K_1 = -10$, $K_2 = 10$.

80 Theorem. *The measure μ_σ satisfies a weak Poincaré inequality with the L^∞ norm, with a compensating function β_σ defined by:*

$$\beta_\sigma(s) = C \exp\left(\frac{d^*}{\sigma}\right) \beta(s),$$

where β is given by the hypothesis. Similarly, μ_σ satisfies a weak inequality with an Orlicz norm and the modified function α_σ given by:

$$\alpha_\sigma(r) = C \beta_\sigma\left(C' \exp\left(-\frac{4}{r}\right)\right) = C \exp\left(\frac{d^*}{\sigma}\right) \beta\left(C' \exp\left(-\frac{4}{r}\right)\right).$$

Finally, there exists a constant A such that the following bound holds:

$$\alpha_\sigma(r) \leq C \exp\left(\frac{d^*}{\sigma}\right) \frac{1}{r^A}.$$

To prove this, we will use a result from Barthe, Cattiaux and Roberto ([BCR05], theorem 3), which gives estimates on the compensating functions for the L^∞ norm. We will then use capacity-measure criteria to derive the result with the Orlicz norm. To state the result we need, we first give some additional notation.

Let m_σ will be a median of μ_σ , and for all x ,

$$B_\sigma(x) = \frac{\int_{m_\sigma}^x e^{\Phi_\sigma(y)} dy \times \int_x^\infty e^{-\Phi_\sigma} dy}{\beta\left(\int_x^\infty e^{-\Phi_\sigma(y)} dy\right)}$$

$$B_\sigma = \sup_{x \geq m_\sigma} B_\sigma(x). \quad (4.7)$$

By symmetry, we also define $b_\sigma(x)$ and b_σ for $x \leq m_\sigma$.

The result from [BCR05] reads:

81 Theorem. *Let $\beta : (0, 1) \rightarrow \mathbb{R}^+$ be non increasing, and B_σ, b_σ be defined by (4.7).*

Then μ_σ satisfies the following weak Poincaré inequality:

$$\mathbf{Var}_{\mu_\sigma}(f) \leq C_\sigma \beta(s) \int |\nabla f|^2 d\mu_\sigma + s \operatorname{osc}(f)^2,$$

where $C_\sigma \leq 12 \max(b_\sigma, B_\sigma)$.

Note that their result is actually stronger, since it also gives a lower bound on the optimal constant C in terms of some quantities very similar to B_σ .

To use this result, we have to bound $B_\sigma(x)$, and this has to be done uniformly in x . We will split \mathbb{R} into two domains, and show that, in some sense, our choice of β already deals with B_σ for large x , so that the crucial region is near the minimum x_1 .

What happens for large x We study the case where $x \geq K_2$ by following the proof of corollary 4 in [BCR05].

82 Lemma. *For all σ , there exists a c_σ such that:*

$$\forall x \notin [K_1, K_2] \quad \beta \left(\frac{2e^{-\Phi_\sigma(x)}}{\Phi'_\sigma(x)} \right) \geq \frac{c_\sigma}{\Phi'_\sigma(x)^2}. \quad (4.8)$$

One may choose $c_\sigma = \frac{1}{\sigma^2}$.

Proof. Recall that the same bound holds for V (cf. hypothesis U3); we try to carry it over to Φ_σ .

The behavior of V near its minimum allows us to get an equivalent for Z_σ using Laplace's method (cf. for example [Die68]); if $V \sim (x - x_1)^b$, we get

$$Z_\sigma \sim C\sigma^{1/b}, \quad (4.9)$$

where C depends only on V . Let us bound the argument in the function β .

$$2 \frac{\exp(-\Phi_\sigma)}{\Phi'_\sigma} = \frac{2\sigma \exp(-V/\sigma)}{Z_\sigma V'} \leq C' \sigma^{1-1/b} \frac{\exp(-V/\sigma)}{V'} \leq C' \sigma^{1-1/b} \frac{\exp(-V)}{V'} \leq \frac{\exp(-V)}{V'}$$

for σ small enough, because b is strictly greater than 1, so that $\sigma^{1-1/b}$ goes to zero. Since β decreases, we get, outside $[K_1, K_2]$:

$$\beta \left(2 \frac{\exp(-\Phi_\sigma)}{\Phi'_\sigma} \right) \geq \beta(\exp(-V)/V') \geq \frac{1}{V'^2} = \frac{1}{\sigma^2 \Phi'_\sigma{}^2}. \quad \square$$

83 Lemma. *For all $x \geq K_2$, we have the following inequalities:*

$$\begin{aligned} \int_{m_\sigma}^x e^{\Phi_\sigma} &\leq \int_{m_\sigma}^{K_2} e^{\Phi_\sigma(y)} dy + 2 \frac{e^{\Phi_\sigma}}{\Phi'_\sigma} \\ \mu_\sigma([x, \infty)) &\leq 2 \frac{e^{-\Phi_\sigma(x)}}{\Phi'_\sigma(x)} \leq 3 \mu_\sigma([x, \infty)). \end{aligned}$$

Proof. For all $x \geq K_2$ and σ small enough (less than $1/(2C_V)$), the hypothesis on V gives us:

$$\frac{|\Phi''_\sigma(x)|}{\Phi'_\sigma(x)^2} = \sigma \frac{|V''|}{V'^2} \leq C_V \sigma \leq \frac{1}{2}.$$

Therefore:

$$\left(\frac{\exp(\Phi_\sigma)}{\Phi'_\sigma} \right)' \geq \frac{1}{2} e^{\Phi_\sigma},$$

This gives the first result by integration. In a similar way, we may integrate the bounds

$$\left(\frac{\exp(-\Phi_\sigma)}{\Phi'_\sigma} \right)' \in \left[\frac{1}{2} e^{-\Phi_\sigma}, \frac{3}{2} e^{-\Phi_\sigma} \right]$$

between $x \geq K_2$ and ∞ , and use the fact that $\exp(-\Phi_\sigma)/\Phi'_\sigma$ goes to zero (in the same way as for $\exp(-V)/(V')$, see the discussion following hypothesis 3), to prove the second claim. \square

We are now in a position to bound $B_\sigma(x)$. We suppose that σ is small enough to guarantee that m_σ lies in $[K_1, K_2]$.

$$\begin{aligned}
B_\sigma(x) &= \mu_\sigma([x, \infty)) \times \frac{1}{\beta(\mu_\sigma([x, \infty)))} \times \int_{m_\sigma}^x e^{\Phi_\sigma(x)} \\
&\leq 2 \frac{e^{-\Phi_\sigma(x)}}{\Phi'_\sigma(x)} \times \frac{1}{\beta\left(2 \frac{e^{-\Phi_\sigma}}{\Phi'_\sigma}\right)} \times \left(\int_{m_\sigma}^{K_2} e^{\Phi_\sigma(y)} dy + 2 \frac{e^{\Phi_\sigma(x)}}{\Phi'_\sigma(x)} \right) && \text{(by lemma 83)} \\
&\leq 2 \frac{e^{-\Phi_\sigma(x)}}{\Phi'_\sigma(x)} \times \frac{\Phi'_\sigma(x)^2}{c_\sigma} \times \left(\int_{m_\sigma}^{K_2} e^{\Phi_\sigma(y)} dy + 2 \frac{e^{\Phi_\sigma(x)}}{\Phi'_\sigma(x)} \right) && \text{(by lemma 82)} \\
&\leq \frac{2}{c_\sigma} \Phi'_\sigma(x) \int_{m_\sigma}^{K_2} e^{V_\sigma(y) - V_\sigma(K_2)} dy + \frac{4}{c_\sigma}. && \text{(because } V(x) \geq V(K_2))
\end{aligned}$$

The hypotheses imply that $V(y) \leq V(K_2)$, whenever $K_1 \leq y \leq K_2$. On the other hand, Φ'_σ is bounded above by C/σ (since V' is supposed to be bounded). Finally, assuming that $\sigma \leq 1$,

$$\forall x \geq K_2, \quad B_\sigma(x) \leq \frac{C'}{c_\sigma \sigma},$$

where C' is independent of σ .

What happens in the well The general strategy here is to bound $B_\sigma(x)$ by studying only the numerator. The denominator can be (very) roughly bounded by $\beta(1/2)$ (which does not depend on σ). The partition function disappears, and we get:

$$B_\sigma(x) \leq C \int_{m_\sigma}^x e^{V(y)/\sigma} dy \times \int_x^\infty e^{-V(y)/\sigma} dy.$$

We need a bound on V near the median: under our hypotheses, since μ_σ converges weakly to δ_{x_1} , the continuity of V in x_1 yields (for σ small enough):

$$\forall x \in [(x_1, m_\sigma)], \quad V(x) \leq d^*/4.$$

Now we can bound the first integral in the following way:

$$\int_{m_\sigma}^x e^{V(y)/\sigma} \leq (K_2 - K_1) \exp\left(\frac{1}{\sigma} \max(s(x), d^*/4)\right),$$

where $d^*/4$ takes care of the case when m_σ is less than x_1 .

We cut the second integral in two parts:

$$\int_x^\infty e^{-V_\sigma(y)} dy \leq \int_x^{K_2} e^{-V_\sigma(y)} dy + \int_{K_2}^\infty e^{-V_\sigma(y)} dy.$$

Since V is strictly increasing after K_2 , we may apply Laplace's method to the second term. In the first one, we use a rough bound on V :

$$\int_x^\infty e^{-V_\sigma(y)} dy \leq (K_2 - K_1) \exp\left(-\frac{i(x)}{\sigma}\right) + C \exp\left(-\frac{V(K_2)}{\sigma}\right).$$

Since $i(x)$ is less than $V(K_2)$, the second term is less than the first one (up to a constant), and there exists C' such that:

$$\int_x^\infty e^{-V(y)} dy \leq C' \exp\left(-\frac{i(x)}{\sigma}\right).$$

Coming back to B_σ , we get:

$$\begin{aligned} B_\sigma(x) &\leq C'' \exp\left(\frac{1}{\sigma} (\max(s(x), d^*/4) - i(x))\right) \\ &\leq C''' \exp\left(\frac{d^*}{\sigma}\right). \end{aligned}$$

Conclusion : An upper bound on β Let us now gather the bounds on $B_\sigma(x)$ we derived in the preceding paragraphs.

84 Lemma. *There exists a C (independent of σ) such that, for all σ ,*

$$B_\sigma = \sup_{x \geq m_\sigma} B_\sigma(x) \leq C \exp\left(\frac{d^*}{\sigma}\right).$$

With this result in hand, we may apply Barthe, Cattiaux and Roberto's result (theorem 81): this proves the first claim of theorem 80.

The modified function α_σ is deduced from β_σ with the help of theorem 97 (see below, in section 4.4).

Finally, the growth hypothesis on β (U4) guarantees that, near 0, β is bounded by a power of $\ln(1/s)$; this immediately implies the last result, and concludes the proof.

4.3 The weak inequality in any dimension

We now turn to the proof of the weak inequality (the bound in lemma 79) in any dimension. We are going to need one more hypothesis on the structure of potential wells, to avoid "pathological" cases.

After that, we proceed in several steps. First we recall our aim and explain the main lines of the proof. During this proof, a certain "path" (in fact, an open set of \mathbb{R}^d) will appear. It will be used to derive a "capacity-measure" inequality. Eventually, we will go from this inequality to the one we seek, using a result from next section.

The last hypothesis on the potential

To write down the last hypothesis we shall make on V , we first need a few more notations.

For all $x \in \mathbb{R}^d$, we call Γ_x the set of paths from x to 0. For each such γ (γ is a continuous function from $[0, 1]$ into \mathbb{R}^d), we call $h(\gamma)$ the "height" of γ , *i.e.* the highest value taken by V along γ :

$$h(\gamma) = \sup_{t \in [0, 1]} V(\gamma(t)).$$

Now suppose we try to go from x to 0 while remaining as low as possible (*i.e.* we try to find a path where V is small). There is a minimum price to pay; whatever path we choose, we will necessarily go at least as high as:

$$h(x) = \inf_{\gamma \in \Gamma_x} h(\gamma).$$

We will call “good paths” the ones that stay below that minimal height:

$$\gamma \text{ is good} \Leftrightarrow h(\gamma) = h(x).$$

A priori, for a given x , a good path from x to 0 need not exist: it may well be the case that, if one tries to find γ such that $h(\gamma) \leq h(x) + 1/n$, one has to go farther and farther as n grows, and that no finite path achieves the infimum bound.

Finally, the height of the “potential barrier” between x and the global minimum will be called $d^*(x)$:

$$d^*(x) = h(x) - V(x),$$

and the height of the biggest barrier will be just d^* :

$$d^* = \sup_x d^*(x).$$

Hypothesis 5. *The potential barriers have a bounded height:*

$$d^* < \infty.$$

Moreover, each point can reach 0 by a “relatively short” good path. More precisely, there exists a function R (a maximal radius), from \mathbb{R}^d to \mathbb{R} , which satisfies the following conditions:

- For all x , the ball centered in zero and of radius $R(|x|)$ contains a good path for x :

$$\forall x, \exists \gamma \in \Gamma_x, \begin{cases} \gamma([0, 1]) & \subset \mathcal{B}_{R(|x|)} \\ \gamma & \text{is good.} \end{cases}$$

- The function R grows like a power of the distance to the origin:

$$R(|x|) \leq c_R |x|^{d_R}.$$

The one-point weak inequality

As was said before, we will not prove in this section a full weak Poincaré inequality, *i.e.* we will not get (4.1) for all r . Instead, we just prove it for a specific value of r , namely the $r = r_t = (\ln t)^{-2} (\ln \ln t)^{-8}$ (*cf.* equation (4.4)). Since $\sigma(t) = c/(\ln t)$, we note that, for any $m > 2$,

$$\begin{aligned} r_t &\geq C \frac{\sigma^2}{\ln(\sigma)^8} \\ &\geq C' \sigma^m, \end{aligned}$$

for some C, C' and σ small enough. Therefore, and since α_t decreases, it suffices to prove an inequality with σ^m instead of r_t .

More precisely, we will get:

85 Theorem. *Let m be such that $2 < m < 1 + m_V$, and let D^* be a constant, $D^* > d^*$. Then there exists a C_m such that, for all σ , the measure μ_σ satisfies the following one-point weak Poincaré inequality*

$$\forall f, \mathbf{Var}_{\mu_\sigma}(f) \leq C_m \exp\left(\frac{D^*}{\sigma}\right) \int |\nabla g|^2 d\mu_\sigma + \sigma^m \|f - m_f\|_\phi^2.$$

where m_f is a median of f under μ_σ .

As was noted before, this entails

$$\alpha_t(r_t) \leq \alpha_t(\sigma^m) \leq C_m \exp(D^* / \sigma),$$

which is the result of lemma 79.

The end of the section is devoted to the proof of the theorem. It can be sketched as follows.

The idea is to use a capacity-measure criterion restricted to certain sets (large enough sets). Intuitively, if a set A has a large μ_σ mass, it must contain points near the origin; and these points are the important ones, for measuring capacity as well as mass. For these sets, located near the origin, everything should behave as in the compact case, and the inequality should depend on σ in the same way as when a Poincaré inequality holds.

