Séminaire BOURBAKI Octobre 2018
71e année, 2018–2019, nO 1151

NORMALITÉ ASYMPTOTIQUE DES VECTEURS PROPRES D'UN GRAPHE RÉGULIER ALÉATOIRE
[d'après Ágnes Backhausz et Balázs Szegedy]

par **Charles BORDENAVE**

INTRODUCTION

Nous présentons ici le résultat principal de Backhausz et Szegedy [4] et nous introduisons les définitions les plus importantes.

Graphe. — Un *graphe* $G = (V, E)$ est la paire formée par un ensemble dénombrable V et un ensemble E formé de parties à deux éléments de V. Les éléments de V et E sont appelés respectivement les *sommets* et les *arêtes* du graphe. Cette définition de graphe correspond aux graphes simples (ni boucles ni arêtes multiples). Le *degré* du sommet $x \in V$ est le nombre d'arêtes $e \in E$ telles que $x \in e$. Pour d entier, un graphe est dit *d-régulier* si tous ses sommets ont degré d.

Soient d, n des entiers positifs. Nous noterons $\mathbb{G}(n, d)$, l'ensemble des graphes $G = (V, E)$ d-réguliers tels que $V = \{1, \ldots, n\}$. Si dn est pair et $2 \leq d \leq n - 1$ alors $\mathbb{G}(n, d)$ n'est pas vide. Si dn est impair ou $d \geq n$, $\mathbb{G}(n, d)$ est vide.

Soit $G = (V, E)$ un graphe. Pour k entier, $x, y \in V$, un *chemin de longueur k de x à y* est une suite $(x_0, \ldots, x_k)$ telle que $x_0 = x$, $x_k = y$ et pour tout $t = 1, \ldots, k$, $\{x_{t-1}, x_t\} \in E$. Le graphe G est *connexe* si pour tous $x, y \in V$ il existe un chemin de x à y. La *distance* entre x et y est la longueur du plus court chemin qui les relie. Un *cycle* est un chemin $(x_0, \ldots, x_k)$ tel que $x_0 = x_k$ et les sommets $(x_1, \ldots, x_k)$ sont tous distincts. Enfin, un *arbre* est un graphe connexe et sans cycle.

Spectre. — On peut associer plusieurs opérateurs à un graphe $G = (V, E)$. Le plus simple est *l'opérateur d'adjacence*, défini dans $\ell^2(V)$ par, pour tous $f \in \ell^2(V)$ et $x \in V$,

$$Af(x) = \sum_{y:\{x,y\}\in E} f(y),$$

où la somme porte sur les sommets y qui partagent une arête avec x. Si le degré des sommets est uniformément borné, l'opérateur A est un opérateur borné et auto-adjoint.

Si le graphe G est d-régulier, l'opérateur $d^{-1}A$ est le noyau de transition de la marche au hasard sur G. Si G est un graphe d-régulier et V est fini alors le vecteur constant $\mathbf{1} \in \ell^2(V)$ défini pour tout $x \in V$ par $\mathbf{1}(x) = 1$ est un vecteur propre de A associé à la valeur propre d.

Il existe de nombreuses relations entre la géométrie du graphe et les propriétés spectrales de son opérateur d'adjacence. Ces dernières années un effort de recherche tout particulier vise à décrire le spectre lorsque le graphe sous-jacent est aléatoire ou quasi-aléatoire. Pour ne citer que quelques références récentes relatives aux graphes réguliers, des outils de combinatoire [19, 11], de la théorie des matrices aléatoires [17, 16, 7, 6, 22], de l'analyse semi-classique [14, 3, 2] ont été utilisés dans ce contexte. Nous allons présenter ici une méthode très originale d'analyse du spectre développée dans [4]. Elle repose sur des liens fructueux entre la théorie de l'information et la convergence locale des graphes (sur les liens entre convergence locale des graphes et spectre, voir [12]). Le résultat principal de Backhausz et Szegedy portera sur les vecteurs propres de A orthogonaux à $\mathbf{1}$ lorsque G est distribué suivant la mesure uniforme sur $\mathbb{G}(n,d)$. Dans la suite de cette introduction, nous allons d'abord énoncer ce résultat sous sa forme la plus simple pour ensuite donner l'énoncé général qui exige une certaine préparation.

Soit $A \in M_n(\mathbb{R})$ une matrice symétrique de taille n. Pour tous $\delta \in [0,1[$ et $\lambda \in \mathbb{R}$, un *quasi-vecteur propre de déviation* δ *et de valeur propre* λ est un vecteur non nul $f \in \mathbb{R}^n$ tel que $\|Af - \lambda f\|_2 \leq \delta \|f\|_2$ où $\|\cdot\|_2$ est la norme euclidienne. Un *quasi-vecteur propre de déviation* δ est un quasi-vecteur propre de déviation δ et de valeur propre λ pour un certain $\lambda \in \mathbb{R}$. Évidemment, un vecteur propre est un quasi-vecteur propre de déviation nulle.

Distribution d'un vecteur. — À tout vecteur $f \in \mathbb{R}^n$, la *mesure empirique de ses entrées* est la mesure de probabilité $\mathrm{distr}(f)$ sur $\mathbb{R}$:

$$\mathrm{distr}(f) = \frac{1}{n} \sum_{x=1}^{n} \delta_{f(x)},$$

où δ est la distribution de Dirac. En termes probabilistes, $\mathrm{distr}(f)$ est la loi de $f(o)$ où o est uniformément distribué sur $\{1,\dots,n\}$. La loi $\mathrm{distr}(f)$ est donc la *loi d'une entrée typique*. Le second moment de $\mathrm{distr}(f)$ est égal à $\|f\|_2^2/n$ de telle sorte que si f est non nul, $\mathrm{distr}(\sqrt{n} f/\|f\|_2)$ a un second moment égal à 1.

Gaussianité d'un vecteur. — Pour tout $\sigma \in \mathbb{R}_+$, on note traditionnellement $N(0,\sigma^2)$ la mesure de probabilité gaussienne sur $\mathbb{R}$ centrée et de variance σ^2. Pour $\sigma = 0$, on pose $N(0,0) = \delta_0$. La *proximité d'un vecteur* $f \in \mathbb{R}^n$ *non nul à un vecteur gaussien centré* est définie comme

$$D(f) = \inf_{\sigma \in [0,1]} \mathrm{d_L}\Big(\mathrm{distr}(\sqrt{n} f/\|f\|_2), N(0,\sigma^2)\Big),$$

où $\mathrm{d_L}$ est la *distance de Lévy-Prohorov* (toute autre distance qui définit la topologie faible conviendrait également). Rappelons que si μ et ν sont deux mesures de probabilités sur un espace métrique (M,d) alors, en notant $B(M)$ sa tribu borélienne, on a

$$\mathrm{d_L}(\mu,\nu) = \inf \{\varepsilon > 0 : \forall A \in B(M),\ \mu(A) \leq \nu(A^\varepsilon) + \varepsilon \text{ and } \nu(A) \leq \mu(A^\varepsilon) + \varepsilon\},$$

où pour tous $A \in B(M)$ et $\varepsilon > 0$, on note A^ε l'ensemble des éléments de M à distance au plus ε de A.

Considérons deux cas extrêmes de vecteurs $f \in \mathbb{R}^n$ « proches » d'un vecteur gaussien au sens que $D(f)$ est petit. Si f est un vecteur de la base canonique alors $\mathrm{distr}(\sqrt{n}f) = 1/n \cdot \delta_{\sqrt{n}} + (1-1/n) \cdot \delta_0$. Il est donc immédiat que lorsque n tend vers l'infini, $\mathrm{distr}(\sqrt{n}f)$ converge faiblement vers $\delta_0 = N(0,0)$. Si f est un vecteur aléatoire gaussien standard dans $\mathbb{R}^n$ (les coordonnées de f dans une base orthogonale sont des variables gaussiennes $N(0,1)$ indépendantes) alors la loi faible des grands nombres implique que lorsque n tend vers l'infini, la variable aléatoire $\mathrm{d_L}(\mathrm{distr}(f), N(0,1))$ converge en probabilité vers 0 et $\|f\|_2/\sqrt{n}$ vers 1. Dans ces deux cas, $D(f)$ tend vers 0 mais pour des raisons différentes. Dans le premier cas, la norme ℓ^2 du vecteur f est localisée sur $o(n)$ coordonnées (ici une seule). Cela se traduit par une perte de masse ℓ^2 dans le sens que $\mathrm{distr}(\sqrt{n}f/\|f\|_2)$, dont le second moment est égal à 1, converge faiblement vers δ_0, dont le second moment est nul. Dans le second cas, en probabilité, il n'y a pas de perte de masse et la mesure empirique $\mathrm{distr}(\sqrt{n}f/\|f\|_2)$ converge vers la loi d'une coordonnée de f.

Plus généralement, le lemme élémentaire suivant décrit la proximité à un vecteur gaussien en terme d'une partie localisée et d'une partie proprement asymptotiquement gaussienne.

Lemme 0.1. — *Soit $\varepsilon > 0$. Il existe $\delta > 0$ tel que pour tout entier $n \geq 1$ si $f \in \mathbb{R}^n$ vérifie $\|f\|_2 = \sqrt{n}$ et $\mathrm{d_L}(\mathrm{distr}(f), N(0,\sigma^2)) \leq \delta$ pour un certain $\sigma \in [0,1]$ alors il existe $\tau \in [0,1]$, $a, b \in \mathbb{R}^n$ tels que f se décompose de la façon suivante*

$$f = \tau a + \sqrt{1-\tau^2} b,$$

avec $\langle a, b\rangle = 0$, $\|a\|_2 = \|b\|_2 = \sqrt{n}$, $|\sigma - \tau| \leq \varepsilon$, $\mathrm{d_L}(\mathrm{distr}(a), N(0,1)) \leq \varepsilon$ et $\mathrm{d_L}(\mathrm{distr}(b), \delta_0) \leq \varepsilon$.

Théorème principal, forme simplifiée. — La forme simplifiée du théorème principal de [4] est la suivante.

Théorème 0.2. — *Soient un entier $d \geq 3$ et $\varepsilon > 0$. Il existe m et δ strictement positifs tels que, pour tout entier $n \geq m$ avec nd pair, si G est un graphe distribué uniformément sur $\mathbb{G}(n,d)$ et A est sa matrice d'adjacence, alors l'événement suivant a une probabilité au moins $1-\varepsilon$: tout quasi-vecteur propre f de déviation δ et orthogonal à $\mathbf{1}$ vérifie $D(f) \leq \varepsilon$.*

En regard du lemme 0.1, le théorème 0.2 affirme donc que si f est un vecteur propre de norme 1 de la matrice d'adjacence d'un graphe d-régulier uniforme avec n sommets alors, lorsque n tend l'infini, f se décompose en une partie de masse ℓ^2 concentrée sur $o(n)$ entrées et le reste est asymptotiquement gaussien. Il est conjecturé que la partie localisée de cette décomposition est triviale :

Conjecture 0.3. — *Soient un entier $d \geq 3$ et $\varepsilon > 0$. Il existe m positif tel que, pour tout entier $n \geq m$ avec nd pair, si G est un graphe distribué uniformément sur $\mathbb{G}(n,d)$*

et A est sa matrice d'adjacence alors l'événement suivant a une probabilité au moins $1-\varepsilon$: tout vecteur propre f orthogonal à $\mathbf{1}$ vérifie $\mathrm{d_L}(\mathrm{distr}(\sqrt{n}f/\|f\|_2), N(0,1)) \leq \varepsilon$.

Il est important de remarquer que dans la conjecture ci-dessus, on se restreint aux vecteurs propres. Nous verrons en effet que pour tout $\sigma \in [0,1]$, il existe une suite de quasi-vecteurs propres f_n de déviation δ_n avec $\delta_n \to 0$ telle que, en probabilité, $\mathrm{d_L}(\mathrm{distr}(\sqrt{n}f_n/\|f_n\|_2), N(0,\sigma^2)) \to 0$. Voir remarque 2.2.

Nous allons maintenant énoncer le résultat principal de [4]. L'objectif est de mesurer la distance d'un vecteur propre f de la matrice d'adjacence de G uniforme sur $\mathbb{G}(n,d)$ à un processus gaussien qui satisfait, avec probabilité 1, l'équation de vecteur propre sur le revêtement universel de G. Nous allons d'abord expliquer tous ces termes.

Arbre régulier infini. — Pour $d \geq 2$ entier, nous noterons $T_d = (V_d, E_d)$ l'*arbre d-régulier infini*. Plus précisément, il sera commode d'utiliser la notation généalogique suivante. L'ensemble des sommets de T_d est le sous-ensemble suivant des suites finies d'entiers, $\varnothing$ représentant la suite vide,

$$V_d = \{\varnothing\} \cup \{1,\ldots,d\} \cup \bigcup_{k=1}^{\infty} \{1,\ldots,d\} \times \{1,\ldots,d-1\}^k.$$

On définit le *parent* de $x \in \{1,\ldots,d\}$ comme $\varnothing$. Pour $k \geq 1$ entier, le parent de $x = (x_1,\ldots,x_{k+1}) \in \mathbb{N}^{k+1} \cap V_d$ est $(x_1,\ldots,x_k) \in \mathbb{N}^k \cap V_d$. Enfin $\varnothing$ n'a pas de parent. L'ensemble des arêtes de T_d est

$$E_d = \{\{x,y\} : x \text{ est parent de } y\}.$$

Nous dirons que y est un *enfant* de x si x est le parent de y. Nous dirons que y est un *descendant* de x si il existe une suite $(x_0,\ldots,x_k)$ avec $x_0 = x$, $x_k = y$ et x_{t-1} est le parent de x_t pour tout entier $1 \leq t \leq k$ (en d'autres termes, x est un suffixe de y).

Pour d pair, l'arbre T_d est isomorphe au graphe de Cayley du groupe libre engendré par $d/2$ générateurs libres et leurs inverses. Pour tout d entier, T_d est aussi le graphe de Cayley du groupe librement engendré par d copies de $\mathbb{Z}/2\mathbb{Z}$ et ses générateurs libres.

Revêtement aléatoire. — Soit $G = (V,E)$ un graphe d-régulier. Nous dirons qu'une application $\pi \colon V_d \to V$ est un *revêtement de G* si π est un homéomorphisme local au sens où pour tout $x \in V_d$, l'image par π des d voisins de x dans T_d sont les d voisins de $\pi(x)$ dans G. Remarquons que $\pi(V_d)$ est alors une composante connexe de G.

Tout graphe d-régulier admet un revêtement. Si V est fini, alors il existe une mesure de probabilité naturelle sur l'ensemble de revêtements de G. Si G est connexe, ce *revêtement aléatoire* π est caractérisé par la propriété que pour tout automorphisme ϕ de T_d, $\pi \circ \phi$ et π ont même loi. Si G n'est pas connexe, pour caractériser de façon unique cette mesure, nous supposons également que $\pi(\varnothing)$ est uniformément distribuée sur V.