Let us fix D^* , strictly bigger than d^* . As was just said, we would like to compare the capacity and measure of large enough sets: let $\kappa > 0$ be the minimum mass we will consider (κ will depend on σ). Let A be a Borel set such that:

$$\mu_\sigma(A) \geq 2\kappa(\sigma).$$

Restricting ourselves to these large sets localizes the problem in some sense. To be more precise, we introduce two radii. The first one, r_σ , is such that:

$$\mu_\sigma(\mathcal{B}_{r_\sigma}) \geq 1 - \kappa.$$

The second one is deduced from it: it is a radius big enough to include good paths (*cf.* hypothesis 5) starting from any point in the small ball \mathcal{B}_{r_σ} .

$$R_\sigma = R(r_\sigma).$$

These two quantities depend on σ and κ ; we will see that, for our choice of κ , r_σ and R_σ will not grow too fast as σ goes to zero.

Let $A = A' \cup A''$, where $A' = A \cap \mathcal{B}_{r_\sigma}$ and A'' is the complement set. Since $\mu_\sigma(A) \geq 2\kappa$ and $\mu_\sigma(A'') \leq \kappa$ (by definition of r_σ), $\mu_\sigma(A') \geq \kappa$, and:

$$\mu_\sigma(A) = \mu_\sigma(A') + \mu_\sigma(A'') \leq 2\mu_\sigma(A').$$

Intuitively, we need only consider the subset A' , because it concentrates enough mass.

At this point, our set A' may still be very complicated. In particular, it could be scattered all over the ball \mathcal{B}_{r_σ} . To avoid this, we will once again restrict ourselves to a subset, trying to keep enough mass in the process.

This is done by cutting \mathcal{B}_{r_σ} into small cubes. The bound on the gradient of V (hypothesis 3) helps us choose a good mesh, such that V does not vary too much inside a little cube.

86 Proposition. *For all η , there exists ϵ (depending only on V and η), such that, on each cube B with radius ϵ ,*

$$\sup_B V - \inf_B V \leq \eta.$$

The parameter η will be chosen later.

So we cut \mathcal{B}_{r_σ} into many little cubes of radius ϵ . This requires a certain number of cubes, which we call n_σ . We then have:

$$\mathcal{B}_{r_\sigma} = B_1 \cup B_2 \dots B_{n_\sigma}. \quad (4.10)$$

In the same way, N_σ will be the number of cubes necessary to cover \mathcal{B}_{R_σ} . We denote by A_i the intersection of A and B_i . We apply the pigeonhole principle to say that one of the A_i 's must be large enough:

$$\exists i_0, \mu_\sigma(A_{i_0}) \geq \frac{1}{n_\sigma} \mu_\sigma(A).$$

To sum up our considerations on sets, for each A , we have found a subset A_{i_0} such that:

- A_{i_0} is a subset of a cube of radius ϵ ,
- A_{i_0} is not too far from the origin ($A_{i_0} \subset \mathcal{B}_{r_\sigma}$)
- A_{i_0} is big enough compared to A : $\mu_\sigma(A_{i_0}) \geq \frac{1}{2n_\sigma} \mu_\sigma(A)$.

In some sense, we need only consider the case when A looks like a ball and is not too far from the origin. We are going to see how this can be used to build a certain path between A_{i_0} and 0, and from this path, deduce a capacity-measure inequality.

Building a path and straightening it out

Recall that our goal is to compare the capacity and the measure of sets, and more precisely to bound the capacity from below and the measure from above.

The capacity is defined by an infimum bound:

$$\text{Cap}_\mu(A) = \inf \left\{ \int |\nabla f|^2 d\mu, \mathbf{1}_A \leq f \leq 1, \mu(\text{supp } f) \leq \frac{1}{2} \right\}, \quad (4.11)$$

where $\text{supp } f$ is the support of f . Note that we only define capacities for sets whose measure is less than $1/2$. This restriction explains why we use function recentered by their median when we deduce functional inequalities from capacity-measure criteria.

Since we seek a bound from below, we consider a function satisfying the conditions, and we try to bound:

$$\int |\nabla f|^2 d\mu.$$

The key idea is to find out a region of \mathbb{R}^d which should contribute a lot to this integral. Since the function f equals 1 near A , and 0 near 0 (the measure of its support being less than $1/2$), there must be a transition between A and 0: this is where the gradient of f appears. Still on the intuitive level, if the integral is to be small, we had better make this transition in a region where μ has less mass, *i.e.* in a zone where V is large. This is the reason why we introduced the good paths: to go from A to zero, a large contribution to the energy should appear along these good paths.

To put these ideas on a firm ground, we will build, starting from A (or more precisely from A_i), an open set \mathcal{C}_A with good regularity properties, and then bound the capacity by integrals over this open set. This construction is depicted in figure 4.2.

Once this set is built, we proceed in two steps. First, for all function f satisfying the conditions of (4.11),

$$\int |\nabla f|^2 d\mu_\sigma \geq \int |\nabla f|^2 \mathbf{1}_{\mathcal{C}_A} d\mu_\sigma$$

On the path, we know by design that V is bounded above by $V(x_A^*) + \eta$. Indeed, V is less than $V(x_A^*)$ along γ , and the size ϵ of the cubes has been chosen so that on each cube, the oscillation

of V is less than η . Therefore, we may compare our integral with an integral with respect to the Lebesgue measure.

$$\int |\nabla f|^2 d\mu_\sigma \geq \frac{1}{Z_\sigma} \exp\left(-\frac{V(x_A^*) - \eta}{\sigma}\right) \int |\nabla f|^2 \mathbf{1}_{\mathcal{C}_A} d\lambda. \quad (4.12)$$

The next step is to bound the latter integral on \mathcal{C}_A . Our only hypotheses is that f must be 1 on A , and 0 near zero. The idea is then to apply a Poincaré inequality to compare the energy to a variance. Unfortunately, though we know that a Poincaré inequality effectively holds under quite general assumptions for a bounded domain in \mathbb{R}^d (this is proved in many textbooks on partial differential equations, see *e.g.* [Eva98], p. 275–276), the explicit constants and their behaviour when the domain changes is not well known. However, there is a case for which we have such explicit estimates, namely the case of convex domains.

87 Theorem (Poincaré inequality for convex domains). *Let L be a convex bounded domain in \mathbb{R}^d . Then λ_L , the normalized Lebesgue measure on L , satisfies a Poincaré inequality, and the constant can be bounded above using only the diameter d_L of the domain:*

$$\mathbf{Var}_{\lambda_L}(f) = \int \left(f - \int f d\lambda_L\right)^2 d\lambda_L \leq \frac{d_L^2}{\pi^2} \int |\nabla f|^2 d\lambda_L.$$

This theorem is proved *e.g.* by Payne and Weinberger, and Bebendorf in [PW60, Beb03]. Note that other bounds in more complicated cases have been derived (see [CL97] for star-shaped domains, or [Che90] for bounds depending on the geometry of the boundary).

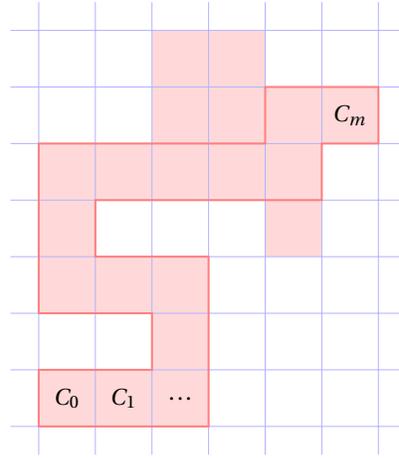
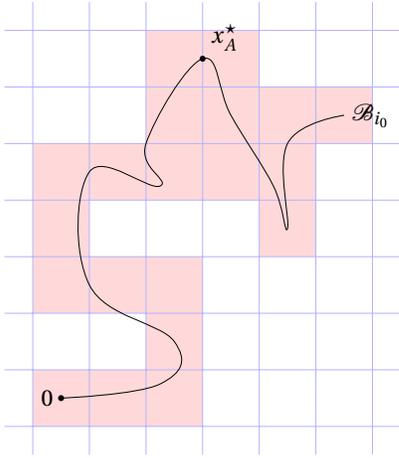
In order to use this result, we try to “straighten out” the set \mathcal{C}_A .

We will build a function ϕ sending \mathcal{C}_A to a tube \mathcal{L}_A . This function will be defined piecewise, on each of the little cubes that \mathcal{C}_A crosses. Let us denote these cubes as C_0, \dots, C_m . It is easy to see that the intersection of \mathcal{C}_A and one of these cubes can only take a finite number of shapes (up to a rotation and/or translation). In $d = 2$ for example, only two different shapes are possible (either a straight tube or a bended one, see figure 4.2). Each of these shapes may be “straightened out” into a tube by a diffeomorphism. We have to be a bit careful in choosing these diffeomorphisms ϕ_j (one for each shape). We will ask two things: they should behave like a rigid motion in the neighborhood of the edges (so we may “glue” two transformations together), and their Jacobian matrix should be sufficiently “nice” (the “niceness” needed will be made precise later). Such a choice is possible; see the figure 4.2 for an explanation of a possible way to find such good functions.

Once this is done, we only have to glue our pieces together. Let us denote the pieces $\mathcal{C}_A \cap C_i$ by T_i . We leave T_0 where it stands, and look at T_1 . We have seen that it may be straightened into a tube, T'_1 : define ϕ on T_1 to be precisely this transformation. Now consider T_2 : we can straighten it by one of our ϕ_j , and then use a rotation and/or a translation to put it next to T'_1 . Since we have asked that the ϕ_j should be rigid motions near the edges, the two pieces of ϕ define a diffeomorphism from $T_1 \cup T_2$ to the straight tube $T'_1 \cup T'_2$. We may iterate the process and eventually we get a diffeomorphism ϕ from \mathcal{C}_A to \mathcal{L}_A . One can see on the figure that a little extra care is needed to deal with the end of the path \mathcal{C}_A — however, adding just one ϕ_j to our set of transformations settles the question.

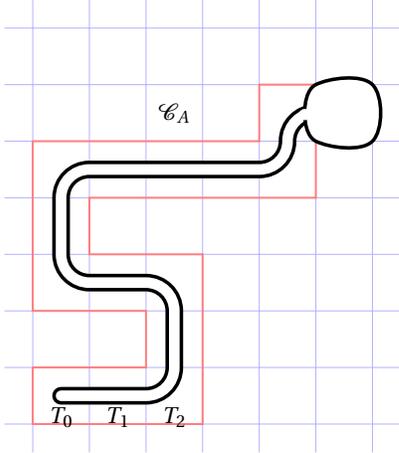
Remember that our goal is to use the Poincaré inequality on the convex set \mathcal{L}_A . For this to work, we need to control some quantities related to the map ϕ .

Figure 4.2: Building the path \mathcal{L}_A

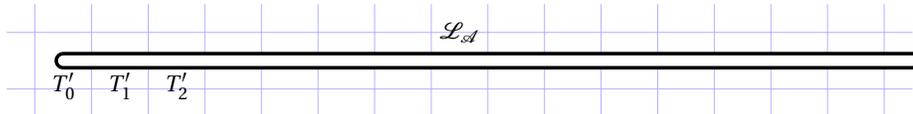


1. We consider a good path starting from the center x_A of the cube \mathcal{B}_{i_0} , and going to the origin. On this path, V reaches its maximum at some x_A^* , and on the colored region, V is bounded above by $V(x_A^*) + \eta$.

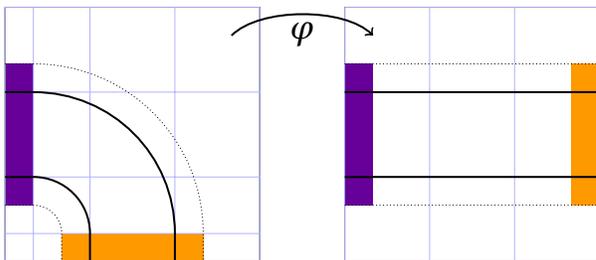
2. We pick a “path of cubes” from 0 to \mathcal{B}_{i_0} which stays entirely within the colored region.



3. Within this path, we draw a smooth tube \mathcal{C}_A . The intersection of \mathcal{C}_A and a given little cube may only take a finite number of shapes (up to a rigid motion); in this 2-dimensional drawing for example, we have either a straight tube (T_1) or a bended one (T_2). For technical reasons, we consider two more shapes at the end of the tube so that \mathcal{B}_{i_0} lies entirely within \mathcal{C}_A .



4. Finally, the tube \mathcal{C}_A is sent onto \mathcal{L}_A , a convex set for which we have an explicit Poincaré inequality.



The bended tube on the left may be straightened as shown. We consider a diffeomorphism which sends the regions between dotted lines on one another, and ask that it should be a rigid motion on the dark regions. Defining the transformation on a set (the region between dotted lines) larger than the tube (the region between plain lines) gives compactness bounds on the Jacobian.

88 Proposition. *There exists a constant C_ϕ , which may depend on ϵ but not on σ , such that, at every point, the Jacobian matrix J_ϕ satisfies*

$$\frac{1}{C_\phi} \leq |\det(J_\phi)| \leq C_\phi,$$

$$\lambda_1(J_\phi^t J_\phi) \geq C_\phi^{-1},$$

where $\lambda_1(M)$ is the smallest eigenvalue of the symmetric matrix M .

Proof. This holds by design of the map ϕ . At each point, ϕ is the composition of a rigid motion (which has no effect on the eigenvalues or the determinant of the Jacobian matrix), and of one of the ϕ_j . For a given ϕ_j , the properties hold: we have designed the ϕ_j as restrictions of diffeomorphisms on larger sets, so the bounds hold by compactness. Since there is a finite number of ϕ_j , we may choose bounds that do not depend on j . This proves that the bounds hold for ϕ . \square

We may now give our “straightening” its rigorous form, namely a change of variables.

89 Proposition. *Let \mathcal{U} and \mathcal{V} be open sets, let ϕ be a diffeomorphism from \mathcal{U} onto \mathcal{V} . If the inequalities in the preceding lemma hold with a constant C_ϕ , then for all continuously differentiable function f on \mathcal{U} , we have:*

$$\int_{\mathcal{U}} |\nabla f|^2 d\lambda \geq C_\phi^{-2} \int_{\mathcal{V}} |\nabla g|^2 d\lambda.$$

where $g = f \circ \phi^{-1}$.