Ce revêtement aléatoire peut être construit explicitement de la façon suivante. Soit $(\sigma_x)_{x\in V_d}$ une famille de permutations aléatoires indépendantes, $\sigma_\varnothing$ permutation uniformément distribuée dans S_d (le groupe symétrique à d éléments) et pour tout $x \in V_d\backslash\{\varnothing\}$, σ_x uniforme dans S_{d-1}. Le revêtement aléatoire est défini itérativement de la façon suivante. Tout d'abord, $\pi(\varnothing)$ est distribué uniformément sur V et pour tout $i \in \{1, \ldots, d\}$, $\pi(i) = v_{\sigma_\varnothing(i)}$ où $\{v_1, \ldots, v_d\}$ sont les voisins de $\pi(\varnothing)$ dans G. Ensuite, par récurrence sur $k \geq 1$, pour tout $x \in \mathbb{N}^k \cap V_d$, on étend π en posant pour tout $i \in \{1, \ldots, d-1\}$, $\pi((x,i)) = v_{\sigma_x(i)}$ où $\{v_1, \ldots, v_{d-1}\}$ sont les $d-1$ voisins de $\pi(x)$ différents de $\pi(y)$, y étant le parent de x.

Distribution d'un vecteur sur le graphe. — La preuve du théorème 0.2 repose sur la possibilité d'associer à la paire formée par un graphe d-régulier $G = (V, E)$ et un vecteur $f \in \mathbb{R}^V$ un processus aléatoire sur T_d qui est invariant par les automorphismes de l'arbre. Cette idée à la fois simple et profonde remonte au moins à Benjamini et Schramm [9]. Plus précisément, soit $G = (V, E) \in \mathbb{G}(n, d)$ un graphe d-régulier avec sommets $V = \{1, \ldots, n\}$. On peut définir la *loi d'un vecteur vue d'un sommet typique du graphe* de la façon suivante. Soient M un espace métrique et $f \in M^V$. On munit M^{V_d} d'une distance qui induit la topologie produit. On définit $\mathrm{distr}_G(f)$ comme la mesure de probabilité sur M^{V_d}

$$\mathrm{distr}_G(f) = \text{loi de } f \circ \pi,$$

où $\pi\colon V_d \to V$ est le revêtement aléatoire de G défini ci-dessus.

Pour $f \in \mathbb{R}^V$, remarquons que $\mathrm{distr}_G(f)$ contient plus d'information que $\mathrm{distr}(f)$. En effet $\mathrm{distr}(f)$ est la loi image de $\mathrm{distr}_G(f)$ par l'application de $\mathbb{R}^{V_d}$ dans $\mathbb{R}$: $g \mapsto g(\varnothing)$. Notons aussi que $\mathrm{distr}_G(f)$ hérite des propriétés de symétrie du revêtement aléatoire π : $\mathrm{distr}_G(f)$ est invariant par les automorphismes de T_d.

Onde aléatoire et onde gaussienne. — Si M est un espace métrique, une variable aléatoire $Y = (Y_x)_{x \in V_d}$ dans M^{V_d} est dite *invariante* si sa loi est invariante par les automorphismes de l'arbre T_d : pour tout automorphisme ϕ de T_d, $Y \circ \phi$ et Y ont même loi.

Si $M = \mathbb{R}$, la variable aléatoire Y est dite *standard* si $\mathbb{E}Y_x = 0$ et $\mathbb{E}Y_x^2 = 1$ pour tout $x \in V_d$. La variable Y est *gaussienne* si toute distribution jointe d'un nombre fini de ses marginales est gaussienne. La variable Y est une *onde aléatoire* de valeur propre λ si elle satisfait avec probabilité 1 l'équation de vecteur propre : pour tout $x \in V_d$,

$$\sum_{y : \{x,y\} \in E_d} Y_y = \lambda Y_x, \tag{1}$$

(c'est-à-dire avec probabilité 1, Y est un vecteur propre généralisé de l'opérateur A).

Il est démontré dans [18, 20] que pour tout $\lambda \in [-2\sqrt{d-1}, 2\sqrt{d-1}]$, il existe une unique onde aléatoire gaussienne standard invariante de valeur propre λ. Nous appellerons cette variable aléatoire simplement l'*onde gaussienne* de valeur propre λ et nous la noterons Y_λ. Nous construirons cette onde gaussienne dans la section 2. Pour $\sigma \in \mathbb{R}$, notons $\mathrm{Gauss}_{\lambda,\sigma}$ la loi de σY_λ. Nous pouvons mesurer la distance d'un vecteur non nul $f \in \mathbb{R}^n$ construit sur le graphe $G \in \mathbb{G}(n, d)$ à une onde gaussienne comme

$$D_G(f) = \inf_{\substack{|\lambda| \leq 2\sqrt{d-1} \\ \sigma \in [0,1]}} \mathrm{d_L}(\mathrm{distr}_G(\sqrt{n} f / \|f\|_2), \mathrm{Gauss}_{\lambda,\sigma}),$$

où $\mathrm{d_L}$ est la distance de Lévy-Prohorov pour les mesures de probabilités sur $\mathbb{R}^{V_d}$ (ou toute autre distance qui définit la convergence faible). À nouveau, la présence de $\sigma < 1$ dans la définition de $D_G(f)$ autorise une concentration de la norme ℓ^2 de f sur $o(n)$ entrées.

Théorème principal. — On peut maintenant énoncer le résultat principal de [4].

Théorème 0.4. — *Soient un entier $d \geq 3$ et $\varepsilon > 0$. Il existe m et δ positifs tels que, pour tout entier $n \geq m$ avec nd pair, si G est un graphe distribué uniformément sur $\mathbb{G}(n, d)$ et A est sa matrice d'adjacence alors l'événement suivant a une probabilité au moins $1 - \varepsilon$: tout quasi-vecteur propre f de déviation δ et orthogonal à $\mathbf{1}$ vérifie $D_G(f) \leq \varepsilon$.*

Organisation. — La suite de ce texte est organisée de la façon suivante. Dans la section 1, nous allons réduire l'énoncé du théorème principal à un énoncé sur une famille d'ondes invariantes standard. La section 2 présente la théorie spectrale des ondes invariantes. Nous allons notamment y montrer l'existence de l'onde gaussienne. La stratégie de la démonstration du résultat principal qui repose sur des inégalités entropiques est présentée dans la section 3. Dans la section 4 nous établirons une première inégalité entropique importante qui sera utilisée dans la section 5. Dans les sections 5 et 6, nous terminerons la démonstration en montrant respectivement que l'onde gaussienne minimise une entropie dans une certaine famille d'ondes invariantes standard et qu'elle maximise cette même entropie parmi les ondes invariantes standard.

Notations. — *Dans toute la suite de ce texte, nous fixons un entier $d \geq 3$.* Pour M espace métrique et V dénombrable, un élément $f \in M^V$ sera noté indifféremment comme un vecteur $(f_x)_{x \in V}$ ou une fonction de V dans M, $x \mapsto f(x)$. L'espérance des variables aléatoires est notée $\mathbb{E}(\cdot)$ et la mesure de probabilité correspondante est $\mathbb{P}(\cdot)$. Enfin, les termes variables aléatoires et processus sont synonymes.

1. PROCESSUS TYPIQUE

Processus typique. — La notion de processus typique fut introduite dans [5]. Nous allons donner une nouvelle version du théorème principal de [4] qui repose sur cette définition. Soient M un espace métrique, μ une mesure de probabilité sur M^{V_d} et Y une variable aléatoire de loi μ. Soit $G \in \mathbb{G}(n, d)$ un graphe d-régulier sur l'ensemble de sommets $V = \{1, \ldots, n\}$. Une mesure de la proximité de μ à l'ensemble des vecteurs construits sur G est

$$\Delta(\mu, G) = \inf\{\mathrm{d_L}(\mathrm{distr}_G(f), \mu) : f \in M^n\}.$$

Définition 1.1. — *Un processus Y sur T_d ou sa loi μ est* typique *si la propriété d'approximation suivante a lieu. Pour tous $\varepsilon > 0$ et $m > 0$, il existe un entier $n \geq m$, tel* *que si G est un graphe aléatoire uniformément distribué sur $\mathbb{G}(n,d)$ alors $\mathbb{P}(\Delta(\mu,G) \leq \varepsilon) \geq 1-\varepsilon$.*

Avec des mots : un processus typique est un processus sur T_d qui peut être approché avec grande probabilité sur une suite de graphes aléatoires finis. Notons qu'un processus typique est nécessairement invariant. Pour illustrer la définition des processus typiques, nous allons en donner une première application.

Théorème de Friedman. — Le spectre de l'opérateur d'adjacence sur T_d est

$$\sigma(T_d) = [-2\sqrt{d-1}, 2\sqrt{d-1}].$$

Le théorème de Friedman [19] affirme que toutes les valeurs propres non triviales d'un graphe aléatoire distribué uniformément sur $\mathbb{G}(n,d)$ sont avec grande probabilité dans un ε-voisinage de $\sigma(T_d)$ si n est suffisamment grand. Plus précisément,

Théorème 1.2. — *Soit $\varepsilon > 0$. Il existe m positif tel que, pour tout entier $n \geq m$ avec nd pair, si G est un graphe distribué uniformément sur $\mathbb{G}(n,d)$ et A est sa matrice d'adjacence, alors l'événement suivant a une probabilité au moins $1-\varepsilon$: pour tout $\delta > 0$, tout quasi-vecteur propre de déviation δ et orthogonal à $\mathbf{1}$ a une valeur propre associée λ qui satisfait $|\lambda| \leq 2\sqrt{d-1} + \varepsilon + \delta$.*

Le théorème de Friedman correspond au cas $\delta = 0$ (les valeurs propres), pour une preuve simplifiée et quantitative voir [11]. Le cas général se déduit facilement du cas $\delta = 0$ en remarquant que $\mathbf{1}^\perp$, l'hyperplan de $\mathbb{R}^n$ orthogonal à $\mathbf{1}$, est invariant par la matrice d'adjacence A. La matrice A étant symétrique, la norme d'opérateur de A restreint à $\mathbf{1}^\perp$ est égale à la valeur absolue de la plus grande valeur propre γ associée à un vecteur propre dans $\mathbf{1}^\perp$ (c'est la seconde plus grande valeur propre de A en valeur absolue). Donc, pour tout $f \in \mathbf{1}^\perp$, $\|Af\|_2 \leq |\gamma| \|f\|_2$. Si, de plus, on a $\|Af - \lambda f\|_2 \leq \delta \|f\|_2$ pour un certain réel λ et $\delta > 0$, alors l'inégalité triangulaire implique $|\lambda| \leq |\gamma| + \delta$. Nous en déduisons bien le théorème 1.2.

En terme de processus typique, nous obtenons le corollaire suivant.

Corollaire 1.3. — *Si Y est une onde typique standard de valeur propre λ alors $\lambda \in \sigma(T_d)$.*

Démonstration. — Le processus Y étant typique, pour tous m et $\varepsilon > 0$, il existe un entier $n \geq m$ tel que si G est uniforme sur $\mathbb{G}(n,d)$, avec probabilité au moins $1/2$, il existe un vecteur $f \in \mathbb{R}^n$ telle que $\mathrm{d_L}(\mathrm{distr}_G(f), \mu) \leq \varepsilon$ où μ est la loi de Y.

Fixons maintenant un graphe $G \in \mathbb{G}(n,d)$. Nous allons construire à la main un quasi-vecteur propre de A, la matrice d'adjacence de G. Pour x réel et $t > 0$, on pose $(x)_t = x \mathbf{1}_{|x| \leq t}$. Étant donné que $\mathbb{E} Y_\varnothing = 0$ et $\mathbb{E} Y_\varnothing^2 = 1$, pour tout $\delta > 0$, il existe $t > 0$ tel que $|\mathbb{E}[(Y_\varnothing)_t]| \leq \delta$ et $|\mathbb{E}[(Y_\varnothing)_t^2] - 1| \leq \delta^2$. On peut aussi supposer que t est un point de continuité de la loi de $Y_\varnothing$. Donc pour tout $\delta > 0$, il existe $t, \varepsilon > 0$ tels que si $f \in \mathbb{R}^n$ vérifie $\mathrm{d_L}(\mathrm{distr}_G(f), \mu) \leq \varepsilon$ alors $|\mathbb{E}[(f(o))_t]| \leq \delta$, $|\mathbb{E}[(f(o))_t^2] - 1| \leq \delta^2$ et $\mathbb{P}(|(Af)(o) - \lambda f(o)| \geq \delta) \leq \delta / t^2$, où o est uniformément distribué sur $\{1, \ldots, n\}$.

Si $0 < \delta \leq 1/2$, on considère le vecteur centré normalisé $g \in \mathbb{R}^n$:

$$g(x) = \frac{(f(x))_t - \mathbb{E}[(f(o))_t]}{\sigma_t},$$

où σ_t^2 est la variance de $(f(o))_t$. On a $|\sigma_t^2 - 1| \leq |\mathbb{E}[(f(o))_t^2] - 1| + (\mathbb{E}[(f(o))_t])^2 \leq 2\delta^2 \leq 1/2$. En outre, par construction, $\|g\|_\infty \leq (t+\delta)/\sqrt{1-2\delta^2} \leq 2t$ si $t \geq 2$. Nous avons aussi $\mathbb{E}g(o) = 0$, ceci est équivalent à g orthogonal à $\mathbf{1}$. Enfin $\mathbb{E}g(o)^2 = 1$, autrement dit $\|g\|_2 = \sqrt{n}$. Ainsi, pour tout $\delta > 0$, il existe $t, \varepsilon > 0$ tels que si $f \in \mathbb{R}^n$ vérifie $\mathrm{d_L}(\mathrm{distr}_G(f), \mu) \leq \varepsilon$ alors il existe un vecteur g orthogonal à $\mathbf{1}$ et de norme $\|g\|_2 = \sqrt{n}$ tel que $\|g\|_\infty \leq t$ et $\{x : |(Ag)(x) - \lambda g(x)| \geq \delta\}$ a au plus $\delta n/t^2$ éléments. On en déduit que $\|Ag - \lambda g\|_2^2 \leq n\delta^2 + (\delta n/t^2)(d + |\lambda|)^2 t^2 \leq \|g\|_2^2(\delta^2 + (d+|\lambda|)^2\delta)$.

En conclusion, si Y est une onde typique de valeur propre λ alors, pour tous m et $\delta > 0$, il existe $n \geq m$ tel que, si G est un uniforme sur $\mathbb{G}(n,d)$, avec probabilité au moins $1/2$, la matrice d'adjacence de G a un quasi-vecteur propre de valeur propre λ, de déviation δ et orthogonal à $\mathbf{1}$. Il reste à appliquer le théorème 1.2. $\square$

Théorème principal, version onde typique. — Nous allons maintenant donner un énoncé qui implique le théorème principal 0.4. Dans cet énoncé, il n'y a plus de quantificateur, il s'agit seulement d'une propriété des ondes typiques.

Nous démarrons avec un théorème de Harangi et Virág [20] sur l'approximation de l'onde gaussienne.

Théorème 1.4. — *Soit $\lambda \in \sigma(T_d)$. L'onde gaussienne de valeur propre λ est typique.*

Nous démontrerons ce résultat en même temps que nous construirons l'onde gaussienne dans la section 2. Dans [4], Backhausz et Szegedy ont démontré la réciproque du théorème 1.4 :

Théorème 1.5. — *Soit $\lambda \in \sigma(T_d)$. Si Y est une onde typique standard de valeur propre λ alors Y est l'onde gaussienne de valeur propre λ.*

Dans le reste de cette section, nous allons démontrer que l'énoncé du théorème 1.5 implique le théorème 0.4. L'argument repose sur des considérations topologiques générales et sur une propriété de *concentration de la mesure uniforme* sur $\mathbb{G}(n,d)$. La démonstration de ce lemme est indépendante du reste de la preuve (elle peut être sautée sans nuire à la compréhension du reste du texte).

Concentration de la mesure. — Soit $\mathcal{M}$ l'ensemble des mesures de probabilité μ sur $\mathbb{R}^{V_d}$ qui sont invariantes par les automorphismes de T_d et telle que $\mathbb{E}Y(\varnothing)^2 \leq 1$ où Y a loi μ. Muni de la distance de Lévy-Prohorov $\mathrm{d_L}$, $\mathcal{M}$ est un espace compact. Soit $G \in \mathbb{G}(n,d)$, on définit l'*ensemble des processus typiques bornés construit sur G* comme

$$T(G) = \{\mathrm{distr}_G(f) : f \in \mathbb{R}^n, \|f\|_2 \leq \sqrt{n}\}.$$

Soit $\mathcal{F}$ l'ensemble des parties fermées de $\mathcal{M}$. Pour tout $G \in \mathbb{G}(n,d)$, on a $T(G) \in \mathcal{F}$. Nous notons $\mathrm{d_H}$ la distance de Hausdorff sur $\mathcal{F}$. L'ensemble $T(G)$ est concentré sous la mesure uniforme sur $\mathbb{G}(n,d)$. Plus précisément, nous allons vérifier le résultat suivant.

Lemme 1.6. — *Pour tout $\varepsilon > 0$, il existe m tel que pour tout entier $n \geq m$ avec nd pair on a, pour un certain $F = F(n, d) \in \mathcal{F}$,*

$$\mathbb{P}(\mathrm{d_H}(T(G), F) \geq \varepsilon) \leq \varepsilon,$$

où G est un graphe aléatoire uniforme sur $\mathbb{G}(n, d)$.

Le théorème suivant est un résultat classique de concentration de la mesure sur $\mathbb{G}(n, d)$. Nous dirons que deux graphes $G, G' \in \mathbb{G}(n, d)$ sont *voisins* si la différence symétrique de leurs ensembles d'arêtes est 4 (c'est-à-dire si G et G' coïncident à la modification près de deux arêtes).

Théorème 1.7. — *Soient $n \geq d + 1$ un entier avec nd pair, $\delta > 0$ et $h \colon \mathbb{G}(n, d) \to \mathbb{R}$ une fonction telle que $|h(G) - h(G')| \leq \delta$ pour tous graphes voisins $G, G' \in \mathbb{G}(n, d)$. Alors, si G est uniformément distribué sur $\mathbb{G}(n, d)$, pour tout $t \geq 0$,*

$$\mathbb{P}(|h(G) - \mathbb{E}h(G)| \geq t) \leq c \exp\left(-\frac{t^2}{2\delta^2 dn}\right),$$

pour une certaine constante $c = c(d)$.

Voir [24, théorème 2.19 – lemme 2.1] pour une preuve. En admettant le théorème 1.7, nous pouvons démontrer le lemme 1.6.

Démonstration du lemme 1.6. — Soit $\varepsilon > 0$. L'espace métrique $(\mathcal{F}, \mathrm{d_H})$ étant compact, il existe une partie finie $\mathcal{R}$ qui est ε-dense, c'est-à-dire telle que $\mathcal{R}^\varepsilon = \mathcal{F}$. Notons r le cardinal de $\mathcal{R}$. Le principe des tiroirs implique qu'il existe $F \in \mathcal{R}$ tel que, si G est uniformément distribué sur $\mathbb{G}(n, d)$,

$$\mathbb{P}(\mathrm{d_H}(T(G), F) < \varepsilon) \geq 1/r.$$

Pour démontrer le lemme, il suffit donc de prouver que pour tout $\varepsilon > 0$ il existe m tel que pour tout entier $n \geq m$ et tout $S \in \mathcal{F}$ on a

$$\mathbb{P}(|\mathrm{d_H}(T(G), S) - \mathbb{E}\mathrm{d_H}(T(G), S)| \geq \varepsilon) \leq \varepsilon. \tag{2}$$

En effet, on choisit $\varepsilon_0 < \min(\varepsilon, 1/r)$ dans l'équation (2), pour en déduire que $\mathbb{E}\mathrm{d_H}(T(G), F) < 2\varepsilon$. On obtient donc l'inclusion d'événements

$$\{\mathrm{d_H}(T(G), F) \geq 3\varepsilon\} \subset \{|\mathrm{d_H}(T(G), F) - \mathbb{E}\mathrm{d_H}(T(G), F)| \geq \varepsilon\}.$$

Il reste à appliquer une nouvelle fois (2) et à ajuster la valeur de ε pour obtenir le lemme 1.6.

Nous passons à la preuve de (2). Soit B_k l'ensemble des sommets de V_d à distance au plus k de $\varnothing$. Si $\mu \in \mathcal{M}$, désignons par μ_{B_k} la marginale de μ sur B_k. De même, si $S \in \mathcal{F}$, S_{B_k} est l'ensemble formé par les marginales des mesures dans S. La topologie produit sur $\mathbb{R}^{V_d}$ étant engendrée par les cylindres, pour montrer (2), il suffit de vérifier que pour tous $\varepsilon > 0$ et $k \in \mathbb{N}$, il existe m tel que pour tout entier $n \geq m$ et $S \in \mathcal{F}$ on a

$$\mathbb{P}(|\mathrm{d_H}(T(G)_{B_k}, S_{B_k}) - \mathbb{E}\mathrm{d_H}(T(G)_{B_k}, S_{B_k})| \geq \varepsilon) \leq \varepsilon. \tag{3}$$

Pour tous $f \in \mathbb{R}^n$ et $x \in \{1, \ldots, n\}$, soit $\mathrm{distr}_{(G,x)}(f)$ la loi de $\mathrm{distr}_G(f)$ conditionnée à ce que le revêtement aléatoire π vérifie $\pi(\varnothing) = x$. Nous avons

$$\mathrm{distr}_G(f) = \frac{1}{n} \sum_{x=1}^{n} \mathrm{distr}_{(G,x)}(f).$$

En outre, $\mathrm{distr}_{(G,x)}(f)_{B_k} = \mathrm{distr}_{(G',x)}(f)_{B_k}$ si les sous-graphes de G et G' engendrés par les sommets qui sont à distance au plus k de x sont égaux. Si G et G' sont voisins, ces sous-graphes sont égaux pour tout x qui n'est pas à distance au plus k d'une arête qui est dans la différence symétrique de G et G'. On obtient que si G et G' sont voisins alors

$$\mathrm{d_L}(\mathrm{distr}_G(f)_{B_k}, \mathrm{distr}_{G'}(f)_{B_k}) \leq 4d(d-1)^{k-1}/n = \delta.$$

En particulier par l'inégalité triangulaire, $|\mathrm{d_H}(T(G)_{B_k}, S_{B_k}) - \mathrm{d_H}(T(G')_{B_k}, S_{B_k})| \leq \delta$ si G et G' sont voisins. Enfin, on applique le théorème 1.7 pour $h(G) = \mathrm{d_H}(T(G)_{B_k}, S_{B_k})$ et on en déduit (3). $\square$

Tous les ingrédients sont réunis pour démontrer que le théorème 1.5 implique le théorème 0.4.

Démonstration du théorème 0.4 (comme conséquence du théorème 1.5)

Si $\mu \in \mathcal{M}$ et $S \in \mathcal{F}$, on pose $\mathrm{d_L}(\mu, S) = \inf\{\mathrm{d_L}(\mu, \nu) : \nu \in S\}$. Soit $S \in \mathcal{F}$ l'ensemble des ondes gaussiennes de T_d.

Raisonnons par contradiction. Si la conclusion du théorème 0.4 est fausse, il existe $\varepsilon_0 \in (0, 1]$, une suite strictement croissante d'entiers $(n_k)_k$ et une suite $(\delta_k)_k$ tendant vers 0 tels que l'événement suivant a une probabilité au moins ε_0 sous la loi uniforme sur $\mathbb{G}(n_k, d)$: il existe un quasi-vecteur propre $f_k = f_k(G)$ de déviation δ_k de norme $\|f_k\|_2 = \sqrt{n}$ et orthogonal à $\mathbf{1}$ tel que $\mathrm{d_L}(\mathrm{distr}_G(f_k), S) > \varepsilon_0$.

Pour tout entier n, soit $F_n = F(n, d)$ où $F(n, d)$ est défini dans le lemme 1.6. D'après le lemme 1.6, quitte à modifier la suite (δ_k) et ε_0, on peut également supposer sans perte de généralité que l'événement $E_k(G)$ suivant a une probabilité au moins ε_0 sous la loi uniforme sur $\mathbb{G}(n_k, d)$: $\mathrm{d_H}(T(G), F_{n_k}) \leq \delta_k$ et il existe un quasi-vecteur propre $f_k = f_k(G)$ de déviation δ_k de norme $\|f_k\|_2 = \sqrt{n}$ et orthogonal à $\mathbf{1}$ tel que $\mathrm{d_H}(\mathrm{distr}_G(f_k), S) > \varepsilon_0$.

Soient $(G_k)_k$ une suite de graphes telle que $E_k(G_k)$ ait lieu pour tout k et f_k le quasi-vecteur propre associé. Étant donné que $\mathrm{distr}_{G_k}(f_k) \in \mathcal{M}$ et $(\mathcal{M}, \mathrm{d_L})$ est compact, quitte à prendre une sous-suite convergente, on peut supposer que $\mathrm{distr}_{G_k}(f_k)$ converge faiblement vers $\mu \in \mathcal{M}$. Soit Y de loi μ. Par construction, Y est une onde invariante qui vérifie $\mathrm{d_H}(\mu, S) \geq \varepsilon_0$ et $\mathbb{E}Y_\varnothing = 0$ (car $\mathbb{E}f_k(o) = 0$ et $\mathbb{E}f_k^2(o) = 1$ avec o uniforme sur $\{1, \ldots, n\}$). En particulier, étant donné que S contient la loi $\delta_0^{\otimes V_d}$, on a $\sigma^2 = \mathbb{E}Y_\varnothing^2 > 0$. Ainsi $Z = \sigma^{-1}Y$ est une onde invariante standard.

Le lemme 1.6 implique qu'une loi ν sur $\mathbb{R}^{V_d}$ est typique si, pour un certain $s > 0$, pour tous $\varepsilon > 0$ et $m > 0$, il existe un entier $n \geq m$ tel que $\mathrm{d_L}(s{\cdot}\nu, F_n) \leq \varepsilon$ où $s{\cdot}\nu$ est la mesure image de ν par l'application $\mathbb{R}^{V_d} \to \mathbb{R}^{V_d} : f \mapsto sf$. Étant donné que $\mathrm{d_H}(T(G_k), F_{n_k})$ et $\mathrm{d_L}(\mathrm{distr}_{G_k}(f_k), \mu)$ tendent vers 0, l'inégalité triangulaire implique donc que Z est une onde typique standard. Le théorème 1.5 et le corollaire 1.3 affirment que Z et $Y = \sigma Z$ sont des ondes gaussiennes. Cela contredit $\mathrm{d_H}(\mu, S) \geq \varepsilon_0$. $\square$

2. THÉORIE SPECTRALE DES ONDES INVARIANTES

Nous étudions ici la théorie spectrale d'une onde invariante standard. Nous construisons aussi l'onde gaussienne et prouvons le théorème 1.4. Pour un opérateur B défini dans $\ell^2(V_d)$, on utilisera la notation matricielle : pour tous $x, y \in V_d$, on pose $B_{xy} = \langle e_x, Be_y \rangle$ où e_x est le vecteur de la base canonique supporté sur x.

2.1. Résolvante et onde gaussienne

Cette partie s'inspire de [18, 20].

Résolvante. — Soit A l'opérateur d'adjacence sur T_d. Son spectre est $\sigma(T_d) = [-2\sqrt{d-1}, 2\sqrt{d-1}]$. La *résolvante* de A est définie pour tout $z \in \mathbb{C} \backslash \sigma(T_d)$ par

$$R(z) = (A - z)^{-1}.$$

On fixe $\lambda \in \mathbb{R}$. Pour $t > 0$, avec $z = \lambda + it$, on a

$$\Im R(z) = t((A - \lambda)^2 + t^2)^{-1}.$$

Il s'agit d'un opérateur positif de norme au plus $1/t$. En outre, tout comme $(A_{xy})_{xy \in V_d}$, le tableau de nombres $(R(z)_{xy})_{x,y \in V_d}$ est invariant par les automorphismes de T_d. Ainsi, pour tout $x \in V_d$, $R(z)_{xx} = R(z)_{\varnothing\varnothing}$ et $R(z)_{\varnothing x} = R(z)_{x\varnothing}$. En utilisant la positivité de $\Im R(z)$, on en déduit que $|\Im R(z)_{\varnothing x}| \leq \Im R(z)_{\varnothing\varnothing}$. En particulier, avec $z = \lambda + it$,

$$\Sigma(t)_{xy} = \frac{\Im R(z)_{xy}}{\Im R(z)_{\varnothing\varnothing}} \tag{4}$$

est de module au plus 1. Pour tout réel λ, par extraction diagonale, il existe donc une suite $(t_k)_k$ tendant vers 0 telle que pour tous $x, y \in V_d$, $\Sigma(t_k)_{xy}$ converge. On pose alors

$$\Sigma_{xy} = \lim_{k \to \infty} \Sigma(t_k)_{xy}.$$

Notons que le tableau de nombres $(\Sigma_{xy})_{x,y \in V_d}$ est invariant par les automorphismes de T_d.

On peut classiquement caractériser le spectre de T_d avec la résolvante appliquée à $e_\varnothing$: $\lambda \in \sigma(T_d)$ est équivalent à l'une des conditions suivantes,

$$\lim_{t \downarrow 0} \|\Im R(\lambda + it) e_\varnothing\|_2 = \infty \quad \text{ou} \quad \liminf_{t \downarrow 0} \Im R(\lambda + it)_{\varnothing\varnothing} > 0. \tag{5}$$

En effet, vérifions la première condition par exemple. Si $\|\Im R(\lambda + it) e_\varnothing\|_2$ est uniformément borné pour $t > 0$ alors, par transitivité, $\|\Im R(\lambda + it) e_x\|_2$ est uniformément borné en $t > 0$ et $x \in V_d$. La positivité de $\Im R(\lambda + it)$ implique donc que la norme d'opérateur de $\Im R(\lambda + it)$ est uniformément bornée en $t > 0$. Cela implique $\lambda \notin \sigma(T_d)$.

Construction de l'onde gaussienne. — Par construction, pour tout $S \subset V_d$, la matrice $\Sigma_S = (\Sigma_{xy})_{x,y\in S}$ est une matrice positive. Par le théorème d'extension de Kolmogorov, il existe un processus gaussien $(Y_x)_{x\in V_d}$ tel que pour tous $x, y \in V_d$, $\mathbb{E}Y_x = 0$ et

$$\mathbb{E}Y_xY_y = \Sigma.$$

Le processus Y est invariant et standard. Nous vérifions maintenant que si $\lambda \in \sigma(T_d)$ alors Y est une onde de valeur propre λ. Pour $t > 0$, soit $Y(t)$ le processus gaussien centré de covariance $\Sigma(t)$: $Y(t)$ converge faiblement vers Y lorsque $t \to 0$. Aussi, avec $z = \lambda + it$, on trouve

$$(\Im R(z))^2_{\varnothing\varnothing}\mathbb{E}((A-\lambda)Y(t))^2_{\varnothing} = ((A-\lambda)^2 t((A-\lambda)^2+t^2)^{-1})_{\varnothing\varnothing} = t-t^3((A-\lambda)^2+t^2)^{-1})_{\varnothing\varnothing} \leq t.$$

Maintenant si $\lambda \in \sigma(T_d)$ alors d'après (5), on a $\liminf_{t\downarrow 0}(\Im R(z))_{\varnothing\varnothing} > 0$. On en déduit que Y est bien une onde.