Proof. Let us define F by $F(x) = |\nabla f|^2(x)$. Then, by a change of variables, and thanks to the upper bound on $|\det(J_\phi)|$,

$$\begin{aligned} \int_{\mathcal{U}} |\nabla f|^2 dx &= \int_{\mathcal{U}} F(x) dx = \int_{\mathcal{V}} F \circ \phi^{-1} \left| \det(J_\phi \circ \phi^{-1})^{-1} \right| dy \\ &\geq \frac{1}{C_\phi} \int_{\mathcal{V}} F \circ \phi^{-1} dy. \end{aligned}$$

Since $f = g \circ \phi$, the gradients are given by:

$$(\nabla f)_x = {}^t J_\phi)_x (\nabla g)_{\phi(x)}.$$

Taking norms, and using the lower bound on the first eigenvalue, we get:

$$(|\nabla f|^2)_x = {}^t \nabla g J_\phi^t J_\phi \nabla g \geq C_\phi^{-1} (|\nabla g|^2)_{\phi(x)}.$$

Rewriting this in y variables,

$$F \circ \phi^{-1}(y) = (|\nabla f|^2)_{\phi^{-1}(y)} \geq C_\phi^{-1} (|\nabla g|^2)_y.$$

Finally,

$$\int_{\mathcal{U}} |\nabla f|^2 \geq \frac{1}{C_\phi^2} \int_{\mathcal{V}} |\nabla g|^2.$$

\square

Putting the last two propositions together, we can show:

90 Proposition. *There exists a C , depending only on ϵ , such that if f satisfies the following conditions:*

1. f is continuously differentiable from \mathcal{C}_A into $[0, 1]$,
2. $\lambda(\{f = 0\}) \geq l_0$,
3. $\lambda(\{f = 1\}) \geq l_1$,

then

$$\int |\nabla f|^2 d\lambda \geq \frac{C}{N_\sigma^2} \min(l_0, l_1).$$

We recall that N_σ is the number of balls of radius ϵ needed to cover the big ball \mathcal{B}_{R_σ} .

Proof. Suppose f satisfies the hypotheses. Define $g = f \circ \phi^{-1}$ as in the preceding proposition. The various bounds needed on the Jacobian matrix of ϕ are provided by proposition 88. The lower bound on $\det(J_\phi)$ implies that g must vanish at least on a set of Lebesgue measure $C_\phi^{-1}l_0$, the same being true for the set where $g = 1$. The change of variables has shown:

$$\int |\nabla f|^2 d\lambda \geq C_\phi^{-2} \int |\nabla g|^2 d\lambda.$$

On the right hand side, we can now use the Poincaré inequality:

$$\int |\nabla g|^2 d\lambda \geq \frac{\lambda(\mathcal{L}_A)}{C_P(\mathcal{L}_A)} \mathbf{Var}_{\lambda_{\mathcal{L}_A}}(g),$$

where $\lambda_{\mathcal{L}_A}$ is the normalized Lebesgue measure on \mathcal{L}_A . The very purpose of our change of variables was to make the domain convex, so we could make use of theorem 87. The constant may therefore be bounded by the square of the diameter of \mathcal{L}_A . Since \mathcal{L}_A results from gluing together at most N_σ little cubes of radius ϵ , the square of the diameter may be bounded by $\epsilon^2 N_\sigma^2$.

We now turn to the variance, and use the information on the sets where g is 0 or 1. We denote by l'_0, l'_1 the respective (non-normalized) measures of these sets, and by m the mean of g ($m = \int g d\lambda_{\mathcal{L}_A} \in [0, 1]$). Then:

$$\lambda(\mathcal{L}_A) \times \mathbf{Var}(g) = \int (g - m)^2 d\lambda \geq m^2 l'_0 + (1 - m)^2 l'_1.$$

The right hand side is easily shown to be greater than $l'_0 l'_1 / (l'_0 + l'_1)$. The latter is bounded below by half the minimum of l'_0 and l'_1 (because the numerator is less than $2 \max(l'_0, l'_1)$). Since $l'_0 \geq C_\phi^{-1}l_0$, and the similar result holds for l'_1 ,

$$\int |\nabla f|^2 d\lambda \geq \frac{C_\epsilon}{N_\sigma^2} \min(l_0, l_1). \quad \square$$

We may now prove the measure-capacity inequality we are looking for. Indeed, recall that our aim is to bound the capacity of a set A from below by a function of its measure. The previous inequality is almost what we want: on the left hand side is (up to a factor, see (4.12) above) the quantity whose infimum gives the capacity (equation (4.11)), and on the right hand side l_0 and l_1 are measures of some sets. It remains to show that these measures may be compared to the measure of A .

The measure-capacity inequality

Let us put together the results from the previous section (equation (4.12) and proposition 90)

$$\begin{aligned} \int |\nabla f|^2 d\mu_\sigma &\geq \frac{1}{Z_\sigma} \exp\left(-\frac{V(x_A^*) - \eta}{\sigma}\right) \int |\nabla f|^2 d\lambda \\ &\geq \frac{C_\epsilon}{Z_\sigma N_\sigma^2} \exp\left(-\frac{V(x_A^*) - \eta}{\sigma}\right) \min(l_0, l_1), \end{aligned} \quad (4.13)$$

where l_0, l_1 are the Lebesgue measure of the following sets:

$$l_0 = \lambda(\{f = 0\} \cap \mathcal{C}_A) \quad l_1 = \lambda(\{f = 1\} \cap \mathcal{C}_A).$$

To bound l_0 , we use the fact that f vanishes on a sufficiently large set (as measured by μ_σ). Since μ_σ concentrates around 0, f should vanish near the origin. More precisely, for a fixed ϵ , we know that for σ small enough, the cube centered in 0 and of radius ϵ concentrates 3/4 of the measure. If this cube is labelled B_0 , we have:

$$\mu_\sigma(\{f = 0\} \cap B_0) \geq \frac{1}{4}.$$

Since V is non negative, μ_σ and λ are easily compared.

$$\begin{aligned} \frac{1}{4} &\leq \mu_\sigma(\{f = 0\} \cap B_0) = \frac{1}{Z_\sigma} \int \mathbf{1}_{f=0} \mathbf{1}_{B_0} \exp\left(-\frac{V}{\sigma}\right) d\lambda \\ &\leq \frac{1}{Z_\sigma} \int \mathbf{1}_{f=0} \mathbf{1}_{B_0} d\lambda \end{aligned}$$

The integral on the right hand side is less than l_0 , therefore:

$$l_0 \geq m_0 = \frac{Z_\sigma}{4}.$$

Let us derive a similar bound, m_1 , for l_1 . On the cube B_{i_0} , $V \geq V(x_A) - \eta$ (recall x_A is the center of B_{i_0}), so:

$$\begin{aligned} \mu_\sigma(A_{i_0}) &= \frac{1}{Z_\sigma} \int \mathbf{1}_{A_{i_0}} \exp\left(-\frac{V}{\sigma}\right) d\lambda \\ &\leq \frac{1}{Z_\sigma} \int \mathbf{1}_{A_{i_0}} \exp\left(-\frac{V(x_A)}{\sigma} + \frac{\eta}{\sigma}\right) d\lambda \\ &\leq \frac{1}{Z_\sigma} \exp\left(\frac{-V(x_A) + \eta}{\sigma}\right) \lambda(A_{i_0}). \end{aligned}$$

Therefore:

$$l_1 \geq \lambda(A_{i_0}) \geq m_1 = Z_\sigma \exp\left(\frac{V(x_A)}{\sigma} - \frac{\eta}{\sigma}\right) \mu_\sigma(A_{i_0}). \quad (4.14)$$

Since we would like to control $\min(l_0, l_1)$, we now have to compare the two bounds m_0 and m_1 . This is possible thanks to the following inequality:

$$\mu_\sigma(A_{i_0}) \leq \frac{1}{Z_\sigma} \exp\left(\frac{-V(x_A) + \eta}{\sigma}\right) \epsilon^d.$$

If we gather almost all terms on the left hand side, we recognize m_1 :

$$m_1 \leq \epsilon^d.$$

Since $m_0 = Z_\sigma/4$, it holds that $m_0 \geq Z_\sigma m_1 \epsilon^{-d}$, and since Z_σ goes to zero, it also holds that $m_1 \geq Z_\sigma m_1 \epsilon^{-d}$, so that both l_0 and l_1 may be bounded below by this quantity:

$$\min(l_0, l_1) \geq \frac{Z_\sigma^2}{\epsilon^d} \exp\left(\frac{V(x_A)}{\sigma} - \frac{\eta}{\sigma}\right) \mu_\sigma(A_{i_0}).$$

Going back to (4.13), we conclude:

$$\begin{aligned} \int |\nabla f|^2 d\mu_\sigma &\geq \frac{C_\epsilon}{Z_\sigma N_\sigma^2} \exp\left(-\frac{V(x_A^*) - \eta}{\sigma}\right) \min(l_0, l_1) \\ &\geq \frac{C'_\epsilon Z_\sigma}{N_\sigma^2} \exp\left(\frac{V(x_A) - V(x_A^*) - 2\eta}{\sigma}\right) \mu_\sigma(A_{i_0}). \end{aligned}$$

By definition of x_A^* , $V(x_A) - V(x_A^*) - 2\eta \geq -d^* - 2\eta \geq -D^*$. On the other hand, A_{i_0} was chosen precisely because it contained enough of A 's mass: $\mu_\sigma(A_{i_0}) \geq (2n_\sigma)^{-1} \mu_\sigma(A)$. Finally, every function f we can choose in the definition of capacity must satisfy:

$$\int |\nabla f|^2 d\mu_\sigma \geq \frac{C_\epsilon Z_\sigma}{2N_\sigma^2 n_\sigma} \exp\left(-\frac{D^*}{\sigma}\right) \mu_\sigma(A).$$

Taking the infimum over all possible f finally yields the following result.

91 Proposition. *Let $\kappa(\sigma)$ be a positive number, less than 1/2. Let n_σ, N_σ be defined as in the discussion near equation (4.10). Then the following bound holds:*

$$\forall A, \mu_\sigma(A) \geq \kappa(\sigma) \quad \implies \quad \mu_\sigma(A) \leq \frac{N_\sigma^2 n_\sigma}{C'_\epsilon Z_\sigma} \exp\left(\frac{D^*}{\sigma}\right) \text{Cap}_{\mu_\sigma}(A). \quad (4.15)$$

Conclusion

The bigger part of the proof has now been done; the last thing we need to check is that the number of balls n_σ and N_σ do not grow too fast as σ decreases. Then we will apply theorem 96 to deduce the one-point inequality of theorem 85 from our measure-capacity inequality.

Recall that we are given a real number m , strictly smaller than $1 + m_V$. Define $\kappa(\sigma) = \exp(-\frac{1}{\sigma^m})$. We want to find an r_σ such that the mass of \mathcal{B}_{r_σ} is greater than $1 - \kappa$. For any set A , we may write:

$$\begin{aligned} \mu_\sigma(A) &= \frac{1}{Z_\sigma} \int \mathbf{1}_A \exp\left(-\frac{V}{\sigma}\right) \\ &= \frac{Z_{2\sigma}}{Z_\sigma} \times \frac{1}{Z_{2\sigma}} \int \mathbf{1}_A \exp\left(-\frac{V}{2\sigma} - \frac{V}{2\sigma}\right) \\ &= \frac{Z_{2\sigma}}{Z_\sigma} \times \int \mathbf{1}_A \exp\left(-\frac{V}{2\sigma}\right) d\mu_{2\sigma}. \end{aligned}$$

If V takes large values on A , we can get a good bound:

$$\mu_\sigma(A) \leq \frac{Z_{2\sigma}}{Z_\sigma} \exp\left(-\frac{\inf_A V}{2\sigma}\right) \mu_{2\sigma}(A).$$

We get rid of the $\mu_{2\sigma}(A)$ by roughly bounding it by 1. Then we use the growth hypothesis on V (hyp. 2), with $A = \mathcal{B}_{r_\sigma}^c$: we get $\inf_A V \geq \ln(r_\sigma)^{m_V} - C$, so that, for any $m'_V < m_V$

$$\inf_A V \geq \ln(r_\sigma)^{m'_V}$$

4.4. A MEASURE-CAPACITY CRITERION FOR ONE-POINT WEAK POINCARÉ INEQUALITIES

whenever r_σ is big enough. We fix an $m' \in]m, 1 + m'_V[$, and choose:

$$r_\sigma = \exp\left(\left(\frac{1}{\sigma}\right)^{(m'-1)/m'_V}\right),$$

which ensures, for σ small enough:

$$\inf_A V \geq \left(\frac{1}{\sigma}\right)^{(m'-1)},$$

$$\mu_\sigma(\mathcal{B}_{r_\sigma}^c) \leq \frac{Z_{2\sigma}}{Z_\sigma} \exp\left(-\frac{1}{2\sigma^{m'}}\right)$$

The asymptotic behavior of Z_σ (cf. annex 4.C) implies that $Z_{2\sigma}/Z_\sigma$ converges, and since $m' > m$,

$$\mu_\sigma(\mathcal{B}_{r_\sigma}^c) \leq \exp\left(-\frac{1}{\sigma^m}\right)$$

for σ small enough. This shows that r_σ satisfies the condition we wanted.

We may now end the proof of the theorem. Coming back to the measure-capacity inequality (4.15), we note that R_σ , n_σ and N_σ all behave like r_σ to a certain power (for R_σ we use hypothesis 5, and n_σ, N_σ are just a number of cubes of fixed radius in the big cubes of side length r_σ and R_σ). Therefore, there exists a C such that

$$\forall A, \mu_\sigma(A) \geq \kappa(\sigma) \quad \Rightarrow \quad \mu_\sigma(A) \leq \frac{r_\sigma^C}{Z_\sigma} \exp\left(\frac{D^\star}{\sigma}\right) \text{Cap}_{\mu_\sigma}(A). \quad (4.16)$$

The value of r_σ and the fact that $m' - 1$ is strictly less than m'_V makes $\exp(D^\star/\sigma)$ the biggest term (recall that $Z_\sigma \sim C\sigma^{d/2}$), so that, up to a slight increase of D^\star ,

$$\forall A, \mu_\sigma(A) \geq \kappa(\sigma) \quad \Rightarrow \quad \mu_\sigma(A) \leq \exp\left(\frac{D^\star}{\sigma}\right) \text{Cap}_{\mu_\sigma}(A).$$

This inequality, thanks to theorem 96 below, implies precisely the one-point weak Poincaré inequality we claimed in theorem 85.

4.4 A measure-capacity criterion for one-point weak Poincaré inequalities

Definitions

In this section we study the interplay between weak Poincaré inequalities and measure-capacity inequalities. Let us start by recalling exactly what a weak Poincaré inequality is.

92 Definition (M. Röckner and F.Y. Wang, [RW01]). *Let μ be a measure and \mathcal{N} be a norm, stronger than the $L^2(\mu)$ norm. The measure μ is said to satisfy a weak Poincaré inequality for the norm \mathcal{N} if there exists a decreasing positive function α , defined on \mathbb{R}_+^\star such that:*

$$\forall f \in L^2(\mu), f \text{ such that } \mu f = 0, \forall r > 0, \mu(f^2) \leq \alpha(r)\mathcal{E}(f, f) + r\mathcal{N}(f)^2$$

If this holds, α will be called a compensating function.

Remark (on means and medians). *The original statement on weak Poincaré inequalities involves functions recentred by their mean value $\mu(f)$, and an L^∞ norm. However, the approach by measure-capacity inequalities developed in [BCR05, BCR] works with functions recentred by their median m_f . When the norm is the sup norm, it is easy to go from one to the other: the three quantities $\text{osc}(f)$, $\|f - m_f\|_\infty$ and $\|f - \mu(f)\|_\infty$ are within (universal) bounds of each other.*

Since we need to work with another norm, we will show that we can still go from $\mathcal{N}(f - m_f)$ to $\mathcal{N}(f - \mu f)$ (cf. equation (4.20) in annex 4.A).