Lemme 2.1. — *Pour tout $\lambda \in \sigma(T_d)$, il existe une onde gaussienne de valeur propre λ.*

Approximation de l'onde gaussienne. — Nous pouvons maintenant démontrer le théorème 1.4.

Démonstration du théorème 1.4. — Soit A l'opérateur d'adjacence de T_d. Pour $t > 0$, soit $Y(t)$ la variable gaussienne sur T_d centrée et de covariance $\Sigma(t)$ donnée par (4). Par construction, l'onde gaussienne est une limite faible $Y(t_k)$ pour une suite $t_k \to 0$. Il suffit donc de montrer que pour tout $t > 0$, le processus $Y(t)$ est typique. D'un autre côté, pour tout $\varepsilon > 0$, le théorème de Weierstrass implique qu'il existe un polynôme Q tel que pour $x \in \sigma(T_d)$,

$$|Q(x)^2 - t/((x-\lambda)^2 + t^2)| \leq \varepsilon.$$

On pose pour tout ε suffisamment petit $\Sigma'_{xy} = Q(A)^2_{xy}/Q(A)^2_{\varnothing\varnothing}$ (l'inégalité $Q(A)^2_{\varnothing\varnothing} > 0$ découle de $t/((x-\lambda)^2 + t^2 > 0$ pour tout $x \in \sigma(T_d)$). Par construction Σ' est positif et la norme d'opérateur de $\Sigma' - \Sigma(t)$ est au plus ε. En conclusion, il suffit donc de montrer l'énoncé suivant : *tout processus μ gaussien centré invariant de covariance $Q(A)^2$ avec Q polynôme est typique.*

Un tel processus μ peut se réaliser comme la loi de $Q(A)Z$ où $Z = (Z_x)_{x\in V_d}$ avec (Z_x) des variables gaussiennes standard indépendantes. Soit $G \in \mathbb{G}(n, d)$ un graphe d-régulier sur $V = \{1, \ldots, n\}$ de matrice d'adjacence A' et $Z' = (Z'_x)_{x\in V}$ des variables gaussiennes standard indépendantes. On considère la variable gaussienne $Y' = Q(A')Z'$. Pour tous $x \in V$ et $k \geq 0$, soit $(G, x)_k$ le sous-graphe de G engendré par les sommets à distance au plus k de x dans G. On note $E_k(x)$ la propriété que $(G, x)_k$ est un arbre. Soit $S \subset V_d$ inclus dans $B_k(\varnothing)$. Par construction, si $E_k(x)$ est vraie et $\pi\colon V_d \to V$ est un revêtement de G tel que $\pi(\varnothing) = x$ alors la variable $(Z'_{\pi(y)})_{y\in S}$ a même loi que $(Z_y)_{y\in S}$. Soit ν la loi de Z. Par la loi faible des grands nombres, il est alors facile de vérifier que pour tout $\delta > 0$, il existe $\varepsilon > 0$ tel que l'événement (sous la loi de Z')

$$\mathrm{d_L}(\mathrm{distr}_G(Z'), \nu) \leq \delta$$

a probabilité au moins $1-\delta$ dès que $|\{x \in V : E_{1/\varepsilon}(x) \text{ est fausse}\}| \leq \varepsilon n$. L'application $\mathbb{R}^{V_d} \to \mathbb{R}^{V_d}$, $f \mapsto Q(A)f$ étant continue, on en déduit que pour tout $\delta' > 0$, il existe $\delta, \varepsilon > 0$ tels que $d_L(\mathrm{distr}_G(Y'), \mu) \leq \delta'$ dès que $d_L(\mathrm{distr}_G(Z'), \nu) \leq \delta$ et $|\{x \in V : E_{1/\varepsilon}(x) \text{ est fausse}\}| \leq \varepsilon n$.

Enfin, pour conclure la preuve que μ est typique, il reste à énoncer cette propriété classique : pour tout $\varepsilon > 0$, pour tout n suffisamment grand, avec probabilité au moins $1-\varepsilon$, si G est uniformément distribué sur $\mathbb{G}(n,d)$, $|\{x \in V : E_{1/\varepsilon}(x) \text{ est fausse}\}| \leq \varepsilon n$ (avec des mots : un graphe d-régulier aléatoire est localement presque comme un arbre). Cette dernière propriété découle par exemple du fait que le nombre de moyen de cycles de taille k dans un graphe d-régulier uniforme est uniformément borné en n (à k et d fixés), voir [24, théorème 2.5]. $\square$

Remarque 2.2 (Construction de quasi-vecteurs propres). — On peut se servir de la résolvante pour construire des quasi-vecteurs propres pour tout $\lambda \in \sigma(T_d)$. Pour $t > 0$, on pose $\psi_t = \Im R(\lambda + it)e_\varnothing$. D'après (5), lorsque $t \to 0$, $\|\psi_t\|_2 \to \infty$. En outre $(A-\lambda)\psi_t = \Im(A-\lambda)(A-\lambda-it)^{-1}e_\varnothing = \Im(e_\varnothing + it(A-\lambda-it)^{-1}e_\varnothing) = t^2((A-\lambda)^2+t^2)^{-1}e_\varnothing$. On en déduit que $\|(A-\lambda)\psi_t\|_2 \leq 1$ pour tout $t > 0$. En conclusion, le vecteur ψ_t est un quasi-vecteur propre de déviation $\delta_t = 1/\|\psi_t\|_2 \to 0$. En tronquant ψ_t, quitte à modifier δ_t, on en déduit qu'il existe un quasi-vecteur propre φ_t de déviation δ_t et de support inclus dans $B_{1/\delta_t}(\varnothing)$ (la boule de rayon $1/\delta_t$).

Maintenant soit $G \in \mathbb{G}(n,d)$ un graphe d-régulier sur $V = \{1,\ldots,n\}$. Si $x \in V$ est tel que la propriété $E_{1/\delta_t}(x)$ définie dans le théorème 1.4 est vraie alors d'après ce qui précède il existe un quasi-vecteur propre φ'_t de déviation δ_t et de support inclus dans la boule de centre x et de rayon $1/\delta_t$ dans G. En particulier, puisque φ'_t a un support de taille indépendante de la taille du graphe, la distance $d_L(\mathrm{distr}_G(\sqrt{n}\varphi'_t/\|\varphi'_t\|_2), \delta_0^{\otimes V_d})$ est arbitrairement proche de 0 lorsque n tend vers l'infini. Or, comme expliqué dans le théorème 1.4, pour tout $\delta > 0$, avec probabilité vers 1 lorsque n tend vers l'infini, si G est uniformément distribué sur $\mathbb{G}(n,d)$, la propriété $E_{1/\delta}(1)$ est vraie. Ainsi, avec grande probabilité, il existe des quasi-vecteurs propres localisés de petite déviation.

D'un autre côté, pour tout $\lambda \in \sigma(T_d)$, d'après le théorème 1.4 et l'argument dans la démonstration du corollaire 1.3, on peut aussi construire des quasi-vecteurs propres proches de l'onde gaussienne de valeur propre λ sur un graphe aléatoire uniforme. En conclusion, dans le contexte du théorème 0.4, en utilisant ces deux exemples opposés, on peut montrer que pour tous $\varepsilon, \delta > 0$, $\sigma \in [0,1]$ et $\lambda \in \sigma(T_d)$, pour tout n suffisamment grand, avec probabilité au moins $1-\varepsilon$, la matrice d'adjacence d'un graphe uniformément distribué sur $\mathbb{G}(n,d)$ admet un quasi-vecteur propre f de déviation δ de valeur propre λ orthogonal à $\mathbf{1}$ tel que $d_L(\mathrm{distr}_G(\sqrt{n}f/\|f\|_2), \mathrm{Gauss}_{\lambda,\sigma}) \leq \varepsilon$.

2.2. Polynômes sans épines

Pour mener des calculs plus fins sur la covariance des ondes invariantes, il est très utile d'introduire une famille de polynômes orthogonaux.

Définition. — Pour $n \geq 0$ entier, le polynôme de Tchebychev de seconde espèce de degré n est défini par l'identité

$$U_n(\cos\theta) = \frac{\sin(n+1)\theta}{\sin\theta}.$$

On pose $U_{-1} = 0$, $U_0 = 1$. On définit alors la famille de polynômes (p_n) avec $n \geq 1$ entier,

$$p_n(\lambda) = \sqrt{\frac{d-1}{d}}U_n\left(\frac{\lambda}{2\sqrt{d-1}}\right) - \frac{1}{\sqrt{d(d-1)}}U_{n-2}\left(\frac{\lambda}{2\sqrt{d-1}}\right),$$

et ses polynômes renormalisés

$$q_n(\lambda) = \frac{p_n(\lambda)}{\sqrt{d(d-1)^{n-1}}} \quad \text{et} \quad Q_n(\lambda) = \sqrt{d(d-1)^{n-1}}p_n(\lambda).$$

Enfin, on pose $p_0 = q_0 = Q_0 = 1$. En utilisant $U_1(\lambda) = 2\lambda$, $U_2(\lambda) = 4\lambda^2 - 1$, on trouve $q_1(\lambda) = \lambda/d$ et $q_2(\lambda) = (\lambda^2 - d)/(d(d-1))$. La relation de récurrence satisfaite par les polynômes de Tchebychev, $U_{n+1}(\lambda) = 2\lambda U_n(\lambda) - U_{n-1}(\lambda)$ pour tout $n \geq 0$, implique que pour tout $n \geq 1$,

$$(d-1)q_{n+1}(\lambda) = \lambda q_n(\lambda) - q_{n-1}(\lambda) \quad \text{et} \quad Q_{n+1}(\lambda) = \lambda Q_n(\lambda) - (d-1)Q_{n-1}(\lambda). \tag{6}$$

Les polynômes Q_n sont souvent appelés les *polynôme sans épines*, voir par exemple [21, 23, 1]. En effet, si A est l'opérateur d'adjacence d'un graphe d-régulier $G = (V, E)$, alors pour tous $x, y \in V$, et $n \geq 1$ entier, $Q_n(A)_{xy}$ est égal au nombre de chemins sans épines de x à y de longueur n (chemins $(x_0, \ldots, x_n)$ tels que $x_{t+1} \neq x_{t-1}$ pour tout $t = 1, \cdots, n-1$). En utilisant (6), cela se vérifie immédiatement par récurrence sur $n \geq 1$.

Orthogonalité. — La *mesure de probabilité de Kesten-McKay* μ_{T_d} est la mesure de support $\sigma(T_d) = [-2\sqrt{d-1}, 2\sqrt{d-1}]$ et densité par rapport à la mesure de Lebesgue donnée par

$$\frac{d\sqrt{4(d-1) - \lambda^2}}{2\pi(d^2 - \lambda^2)}.$$

Cette mesure est la *mesure spectrale* de l'opérateur d'adjacence A sur T_d définie par la propriété : pour tout entier $n \geq 0$

$$\int \lambda^n d\mu_{T_d}(\lambda) = (A^n)_{\varnothing\varnothing}.$$

Notons aussi que les termes diagonaux de la résolvante $R(z) = (A - z)^{-1}$ sont la transformation de Cauchy-Stieltjes de μ_{T_d}, pour tout $z \in \mathbb{C} \backslash \sigma(T_d)$

$$R(z)_{\varnothing\varnothing} = \int \frac{d\mu_{T_d}(\lambda)}{\lambda - z}.$$

En particulier, pour $\lambda \in \sigma(T_d)$,

$$\lim_{t \downarrow 0} \Im R(\lambda + it)_{\varnothing\varnothing} = \frac{1}{\pi}\frac{d\mu_{T_d}(\lambda)}{d\lambda} = \frac{d\sqrt{4(d-1) - \lambda^2}}{2(d^2 - \lambda^2)}. \tag{7}$$

Dans T_d il n'existe qu'un seul chemin sans épines entre deux sommets. On en déduit que pour tous $n, m \geq 1$ entiers

$$\int Q_n Q_m d\mu_{T_d} = \langle e_\varnothing, Q_n(A)Q_m(A)e_\varnothing\rangle = \langle Q_n(A)e_\varnothing, Q_m(A)e_\varnothing\rangle = d(d-1)^{n-1}\mathbf{1}_{n=m}.$$

En particulier la suite $(p_n)_{n\geq 0}$ est une *famille orthonormale* pour la mesure μ_{T_d}. Cela permet de donner une définition plus analytique des polynômes sans épines à partir de la résolvante $R(z) = (A-z)^{-1}$. On développe en série entière $\lambda \mapsto 1/(\lambda - z)$ dans la famille orthonormale (p_n). Ce développement est convergent pour $|z|$ suffisamment grand pour tout $\lambda \in \sigma(T_d)$. On trouve donc, pour z dans un voisinage de l'infini,

$$R(z) = \sum_{n=0}^{\infty} p_n(A) \int \frac{p_n(\lambda)}{\lambda - z} d\mu_{T_d}(\lambda).$$

En passant par les chemins sans épines, on obtient, si $|x|$ est la distance de $x \neq \varnothing$ à $\varnothing$,

$$R(z)_{\varnothing x} = \frac{1}{\sqrt{d(d-1)^{|x|-1}}} \int \frac{p_{|x|}(\lambda)}{\lambda - z} d\mu_{T_d}(\lambda) = \int \frac{q_{|x|}(\lambda)}{\lambda - z} d\mu_{T_d}(\lambda).$$

Cette formule est également vraie pour $x = \varnothing$. Par analyticité, l'identité s'étend à tout $z \notin \sigma(T_d)$. On prend la partie imaginaire et on fait tendre z vers $\lambda \in \mathbb{R}$. D'après (7), on en déduit la formule pour tous $|\lambda| < 2\sqrt{d-1}$ et $x \in V_d$,

$$q_{|x|}(\lambda) = \frac{\lim_{t\downarrow 0} \Im R(\lambda + it)_{\varnothing x}}{\lim_{t\downarrow 0} \Im R(\lambda + it)_{\varnothing\varnothing}}.$$

Ainsi, $\Sigma_{\varnothing x} = q_{|x|}(\lambda)$, où Σ est la covariance de l'onde gaussienne de valeur propre λ. Par continuité, cette dernière identité s'étend pour tout $\lambda \in \sigma(T_d)$.

Calcul de la covariance. — On déduit de (6) que, pour tout $\lambda \in \mathbb{R}$, le vecteur $\varphi \in \mathbb{R}^{V_d}$ défini par

$$\varphi(x) = q_{|x|}(\lambda),$$

où $|x|$ est la distance de x à $\varnothing$, satisfait l'équation de vecteur propre (1). En outre, si X est une onde invariante sur T_d alors $\mathbb{E}X_x X_y$ ne dépend que de la distance entre x et y. Il est immédiat de vérifier que le vecteur $\mathbb{E}XX^\intercal e_x$ satisfait l'équation de vecteur propre de (1) : si $g = \mathbb{E}XX^\intercal f$ pour $f \in \ell^2(V)$ alors $(\lambda I - A)g = \mathbb{E}(\lambda I - A)XX^\intercal f = 0$. Par récurrence, on obtient que $\varphi = \mathbb{E}XX^\intercal e_\varnothing$. Le lemme suivant en découle.

Lemme 2.3. — *Soit X une onde invariante standard de valeur propre λ. Pour tous $x, y \in V_d$, on a $\mathbb{E}X_x X_y = q_{|x-y|}(\lambda)$ où $|x-y|$ est la distance de x, y dans T_d.*

En combinant avec le lemme 2.1, nous avons démontré le résultat suivant.