This is equivalent to the slightly modified definition:

93 Proposition. *A weak Poincaré inequality holds if and only if:*

$$\forall r > 0, \exists c_r, \forall f \in L^2(\mu), \quad \mu(f) = 0 \implies \mu(f^2) \leq c_r \mathcal{E}_\mu(f, f) + r \mathcal{N}(f)^2. \quad (4.17)$$

If the inequality holds for a given couple (r, c_r) , we will say that μ satisfies a one-point weak Poincaré inequality.

Therefore the weak Poincaré inequality holds if and only if a one-point inequality holds for each point r .

Proof. The only thing to check is that we can deduce the inequality of the definition from (4.17). To each r , we associate c_r according to (4.17). Then we just define $\alpha(r) = \inf\{c_s; s \leq r\}$. The function α is decreasing. Now let f be a function in L^2 and $r > 0$. For any ϵ , we may find an $s \leq r$ such that:

$$c_s \leq \alpha(r) + \epsilon.$$

If we apply (4.17) with this s , we get (since $s \leq r$):

$$\begin{aligned} \mu(f^2) &\leq c_s \mathcal{E}(f) + s \mathcal{N}(f)^2 \\ &\leq \alpha(r) \mathcal{E}(f) + r \mathcal{N}(f)^2 + \epsilon \mathcal{E}(f). \end{aligned}$$

Since this is true for any ϵ , we may let it go to zero, and we have found a function α . \square

We will be specifically interested in these inequalities for one special norm. We now define this norm and recall some of its properties, without proofs. For a short introduction (with the results we need here), see e.g. [Ale04]; for an extensive treatment we refer to [RR91].

Let ϕ, ψ be defined on \mathbb{R}_+ by $\psi(x) = x \log(1+x)$, $\phi(x) = \psi(x^2)$. For any measurable f , define the *Orlicz norm* (usually called the *Luxembourg norm*; there is another natural norm on the Orlicz space, which will not be needed here) of f to be:

$$\|f\|_\phi = \inf \left\{ \lambda, \int \phi \left(\frac{|f|}{\lambda} \right) \leq 1 \right\}.$$

Note that, with this definition, $\|1\|_\phi$ need not be equal to 1. The set of functions f for which this norm is finite is denoted L_ϕ , it is a vector space, and it is complete for the Orlicz norm. In the same way, if ψ^*, ϕ^* are the convex dual functions of ψ, ϕ , we may define the corresponding Orlicz spaces. It is easily seen that for every positive f , $\|f^2\|_\psi = \|f\|_\phi^2$. The dual functions allow us to state the following Hölder-like property.

94 Proposition (Hölder-Orlicz). *If f, g are two measurable functions, respectively in L_ψ and L_{ψ^*} , then fg is in L^1 , and*

$$\left| \int fg d\mu \right| \leq 2 \|f\|_\psi \|g\|_{\psi^*}.$$

4.4. A MEASURE-CAPACITY CRITERION FOR ONE-POINT WEAK POINCARÉ INEQUALITIES

The constant 2 is necessary because we work with Luxembourg norms. To conclude this account on Orlicz norm, we recall here the norm of an indicator function:

95 Proposition. *Let A be a measurable set. Then $\mathbf{1}_A$ is in the Orlicz space L_ψ and:*

$$\|\mathbf{1}_A\|_{\psi^\star} = \hat{\psi}(\mu(A)),$$

where $\hat{\psi}(x) = \frac{1}{(\psi^\star)^{-1}(1/x)}$. Moreover, for all x sufficiently small, and for $\psi(x) = x \log(1+x)$, we have the following bound:

$$\hat{\psi}(x) \leq \frac{2}{\log(1/x)}.$$

Proof. Once again we refer to [Ale04, RR91] for the first result. The explicit bound on $\hat{\psi}$ follows easily from the bound $\psi^\star \leq xe^x$ and the definition of $\hat{\psi}$. \square

Measure-capacity inequalities for large sets and one-point inequalities

Here we show the result which was used in the preceding section: if we can compare the measure and the capacity of large sets, we can deduce a one-point weak inequality.

96 Theorem. *Suppose that there exists $\kappa < 1/2$, and a real constant C_κ such that, for every set A whose measure is larger than κ , we have:*

$$\text{Cap}_\mu(A) \geq C_\kappa \mu(A). \tag{4.18}$$

Then μ satisfies the one-point weak Poincaré inequality:

$$\text{Var}_\mu(g) \leq \frac{c}{C_\kappa} \int |\nabla g|^2 d\mu + \kappa \text{osc}^2(g),$$

where c is universal. We may replace the L^∞ norm by an Orlicz norm, in which case the inequality reads:

$$\text{Var}_\mu(g) \leq \frac{c}{C_\kappa} \int |\nabla g|^2 d\mu + 4\hat{\psi}(\kappa) \|g - m_g\|_\psi^2.$$

Remark. Note that if (4.18) holds for all sets, regardless of their measure, then μ satisfies a (strong) Poincaré inequality (since we may take $\kappa = 0$). This is well-known, cf. [BCR] and references therein. This characterization of a functional inequality in terms of a relation between measures and capacities of sets is in fact more general, and provides a way to compare many functional inequalities. For a detailed account on these questions, and links with isoperimetric properties, we refer to [BCR] (especially section 5).

Proof. We follow the proof of theorem 2 in [BCR05] (which deals with the (full) weak inequality).

Let f be a function and m a median for f . We cut the space in half, according to whether f is greater than m or not; we denote by Ω_+, Ω_- the two sets. The integral may be written as:

$$\text{Var}_\mu(f) \leq \int (f - m)^2 d\mu = \int_{\Omega_+} (f - m)^2 d\mu + \int_{\Omega_-} (f - m)^2 d\mu.$$

We will show how to deal with the leftmost integral, the other one being similar.

Let us write $g = f - m$, and define c by:

$$c = \inf\{t \geq 0, \mu(g^2 > t) < \kappa\}.$$

If c is zero, then $\mu(g > 0)$ is less than κ , and:

$$\int_{\Omega_+} g^2 d\mu \leq \begin{cases} \kappa \sup g^2 \text{ in the } L^\infty \text{ case,} \\ \hat{\psi}(\kappa) \|f - m\|_\phi^2 \text{ in the Orlicz case,} \end{cases}$$

so the inequalities we are looking for hold in the half-space Ω_+ .

Thus we need only consider the case where c is strictly positive. By a continuity argument (μ will always have a density), we can find a set Ω_0 such that $\mu(\Omega_0) = \kappa$ and $\{g^2 > c\} \subset \Omega_0 \subset \{g^2 \geq c\}$. We fix a $\rho > 1$, and introduce the level sets $\Omega_k = \{g^2 \geq \frac{c}{\rho^k}\}$. We decompose the integral over these sets:

$$\begin{aligned} \int_{\Omega_+} g^2 &= \int_{\Omega_0} g^2 d\mu + \sum_{k \geq 1} \int_{\Omega_k \setminus \Omega_{k-1}} g^2 d\mu \\ &\leq \int_{\Omega_0} g^2 d\mu + \sum_{k \geq 1} \frac{c}{\rho^{k-1}} (\mu(\Omega_k) - \mu(\Omega_{k-1})) \end{aligned}$$

Put $\mu_k = \mu(\Omega_k)$, and apply the Abel transform to the sum:

$$\begin{aligned} \sum_{k \geq 1} \frac{1}{\rho^{k-1}} (\mu_k - \mu_{k-1}) &= \sum_{k \geq 1} \frac{\mu_k}{\rho^{k-1}} - \sum_{k \geq 0} \frac{\mu_k}{\rho^k} \\ &= \sum_{k \geq 1} \mu_k \left(\frac{1}{\rho^{k-1}} - \frac{1}{\rho^k} \right) - \mu_0. \end{aligned}$$

This is where we do not follow [BCR05]: since we simply suppose an inequality between capacity and measure, we can get rid of the μ_0 and write

$$\sum_{k \geq 1} \frac{1}{\rho^{k-1}} (\mu_k - \mu_{k-1}) \leq (\rho - 1) \sum_{k \geq 1} \frac{\mu_k}{\rho^k}.$$

The rest of the proof follows the same line as in [BCR05] — at this point, we use the measure-capacity inequality on each set Ω_k . They are designed to have their measure bigger than κ , so that we may apply our hypothesis:

$$\mu_k \leq \frac{1}{C_\kappa} \text{Cap}(\Omega_k).$$

Now, to bound the capacity from above, we apply the definition with well-chosen functions g_k :

$$g_k = \min \left(1, \left(\frac{g - \sqrt{c\rho^{-k-1}}}{\sqrt{c\rho^{-k}} - \sqrt{c\rho^{-k-1}}} \right)_+ \right)$$

This entails:

$$\begin{aligned} \mu_k &\leq \frac{1}{C_\kappa} \int |\nabla g_k|^2 d\mu \\ &\leq \frac{\rho^{k+1}}{C_\kappa c (\sqrt{\rho} - 1)^2} \int_{\Omega_k \setminus \Omega_{k-1}} |\nabla g|^2 d\mu. \end{aligned}$$

Summing over k , we get:

$$\int_{\Omega_+} g^2 d\mu \leq \int_{\Omega_0} g^2 d\mu + \frac{\rho(\rho-1)}{C_\kappa (\sqrt{\rho}-1)^2} \int |\nabla g|^2 d\mu.$$

4.4. A MEASURE-CAPACITY CRITERION FOR ONE-POINT WEAK POINCARÉ INEQUALITIES

We may now choose ρ ; the (non optimal) choice $\rho = 4$ gives:

$$\int_{\Omega_+} g^2 d\mu \leq \int_{\Omega_0} g^2 d\mu + \frac{12}{C_\kappa} \int |\nabla g|^2 d\mu.$$

The only thing left to do is to take care of the integral on Ω_0 . This is done with an Hölder-like inequality. In the Orlicz norm case, for example, we write:

$$\begin{aligned} \int_{\Omega_0} g^2 d\mu &\leq 2 \|g^2\|_\psi \| \mathbf{1}_{\Omega_0} \|_{\psi^*} \\ &\leq 2 \| (f - m)^2 \|_\psi \hat{\psi}(\kappa) \leq 2 \hat{\psi}(\kappa) \|f - m\|_\phi^2. \end{aligned}$$

thanks to the Hölder-Orlicz inequality and the relation between ϕ and ψ (see the beginning of this section). This concludes the proof (the extra factor 2 comes from the bound on Ω_- , and we can set the universal constant $c = 24$). \square

Weak inequalities for different norms

To conclude this section, let us state a corollary to the previous result, and prove that weak Poincaré inequalities for many different norms are in fact equivalent. Moreover, if a compensating function is known for one norm, we can immediately deduce a function for another norm; this result was used in the one dimensional case (section 4.2) where the explicit Hardy-like criteria were known for the L^∞ norm.

97 Theorem. *Let ϕ, ψ be two Young functions, with $\phi(x) = \psi(x^2)$. A measure μ satisfies a weak Poincaré inequality with the L^∞ norm if and only if it satisfies one with the Orlicz norm $\|\cdot\|_\phi$.*

Moreover, if β is a compensating function for the L^∞ norm, then the following function may be chosen for the Orlicz norm:

$$\alpha(s) = \frac{c}{4} \beta \left(\frac{1}{4} \hat{\psi}^{-1} \left(\frac{s}{2} \right) \right),$$

where c is universal ($c = 24$ works).

Proof. First, let us introduce a few notations. We will denote by $\mathbf{M-C}(\kappa, C(\kappa))$ the following comparison between measure and capacity:

$$\forall A, \mu(A) > \kappa \implies \text{Cap}(A) \geq C(\kappa) \mu(A).$$

Similarly, $\mathbf{PWP}(r, C(r), \mathcal{N})$ will denote the one-point weak Poincaré inequality for a norm \mathcal{N} with constants $r, C(r)$, and $\mathbf{WP}(\alpha, \mathcal{N})$ will be the (full) weak inequality, with a norm \mathcal{N} and a compensating function α . In the previous section, we showed:

$$\begin{aligned} \mathbf{M-C}(\kappa, C(\kappa)) &\implies \mathbf{PWP} \left(\kappa, \frac{c}{C_\kappa}, \|\cdot\|_\infty \right), \\ \mathbf{M-C}(\kappa, C(\kappa)) &\implies \mathbf{PWP} \left(2\hat{\psi}(\kappa), \frac{c}{C_\kappa}, \|\cdot\|_\phi \right). \end{aligned}$$

Going the other way around is easy. Indeed, suppose that $\mathbf{PWP}(r, C(r), \|\cdot\|_\infty)$ holds. Let A be a set whose measure is less than $1/2$, but greater than $4r$. Let g be any function which may appear in the definition of the capacity of A (cf. (4.11)), and let m_g be a median of g . Then:

$$\mathbf{Var}_\mu g \leq C_r \int |\nabla g|^2 d\mu + r \|g - m_g\|_\infty.$$

Without loss of generality, we suppose that $0 \leq g \leq 1$, so that the L^∞ norm is bounded by 1. Moreover, r is less than $\mu(A)/4$, and the variance on the left hand side is bounded below by $(1/2) \min(\mu(A), 1/2) \geq (\mu(A)/2)$ (by the same argument used previously, during the proof of proposition 90). This entails:

$$\frac{\mu(A)}{2} \leq C_r \int |\nabla g|^2 d\mu + \frac{\mu(A)}{4}.$$

This immediately implies the measure capacity inequality **M-C**($4r, 4/C_r$).

If we now try to derive an inequality with an Orlicz norm starting from one with an L^∞ norm, we just translate them in terms of measure and capacity:

$$\begin{aligned} \mathbf{PWP}(r, C_r, \|\cdot\|_\infty) &\implies \mathbf{M-C}(4r, 4/C_r) \\ &\implies \mathbf{PWP}(2\hat{\psi}(4r), \frac{cC_r}{4}, \|\cdot\|_\phi). \end{aligned}$$

If we are looking for a full weak Poincaré inequality, we fix an s , and define $r = (1/4)\hat{\psi}^{-1}(s/2)$. We may then apply **PWP**($r, \beta(r), \|\cdot\|_\infty$) to obtain:

$$\mathbf{PWP}(s, c\beta(r)/4, \|\cdot\|_\phi).$$

Since s is arbitrary, this concludes the proof. \square

4.A Orlicz norms, entropy and centering

The proof of weak Poincaré inequalities starting from measure-capacity comparisons for an Orlicz norm leads us to consider norms of functions recentered by their median. In fact, what one obtains when applying these criteria is of the form:

$$\mathbf{Var}_\mu f^2 \leq \beta(s)\mathcal{E}(f) + s\|f - m_f\|_\phi^2,$$

where m_f is a median for f . The aim of this section is to bound this term by more tractable quantities (we will use an entropy and a moment).

More precisely we prove the following result:

98 Lemma. *Let ϕ be the Young function $x \mapsto x^2 \log(1 + x^2)$. There exists a C such that, for any positive f and any probability measure μ , the following holds:*

$$\|f - m_f\|_\phi^2 \leq C(\mathbf{Ent}_\mu(f^2) + 3\mathbf{E}_\mu(f^2)).$$

The proof is done in several steps, and borrows several arguments from [BG99]. First of all, we get rid of the median and replace it by a mean value.

$$\begin{aligned} \|f - m_f\|_\phi &\leq \|f - \mu f\|_\phi + \|\mu f - m_f\|_\phi \\ &\leq \|f - \mu f\|_\phi + |\mu f - m_f|. \end{aligned} \tag{4.19}$$

Let us consider the last term.

$$\mu f - m_f = \int f(x) d\mu - m_f = \int (f - m_f)_+ d\mu - \int (f - m_f)_- d\mu,$$

where the integrals are both positive. The absolute value of the left hand side may then be bounded above:

$$|\mu f - m_f| \leq \max\left(\int (f - m_f)_+ d\mu, \int (f - m_f)_- d\mu\right)$$

Each of the arguments in the max can be controlled by Hölder's inequality.

$$\begin{aligned} \int (f - m_f)_+ d\mu &= \int (f - m_f) \mathbf{1}_{f > m_f} d\mu \leq \|f - m_f\|_2 \|\mathbf{1}_{f > m_f}\|_2 \\ &\leq \frac{1}{\sqrt{2}} \|f - m_f\|_2 && \text{(since } \mu(f > m_f) < 1/2) \\ &\leq \frac{1}{\sqrt{2}} \frac{\sqrt{5}}{2} \|f - m_f\|_\phi && \text{(cf. [BG99], lemma 4.3)} \end{aligned}$$

Coming back to (4.19), we get:

$$\|f - m_f\|_\phi \leq \|f - \mu f\|_\phi + |\mu f - m_f| \leq \|f - \mu f\|_\phi + \sqrt{\frac{5}{8}} \|f - m_f\|_\phi.$$

Since $\sqrt{\frac{5}{8}} \leq 1$, we may put it on the other side to get:

$$\|f - m_f\|_\phi \leq C \|f - \mu f\|_\phi \tag{4.20}$$

where $C = (1 - \sqrt{\frac{5}{8}})^{-1}$ is universal.

The next step is to bound the Orlicz norm by an entropy. Once again, we use a result from Bobkov and Götze ([BG99]):

$$\|f - \mu f\|_\phi^2 \leq \frac{3}{2} \sup_{a \in \mathbb{R}} \mathbf{Ent}_\mu((f + a)^2).$$

Since we would like to deal only with the entropy of f^2 , we try to compare the entropies of translated functions. Rothaus' lemma tells us:

$$\mathbf{Ent}_\mu((f + a)^2) \leq \mathbf{Ent}_\mu(\tilde{f}^2) + 2\mathbf{Var}_\mu(f),$$

where \tilde{f} is the centered function $f - \mu f$. The only thing left to do is to bound the entropy of the square of this centered function. This is done in the following lemma.

99 Lemma. *Let f be a positive function, and $\tilde{f} = f - \mu f$. Then the following holds:*

$$\mathbf{Ent}_\mu(\tilde{f}^2) \leq \mathbf{Ent}_\mu(f^2) + \int f^2 d\mu.$$

Proof. Both sides of the equation are homogeneous (of order two), so we may as well suppose $\int f^2 d\mu = 1$. We rewrite the left hand side.

$$\begin{aligned} \mathbf{Ent}_\mu(\tilde{f}^2) &= \int \tilde{f}^2 \log(\tilde{f}^2) d\mu - \mathbf{E}_\mu(\tilde{f}^2) \log(\mathbf{E}_\mu(\tilde{f}^2)) \\ &= \int \tilde{f}^2 \log(\tilde{f}^2) d\mu - \mathbf{Var}_\mu(f) \log(\mathbf{Var}_\mu(f)). \end{aligned}$$

The second term is easily dealt with. Indeed, since $\int f^2 = 1$, $\mathbf{Var}_\mu f$ must be between 0 and 1. Since $x \mapsto |x \log(x)|$ is bounded by $1/e$ on this interval, one can write:

$$\mathbf{Ent}_\mu(\tilde{f}^2) \leq \int \tilde{f}^2 \log(\tilde{f}^2) d\mu + \frac{1}{e}.$$

We decompose the integral in two parts, according to whether f is less than 1 or not.

$$\begin{aligned} \mathbf{Ent}_\mu(\tilde{f}^2) &\leq \int \tilde{f}^2 \log(\tilde{f}^2) \mathbf{1}_{|\tilde{f}| \leq 1} d\mu + \int \tilde{f}^2 \log(\tilde{f}^2) \mathbf{1}_{|\tilde{f}| > 1} d\mu + \frac{1}{e} \\ &\leq \int \tilde{f}^2 \log(\tilde{f}^2) \mathbf{1}_{|\tilde{f}| > 1} d\mu + \frac{1}{e}, \end{aligned}$$

since the first term is less than 0. Now, on the set where $|\tilde{f}|$ exceeds one, f must be above its mean: f is indeed positive, and since $\int f^2 d\mu = 1$, μf must be in $[0, 1]$. So $|f - \mu f|$ may be greater than 1 only when f itself is greater than 1. This shows that, on $\{|\tilde{f}| > 1\}$,

$$1 \leq \tilde{f} = f - \mu f \leq f.$$

Since $x \mapsto x \log(x)$ increases on $[1, \infty)$, we have:

$$\begin{aligned} \mathbf{Ent}_\mu(\tilde{f}^2) &\leq \int \tilde{f}^2 \log(\tilde{f}^2) \mathbf{1}_{|\tilde{f}| > 1} d\mu + \frac{1}{e} \\ &\leq \int f^2 \log(f^2) \mathbf{1}_{|\tilde{f}| > 1} d\mu + \frac{1}{e}. \end{aligned}$$

At this point, remark that on $\{f > 1\}$, $f^2 \log(f^2)$ is positive, and since $\mathbf{1}_{|\tilde{f}| > 1} \leq \mathbf{1}_{f > 1}$,

$$\begin{aligned} \mathbf{Ent}_\mu(\tilde{f}^2) &\leq \int f^2 \log(f^2) \mathbf{1}_{f > 1} d\mu + \frac{1}{e} \\ &\leq \mathbf{Ent}(f^2) - \int f^2 \log(f^2) \mathbf{1}_{f < 1} d\mu + \frac{1}{e} \\ &\leq \mathbf{Ent}(f^2) + \frac{2}{e}. \end{aligned}$$

Since $\frac{2}{e} \leq 1$, the proof is complete. □

Gathering our results, we have shown that:

$$\begin{aligned} \|f - m_f\|_\phi^2 &\leq C \|f - \mu_f\|_\phi^2 && \text{(inequality (4.20))} \\ &\leq \frac{3C}{2} \sup_{a \in \mathbb{R}} \mathbf{Ent}((f + a)^2) && \text{(Bobkov and Götze's lemma)} \\ &\leq \frac{3C}{2} (\mathbf{Ent}(\tilde{f}^2) + 2\mathbf{Var}_\mu(f)) && \text{(Rothaus' lemma)} \\ &\leq \frac{3C}{2} (\mathbf{Ent}(f^2) + 3\mathbf{E}_\mu(f^2)). && \text{(lemma 99)} \end{aligned}$$

The last line is precisely the result we claimed in lemma 98.

4.B A moment bound

In this annex we prove lemma 76. The proof mainly follows the one in Miclo's doctoral dissertation, with a few changes to accommodate our hypotheses.

Outline of the proof

We need to introduce some notation.

For $\epsilon > 0$, we denote by \mathcal{L}_ϵ the generator of the diffusion at fixed temperature ϵ :

$$\mathcal{L}_\epsilon = \frac{\epsilon}{2} \Delta - \frac{1}{2} \nabla V \nabla \cdot.$$

We will need a smooth version of a step function; we call it f and suppose that it satisfies:

$$f(x) = \begin{cases} 0 & \text{if } x \leq 0, \\ \exp(-\exp(\frac{1}{x})) & \text{on } [0, 1], \\ 1 & \text{on } [2, \infty[. \end{cases}$$

We recall the hypotheses on V :

- It goes to infinity at infinity,
- its gradient ∇V is bounded, and
- its Laplacian ΔV is negative for large x .

Note that, since V is continuous, there must be an R such that ΔV is negative whenever $V(x) \geq R$.

Finally, let g be an increasing function, going to zero at zero.

The idea of the proof is that, as time goes by, the value of V at X_t has a typical scale, namely $\frac{1}{g(\sigma(t))}$, for a function g to be made precise later, so that when we try to estimate $\mathbf{E}(V^p(X_t))$, we only have to take into account the small values of V .

More precisely, let $\rho_\epsilon(\cdot) = f(g(\epsilon)V(\cdot) - (R+1))$. This is a smooth approximation of $\mathbf{1}_{V \geq \frac{R}{g(\epsilon)}}$.

We may bound the expectation of $V^p(X_t)$:

$$\begin{aligned} \mathbf{E}[V^p(X_t)] &= \mathbf{E}[V^p \rho_{\sigma(t)}(X_t)] + \mathbf{E}[V^p (1 - \rho_{\sigma(t)})(X_t)] \\ &\leq \mathbf{E}[V^p \rho_{\sigma(t)}(X_t)] + \left(\frac{R+3}{g(\sigma(t))} \right)^p. \end{aligned} \quad (4.21)$$

To bound the first term, we use the explicit expression of the generator. Intuitively, we write, for $h_t = V^p \rho_{\sigma(t)}$:

$$\frac{d}{dt}(P_t h_t) = P_t \mathcal{L}_{\sigma(t)} h_t + P_t \left(\frac{d}{dt} h_t \right),$$

and integrate between two times t and t' . To ensure that everything exists, we use the stopping time $T_k = \inf\{t, V(X_t) \geq k\}$. We get:

$$\begin{aligned} \mathbf{E}[h_{t \wedge T_k}(X_{t \wedge T_k})] &= \mathbf{E}[h_{t' \wedge T_k}(X_{t' \wedge T_k})] \\ &+ \mathbf{E} \left[\int_{t' \wedge T_k}^{t \wedge T_k} \mathcal{L}_{\sigma(s)}(h_s)(X_s) ds \right] \\ &+ \mathbf{E} \left[\int_{t' \wedge T_k}^{t \wedge T_k} \sigma'(s) g'(\sigma(s)) f'(g(\sigma(s))V(X_s) - (R+1)) V^{p+1}(X_s) ds \right]. \end{aligned}$$

Since V is positive, f and g increasing and σ decreases, the whole last term is negative. We try to estimate the second one, and study $\mathcal{L}_{\sigma(s)} h_s(X_s)$.

100 Lemma. *Let us define $\varphi : x \mapsto x \log^2(x)$. There exists an M and a time t' (which may depend on p and on the initial law) such that:*

$$\forall t \geq t', \forall x, \quad \mathcal{L}_{\sigma(t)}(h_t)(X_t) \leq \exp\left(-\frac{M}{\varphi(\sigma(t)g(\sigma(t)))}\right).$$

We postpone the proof and finish the argument. The inequality dictates the choice of g : $g = \ln(1/\cdot)^{-3}$ guarantees

$$\begin{aligned} \sigma(t)g(\sigma(t)) &= \frac{1}{\ln(t)(\ln \ln(t))^3}, \\ \varphi(\sigma(t)g(\sigma(t))) &= \frac{\ln^2(1/\ln(t)(\ln \ln(t))^3)}{\ln(t)(\ln \ln(t))^3} = \frac{\ln^2(\ln(t)(\ln \ln(t))^3)}{\ln(t)(\ln \ln(t))^3}. \end{aligned}$$

Indeed, the upper bound on the generator then becomes

$$\begin{aligned} \mathcal{L}_{\sigma(t)}(h_t)(X_t) &\leq \exp\left(-\frac{M}{\varphi(\sigma(t)g(\sigma(t)))}\right) \\ &\leq \exp\left(-M \ln(t) \times \frac{(\ln \ln(t))^3}{\ln^2(\ln(t)(\ln \ln(t))^3)}\right). \end{aligned}$$

Since the ratio $(\ln \ln(t))^3 / (\ln^2(\ln(t)(\ln \ln(t))^3))$ goes to infinity, it eventually exceeds $2/M$, so that for t big enough,

$$\mathcal{L}_{\sigma(t)}(h_t)(X_t) \leq \exp(-2 \ln(t)).$$

Going back to the bound on the expected value we were looking for, the two previous arguments imply:

$$\mathbf{E}[h_{t \wedge T_k}(X_{t \wedge T_k})] \leq \mathbf{E}[h_{t' \wedge T_k}(X_{t' \wedge T_k})] + \int_{t'}^{\infty} \exp(-2 \ln(t)).$$

Everything was done in order to make the last integral finite. We can then let k go to infinity, and since t' is fixed, we get the existence of a constant M_p (which depends on p and on the initial law) such that:

$$\mathbf{E}[h_t(X_t)] \leq M.$$

Plugging this back into inequality (4.21) yields:

$$\mathbf{E}[V^p(X_t)] \leq M + \left(\frac{R+3}{g(\sigma(t))}\right)^p.$$

The expression of g shows that, for a new constant M :

$$\mathbf{E}[V^p(X_t)] \leq M(\sigma(t) \ln(t)(\ln \ln(t))^3)^p,$$

and the result is proved.

An estimate on the generator

We now turn to the proof of lemma 100. We have to bound $\mathcal{L}_\epsilon(\rho_\epsilon V^p)(x)$, and our first step will be to give a more explicit expression of this quantity. We will need the derivatives of $\rho_\epsilon(x)$. To alleviate notations, we will write $y = y(x, \epsilon) = g(\epsilon)V(x) - (R+1)$.

$$\begin{aligned} \rho_\epsilon(x) &= f(g(\epsilon)V(x) - (R+1)) = f(y), \\ \nabla \rho_\epsilon(x) &= g(\epsilon) f'(y) \nabla V(x), \\ \Delta \rho_\epsilon(x) &= g(\epsilon)^2 f''(y) |\nabla V|^2 + g(\epsilon) f'(y) \Delta V. \end{aligned}$$

The quantity we would like to estimate is

$$\mathcal{L}_\epsilon(\rho_\epsilon V^p)(x) = \rho_\epsilon \mathcal{L}_\epsilon V^p(x) + \epsilon \langle \nabla \rho_\epsilon, \nabla V^p \rangle(x) + V^p \mathcal{L}_\epsilon \rho_\epsilon(x)$$

We consider three cases, according to the value of $V(x)g(\epsilon)$.

V is small: $V(x)g(\epsilon) \in [0, R+1]$ On this interval, ρ_ϵ vanishes, so $\mathcal{L}_\epsilon(\rho_\epsilon)$ is zero.