Corollaire 2.4. — *Pour tout $\lambda \in \sigma(T_d)$, il existe une unique onde gaussienne de valeur propre λ.*

2.3. Rang de la matrice de covariance

Soit $S \subset V_d$ un sous-ensemble fini connexe de sommets de l'arbre T_d. Pour tout entier k, nous noterons $B_k(S)$ le voisinage de S dans T_d de rayon k. On définit l'*intérieur de* S comme $\mathrm{int}(S) = \{x \in S : B_1(x) \subset S\}$ et son *bord* comme $\partial S = S \backslash \mathrm{int}(S)$.

Pour $\lambda \in \mathbb{R}$, on note $W_\lambda(S)$ le sous-espace vectoriel de $\mathbb{R}^S$ des vecteurs $Y \in \mathbb{R}^S$ qui vérifient l'équation de vecteur propre (1) pour tout $x \in \mathrm{int}(S)$. Nous allons commencer par un résultat classique sur la dimension de $W_\lambda(S)$.

Lemme 2.5. — *Soit $\lambda \in \mathbb{R}$. Pour tout $S \subset V_d$ fini, on a*

$$\dim(W_\lambda(S)) = |\partial S|.$$

Démonstration. — La fonction $S \mapsto \dim(W_\lambda(S))$ étant invariante par automorphisme, on peut supposer que $\varnothing \in S$. On peut aussi supposer S connexe. En effet, $W_\lambda(S \cup S')$ est la somme directe de $W_\lambda(S)$ et $W_\lambda(S')$ si aucun sommet de S ne partage d'arête avec un sommet de S'. Pour démontrer le lemme, nous raisonnons par récurrence sur la taille de S. Si $S = \{\varnothing\}$, $\partial S = S$ et l'énoncé est vrai. Pour $n \geq 2$ entier, nous supposons l'énoncé vrai pour tous les ensembles connexes de taille $n-1$. Soit S tel que $|S| = n$. Il existe U connexe de taille $n-1$ tel que $S = U \cup \{x\}$ avec $\varnothing \in U$ et $y \in U$, y parent de x. En particulier, $y \in \partial U$, $\dim(W_\lambda(U)) = |\partial U|$ et, T_d étant un arbre, aucun enfant de x n'est dans U. Si tous les enfants de y ne sont pas dans S alors $\partial S = \partial U \cup \{x\}$ et $\dim(W_\lambda(S)) = \dim(W_\lambda(U)) + 1$ ($W_\lambda(S)$ est l'espace vectoriel engendré par $W_\lambda(U)$ et e_x). Dans l'autre cas possible, U contient les $d-1$ autres voisins de y (parent et enfants). On trouve alors $\partial S = (\partial U \cup \{x\}) \backslash \{y\}$ et $\dim(W_\lambda(S)) = \dim(W_\lambda(U))$ (l'équation de vecteur propre (1) devant être satisfaite au sommet y, tout vecteur $f \in W_\lambda(U)$ s'étend de façon unique en x). Donc, dans les deux cas, on a bien $\dim(W_\lambda(S)) = |\partial S|$. □

Nous allons maintenant étudier le rang de la matrice de covariance d'une marginale d'une onde invariante. Nous démarrons avec la remarque suivante.

Lemme 2.6. — *Soit X une onde invariante. Pour tous $x \neq \varnothing \in V_d$, $i \in \{2, \dots, d-1\}$, on a $\mathbb{E} X_\varnothing (X_{(x,i)} - X_{(x,1)}) = 0$.*

De même pour tout $i \in \{2, \dots, d\}$, on a $\mathbb{E} X_\varnothing (X_i - X_1) = 0$.

Démonstration. — Il suffit d'observer qu'il existe un automorphisme de T_d qui laisse $\varnothing$ invariant mais qui permute (x,i) et $(x,1)$. □

Notre objectif est maintenant de démontrer le résultat suivant.

Lemme 2.7. — *Soient $\lambda \in \sigma(T_d)$ et $S \subset V_d$ fini tel que $S = \bigcup_{x \in \mathrm{int}(S)} B_1(x)$ ou $|S| \leq 2$. Soit X une onde invariante standard de valeur propre λ et $\Sigma_S = (\mathbb{E}[X_x X_y])_{x,y \in S}$ la matrice de covariance de X restreinte à S. L'image de Σ_S est $W_\lambda(S)$.*

Démonstration. — Soit X une onde standard de loi μ. Rappelons que l'image de Σ_S est contenue dans $W_\lambda(S)$: si $g = \mathbb{E}XX^\intercal f$ alors $(\lambda I - A)g = \mathbb{E}(\lambda I - A)XX^\intercal f = 0$. Il suffit donc de montrer que le rang Σ_S est (au moins) égal à $\dim(W_\lambda(S)) = |\partial S|$. Pour ce faire, si $m = |\partial S|$, nous allons trouver une transformation linéaire $T : \mathbb{R}^S \to \mathbb{R}^m$ telle que la matrice de covariance de $T((X_x)_{x\in S})$ est diagonale par blocs avec des blocs inversibles.

On considère les variables aléatoires suivantes pour $x \neq \varnothing$,

$$X_\varnothing\,, \quad A_\varnothing = (X_i - X_1)_{i=2,\ldots,d} \quad \text{et} \quad A_x = (X_{(x,i)} - X_{(x,1)})_{i=2,\ldots,d-1}. \tag{8}$$

Le lemme 2.6 implique que toutes ces variables sont décorrélées : pour tous $x \neq y$, i, j $\mathbb{E}(A_x)_i(A_y)_j = 0$. On vérifie maintenant que le rang de la matrice de covariance $\mathbb{E}A_xA_x^\intercal$ est maximal. Soit $x \neq \varnothing$. Le lemme 2.6 implique que pour tous $i \neq j \in \{2, \cdots, d-1\}$, $\mathbb{E}X_{(x,i)}(X_{(x,j)} - X_{(x,1)}) = 0$. Par linéarité, on trouve $\mathbb{E}(A_x)_i(A_x)_j = a$ et $\mathbb{E}((A_x)_i)^2 = 2a$ avec $a = 1 - \mathbb{E}X_1X_2$. Cela implique que pour tout $x \neq \varnothing$, $\mathbb{E}A_xA_x^\intercal = a(I + J)$ où I est la matrice identité et $J = \mathbf{1}\mathbf{1}^\perp$ (matrices de dimension $(d-1)\times(d-1)$). De même, pour $x = \varnothing$, le lemme 2.6 implique que $\mathbb{E}A_\varnothing A_\varnothing^\intercal = a(I + J)$ (de dimension $d \times d$).

Nous déterminons maintenant la valeur de a. L'équation de vecteur propre (1) en $\varnothing$ implique que

$$\lambda^2 \mathbb{E}X_\varnothing^2 = \lambda^2 = \mathbb{E}\left(\sum_{i=1}^d X_i\right)^2 = d + d(d-1)\mathbb{E}X_1X_2.$$

On en déduit que $\mathbb{E}X_1X_2 = (\lambda^2 - d)/(d(d-1))$. On trouve $a = (d^2 - \lambda^2)/(d(d-1)) > 0$ si et seulement si $|\lambda| < d$. On a donc vérifié que pour tout $x \in V_d$, la matrice $\mathbb{E}A_xA_x^\intercal$ est définie positive.

On peut maintenant compléter la preuve. Si $|S| \in \{1, 2\}$ l'énoncé est trivial. Sinon, on peut supposer que $\varnothing \in S$. L'ensemble de variables $(X_\varnothing, (A_x)_{x\in\mathrm{int}(S)}) \in \mathbb{R}^m$ a une matrice de covariance de rang plein et est une combinaison linéaire de $(X_x)_{x\in S}$. En comptant (par exemple comme dans le lemme 2.5), on trouve $m = |\partial S|$. □

Remarque 2.8. — Avec plus de soin, on peut enlever la restriction $S = \bigcup_{x\in\mathrm{int}(S)} B_1(x)$ dans l'énoncé du lemme 2.7.

3. UNE STRATÉGIE ENTROPIQUE

Dans cette section, nous présentons les grandes lignes de la démonstration du théorème 1.5.

Théorème principal, version onde typique lisse. — Avant de nous lancer dans le cœur de la preuve du théorème principal, nous procédons à une ultime réduction.

Nous dirons qu'une loi μ sur $\mathbb{R}^{V_d}$ est *lisse* si pour tout ensemble fini S tel que $S = \bigcup_{x \in \mathrm{int}(S)} B_1(x)$ ou $|S| \leq 2$, la loi marginale de μ sur S admet une densité C^∞ et strictement positive sur le sous-espace vectoriel de $\mathbb{R}^S$ engendré par son support (muni de la mesure de Lebesgue). La restriction sur les ensembles S est purement technique et provient de l'énoncé du lemme 2.7.

THÉORÈME 3.1. — *Soit $\lambda \in \sigma(T_d) = [-2\sqrt{d-1}, 2\sqrt{d-1}]$. Si Y est une onde typique standard lisse de valeur propre λ alors Y est l'onde gaussienne de valeur propre λ.*

Par un argument de convolution, il est facile de vérifier que le théorème 3.1 implique le théorème 1.5. En effet, soit Y une onde typique standard et Z une onde gaussienne indépendante de même valeur propre $\lambda \in \sigma(T_d)$. Soit $S \subset V_d$ fini tel que $S = \bigcup_{x \in \mathrm{int}(S)} B_1(x)$. D'après le lemme 2.7, les supports des lois de Z et de Y restreintes à S engendrent le même espace vectoriel, $W_\lambda(S)$. Ainsi, pour tout $\sigma \in (0,1]$, $Y(\sigma) = \sigma Z + \sqrt{1-\sigma^2} Y$ est une onde standard dont $\mu_{\sigma,S}$, la loi restreinte à $W_\lambda(S)$, est la convolution d'une loi gaussienne de covariance définie positive et d'une autre loi. En particulier, $\mu_{\sigma,S}$ est lisse. D'un autre côté, pour tout $\sigma \in [0,1]$, en répétant la preuve du théorème 1.4, $Y(\sigma)$ est une onde typique. Maintenant, étant donné que toute limite faible d'une loi gaussienne dans $\mathbb{R}^n$ est gaussienne, il est clair en considérant la limite $\sigma \to 0$ que le théorème 3.1 implique le théorème 1.5.

Entropie. — Soit μ une mesure de probabilité sur un espace mesuré (E, λ). Si μ est absolument continue par rapport à λ et de densité f, l'*entropie différentielle* de μ est (si l'intégrale est bien définie)

$$\mathfrak{H}_E(\mu) = -\int_E f(x) \log f(x)\, d\lambda(x).$$

Si Y a loi μ, on notera cette quantité indifféremment $\mathfrak{H}_E(\mu)$ ou $\mathfrak{H}_E(Y)$. On omettra souvent l'indice E dans la notation $\mathfrak{H}_E(\cdot)$. Si E est un espace dénombrable muni de la mesure de comptage alors $\mathfrak{H}(\mu)$ coïncide avec l'*entropie de Shannon*

$$\mathfrak{H}_E(\mu) = -\sum_{x \in E} \mu(x) \log \mu(x).$$

Soient μ une mesure de probabilité dans $\mathbb{R}^n$ et E l'espace vectoriel engendré par le support de μ. Si μ est absolument continue par rapport à la mesure de Lebesgue sur E, on pose

$$\mathfrak{H}_{\mathrm{supp}}(\mu) = \mathfrak{H}_E(\mu).$$

Il est facile de vérifier que si la densité de μ sur E est uniformément bornée et $\mathbb{E}\|X\|_2^2 < \infty$ avec X de loi μ alors $\mathfrak{H}_{\mathrm{supp}}(\mu)$ est bien défini.

Stratégie de la démonstration. — On considère les ensembles de V_d,

$$S^\star = B_1(\varnothing) \quad \text{et} \quad S^- = \{\varnothing, 1\}.$$

La preuve se base sur l'étude pour une onde invariante lisse X de son *entropie invariante* définie comme :

$$\mathfrak{D}(X) = \mathfrak{H}_{\mathrm{supp}}(X_{S^\star}) - (d/2)\mathfrak{H}_{\mathrm{supp}}(X_{S^-}), \tag{9}$$

avec $X_S = (X_x)_{x\in S}$ pour tout $S \subset V_d$.

Dans [4], les auteurs démontrent la propriété suivante de maximalité de l'entropie d'une gaussienne.

Théorème 3.2. — *Soit $\lambda \in \sigma(T_d)$. Si X est une onde invariante standard lisse et Y l'onde gaussienne de même valeur propre λ. Alors*

$$\mathfrak{D}(X) \leq \mathfrak{D}(Y),$$

avec égalité si et seulement si $X_{S^\star}$ a même loi que $Y_{S^\star}$.

D'un autre côté, Backhausz et Szegedy ont établi le résultat suivant. Nous dirons qu'un processus X sur $\mathbb{R}^{V_d}$ est 2-*markovien* si pour toute arête $e = \{a, b\}$ de T_d, les variables $(X_x)_{x\in V_a}$ et $(X_x)_{x\in V_b}$ sont indépendantes conditionnellement à (X_a, X_b) avec $V_a \cup V_b$ la partition de V_d en les deux composantes connexes du sous-graphe de T_d où seule l'arête e est retirée. Cette propriété s'interprète comme une propriété de Markov spatiale. En utilisant le lemme 2.6, il ne sera pas difficile de vérifier que l'onde gaussienne est 2-markovienne.

Théorème 3.3. — *Soit $\lambda \in \sigma(T_d)$. Si X est une onde typique standard lisse et Y l'onde gaussienne de même valeur propre λ alors*

$$\mathfrak{D}(X) \geq \mathfrak{D}(Y),$$

avec égalité si et seulement X est 2-markovien.

Les théorèmes 3.2 et 3.3 impliquent le théorème 3.1. Il suffit en effet d'observer qu'une onde invariante 2-markovienne X est caractérisée par la loi de $X_{S^\star}$.

Dans la section 5, nous démontrerons le théorème 3.3 grâce à des inégalités entropiques générales satisfaites par les processus typiques que nous présentons dans la section 4. La démonstration du théorème 3.2 dans la section 6 se basera sur l'étude de l'évolution de l'entropie invariante le long du flot de la chaleur.

4. INÉGALITÉ ENTROPIQUE POUR LES ONDES TYPIQUES

Si $f \in F^{V_d}$ et $S \subset V_d$ alors nous noterons $f_S = (f_x)_{x\in S}$. Nous noterons pour tout entier $k \geq 0$,

$$S_k^\star = B_k(S^*) \quad \text{et} \quad S_k^- = B_k(S^-).$$

4.1. Inégalité entropique étoile-arête discrète

La preuve du théorème 3.3 repose sur l'inégalité entropique suivante.

Théorème 4.1. — *Soit F un ensemble fini. Si Y est un processus typique sur F^{V_d} alors pour tout $k \geq 0$, on a*

$$\mathfrak{H}(Y_{S_k^\star}) - (d/2)\mathfrak{H}(Y_{S_k^-}) \geq 0.$$

La démonstration de ce théorème repose sur des idées présentes dans [13, 5].

Appariement et coloriage. — Soit V un ensemble fini avec $|V|$ pair, un *appariement* est une involution sur V sans point fixes. Un appariement peut s'identifier à un graphe (V, E_σ) sur V formé de $|V|/2$ arêtes disjointes. Le nombre d'appariements d'un ensemble à n éléments est (avec n pair)

$$n!! = (n-1)(n-3)\cdots 1 = \frac{n!}{2^{n/2}(n/2)!}$$

(ce nombre est aussi égal au moment d'ordre n d'une loi gaussienne standard sur $\mathbb{R}$).