V is large. Let λ be a strictly positive real, to be fixed later on. We consider the case where $V(x)g(\epsilon) \in [R+1+\lambda, \infty)$, which may be rewritten as: $y \in [\lambda, \infty)$. We develop the expression of $\mathcal{L}_\epsilon(\rho_\epsilon V^p)$.

$$\mathcal{L}_\epsilon(\rho_\epsilon V^p)(x) = \rho_\epsilon \mathcal{L}_\epsilon V^p(x) + \epsilon g(\epsilon) f'(y) \times p V^{p-1} |\nabla V|^2 + V^p \left(\frac{1}{2} \epsilon \Delta \rho_\epsilon - \frac{1}{2} \langle \nabla \rho_\epsilon, \nabla V \rangle \right).$$

We compute the derivatives of ρ_ϵ and put together the terms involving $|\nabla V|^2$.

$$\begin{aligned} \mathcal{L}_\epsilon(\rho_\epsilon V^p)(x) &= \rho_\epsilon \mathcal{L}_\epsilon V^p(x) + \left(\epsilon g(\epsilon) f'(y) p V^{p-1} + V^p \left(\frac{1}{2} \epsilon g(\epsilon)^2 f''(y) - \frac{1}{2} g(\epsilon) f'(y) \right) \right) |\nabla V|^2 \\ &\quad + \frac{1}{2} \epsilon g(\epsilon) f'(y) V^p \Delta V \\ &= A + B + C. \end{aligned}$$

Since $V \times g(\epsilon) \geq R$, $V \geq R$. We already noted that R may be chosen so that, if V is bigger than R , ΔV is less than zero, and this makes the third term C negative. The term B can be rewritten as:

$$\begin{aligned} B &= \left(\epsilon g(\epsilon) f'(y) p V^{p-1} + V^p \left(\frac{1}{2} \epsilon g(\epsilon)^2 f''(y) - \frac{1}{2} g(\epsilon) f'(y) \right) \right) |\nabla V|^2 \\ &= V^p g(\epsilon) \left(\left(\frac{p\epsilon}{V} - \frac{1}{2} \right) f'(y) + \frac{1}{2} \epsilon g(\epsilon) f''(y) \right) |\nabla V|^2. \end{aligned} \tag{4.22}$$

We add another condition on f : it should be concave when y is near 2 (e.g. on $[\frac{3}{2}, 2]$). On $[\lambda, 3/2]$, f''/f' is bounded — let M be a bound (λ will be chosen so that other useful bounds hold on f on $[0, \lambda]$, cf. infra). This entails:

$$\forall y \geq \lambda, \quad f''(y) \leq M f'(y).$$

Coming back to B , we deduce:

$$B \leq \left(\frac{p\epsilon}{V} + \frac{M\epsilon g(\epsilon)}{2} - \frac{1}{2} \right) f'(y) g(\epsilon) V^p |\nabla V|^2.$$

The term between brackets is negative, uniformly in x as soon as ϵ is small enough.

Finally, the first term $A = \mathcal{L}_\epsilon V^p$ is also negative:

$$\begin{aligned} A &= \frac{\epsilon}{2} \Delta(V^p) - \frac{1}{2} \langle \nabla V, \nabla(V^p) \rangle \\ &= \frac{\epsilon}{2} (p(p-1)V^{p-2} |\nabla V|^2 + pV^{p-1} \Delta V) - \frac{p}{2} V^{p-1} |\nabla V|^2 \\ &\leq \frac{p}{2} |\nabla V|^2 V^{p-2} ((p-1)\epsilon - V). \end{aligned}$$

Once more, the term between brackets is negative when ϵ is small (because $V \geq R$). To conclude, for any λ , there exists an ϵ_0 such that:

$$\forall \epsilon < \epsilon_0, \forall x, \quad V(x)g(\epsilon) \geq R+1+\lambda \implies \mathcal{L}_\epsilon(\rho_\epsilon V^p) \leq 0.$$

V is of the order of $R/g(\epsilon)$. This last case is that where $g(\epsilon)V(x) \in [R+1, R+1+\lambda]$. Let us reuse the decomposition $\mathcal{L}_\epsilon(\rho_\epsilon V^p) = A + B + C$ from the previous paragraph. The same reasoning applies for A and C , and they are both negative, so it suffices to get a bound on B . From (4.22):

$$B = \left(\left(\frac{p\epsilon}{V} - \frac{1}{2} \right) f'(y) + \frac{1}{2} \epsilon g(\epsilon) f''(y) \right) g(\epsilon) V^p |\nabla V|^2.$$

If we choose R sufficiently big and ϵ small enough, the quantity between brackets in front of $f'(y)$ is less than $(-1/4)$.

$$B \leq \left(-\frac{1}{4} f'(y) + \frac{1}{2} \epsilon g(\epsilon) f''(y) \right) g(\epsilon) V^p |\nabla V|^2.$$

Recall that $f = \exp(-\tau)$, where $\tau(y) = \exp(1/y)$, for any $y \in (0, 1]$. This implies:

$$\begin{aligned} B &\leq \left(\frac{1}{4} \tau' f + \frac{1}{2} \epsilon g(\epsilon) (-\tau'' f + (\tau')^2 f) \right) g(\epsilon) V^p |\nabla V|^2 \\ &\leq \frac{1}{2} \left(\frac{1}{2} \tau' f + \epsilon g(\epsilon) (\tau'(y))^2 f(y) \right) g(\epsilon) V^p |\nabla V|^2 \end{aligned}$$

Define $h_\epsilon = \frac{1}{2} \tau' f + \epsilon g(\epsilon) \tau'^2 f$. We study it by differentiating:

$$h'_\epsilon = \left(\frac{1}{2} \tau'' - \frac{1}{2} \tau'^2 + 2\epsilon g(\epsilon) \tau' \tau'' - \epsilon g(\epsilon) \tau'^3 \right) f.$$

The explicit expression of τ ensures:

$$\exists \lambda \forall y \in [0, \lambda] \quad 0 \leq \tau''(y) \leq \frac{1}{4} \tau'^2(y).$$

This λ does not depend on ϵ . This can be used to bound h'_ϵ from below:

$$\begin{aligned} h'_\epsilon(y) &\geq \left(-\frac{1}{2} \tau'^2(y) + \frac{1}{2} \epsilon g(\epsilon) \tau'(y)^3 - \epsilon g(\epsilon) \tau'(y)^3 \right) f(y) \\ &\geq \left(-\frac{1}{2} - \frac{1}{2} \epsilon g(\epsilon) \tau'(y) \right) \tau'(y)^2 f(y). \end{aligned}$$

Let $y_{1,\epsilon}$ be the solution of the equation: $-1 - \epsilon g(\epsilon) \tau'(y) = 0$. When ϵ is small, $y_{1,\epsilon}$ will be less than λ , and the monotonicity of τ' will give:

$$\forall y \leq y_{1,\epsilon}, \quad h'_\epsilon(y) \geq 0.$$

Similarly, h'_ϵ can be bounded above:

$$\begin{aligned} h'_\epsilon &\leq \left(\frac{1}{8} \tau'^2(y) - \frac{1}{2} \tau'^2(y) - \epsilon g(\epsilon) \tau'^3(y) \right) f(y) \\ &\leq \left(-\frac{3}{8} - \epsilon g(\epsilon) \tau'(y) \right) \tau'(y)^2 f(y). \end{aligned}$$

Now, let $y_{2,\epsilon}$ be the root of $-\frac{3}{8} - \epsilon g(\epsilon) \tau'(y) = 0$. Once more, when ϵ is small, $y_{2,\epsilon}$ falls within $[0, \lambda]$. We deduce:

$$\forall y \in [y_{2,\epsilon}, \lambda], \quad h'_\epsilon(y) \leq 0.$$

We now know the h_ϵ increases on $[0, y_{1,\epsilon}]$, and decreases on $[y_{2,\epsilon}, \lambda]$, so that its maximum must be reached somewhere between these two points. More precisely, whenever ϵ is less than some ϵ_0 , it holds that

$$\exists y_\epsilon \in [y_{1,\epsilon}, y_{2,\epsilon}], \forall y \in [0, \lambda], \quad h_\epsilon(y) \leq h_\epsilon(y_\epsilon).$$

The bounds on y_ϵ , the fact that τ decreases and the equations defining $y_{1,\epsilon}, y_{2,\epsilon}$ allow us to conclude:

$$\begin{aligned} \forall y \leq \lambda, \quad h_\epsilon(y) &\leq \left(\frac{1}{2} \tau'(y_\epsilon) + \epsilon g(\epsilon) \tau'(y_\epsilon)^2 \right) f(y_\epsilon) \\ &\leq \left(\frac{1}{2} \tau'(y_{2,\epsilon}) + \epsilon g(\epsilon) \tau'(y_{1,\epsilon})^2 \right) f(y_{2,\epsilon}) \\ &\leq \left(-\frac{3}{16\epsilon g(\epsilon)} + \frac{1}{\epsilon g(\epsilon)} \right) f(y_{2,\epsilon}) \\ &\leq \frac{1}{\epsilon g(\epsilon)} f(y_{2,\epsilon}). \end{aligned}$$

It remains to estimate $f(y_{2,\epsilon}) = \exp(-\tau(y_{2,\epsilon}))$. Since $y_{2,\epsilon}$ is defined as a solution of an equation involving τ' , we would like to compare τ and τ' . The explicit expression of τ easily implies:

$$\ln(|\tau'(y)|) = \ln(y^{-2}) + \frac{1}{y} \geq \frac{1}{y},$$

therefore:

$$\tau(y) = y^2 |\tau'(y)| \geq \frac{|\tau'(y)|}{\ln^2(|\tau'(y)|)}$$

Applying this for $y = y_{2,\epsilon}$, for which $|\tau'(y)| = 3/(8\epsilon g(\epsilon))$, entails:

$$\begin{aligned} \tau(y_{2,\epsilon}) &\geq \frac{3}{8\epsilon g(\epsilon) \ln^2(8\epsilon g(\epsilon)/3)} \\ &\geq \frac{3}{8\epsilon g(\epsilon) \ln^2(\epsilon g(\epsilon))}. \end{aligned}$$

Turning back to f , and defining $\varphi : x \mapsto x \ln^2(x)$, and $M = 3/8$, we have:

$$f(y_{2,\epsilon}) = \exp(-\tau(y_{2,\epsilon})) \leq \exp\left(-\frac{M}{\varphi(\epsilon g(\epsilon))}\right).$$

We now come back to the upper bound on B , and plug in the last equation.

$$B \leq \frac{1}{2} \times \frac{1}{\epsilon g(\epsilon)} \exp\left(-\frac{M}{\varphi(\epsilon g(\epsilon))}\right) g(\epsilon) V^p |\nabla V|^2.$$

Since we suppose that $V(x)g(\epsilon)$ belongs to $[R+1, R+2]$, we may bound V^p by $(R+2)^p g(\epsilon)^{-p}$. We also supposed that ∇V is bounded, so that there exists an M' such that:

$$B \leq \frac{M'}{\epsilon g(\epsilon)^p} \exp\left(-\frac{M}{\varphi(\epsilon g(\epsilon))}\right).$$

Up to a slight change of the constant M in the exponential, we may neglect the pre-exponential term and write:

$$B \leq M'' \exp\left(-\frac{M}{\varphi(\epsilon g(\epsilon))}\right)$$

This concludes the proof.

4.C Regularity results and estimates on the process

An equivalent of the partition function

We recall here Laplace's method, which enable us to study the asymptotic behaviour of the partition function, *i.e.* the constant $Z_\sigma = \int \exp(-V/\sigma) dx$.

101 Theorem. *Let V be a function from \mathbb{R}^d to \mathbb{R} , satisfying hypotheses 1 and 2 (V has a unique, well behaved, global minimum, and V goes to infinity at infinity rapidly enough). Then Z_σ exists, and the following holds:*

$$Z_\sigma \underset{\sigma \rightarrow 0}{\sim} \frac{(2\pi\sigma)^{d/2}}{\sqrt{\det \mathbf{Hess} V}}.$$

To prove this classical result, we cut the integral in two parts, the main one (near the origin) and a remainder. Before we proceed, let us remark that, up to a change of coordinates, we may as well suppose that $\mathbf{Hess}(V)_0$ is a diagonal matrix, and we have Taylor's formula:

$$V(x) = \frac{1}{2} \sum_i \lambda_i x_i^2 + \epsilon(x) \sum_i x_i^2,$$

where $\epsilon(x)$ goes to zero at 0. We choose an r such that, on $B = [-r, r]^d$, $\epsilon(x) \leq \frac{1}{4}(\inf \lambda_i)$.

Let us begin by the negligible part, outside of B . Since V goes to infinity, and 0 is the unique global minimum, there exists an $\eta > 0$ such that $V(x) \geq \eta$ outside B . We introduce an $\exp(-V)$ in the integral (the growth hypothesis makes it integrable), and use this bound:

$$\begin{aligned} \int_{x \notin B} \exp(-V/\sigma) dx &= \int_{x \notin B} \exp(-V) \exp(-(1/\sigma - 1)V(x)) dx \\ &\leq \int_{x \notin B} \exp(-V) dx \exp(-(1/\sigma - 1)\eta) \\ &\leq Z_1 \exp(-(1/\sigma - 1)\eta). \end{aligned}$$

Let us turn to the main term. We divide it by $\sigma^{d/2}$ (so that we only have to find a limit). We change variables and use $x = \phi_\sigma(y)$ defined by $x_i = y_i \sqrt{\sigma/\lambda_i}$.

$$\begin{aligned} \sigma^{-d/2} \int_B \exp(-V/\sigma) dx_1 \cdots dx_n &= \sigma^{-d/2} \int \mathbf{1}_{x \in B} \exp\left(-\frac{1}{2} \sum_i \frac{\lambda_i}{\sigma} x_i^2 + \frac{\epsilon(x)}{\sigma} \sum_i x_i^2\right) dx \\ &= \frac{1}{\sqrt{\lambda_1 \cdots \lambda_n}} \int \mathbf{1}_{\phi_\sigma(y) \in B} \exp\left(-\frac{1}{2} \sum_i y_i^2 + \epsilon(\phi_\sigma(y)) \sum_i \frac{y_i^2}{\lambda_i}\right) dy. \end{aligned}$$

The function inside the integral converges pointwise to $\exp(-\sum y_i^2)$ when σ goes to zero (because $\phi_\sigma(y)$ goes to zero for a fixed y). It is bounded from above by the integrable function $\exp(-\frac{1}{4} \sum y_i^2)$ (because $\phi_\sigma(y) \in B$), and we may apply Lebesgue's dominated convergence:

$$\sigma^{-d/2} \int_B \exp(-V/\sigma) dx_1 \cdots dx_n \xrightarrow{\sigma \rightarrow 0} \frac{(2\pi)^{d/2}}{\sqrt{\lambda_1 \cdots \lambda_n}}$$

With the bound on the remainder, this gives the equivalent of Z_σ .

Finiteness of the entropy and regularity

We begin by proving that the relative entropy I_t is finite. To do this, we study directly the explicit density, which we know thanks to a Girsanov transform. We follow a proof from [Roy99], with a few minor changes to deal with the non-homogeneity in time.

Recall that the process X is defined by the following SDE:

$$dX_t = \sqrt{\sigma(t)}dB_t - \frac{1}{2}\nabla V(X_t)dt.$$

If we define a new reference martingale $M_t = \int_0^t \sqrt{\sigma(s)}dB_s$, we may define X as the solution to the SDE:

$$dX_t = dM_t - \frac{1}{2}\nabla V(X_t)dt.$$

Note that M_t is just a Brownian motion under a (deterministic) change of time — if we define $\tau(t) = \int_0^t \sigma(s)ds$, $M_{\tau^{-1}(t)}$ is a Brownian motion. To find the density of the law of X_t with respect to its equilibrium measure μ_t , we decompose it in three terms:

$$\frac{d\mathcal{L}(X_t)}{d\mu_t} = \frac{d\mathcal{L}(X_t)}{d\mathcal{L}(M_t)} \times \frac{d\mathcal{L}M_t}{d\lambda} \times \frac{d\lambda}{d\mu_t}.$$

To compute the first term, we use the (trajectorial) density of $X_{[0,t]}$ with respect to $M_{[0,t]}$, which is given by Girsanov's theorem:

$$\begin{aligned} F &= \exp\left(-\frac{1}{2}\int \nabla V(M_s)dM_s - \frac{1}{2}\int_0^t \frac{|\nabla V|^2}{4}(M_s)d\langle M \rangle_s\right) \\ &= \exp\left(-\frac{1}{2}\int \nabla V(M_s)dM_s - \frac{1}{8}\int_0^t |\nabla V|^2(M_s)\sigma(s)ds\right). \end{aligned}$$

To get rid of the martingale term in the exponential, we apply Itô's formula to V and the martingale M :

$$V(M_t) = V(x) + \int_0^t \nabla V(M_s)dM_s + \frac{1}{2}\int_0^t \Delta V(M_s)d\langle M \rangle_s.$$

The functional F may thus be rewritten:

$$F = \exp\left(\frac{1}{2}V(x) - \frac{1}{2}V(M_t) + \int_0^t \left(\frac{1}{4}\Delta V(M_s) - \frac{1}{8}|\nabla V|^2(M_s)\right)\sigma(s)ds\right).$$

The three densities we are looking for are:

$$\begin{aligned} \frac{d\mathcal{L}(X_t)}{d\mathcal{L}(M_t)}(M_t) &= f(M_t) = \mathbf{E}[F|\mathcal{F}_{\{t\}}] \\ \frac{d\mathcal{L}(M_t)}{d\lambda}(y) &= \exp(-2v_t(y)) = (2\pi\tau(t))^{-d/2} \exp\left(-\frac{(x-y)^2}{2\tau(t)}\right) \\ \frac{d\lambda}{d\mu_t}(y) &= Z_{\sigma(t)} \exp\left(\frac{V(y)}{\sigma(t)}\right), \end{aligned}$$

where x is the starting point X_0 . We take the product of these terms; the last two quantities may be put into the conditional expectation, so that the density we are looking for (say G) may be written as:

$$G(M_t) = Z_{\sigma(t)} \mathbf{E}\left[F \exp\left(\frac{V(M_t)}{\sigma(t)} - 2v_t(M_t)\right) | \mathcal{F}_{\{t\}}\right]$$

Let us now define $\gamma : x \mapsto x \log(x)$, and start to study I_t . By definition, $I_t = \int \gamma(G(y)) d\mu_t(y)$. Since G is best expressed as a conditional expectation, we rewrite I_t :

$$\begin{aligned} I_t &= \mathbf{E} \left[\gamma(G(M_t)) \frac{d\mu_t}{d\mathcal{L}(M_t)} \right] \\ &= \mathbf{E} \left[\gamma(G(M_t)) \frac{1}{Z_{\sigma(t)}} \exp \left(-\frac{V(M_t)}{\sigma(t)} + 2v_t(M_t) \right) \right]. \end{aligned} \quad (4.23)$$

Since γ is convex, we may apply Jensen's conditional inequality to $\gamma(G(M_t))$, and develop γ :

$$\begin{aligned} \gamma(G(M_t)) &\leq \mathbf{E} \left[\gamma \left(Z_{\sigma(t)} F \exp \left(\frac{V}{\sigma(t)} - 2v_t \right) \right) \middle| \mathcal{F}_{\{t\}} \right] \\ &\leq \mathbf{E} \left[Z_{\sigma} F \exp \left(\frac{V}{\sigma(t)} - 2v_t \right) \left(\log Z_{\sigma(t)} + \log F + \frac{V}{\sigma(t)} - 2v_t \right) \middle| \mathcal{F}_{\{t\}} \right]. \end{aligned}$$

Multiply both sides by $\left(Z_{\sigma(t)}^{-1} \exp(-V/\sigma(t) + 2v_t) \right)$, and take the expected value; the left hand side becomes I_t (thanks to (4.23)), the conditioning disappears and we get:

$$I_t \leq \mathbf{E} \left[F \left(\log Z_{\sigma(t)} + \log F + \frac{V(M_t)}{\sigma(t)} - 2v_t(M_t) \right) \right]$$

Recall that F is a density, so that $\mathbf{E}[F] = 1$, and we may take the constant $Z_{\sigma(t)}$ out of the expectation. We add and subtract $(2/\sigma(t)) \log(F)$ inside the integral — this will help us get rid of the term $V(M_t)/\sigma(t)$:

$$I_t \leq \log(Z_{\sigma(t)}) - \left(\frac{2}{\sigma(t)} - 1 \right) \mathbf{E} [F \log F] + \mathbf{E} \left[F \left(\frac{2}{\sigma(t)} \log F + \frac{V(M_t)}{\sigma} - 2v_t(M_t) \right) \right].$$

Since $x \log x$ is bounded below, and we may suppose $2/\sigma(t) - 1 > 0$, the second term is bounded from above (for any finite time t). The same is true for the first term. The only thing to check is that the last term is finite; let us call this term A . Since F is given by an exponential, A is given by:

$$A = \mathbf{E} \left[F \left(\frac{1}{\sigma(t)} V(x) + \frac{1}{4\sigma(t)} \int_0^t (2\Delta V - |\nabla V|^2)(M_s) \sigma(s) ds - 2v_t(M_t) \right) \right].$$

Let us consider the quantity between brackets. The first term is finite and does not depend on M_t . The integral is bounded above by something also independant of M_t (indeed, $2\Delta V - |\nabla V|^2$ is uniformly bounded from above, because ΔV is negative outside a compact set). The only thing left to check is that:

$$\mathbf{E}[F(-2v_t(M_t))] < \infty.$$

We have already seen the explicit value of v_t :

$$\exp(-2v_t(y)) = (2\pi\tau(t))^{-d/2} \exp \left(-\frac{(y-x)^2}{2\tau(t)} \right).$$

Taking logarithms, we see that:

$$\begin{aligned} -2v_t(y) &= -\frac{d}{2} \log(2\pi\tau(t)) - \frac{(y-x)^2}{2\tau(t)} \\ &\leq -\frac{d}{2} \log(2\pi\tau(t)). \end{aligned}$$

Since this bound does not depend on y , $\mathbf{E}[-F \times (2v_t(M_t))]$ is finite. This concludes the proof.

Table des matières

1	Introduction	1
1.1	À propos des inégalités fonctionnelles	3
1.2	Le recuit simulé	9
1.3	Mécanique statistique et modèle d'Ising	13
I	Mécanique statistique	29
2	Un modèle d'Ising à spins réels	33
2.1	Notations et hypothèses	33
2.2	Mesures de Gibbs tempérées	37
2.3	Existence de mesures de Gibbs tempérées	42
2.4	Unicité : le critère de Dobrushin	44
2.5	Unicité : le critère de Dobrushin–Shlosman	50
2.6	Construction de la dynamique	54
3	Inégalités fonctionnelles et mesures de Gibbs	61
3.1	Unicité de la mesure de Gibbs grâce à Sobolev logarithmique	61
3.2	Vers l'inégalité de Poincaré pour l'unicité de la mesure de Gibbs	68
3.3	Des inégalités fonctionnelles aux conditions de mélange	81
3.4	Du mélange polynomial au trou spectral uniforme.	92
A	Topologie et mesures sur l'espace des configurations	99
B	Décomposition d'une fonction strictement convexe à l'infini	101
II	Recuit simulé	103
4	Annealing diffusions in a slowly growing potential	107
4.1	The convergence of the process (main line of the proof)	110
4.2	The one-dimensional case	116
4.3	The weak inequality in any dimension	121
4.4	A measure-capacity criterion for one-point weak Poincaré inequalities	131
4.A	Orlicz norms, entropy and centering	136
4.B	A moment bound	138
4.C	Regularity results and estimates on the process	144

Table des matières

147

Bibliographie

149

Bibliographie

- [AA96] A. V. ANTONIOUK & A. V. ANTONIOUK – « Decay of correlations and uniqueness of Gibbs lattice systems with nonquadratic interaction », *J. Math. Phys.* **37** (1996), no. 11, p. 5444–5454.
- [ABC⁺00] C. ANÉ, S. BLACHÈRE, D. CHAFAÏ, P. FOUGÈRES, I. GENTIL, F. MALRIEU, C. ROBERTO & G. SCHEFFER – *Sur les inégalités de Sobolev logarithmiques*, Panoramas et synthèses, no. 10, SMF, 2000.
- [AH87] M. AIZENMAN & R. A. HOLLEY – « Rapid convergence to equilibrium of stochastic Ising models in the Dobrushin Shlosman regime », in *Percolation theory and ergodic theory of infinite particle systems (Minneapolis, Minn., 1984–1985)*, IMA Vol. Math. Appl., vol. 8, Springer, New York, 1987, p. 1–11.
- [Aiz80] M. AIZENMAN – « Translation invariance and instability of phase coexistence in the two-dimensional Ising system », *Comm. Math. Phys.* **73** (1980), no. 1, p. 83–94.
- [AKRT00] S. ALBEVERIO, Y. G. KONDRATIEV, M. RÖCKNER & T. V. TSIKALENKO – « A priori estimates for symmetrizing measures and their applications to Gibbs states », *J. Funct. Anal.* **171** (2000), no. 2, p. 366–400.
- [Ale04] J. ALEXOPOULOS – « A brief introduction to n -functions and Orlicz function spaces », <http://www.personal.kent.edu/~jalexopo/Lectures.pdf>, 2004.
- [Arr95] J. M. ARRIETA – « Neumann eigenvalue problems on exterior perturbations of the domain », *J. Differential Equations* **118** (1995), no. 1, p. 54–103.
- [Bak94] D. BAKRY – « L'hypercontractivité et son utilisation en théorie des semigroupes », in *Lectures on probability theory (Saint-Flour, 1992)*, Lecture Notes in Math., vol. 1581, Springer, Berlin, 1994, p. 1–114.
- [BCR] F. BARTHE, P. CATTIAUX & C. ROBERTO – « Interpolated inequalities between exponential and gaussian, orlicz hypercontractivity and application to isoperimetry », *Revista Mat. Iberoamericana*, To appear.
- [BCR05] — , « Concentration for independent random variables with heavy tails », *AMRX Appl. Math. Res. Express* (2005), no. 2, p. 39–60.
- [BÉ85] D. BAKRY & M. ÉMERY – « Diffusions hypercontractives », in *Séminaire de probabilités, XIX, 1983/84*, Lecture Notes in Math., vol. 1123, Springer, Berlin, 1985, p. 177–206.
- [Beb03] M. BEBENDORF – « A note on the Poincaré inequality for convex domains », *Zeitschrift für Analysis und Ihre Anwendungen* **22** (2003), p. 751–756.

- [BG99] S. G. BOBKOV & F. GÖTZE – « Exponential integrability and transportation cost related to logarithmic Sobolev inequalities », *Journal of Functional Analysis* (1999), no. 163, p. 1–28.
- [BH99a] T. BODINEAU & B. HELFFER – « Correlations, spectral gaps and log-Sobolev inequalities for unbounded spins systems », in *Differential equations and mathematical physics* (Birmingham), International Press, 1999, p. 27–42.
- [BH99b] — , « The log-Sobolev inequality for unbounded spin systems », *Journal of Functional Analysis* **166** (1999), p. 168–178.
- [BHK82] J. BELLISSARD & R. HØEGH-KROHN – « Compactness and the maximal Gibbs state for random Gibbs fields on a lattice », *Comm. Math. Phys.* **84** (1982), no. 3, p. 297–327.
- [BM02] T. BODINEAU & F. MARTINELLI – « Some new results on the kinetic Ising model in a pure phase », *J. Statist. Phys.* **109** (2002), no. 1-2, p. 207–235.
- [Bod05] T. BODINEAU – « Translation invariant Gibbs states for the Ising model », à paraître dans *Probability Theory and Related Fields*, 2005.
- [Bra05] R. C. BRADLEY – « Basic properties of strong mixing conditions. a survey and some open questions », *Probability Surveys* **2** (2005), p. 107–144.
- [BS02] E. BOLTHAUSEN & A.-S. SZNITMAN – *Ten lectures on random media*, DMV Seminar, vol. 32, Birkhäuser Verlag, Basel, 2002.
- [Cat04] P. CATTIAUX – « A pathwise approach of some classical inequalities », *Potential Anal.* **20** (2004), no. 4, p. 361–394.
- [Cat05] — , « Hypercontractivity for perturbed diffusion semigroups », *Ann. Fac. Sci. Toulouse Math. (6)* **14** (2005), no. 4, p. 609–628.
- [CG06] P. CATTIAUX & A. GUILLIN – « On quadratic transportation cost inequalities », To appear in *Journal de Mathématiques pures et Appliquées*, 2006.
- [CGG05] P. CATTIAUX, I. GENTIL & A. GUILLIN – « Weak logarithmic Sobolev inequalities and entropic convergence », <http://hal.ccsd.cnrs.fr/ccsd-00013700>, 2005.
- [CH53] R. COURANT & D. HILBERT – *Methods of mathematical physics*, vol. 1, Interscience, New York, 1953, Translated and rev. from the German original.
- [Che90] R. CHEN – « Neumann eigenvalue estimate on a compact Riemannian manifold », *Proc. Amer. Math. Soc.* **108** (1990), no. 4, p. 961–970.
- [CL97] R. CHEN & P. LI – « On Poincaré type inequalities », *Trans. Amer. Math. Soc.* **349** (1997), no. 4, p. 1561–1585.
- [COPP78] M. CASSANDRO, E. OLIVIERI, A. PELEGRINOTTI & E. PRESUTTI – « Existence and uniqueness of DLR measures for unbounded spin systems », *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete* (1978), no. 41, p. 313–334.
- [Die68] J. DIEUDONNÉ – *Calcul infinitésimal*, Hermann, Paris, 1968.
- [DMM99] P. DEL MORAL & L. MICLO – « On the convergence and applications of generalized simulated annealing », *SIAM J. Control Optim.* **37** (1999), no. 4, p. 1222–1250 (electronic).