Soit F un ensemble fini dont les éléments de F seront appelés des *couleurs*. Un vecteur $f \in F^V$ est un *coloriage* de V. On rappelle que $\mathrm{distr}(f) = (1/n)\sum_{x\in V}\delta_{f(x)}$ est la loi d'une entrée typique de f. Pour toute mesure μ sur F on notera

$$H(\mu, n) = |\{f \in F^n : \mathrm{distr}(f) = \mu\}|.$$

On notera de même la mesure de probabilité sur F^2,

$$\mathrm{distr}_\sigma(f) = \frac{1}{n}\sum_{x\in V}\delta_{(f(x), f(\sigma(x)))}. \tag{10}$$

Notons que cette mesure est *symétrique* au sens où elle est invariante par l'application $(a, b) \mapsto (b, a)$. Aussi, les deux marginales de $\mathrm{distr}_\sigma(f)$ sont égales à $\mathrm{distr}(f)$. Le résultat suivant est extrait de [5, lemme 4.1].

Lemme 4.2. — *Soient V, F des ensembles finis avec $|V| = n$ pair et $f \in F^V$. Soit μ une mesure de probabilité sur F^2 symétrique de loi marginale $\mu_1 = \mathrm{distr}(f)$. Alors le nombre d'appariements de V tels que $\mathrm{distr}_\sigma(f) = \mu$ est égal à*

$$n!!H(\mu, n/2)/H(\mu_1, n).$$

Démonstration. — Soient A_f l'ensemble des appariements σ de V tels que $\mathrm{distr}_\sigma(f) = \mu$ et M l'ensemble des vecteurs $h \in (F^2)^n$ tels que $h(x) = (g(x), g(\sigma(x))$ pour $g \in F^n$ et σ appariement avec $\mathrm{distr}_\sigma(g) = \mu$. On calcule $|M|$ de deux façons. Tout d'abord, on remarque que $\mathrm{distr}(f) = \mathrm{distr}(g)$ implique que $f = g \circ \theta$ pour une certaine permutation θ sur V. En particulier $|A_f| = |A_g|$ et on trouve $M = H(\mu_1, n)|A_f|$. D'un autre côté, un élément de M peut aussi se construire en choisissant d'abord un appariement σ et ensuite en positionnant les paires de couleurs sur les arêtes. On obtient $M = n!!H(\mu, n/2)$. $\square$

Nous aurons aussi besoin d'un théorème standard de Shannon, voir [15, théorème 3.1.2].

Lemme 4.3. — *Soient F un ensemble fini et μ une mesure de probabilité sur F. Pour $\varepsilon > 0$, on pose $H_\varepsilon(\mu, n) = |\{f \in F^n : \mathrm{d_L}(\mathrm{distr}(f), \mu) \leq \varepsilon\}|$. On a*

$$\mathfrak{H}(\mu) = \lim_{\varepsilon \downarrow 0} \liminf_{n\to\infty} \frac{1}{n} \log H_\varepsilon(\mu, n) = \lim_{\varepsilon \downarrow 0} \limsup_{n\to\infty} \frac{1}{n} \log H_\varepsilon(\mu, n).$$

Inégalité entropique étoile-arête. — Nous allons maintenant démontrer le théorème 4.1 Nous démarrons par le cas $k = 0$.

Théorème 4.4. — *Soit F un ensemble fini. Si Y est un processus typique sur F^{V_d} alors*

$$\mathfrak{H}(Y_{S^\star}) - (d/2)\mathfrak{H}(Y_{S^-}) \geq 0.$$

Démonstration. — Un graphe coloré sur F est une paire (G, f) avec $G = (V, E) \in \mathbb{G}(n, d)$ et $f \in F^V$. L'ensemble des arêtes orientées de G est $\vec{E} = \{(u, v) : \{u, v\} \in E\}$, il est muni de l'appariement $\sigma_G(u, v) = (v, u)$. On étend f à $F^{\vec{E}}$ en posant $\vec{f}(u, v) = f(u)$ pour tout $(u, v) \in \vec{E}$. La loi d'une couleur typique est $\mathrm{distr}(f) = \mathrm{distr}(\vec{f})$, la loi d'une arête orientée typique est définie comme $\mathrm{distr}_{\sigma_G}(\vec{f})$.

Nous allons encoder la boule de rayon 1 autour d'un sommet par un nouveau graphe coloré. On définit le graphe coloré $(G, f^\star)$ sur l'ensemble de couleurs $(F^\star)^V$ en posant $F^\star = F^{d+1}$ et pour tout $x \in V$, $f^\star(x) = (f(x), f(x_i)_{i=1,\ldots,d})$ où les (x_i) sont les couleurs des voisins de x classés par ordre lexicographique. On note $f^\star(x) = (f^\star(x)_0, f^\star(x)_1, \ldots, f^\star(x)_d)$.

De même, on définit $Y^\star$ comme le processus sur $(F^\star)^{V_d}$ défini par $Y^\star_x = (Y_x, (Y_y)_{y:\{x,y\}\in E})$ où E est l'ensemble des arêtes de T_d. Enfin, on note $\mu^\star$ la loi de $Y^\star_\varnothing$ et μ^- la loi de $(Y_\varnothing, Y_1)$. On a $\mathfrak{H}(\mu^\star) = \mathfrak{H}(\mu_{S^\star})$ et $\mathfrak{H}(\mu^-) = \mathfrak{H}(\mu_{S^-})$.

Pour tout $\varepsilon > 0$, soit $C(n, \varepsilon)$ l'ensemble des graphes colorés (G, f) avec $G \in \mathbb{G}(n, d)$ et $f \in F^V$ tels que $\mathrm{d_L}(\mathrm{distr}_{\sigma_G}(\vec{f}), \mu^-) \leq \varepsilon$ et $\mathrm{d_L}(\mathrm{distr}(f^\star), \mu^\star) \leq \varepsilon$. La définition d'un processus typique implique

$$\liminf_{\varepsilon\to 0} \limsup_{n} \frac{|C(n, \varepsilon)|}{|\mathbb{G}(n, d)|} \geq 1. \tag{11}$$

Un résultat classique de Bender et Canfield [8] affirme que pour tout entier $d \geq 2$, si nd est pair et dans la limite $n \to \infty$,

$$|\mathbb{G}(n, d)| \sim \frac{n!!}{(d!)^n} e^{\frac{1-d^2}{4}}. \tag{12}$$

En utilisant une idée déjà présente dans Bender et Canfield, nous allons maintenant estimer $|C(n, \varepsilon)|$. Nous noterons $o(1)$ une fonction de (n, ε) qui tend vers 0 lorsque n tend vers l'infini puis ε vers 0. Soit $f^\star \in (F^\star)^n$ tel que $\mathrm{d_L}(\mathrm{distr}(f^\star), \mu^\star) \leq \varepsilon$. D'après le lemme 4.3, il y a $\exp(n(1 + o(1))\mathfrak{H}(\mu^\star))$ possibilités pour $f^\star$.

Le *modèle de configuration* introduit dans [8, 10] est une représentation d'un graphe à une numérotation des arêtes près par un appariement. On définit l'ensemble

$\vec{E} = V \times \{1, \ldots, d\}$, pour $x \in V$, on pose $\vec{E}(x) = \{(x,i) : i = 1, \ldots, d\}$. On conçoit un élément de $(x,i) \in \vec{E}(x)$ comme une demi-arête orientée dont le sommet de départ est x. À tout appariement σ sur $\vec{E}$, on peut associer un graphe $G(\sigma) = (V, E_\sigma)$ avec E_σ l'ensemble des $\{x,y\}$ tels que $x \neq y$ et $\vec{E}(x) \cap \vec{E}(y)$ non vide. On aura $G(\sigma) \in \mathbb{G}(n,d)$ si et seulement si pour tout $x \neq y$, $\vec{E}(x) \cap \vec{E}(x) = \emptyset$ et $|\vec{E}(x) \cap \vec{E}(y)| \in \{0,1\}$. De plus, en comptant les possibilités de numérotations des demi-arêtes, on a pour tout $G \in \mathbb{G}(n,d)$,

$$|\{\sigma : G(\sigma) = G\}| = (d!)^n. \tag{13}$$

Pour $f^\star \in (F^\star)^n$ fixé, on définit la fonction $f^- \colon \vec{E} \to F^2$ par $f^-(x,i) = (f^\star(x)_0, f^\star(x)_i)$. On obtient ainsi une couleur sur les demi-arêtes. On cherche à apparier une demi-arête de couleur (a,b) avec une demi-arête de couleur (b,a). D'après le lemme 4.2, le nombre d'appariements possibles est

$$n!! \frac{H(\mathrm{distr}(f^-)^2, (nd)/2)}{H(\mathrm{distr}(f^-), nd)},$$

où pour une mesure ν sur F^2 on a posé $\nu^2((a,b),(c,d)) = \nu(a,b)\mathbf{1}_{(c,d)=(b,a)}$. On note que $\mathfrak{H}(\nu^2) = \mathfrak{H}(\nu)$. Par une nouvelle application du lemme 4.3 et (13), on en déduit

$$|C(n,\varepsilon)| \leq \frac{n!!}{(d!)^n} \exp\left(n(1+o(1))\left(\mathfrak{H}(\mu_\star) - \frac{d}{2}\mathfrak{H}(\mu_-)\right)\right).$$

Il reste à combiner cette dernière inégalité à (11)-(12) pour conclure la démonstration. $\square$

Le théorème 4.1 découle du cas $k = 0$.

Démonstration du théorème 4.1. — Soit $f \in F^{V_d}$ un coloriage de T_d. Nous encodons la boule de rayon $k+1$ autour d'un sommet par une nouvelle coloration. Soit $\vec{E}$ l'ensemble des arêtes orientées de T_d. Nous allons d'abord définir la couleur d'une arête orientée. Dans le langage de la combinatoire, on définit la couleur de $(x,y) \in \vec{E}$ comme la paire d'arbres enracinés non étiquetés de profondeur k associée aux sous-arbres issus de (x,y) et (y,x). Plus précisément, pour tout $e = (x,y) \in \vec{E}$, on fixe un isomorphisme de T_d tel que $\varphi_e(\varnothing) = x$ et $\varphi_e(1) = y$. On définit $T(e)$ comme la fonction $f \circ \varphi_e$ restreinte à $B'_k(1)$ les sommets descendants de 1 et qui sont à distance au plus $k+1$ de $\varnothing$ (c'est le sous-arbre issu de 1 de profondeur k). Les automorphismes φ de T_d tels que $\varphi(\varnothing) = \varnothing$ et $\varphi(1) = 1$ définissent un sous-groupe et une relation d'équivalence sur $F^{B'_k(1)}$. L'ensemble des classes d'équivalence est un ensemble fini que l'on note $F_\sim$. Pour tout $e \in \vec{E}$, on définit $t(e)$ comme la classe d'équivalence de $T(e)$. Enfin, on pose $g_f \in (F^d_\sim)^{V_d}$ avec $g_f(x) = (t(x,x_i))_{i=1,\ldots,d}$ où les x_i sont les voisins de x classés par ordre lexicographique.

Soient Y un processus typique et $Z = g_Y \in (F^d_\sim)^{V_d}$ le processus de coloriage ci-dessus associé à Y. Par construction, Z est typique et d'après le théorème 4.1,

$$\mathfrak{H}(Z_{S^\star}) - (d/2)\mathfrak{H}(Z_{S^-}) \geq 0.$$

On remarque que pour tout $f \in F^{V_d}$, $(g_f(x))_{x\in S^\star}$ contient toutes les couleurs dans $S_k^\star = B_k(S^\star)$ et $(g_f(x))_{x\in S^-}$ toutes les couleurs dans $B_k(e)$. Cependant à cause de la relation d'équivalence, l'indice des sommets dans ces boules est inconnu. En utilisant l'invariance de Y par les automorphismes de T_d, nous avons $\mathfrak{H}(Y_{S_k^\star}) = \mathfrak{H}(Z_{S^\star}) + d\log m$ et $\mathfrak{H}(Y_{S_k^-}) = \mathfrak{H}(Z_-) + 2\log m$ où m est le nombre d'automorphismes de l'arbre $(d-1)$-aire de profondeur k (l'arbre de profondeur k issu de 1 dans T_d). En effet, on peut construire $Y_{S_k^\star}$ en échantillonnant les classes d'équivalence des sous-arbres suivant la loi de $Z_{S^\star}$ et ensuite en choisissant un étiquetage pour les sommets issus de $i = 1,\dots,d$ grâce à d automorphismes indépendants uniformes. De même pour $Y_{S_k^-}$. Ainsi $\mathfrak{H}(Z_{S^\star}) - (d/2)\mathfrak{H}(Z_{S^-}) = \mathfrak{H}(Y_{S_k^\star}) - (d/2)\mathfrak{H}(Y_{S_k^-})$. Cela conclut la preuve. □

Nous aurons en fait besoin d'un raffinement du théorème 4.4 pour un processus qui est obtenu à partir de tirages indépendants sur un processus typique.

Corollaire 4.5. — *Soient E un espace métrique complet séparable, F un ensemble fini, $\mathcal{P}(F)$ l'ensemble des mesures de probabilité sur F et $w\colon E \to \mathcal{P}(F)$ une application continue. Soit (X,Y) un processus sur $(E\times F)^{V_d}$ tel que pour tout $x \in V_d$ la variable Y_x sachant $(X,(Y_y)_{y\neq x})$ a distribution $w(X_x)$. Si X est typique alors Y est typique et pour tout k entier,*

$$\mathfrak{H}(Y_{S_k^\star}) - (d/2)\mathfrak{H}(Y_{S_k^-}) \geq \mathbb{E}[\mathfrak{H}(Y_\varnothing|X_\varnothing)],$$

où $\mathfrak{H}(Y_\varnothing|X_\varnothing)$ est l'entropie de la loi conditionnelle de $Y_\varnothing$ sachant $X_\varnothing$.

Démonstration. — Nous esquissons la démonstration pour le cas E fini, le cas général se déduit du même argument en utilisant l'uniforme continuité de w sur les compacts et la séparabilité de E. Soit μ la loi de Y. Les démonstrations des théorèmes 4.4 et 4.1 et (11) montrent en fait que

$$\mathfrak{H}(Y_{S_k^\star}) - \frac{d}{2}\mathfrak{H}(Y_{S_k^-}) \geq \lim_{\varepsilon\to 0}\limsup_{n\to\infty}\frac{1}{n}\log\frac{|C_k(n,\varepsilon)|}{|\mathbb{G}(n,d)|},$$

où pour tout $\varepsilon > 0$, $C_k(n,\varepsilon)$ est l'ensemble des graphes colorés (G,f) avec $G \in \mathbb{G}(n,d)$ et $f \in F^n$ tels que $\mathrm{d_L}(\mathrm{distr}_G(f)_{S_k^\star}, \mu_{S_k^\star}) < \varepsilon$ où, pour ν loi d'un processus sur V_d, $\nu_{S_k^\star}$ est sa loi marginale sur $S_k^\star$.