- [Dob70] R. L. DOBRUSHIN – « Prescribing a system of random variables by conditional distributions », *Theory of Probability and its Applications* **15** (1970), no. 3, p. 458–486.
- [Dob72] —, « The Gibbs state that describes the coexistence of phases for a three dimensional Ising model », *Teor. Verojatnost. i Primenen.* **17** (1972), p. 619–639.
- [DR79] H. DOSS & G. ROYER – « Processus de diffusion associé aux mesures de Gibbs sur $\mathbb{R}^{\mathbb{Z}^d}$ », *Z. Wahrsch. Verw. Gebiete* **46** (1978/79), no. 1, p. 107–124.
- [DS81] R. L. DOBRUSHIN & S. B. SHLOSMAN – « Phases corresponding to minima of the local energy », *Selecta Math. Soviet.* **1** (1981), no. 4, p. 317–338.
- [DS85a] —, « Completely analytical Gibbs fields », in *Statistical physics and dynamical systems (Köszeg, 1984)*, Progr. Phys., vol. 10, Birkhäuser Boston, Boston, MA, 1985, p. 371–403.
- [DS85b] —, « Constructive criterion for the uniqueness of Gibbs field », in *Statistical physics and dynamical systems (Köszeg, 1984)*, Progr. Phys., vol. 10, Birkhäuser Boston, Boston, MA, 1985, p. 347–370.
- [DS87] —, « Completely analytical interactions : constructive description », *J. Statist. Phys.* **46** (1987), no. 5-6, p. 983–1014.
- [DSC96] P. DIACONIS & L. SALOFF-COSTE – « Logarithmic Sobolev inequalities for finite Markov chains », *Ann. Appl. Probab.* **6** (1996), no. 3, p. 695–750.
- [Dud02] R. M. DUDLEY – *Real analysis and probability*, Cambridge Studies in Advanced Mathematics, vol. 74, Cambridge University Press, Cambridge, 2002.
- [Eva98] L. C. EVANS – *Partial differential equations*, Graduate Studies in Mathematics, vol. 19, American Mathematical Society, Providence, RI, 1998.
- [FKG71] C. M. FORTUIN, P. W. KASTELEYN & J. GINIBRE – « Correlation inequalities on some partially ordered sets », *Comm. Math. Phys.* **22** (1971), p. 89–103.
- [FW86] H. FÖLLMER & A. WAKOLBINGER – « Time reversal of infinite-dimensional diffusions », *Stochastic Process. Appl.* **22** (1986), no. 1, p. 59–77.
- [FW98] M. I. FREIDLIN & A. D. WENTZELL – *Random perturbations of dynamical systems*, Grundlehren der Mathematischen Wissenschaften, vol. 260, Springer-Verlag, New York, 1998.
- [Geo88] H.-O. GEORGII – *Gibbs measures and phase transitions*, Studies in mathematics, no. 9, de Gruyter, 1988.
- [GHM01] H.-O. GEORGII, O. HÄGGSTRÖM & C. MAES – « The random geometry of equilibrium phases », in *Phase transitions and critical phenomena, Vol. 18*, Phase Transit. Crit. Phenom., vol. 18, Academic Press, San Diego, CA, 2001, p. 1–142.
- [GJ87] J. GLIMM & A. JAFFE – *Quantum physics : A functional integral point of view*, 2 éd., Springer-Verlag, 1987.
- [GR01] I. GENTIL & C. ROBERTO – « Spectral gaps for spin systems : some non-convex phase examples », *Journal of Functional Analysis* (2001), no. 180, p. 66–84.
- [Gro75] L. GROSS – « Logarithmic Sobolev inequalities », *American Journal of Mathematics* **97** (1975), no. 4, p. 1061–1083.

- [GZ03] A. GUIONNET & B. ZEGARLIŃSKI – «Lectures on logarithmic Sobolev inequalities», in *Séminaire de Probabilités XXXVI*, Lecture Notes in Mathematics, vol. 1801, Springer-Verlag, 2003.
- [Haj88] B. HAJEK – «Cooling schedules for optimal annealing», *Mathematics of Operations Research* **13** (1988), p. 311–329.
- [HCS87] C.-R. HWANG, T.-S. CHIANG & S.-J. SHEU – «Diffusion for global optimization in \mathbb{R}^n », *Siam Journal on Control and Optimization* **25** (1987), p. 737–753.
- [Hel99] B. HELFFER – «Remarks on decay of correlations and Witten Laplacians. III. Application to logarithmic Sobolev inequalities», *Ann. Inst. H. Poincaré Probab. Statist.* **35** (1999), no. 4, p. 483–508.
- [Hel02] —, *Semiclassical analysis, Witten laplacians, and statistical mechanics*, Series on Partial Differential Equations and Applications, World Scientific, 2002.
- [Hel03] —, «Spectral theory and applications. an elementary introductory course», notes de cours de DEA, <http://www.math.u-psud.fr/~helfffer/deamain2002.ps>, 2003.
- [Hig81] Y. HIGUCHI – «On the absence of non-translation invariant Gibbs states for the two-dimensional Ising model», in *Random fields, Vol. I, II (Esztergom, 1979)*, Colloq. Math. Soc. János Bolyai, vol. 27, North-Holland, Amsterdam, 1981, p. 517–534.
- [HJJ03] D. HENDERSON, S. JACOBSON & A. JOHNSON – «The theory and practice of simulated annealing», in *Handbook on Metaheuristics* (F. Glover & G. Kochenberger, eds.), Kluwer Academic Publishers, Norwell MA, 2003.
- [HKS89] R. A. HOLLEY, S. KUSUOKA & D. W. STROOCK – «Asymptotics of the spectral gap with applications to the theory of simulated annealing», *Journal of Functional Analysis* **83** (1989), no. 2, p. 333–347.
- [Hol85] R. A. HOLLEY – «Possible rates of convergence in finite range, attractive spin systems», in *Particle systems, random media and large deviations (Brunswick, Maine, 1984)*, Contemp. Math., vol. 41, Amer. Math. Soc., Providence, RI, 1985, p. 215–234.
- [HS88] R. A. HOLLEY & D. W. STROOCK – «Simulated annealing via Sobolev inequalities», *Communications in Mathematical Physics* **115** (1988), no. 4, p. 553–569.
- [HS94] B. HELFFER & J. SJÖSTRAND – «On the correlation for Kac-like models in the convex case», *J. Statist. Phys.* **74** (1994), no. 1-2, p. 349–409.
- [Hwa80] C.-R. HWANG – «Laplace's method revisited : weak convergence of probability measures», *Annals of Probability* **8** (1980), p. 1177–1182.
- [Jac92] S. JACQUOT – «Comportement asymptotique de la seconde valeur propre des processus de Kolmogorov», *J. Multivariate Anal.* **40** (1992), no. 2, p. 335–347.
- [Jac94] —, «Simulated annealing on Wiener space», *Stochastics Stochastics Rep.* **51** (1994), no. 3-4, p. 159–194.
- [Jac96] —, «Simulated annealing for stochastic semilinear equations on Hilbert spaces», *Stochastic Process. Appl.* **64** (1996), no. 1, p. 73–91.
- [JR95] S. JACQUOT & G. ROYER – «Ergodicity of stochastic plates», *Probab. Theory Related Fields* **102** (1995), no. 1, p. 19–44.

- [Kal02] O. KALLENBERG – *Foundations of modern probability*, second éd., Probability and its Applications (New York), Springer-Verlag, New York, 2002.
- [KGV83] S. KIRKPATRICK, C. D. GELATT, JR. & M. P. VECCHI – « Optimization by simulated annealing », *Science* **220** (1983), no. 4598, p. 671–680.
- [Kun02] A. KUNZ – « On extremes of multidimensional stationary processes in euclidian norm », 2002, Prépublication, <http://www-m4.ma.tum.de/pers/kunz/pub/MxDf.pdf>.
- [LBB04] G. LUGOSI, S. BOUCHERON & O. BOUSQUET – « Concentration inequalities », in *Advanced lectures in machine learning* (O. Bousquet, U. von Luxburg & G. Rätsch, éd.), Springer, 2004, p. 208–240.
- [LD00] J.-L. LIONS & R. DAUTRAY – *Mathematical analysis and numerical methods for science and technology*, vol. 2, Springer, 2000.
- [Led99] M. LEDOUX – « Concentration of measure and logarithmic Sobolev inequalities », in *Séminaire de Probabilités, XXXIII*, Lecture Notes in Math., vol. 1709, Springer, Berlin, 1999, p. 120–216.
- [Led01] — , « Logarithmic Sobolev inequalities for unbounded spin systems revisited », in *Séminaire de Probabilités, XXXV*, Lecture Notes in Math., vol. 1755, Springer, Berlin, 2001, p. 167–194.
- [Lig85] T. M. LIGGETT – *Interacting particle systems*, Springer-Verlag, New York, 1985.
- [LML72] J. L. LEBOWITZ & A. MARTIN-LÖF – « On the uniqueness of the equilibrium state for Ising spin systems », *Comm. Math. Phys.* **25** (1972), p. 276–282.
- [LO00] R. LATAŁA & K. OLESZKIEWICZ – « Between Sobolev and Poincaré », in *Geometric aspects of functional analysis*, Lecture Notes in Math., vol. 1745, Springer, Berlin, 2000, p. 147–168.
- [Loc00] M. LOCATELLI – « Simulated annealing algorithms for continuous global optimization : Convergence conditions », *Journal of Optimization Theory and Applications* (2000), no. 104, p. 121–133.
- [LP76] J. L. LEBOWITZ & E. PRESUTTI – « Statistical mechanics of systems of unbounded spins », *Comm. Math. Phys.* **50** (1976), no. 3, p. 195–218.
- [LP81] — , « Erratum : “Statistical mechanics of systems of unbounded spins” [Comm. Math. Phys. **50** (1976), no. 3, 195–218; MR **56** #4579] », *Comm. Math. Phys.* **78** (1980/81), no. 1, p. 151.
- [LR69] O. LANFORD & D. RUELLE – « Observables at infinity and states with short range correlation in statistical mechanics », *Communications in mathematical physics* (1969), p. 194–215.
- [Mal75] V. A. MALYSHEV – « Phase transition in classical heisenberg ferromagnets with arbitrary parameters of anisotropy », *Communications in Mathematical Physics* **40** (1975), p. 75–82.
- [Mar06] K. MARTON – « An explicit bound on the logarithmic Sobolev constant of weakly dependent random variables », <http://arxiv.org/abs/math/0605397>, 2006.

- [Mat04] O. MATTE – « Spectral analysis in large and infinite dimension with applications to statistical mechanics », Thèse, J. Gutenberg Universität-Mainz, 2004.
- [Mic91] L. MICLO – « Thèse de doctorat », Thèse, Université Paris VI, 1991.
- [Mic92] — , « Recuit simulé sur \mathbb{R}^n . Étude de l'évolution de l'énergie libre », *Annales de l'Institut Henri Poincaré* **28** (1992), no. 2, p. 235–266.
- [Mic95] — , « Une étude des algorithmes de recuit simulé sous-admissibles », *Ann. Fac. Sci. Toulouse Math. (6)* **4** (1995), no. 4, p. 819–877.
- [MO94] F. MARTINELLI & E. OLIVIERI – « Approach to equilibrium of Glauber dynamics in the one phase region. I. The attractive case », *Comm. Math. Phys.* **161** (1994), no. 3, p. 447–486.
- [Muc72] B. MUCKENHOUPT – « Weighted norm inequalities for the Hardy maximal function », *Trans. Amer. Math. Soc.* **165** (1972), p. 207–226.
- [Ons44] L. ONSAGER – « Crystal statistics i : A two-dimensional model with order disorder transition », *Physical Reviews* (1944), no. 65, p. 117–149.
- [Pic79] P. PICCO – « Thèse de troisième cycle », Thèse, Université de Marseille, 1979.
- [Pru86] B. PRUM – *Processus sur un réseau et mesures de Gibbs*, Techniques stochastiques, Masson, Paris, 1986.
- [PS75] S. PIROGOV & Y. G. SINAI – « Phase diagram of classical systems, i », *Theoretical Mathematical Physics* (1975), no. 25, p. 1185–1192.
- [PW60] L. E. PAYNE & H. F. WEINBERGER – « An optimal Poincaré inequality for convex domains », *Archive for Rational Mechanics and Analysis* **5** (1960), p. 286–292.
- [Roy89] G. ROYER – « A remark on simulated annealing of diffusion processes », *SIAM Journal on Control and Optimization* **27** (1989), no. 6, p. 1403–1408.
- [Roy99] — , *Une initiation aux inégalités de Sobolev logarithmiques*, Cours spécialisés, no. 5, SMF, 1999.
- [Roy79] — , « Processus de diffusion associé à certains modèles d'Ising à spins continus », *Z. Wahrsch. Verw. Gebiete* **46** (1978/79), no. 2, p. 165–176.
- [RR91] M. M. RAO & Z. D. REN – *Theory of Orlicz spaces*, Monographs and Textbooks in Pure and Applied Mathematics, vol. 146, Marcel Dekker Inc., New York, 1991.
- [RS80] M. REED & B. SIMON – *Methods of modern mathematical physics. I*, 2 éd., Academic Press Inc. [Harcourt Brace Jovanovich Publishers], New York, 1980.
- [Rue72] D. RUELLE – « On the use of "small external fields" in the problem of Physics **69** (1972), p. 364–374.
- [RW01] M. RÖCKNER & F.-Y. WANG – « Weak Poincaré inequalities and L^2 convergence rates of Markov semigroups », *Journal of Functional Analysis* **185** (2001), p. 564–603.
- [Sin82] Y. G. SINAI – *Theory of phase transition. rigorous results.*, Pergamon Press, 1982.
- [SW87] R. SWENDSEN & J. WANG – « Nonuniversal critical dynamics in monte carlo simulations », *Phys. Rev. Lett.* **58** (1987), no. 2, p. 86–88.

- [SZ92a] D. W. STROOCK & B. ZEGARLIŃSKI – «The equivalence of the logarithmic Sobolev inequality and the Dobrushin-Shlosman mixing condition», *Communications in Mathematical Physics* **144** (1992), p. 303–323.
- [SZ92b] — , «The logarithmic Sobolev inequality for discrete spin systems on a lattice», *Comm. Math. Phys.* **149** (1992), no. 1, p. 175–193.
- [Vil03] C. VILLANI – *Topics in optimal transportation*, Graduate Studies in Mathematics, vol. 58, American Mathematical Society, Providence, RI, 2003.
- [Wlo87] J. WLOKA – *Partial differential equations*, Cambridge University Press, Cambridge, 1987.
- [Wu06] L. WU – «Poincaré and transportation inequalities for Gibbs measure under the Dobrushin uniqueness condition», to appear in *Annals of Probability*, 2006.
- [Yos97] N. YOSHIDA – «Relaxed criteria of the Dobrushin–Shlosman mixing condition», *Journal of Statistical Physics* **87** (1997), p. 293–309.
- [Yos99] — , «The log-Sobolev inequality for weakly coupled lattice fields», *Probability Theory and Related Fields* (1999), no. 115, p. 1–40.
- [Yos00] — , «Application of the log-Sobolev inequality to the stochastic dynamics of unbounded spin systems on the lattice», *Journal of Functional Analysis* **173** (2000), p. 74–102.
- [Yos01] — , «The equivalence of the log-Sobolev inequality and a mixing condition for unbounded spin systems on the lattice», *Annales de l'Institut H. Poincaré* **37** (2001), no. 2, p. 223–243.
- [Zah98] M. ZAHRADNÍK – «A short course on the Pirogov-Sinai theory», *Rend. Mat. Appl. (7)* **18** (1998), no. 3, p. 411–486.
- [Zeg96] B. ZEGARLIŃSKI – «The strong decay to equilibrium for the stochastic dynamics of unbounded spin systems on a lattice», *Comm. Math. Phys.* **175** (1996), no. 2, p. 401–432.