Soit ν la loi de X. En utilisant le lemme 4.3, pour tout $\varepsilon > 0$, pour tout entier ℓ suffisamment grand et pour tout $a \in E$, le nombre de vecteurs $f \in F^\ell$ tels que $\mathrm{d_L}(\mathrm{distr}(f), w(a)) < \varepsilon$ est au moins $\exp(\ell(1-\varepsilon)\mathfrak{H}(w(a)))$. Il n'est pas difficile d'en déduire que pour tout $\varepsilon > 0$, il existe $\varepsilon' > 0$ et m tel que si $n \geq m$, $G \in \mathbb{G}(n,d)$ et $g \in \mathbb{R}^n$ sont tels que $\mathrm{d_L}(\mathrm{distr}_G(g)_{S_k^\star}, \nu_{S_k^\star}) < \varepsilon'$ alors il existe au moins $\exp(n(1-\varepsilon)\mathbb{E}[\mathfrak{H}(Y_\varnothing|X_\varnothing)])$ vecteurs $f \in F^n$ tels que $\mathrm{d_L}(\mathrm{distr}_G(f)_{S_k^\star}, \mu_{S_k^\star}) < \varepsilon$ (on applique la propriété précédente pour chaque k-voisinage possible d'un graphe d-régulier coloré par E, avec ℓ le nombre de $x \in \{1,\dots,n\}$ tels que g restreint à la boule de centre x et de rayon k dans G soit égal à ce k-voisinage).

Enfin, le processus X étant typique, pour tout $\varepsilon' > 0$, le nombre de graphes $G \in \mathbb{G}(n,d)$ tels qu'il existe $g \in \mathbb{R}^n$ avec $\mathrm{d_L}(\mathrm{distr}_G(g)_{S_k^\star}, \nu_{S_k^\star}) \leq \varepsilon'$ est au moins $|\mathbb{G}(n,d)|/2$ pour une suite infinie d'entiers n. On en déduit

$$\limsup_{n\to\infty} \frac{1}{n} \log \frac{|C_k(n,\varepsilon)|}{|\mathbb{G}(n,d)|} \geq (1-\varepsilon)\mathbb{E}[\mathfrak{H}(Y_\varnothing|X_\varnothing)].$$

Il reste à faire tendre ε vers 0. □

4.2. Inégalité entropique étoile-arête pour les ondes

Soit Y un processus standard invariant et lisse. Pour tout entier $k \geq 0$, nous généralisons l'entropie invariante (9) de la façon suivante

$$\mathfrak{D}_k(Y) = \mathfrak{H}_{\mathrm{supp}}(Y_{S_k^\star}) - (d/2)\mathfrak{H}_{\mathrm{supp}}(Y_{S_k^-}). \tag{14}$$

Théorème 4.6. — *Soit $k \geq 0$ entier. Si est Y une onde standard typique lisse de valeur propre $\lambda \in \sigma(T_d)$ alors*

$$\mathfrak{D}_k(Y) \geq 0.$$

Nous expliquons comment déduire le théorème 4.6 du théorème 4.1 par un argument de discrétisation. Soit $\delta > 0$ avec $2/\delta^2$ entier, on partitionne $[-1/\delta, 1/\delta]$ en intervalles de longueurs égales à δ. Pour $x \in [-1/\delta, 1/\delta]$, on peut ainsi définir x_δ comme le centre de l'intervalle auquel il appartient. Pour $x \in \mathbb{R}\backslash[-1/\delta, 1/\delta]$ on pose $x_\delta = 0_\delta$. Pour tout entier $n \geq 2$ et $x \in \mathbb{R}^n$, on peut également définir x_δ en appliquant la partition en intervalles à chacune des coordonnées. Notons que x_δ appartient à R_δ^n une partie finie de $\mathbb{R}^n$ (de cardinal $(2/\delta^2)^n$). Si X est une variable sur $\mathbb{R}^n$ avec une densité, on peut classiquement relier l'entropie de Shannon de X_δ à l'entropie de X, voir [15, théorème 8.3.1] et [4, lemme 8.1] :

Lemme 4.7. — *Soit X une variable aléatoire dans $\mathbb{R}^n$ de densité C^1 telle que $\mathbb{E}\|X\|_2^2 < \infty$. Alors*

$$\lim_{\delta\to 0} \Big(\mathfrak{H}_{R_\delta^n}(X_\delta) + n\log\delta\Big) = \mathfrak{H}_{\mathbb{R}^n}(X).$$

On en déduit également l'énoncé suivant.

Corollaire 4.8. — *Soient X une variable aléatoire dans $\mathbb{R}$ avec $\mathbb{E}|X|^2 < \infty$ et Y une variable gaussienne $N(0,\sigma^2)$ indépendante avec $\sigma > 0$. Alors,*

$$\lim_{\delta\to 0} (\mathbb{E}\{\mathfrak{H}_{R_\delta}((X+Y)_\delta|X)\} + \log\delta) = \mathfrak{H}_{\mathbb{R}}(N(0,\sigma^2)),$$

où $\mathfrak{H}(Z|X)$ est l'entropie de la loi conditionnelle de Z sachant X.

Démonstration. — Si X est presque sûrement égal à 0 alors l'énoncé est une conséquence du lemme 4.7. Sinon on écrit

$$\mathbb{E}\mathfrak{H}_{R_\delta}((X+Y)_\delta|X) = \mathbb{E}\mathbf{1}_{|X|<1/(2\delta)}\mathfrak{H}_{R_\delta}((X+Y)_\delta|X) + \mathbb{E}\mathbf{1}_{|X|\geq 1/(2\delta)}\mathfrak{H}_{R_\delta}((X+Y)_\delta|X).$$

On rappelle que $|R_\delta| = 2/\delta^2$. On a donc $\mathfrak{H}_{R_\delta}((X+Y)_\delta|X) \leq -\log(2/\delta^2)$. En outre $\mathbb{P}(|X| \geq 1/(2\delta)) \leq 4\delta^2\mathbb{E}|X|^2$. Il suffit donc de montrer que

$$\sup_{x\in(-1/(2\delta),1/(2\delta))} |\mathfrak{H}_{R_\delta}((x+Y)_\delta) - \mathfrak{H}_{R_\delta}(Y_\delta)|$$

tend vers 0 avec δ. Soit $x \in (-1/(2\delta), 1/(2\delta))$. Nous observons que si $x \in R_\delta$ alors $(x+Y)_\delta = x + Y_\delta$ si $|Y| \leq 1/(2\delta)$. On en déduit que si $x \in R_\delta$,

$$|\mathfrak{H}_{R_\delta}((x+Y)_\delta) - \mathfrak{H}_{R_\delta}(Y_\delta)| \leq 2\mathbb{P}(|Y| \geq 1/(2\delta)) \log|R_\delta|$$

qui tend vers 0 avec δ. Si $x \notin R_\delta$, soit $\bar{x} \in R_\delta$ tel que $|\bar{x} - x| \leq \delta$. La distance en variation totale entre $N(x,\sigma^2)$ et $N(\bar{x},\sigma^2)$ est d'ordre de grandeur δ. On rappelle que si Z, Z' sont des variables aléatoires sur un ensemble discret F et de distance en variation totale D, on a

$$|\mathfrak{H}_F(Z) - \mathfrak{H}_F(Z')| \leq -D \log \frac{D}{|F|},$$

voir [15, théorème 17.3.3]. Cela complète l'argument. □

Les ingrédients pour démontrer le théorème 4.6 sont maintenant réunis.

Démonstration du théorème 4.6. — Soient $\delta > 0$ et $\sigma > 0$. Soit Z un processus gaussien sur $\mathbb{R}^{V_d}$ avec $(Z_x)_{x\in V_d}$ variables gaussiennes $N(0,\sigma^2)$ indépendantes. On applique le corollaire 4.5 à $(Y+Z)_\delta$, on trouve

$$\mathfrak{H}\Big(((Y+Z)_\delta)_{S_k^\star}\Big) - (d/2)\mathfrak{H}\Big(((Y+Z)_\delta)_{S_k^-}\Big) \geq \mathbb{E}\Big[\mathfrak{H}\Big((Y_\varnothing + Z_\varnothing)_\delta\Big)|Y_\varnothing)\Big].$$

On note que pour tout entier $k \geq 0$, on a $|S_k^\star| - (d/2)|S_k^-| = 1$. On applique le lemme 4.7 (avec $n = |S_k^\star|$ et $n = |S_k^-|$) et le corollaire 4.8. Dans la limite δ tend vers 0, les termes en $\log\delta$ s'annulent et on trouve

$$\mathfrak{H}\Big((Y+Z)_{S_k^\star}\Big) - (d/2)\mathfrak{H}\Big((Y+Z)_{S_k^-}\Big) \geq \mathfrak{H}(N(0,\sigma^2)). \tag{15}$$

La loi du processus Z est invariante par transformation orthogonale. Ainsi, si $S \subset V_d$ est fini et E_S est l'espace vectoriel engendré par le support de Y_S, la variable Z_S se décompose comme la somme indépendante d'une variable gaussienne dans E_S de loi $N(0,\sigma^2 I_{E_S})$ et d'une variable gaussienne dans $(E_S)^\perp$ de loi $N(0,\sigma^2 I_{E_S^\perp})$ (où pour Σ matrice définie positive de taille n, $N(0,\Sigma)$ est la loi gaussienne centrée dans $\mathbb{R}^n$ de covariance Σ) . On a donc

$$\mathfrak{H}\Big((Y+Z)_S\Big) = \mathfrak{H}\Big((Y+Z)_{|E_S}\Big) + \dim(E_S^\perp)\mathfrak{H}(N(0,\sigma^2)).$$

Le processus Y étant lisse, grâce au théorème de convergence dominée, on trouve pour $S \in \{S_k^\star, S_k^-\}$,

$$\lim_{\sigma\to 0} \mathfrak{H}\Big((Y+Z)_{|E_S}\Big) = \mathfrak{H}_{\mathrm{supp}}\Big(Y_S\Big).$$

En outre, pour tout $k \geq 0$, d'après le lemme 2.7,

$$\dim(E_{S_k^\star}^\perp) - (d/2)\dim(E_{S_k^-}^\perp) - 1 = |\partial S_k^\star| - (d/2)|\partial S_k^-| = 0,$$

où on a utilisé que $|S_k^\star| - (d/2)|S_k^-| = 1$ et $|\partial S_k^\star| = (d/2)|\partial S_k^-|$. Il reste donc à prendre la limite $\sigma \to 0$ dans (15). $\square$

5. MINIMUM DE L'ENTROPIE INVARIANTE

Dans cette section nous démontrons le théorème 3.3. Ce résultat est une conséquence de l'inégalité entropique de la section précédente, de considérations générales sur l'entropie et du calcul explicite de $\mathfrak{D}_k(Y) = \mathfrak{H}_{\text{supp}}(Y_{S_k^\star}) - (d/2)\mathfrak{H}_{\text{supp}}(Y_{S_k^-})$ pour Y l'onde gaussienne. Nous commençons par ce dernier point.

5.1. Entropie de l'onde gaussienne

Backhausz et Szegedy ont démontré l'énoncé suivant.

THÉORÈME 5.1. — *Soient $\lambda \in \sigma(T_d)$ et Y l'onde gaussienne de valeur propre λ. On a*

$$\lim_{k\to\infty} \mathfrak{D}_k(Y) = 0.$$

Une interprétation erronée mais éclairante de ce théorème est la suivante. L'arbre T_d restreint à $S_k^\star$ peut se décomposer comme $\varnothing$ et l'union de d arbres isomorphes par un automorphisme de T_d, chacun correspondant à l'arbre de profondeur k engendré par les sommets de $S_k^\star$ qui sont les descendants du sommet $i = 1, \ldots, d$. On peut également écrire T_d restreint à S_k^- comme l'union de 2 tels arbres de profondeur k. Si l'onde gaussienne restreinte à ces arbres était indépendante alors on aurait $\mathfrak{H}_{\text{supp}}(Y_{S_k^\star})$ égal à $(d/2)\mathfrak{H}_{\text{supp}}(Y_{S_k^-})$. La réalité est plus subtile mais la démonstration du théorème 5.1 cache un phénomène similaire sur des sous-espaces vectoriels de grandes dimensions de $\mathbb{R}^{S_k^\star}$ et $\mathbb{R}^{S_k^-}$.

Nous allons seulement esquisser la démonstration du théorème 5.1. On démarre avec la formule : pour $X \in \mathbb{R}^n$ une variable gaussienne dont la matrice de covariance Σ est de rang m,

$$\mathfrak{H}_{\text{supp}}(X) = \frac{1}{2}\log((2\pi e)^m \det{}'(\Sigma)),$$

où $\det'(\Sigma)$ est le produit des m valeurs propres non nulles de Σ. On observe que pour tout $k \geq 1$, $|\partial S_k^\star| = (d/2)|\partial S_k^-|$. Ainsi, grâce au lemme 2.7, on obtient pour $k \geq 1$,

$$2\mathfrak{D}_k(Y) = \log(\det{}'(\Sigma_k^\star)) - (d/2)\log(\det{}'(\Sigma_k^-)),$$

où $\Sigma_k^\star$, Σ_k^- sont les covariances respectives de $Y_{S_k^\star}$ et $Y_{S_k^-}$. Il faut donc démontrer que $\log(\det'(\Sigma_k^\star)) - (d/2)\log(\det'(\Sigma_k^-))$ tend vers 0 avec k tendant vers l'infini.

Dans le lemme 2.3, nous avons donné une formule pour les entrées de $\Sigma_k^\star$ et Σ_k^-. En utilisant que ces matrices sont invariantes par les automorphismes de T_d et que leurs images sont $W_\lambda(S_k^\star)$ et $W_\lambda(S_k^-)$, il est en fait possible de diagonaliser explicitement ces matrices.

Nous expliquons maintenant comment réaliser cette diagonalisation pour $\Sigma_k^\star$. Le cas de Σ_k^- est similaire. On définit

$$F = \Big\{ f \in \mathbb{R}^{S_k^\star} : f(x) = f(y) \text{ si } |y| = |x| \Big\},$$

où $|x|$ est la distance à $\varnothing$. Pour tout $x \in \mathrm{int}(S_k^\star)$, on introduit également le sous-espace vectoriel E_x de $\mathbb{R}^{S_k^\star}$ composé des vecteurs f qui vérifient les propriétés suivantes :

(i) $f(y) = 0$ si y n'est pas un descendant de x ;

(ii) pour tout $i \in \{1, \dots, d-1\}$, $y, y' \in \Sigma_k^\star$, $f(y) = f(y')$ si y et y' sont des descendants de (x, i) ;

(iii) $f \in F^\perp$.

Lemme 5.2. — *Nous avons les propriétés suivantes :*

1. L'espace vectoriel $\mathbb{R}^{S_k^\star}$ admet la décomposition en somme directe

$$\mathbb{R}^{S_k^\star} = F \oplus \bigoplus_{x \in \mathrm{int}(S_k^\star)} E_x.$$

2. Les espaces vectoriels F et $(E_x), x \in \mathrm{int}(\Sigma_k^\star)$, sont invariants par $\Sigma_k^\star$.

Démonstration. — L'orthogonalité de F et E_x est la propriété (iii) de E_x. Soient $x \neq x'$. Si ni x ni x' ne sont ancêtres l'un de l'autre alors l'orthogonalité de E_x et $E_{x'}$ vient de la propriété (i) (les vecteurs de E_x et $E_{x'}$ ont supports disjoints). Si x est l'ancêtre de x' alors la propriété (ii) implique que les vecteurs de E_x sont constants sur le support des vecteurs de $E_{x'}$. Enfin, la propriété (iii) de $E_{x'}$ implique l'orthogonalité aux vecteurs constants sur le support des vecteurs de $E_{x'}$. Un calcul immédiat de dimension implique la première assertion du lemme.

Rappelons que les entrées de la matrice $\Sigma_k^\star$ ne dépendent que de la distance entre les sommets (voir lemme 2.3). L'invariance est une conséquence de cette seule propriété de $\Sigma_k^\star$. Tout d'abord, l'assertion $\Sigma_k^\star F = F$ est facile à vérifier : si deux sommets x, y sont à la même distance, disons i, de $\varnothing$ alors pour tout entier $j \geq i$, x et y ont le même nombre de descendants à distance j de $\varnothing$. Nous pouvons aussi vérifier l'invariance de E_x. Tout d'abord (ii) est à nouveau vérifiée car les entrées de $\Sigma_k^\star$ ne dépendent que de la distance entre les sommets. Pour (iii), F étant invariant et $\Sigma_k^\star$ symétrique, $F^\perp$ est également invariant. Il reste à vérifier (i) : tout sommet qui n'est pas un descendant x est à distance égale de tous les descendants de x à une distance donnée de $\varnothing$. Ainsi, la propriété (iii) de E_x implique que les vecteurs dans $\Sigma_k^\star E_x$ vérifient (i). □

La diagonalisation de $\Sigma_k^\star$ est maintenant à notre portée. Expliquons par exemple commet calculer les valeurs propres de $\Sigma_k^\star$ restreinte à F. D'après le lemme 2.7, l'image de $\Sigma_k^\star$ est $W_\lambda(S_k^\star)$. En particulier, la dimension de $F \cap W_\lambda(S_k^\star)$ étant 1, les valeurs propres de $\Sigma_k^\star$ restreinte à F sont égales à 0 avec multiplicité k et λ_F avec multiplicité 1 pour un certain $\lambda_F \geq 0$. On en déduit que

$$\lambda_F = \mathrm{tr}((\Sigma_k^\star)_{|F}).$$

Pour calculer cette trace, on choisit la base orthogonale naturelle de F,

$$f_i = (d(d-1)^{i-1})^{-1/2}\mathbf{1}_i$$

où pour $i \in \{0,\dots,k\}$, $\mathbf{1}_i$ est le vecteur égal à 1 sur tous les sommets à distance i de $\varnothing$. En utilisant la notation du lemme 2.3, on obtient

$$\lambda_F = \sum_{i=0}^{k} \langle f_i, \Sigma_k^* f_i\rangle = \sum_{i=0}^{k}\sum_{j=0}^{k} N(i,2j) q_{2j}(\lambda),$$

où $N(i,2j)$ est le nombre de sommets à distance i de $\varnothing$ et à distance $2j$ d'un sommet fixé à distance i de $\varnothing$.

Une fois ces calculs explicites sur les valeurs propres menés, pour terminer la preuve du théorème 5.1, il reste à utiliser les propriétés des polynômes de Tchebychev pour estimer l'asymptotique en j tendant vers l'infini de $q_{2j}(\lambda)$ avec $\lambda \in \sigma(T_d)$ fixé. Voir [4, Section 9] pour l'argument complet.

5.2. Fin de la démonstration du théorème 3.3

Soit Y une onde invariante standard lisse. On pose pour tout $S \subset V_d$ tel que $\varnothing \in S$ et $S = \bigcup_{x\in \mathrm{int}(S)} B_1(x)$,

$$\widetilde{\mathfrak{H}}_{\mathrm{supp}}(Y_S) = \mathfrak{H}(Y_\varnothing, (A_x)_{x\in\mathrm{int}(S)}),$$

où les variables aléatoires $(A_x)_{x\in V_d}$ sont définies par (8). On pose également

$$\widetilde{\mathfrak{D}}_k(Y) = \widetilde{\mathfrak{H}}_{\mathrm{supp}}(Y_{S_k^\star}) - (d/2)\widetilde{\mathfrak{H}}_{\mathrm{supp}}(Y_{S_k^-}).$$

Comme sa notation l'indique $\widetilde{\mathfrak{D}}_k(Y)$ sert à l'analyse de $\mathfrak{D}_k(Y)$. Rappelons en effet que si X est une variable dans $\mathbb{R}^n$ et si T une application linéaire inversible de $\mathbb{R}^n$, on a

$$\mathfrak{H}(TX) = \mathfrak{H}(X) + \log|\det(T)|. \tag{16}$$

Ainsi, si X, Y sont des ondes invariantes standard lisses, d'après le lemme 2.7, on trouve

$$\widetilde{\mathfrak{H}}_{\mathrm{supp}}(Y_S) - \widetilde{\mathfrak{H}}_{\mathrm{supp}}(X_S) = \mathfrak{H}_{\mathrm{supp}}(Y_S) - \mathfrak{H}_{\mathrm{supp}}(X_S).$$

En particulier,

$$\widetilde{\mathfrak{D}}_k(Y) - \widetilde{\mathfrak{D}}_k(X) = \mathfrak{D}_k(Y) - \mathfrak{D}_k(X). \tag{17}$$

Le lemme suivant repose sur des inégalités basiques en théorie de l'information.

Lemme 5.3. — *Soit Y une onde invariante standard lisse sur T_d. Pour tout entier $k \geq 1$, on a*

$$\widetilde{\mathfrak{D}}_k(Y) \leq \widetilde{\mathfrak{D}}_{k-1}(Y).$$

En outre si pour tout $k \geq 1$, $\widetilde{\mathfrak{D}}_k(Y) = \widetilde{\mathfrak{D}}_{k-1}(Y)$ alors l'onde Y est 2-markovienne. Enfin, l'onde gaussienne est 2-markovienne.

Démonstration. — On pose $B_x = A_x$ si $x \neq \varnothing$, $B_\varnothing = (Y_\varnothing, A_\varnothing)$ et pour $S \subset V_d$, $B_S = (B_x)_{x\in \mathrm{int}(S)}$. De telle sorte que nous trouvons $\widetilde{\mathfrak{H}}_{\mathrm{supp}}(Y_S) = \mathfrak{H}(B_S)$ pour $S \in \{S_k^\star, S_k^-\}$.

Pour $k \geq 1$ et $i \in \{1, \dots, d\}$, on considère l'ensemble S_k^i formé des éléments de $S_k^\star \backslash S_{k-1}^\star$ qui sont des descendants de i. En particulier, nous avons les unions disjointes suivantes : $S_k^- = S_{k-1}^\star \cup S_k^1$ et $S_k^\star = S_{k-1}^\star \cup S_k^1 \cup \cdots \cup S_k^d$. On rappelle que pour tout triplet de variables aléatoires (T, U, V), on a $\mathfrak{H}(T, U, V) \leq \mathfrak{H}(T, U) + \mathfrak{H}(T, V) - \mathfrak{H}(T)$ (voir [15, Problem 2.29]). On obtient

$$\begin{aligned}
\mathfrak{H}(B_{S_k^\star}) &= \mathfrak{H}(B_{S_{k-1}^\star}, B_{S_k^1}, \dots, B_{S_k^d}) \\
&\leq \mathfrak{H}(B_{S_{k-1}^\star}, B_{S_k^1}, \dots, B_{S_k^{d-1}}) + \mathfrak{H}(B_{S_{k-1}^\star}, B_{S_k^d}) - \mathfrak{H}(B_{S_{k-1}^\star}) \\
&= \mathfrak{H}(B_{S_{k-1}^\star}, B_{S_k^1}, \dots, B_{S_k^{d-1}}) + \widetilde{\mathfrak{H}}_{\mathrm{supp}}(Y_{S_k^-}) - \widetilde{\mathfrak{H}}_{\mathrm{supp}}(Y_{S_{k-1}^\star}),
\end{aligned}$$

où à la dernière ligne on a utilisé l'invariance de Y : $\mathfrak{H}(B_{S_{k-1}^\star}, B_{S_k^d}) = \mathfrak{H}(B_{S_k^-})$. En répétant cette inégalité $(d-1)$ fois, on arrive à l'inégalité

$$\widetilde{\mathfrak{H}}_{\mathrm{supp}}(Y_{S_k^\star}) \leq d\widetilde{\mathfrak{H}}_{\mathrm{supp}}(Y_{S_k^-}) - (d-1)\widetilde{\mathfrak{H}}_{\mathrm{supp}}(Y_{S_{k-1}^\star}). \tag{18}$$

De même, avec $T_k = S_{k-1}^\star \backslash S_{k-1}^-$, on a l'union disjointe $S_k^- = S_{k-1}^- \cup T_k \cup S_k^1$ et

$$\begin{aligned}
\widetilde{\mathfrak{H}}_{\mathrm{supp}}(Y_{S_k^-}) &= \mathfrak{H}(B_{S_{k-1}^-}, B_{T_k}, B_{S_k^1}) \\
&\leq \mathfrak{H}(B_{S_{k-1}^-}, B_{T_{k-1}}) + \mathfrak{H}(B_{S_{k-1}^-}, B_{S_k^1}) - \mathfrak{H}(B_{S_{k-1}^-}) \\
&= 2\mathfrak{H}_{\mathrm{supp}}(X_{S_{k-1}^\star}) - \widetilde{\mathfrak{H}}_{\mathrm{supp}}(Y_{S_{k-1}^-}),
\end{aligned}$$

où on a utilisé l'invariance de Y et $S_{k-1}^- \cup S_k^1 = B_k(1)$ est isomorphe à $S_{k-1}^\star = S_{k-1}^- \cup T_k$. Ainsi, si on additionne $(d/2)$ fois cette dernière inégalité à (18), on arrive précisément à $\widetilde{\mathfrak{D}}_k(Y) \leq \widetilde{\mathfrak{D}}_{k-1}(Y)$ comme annoncé.

Pour le cas d'égalité, on rappelle que si $\mathfrak{H}(T, U, V) = \mathfrak{H}(T, U) + \mathfrak{H}(T, V) - \mathfrak{H}(T)$ alors U et V sont indépendants conditionnellement à T (c'est une conséquence de [15, théorème 2.6.5]). En utilisant l'invariance de Y, il n'est alors pas difficile de vérifier la propriété 2-markovienne à l'intérieur de $S_k^\star$ par récurrence sur $k \geq 0$. Enfin, il est immédiat de vérifier que pour l'onde gaussienne, nous avons l'égalité $\widetilde{\mathfrak{D}}_k(Y) = \widetilde{\mathfrak{D}}_{k-1}(Y)$. En effet, d'après le lemme 2.6, les variables (B_x) sont décorrélées et donc indépendantes dans le cas gaussien. □

On peut maintenant terminer la démonstration du théorème 3.3.

Démonstration du théorème 3.3. — Soient X une onde invariante lisse standard et Y l'onde gaussienne correspondante. D'après le lemme 5.3, $\widetilde{\mathfrak{D}}_k(Y) = \widetilde{\mathfrak{D}}_0(Y)$ pour tout entier $k \geq 1$. D'après (17) et le lemme 5.3 appliqué à X, on trouve

$$\begin{aligned}
\mathfrak{D}(X) - \mathfrak{D}(Y) = \widetilde{\mathfrak{D}}_0(X) - \widetilde{\mathfrak{D}}_0(Y) &\geq \widetilde{\mathfrak{D}}_1(X) - \widetilde{\mathfrak{D}}_1(Y) \\
&\geq \cdots \geq \widetilde{\mathfrak{D}}_k(X) - \widetilde{\mathfrak{D}}_k(Y) = \mathfrak{D}_k(X) - \mathfrak{D}_k(Y).
\end{aligned}$$

Si X est typique, d'après le théorème 4.6, on en déduit que $\mathfrak{D}(X)-\mathfrak{D}(Y) \geq -\mathfrak{D}_k(Y)$. On fait tendre k vers l'infini. En conjonction avec le théorème 5.1, on obtient

$$\mathfrak{D}(X) - \mathfrak{D}(Y) \geq 0.$$

C'est la première assertion du théorème 3.3. En cas d'égalité alors toutes les inégalités ci-dessus sont des égalités. On en déduit d'après le lemme 5.3 que l'onde X est 2-markovienne. □

6. MAXIMUM DE L'ENTROPIE INVARIANTE

Dans cette section, nous présentons la démonstration du théorème 3.2. Nous avons le résultat suivant sur la croissance de l'entropie invariante le long de l'équation de la chaleur.

THÉORÈME 6.1. — *Soient $\lambda \in \sigma(T_d)$, X une onde invariante lisse de valeur propre λ et Y une onde gaussienne indépendante de même valeur propre λ. Pour tout $t \geq 0$, on pose $X(t) = X + \sqrt{2t}Y$ et on considère la fonction $t \mapsto \mathfrak{D}(X(t))$. Alors pour tout $t \geq 0$, on a $\mathfrak{D}(X(t))' \geq 0$. En outre, si $\mathfrak{D}(X(0))' = 0$ alors $X_{S^\star}$ et $Y_{S^\star}$ ont même loi.*

Le théorème 3.2 est une conséquence facile du théorème 6.1.

Démonstration du théorème 3.2. — On remarque que l'entropie invariante est invariante par changement d'échelle : pour tout $s \neq 0$ et tout processus X invariant non nul,

$$\mathfrak{D}(sX) = \mathfrak{D}(X).$$

En effet, pour tout $S \subset V_d$ fini, d'après (16), on a $\mathfrak{H}_{\text{supp}}(sX_S) = \mathfrak{H}_{\text{supp}}(X_S) + m_S \log s$, où m_S est la dimension de l'espace vectoriel engendré par le support de X_S. On rappelle d'après le lemme 2.7 que $m_{S^\star} = d$ et $m_{S^-} = 2$. On en déduit bien que $\mathfrak{D}(sX) = \mathfrak{D}(X)$.

Ainsi, avec les notations du théorème 6.1, si X est une onde invariante lisse standard, alors $\bar{X}(t) = X(t)/\sqrt{1+2t}$ est une onde invariante standard lisse et $\mathfrak{D}(\bar{X}(t)) \geq \mathfrak{D}(X)$. En faisant tendre t vers l'infini, on obtient bien $\mathfrak{D}(Y) \geq \mathfrak{D}(X)$. C'est le premier énoncé du théorème 3.2. En cas d'égalité, d'après le théorème 6.1, nous en déduisons que $X_{S^\star}$ et $Y_{S^\star}$ ont nécessairement la même loi. □

Décroissance de l'entropie invariante. — Pour démontrer le théorème 6.1, nous allons commencer par exprimer la densité de $X_{S^\star}$ dans une base de $W_\lambda(S^\star)$ qui exploite au mieux les symétries des variables.

Soit Σ_0 la matrice de covariance commune de $(X_\varnothing(t), X_1(t), \ldots, X_{d-1}(t))$ pour tout $t \geq 0$ donnée par le lemme 2.3. Cette matrice est inversible et on pose $Z(t) = \Sigma_0^{-1/2}(X_\varnothing(t), X_1(t), \ldots, X_{d-1}(t))$. Par construction, $\mathbb{E}Z_i(t)Z_j(t) = \mathbf{1}_{i=j}$ pour tous $i, j \in \{1, \ldots, d\}$. On note $(v_\varnothing, v_1, \ldots, v_{d-1})$ les vecteurs colonnes de $\Sigma_0^{1/2}$ et on pose

$v_d = \lambda v_\varnothing - v_1 - \cdots - v_{d-1}$, de telle sorte que $X_x(t) = \langle v_x, Z(t)\rangle$ et $\mathbb{E}X_x(t)X_y(t) = \langle v_x, v_y\rangle$ pour tous $x, y \in S^\star$. D'après le lemme 2.3, pour tous $i \neq j \in \{1, \dots, d\}$,

$$\|v_\varnothing\|_2 = \|v_i\|_2 = 1\,, \quad \langle v_i, v_j\rangle = \frac{\lambda^2 - d}{d(d-1)} \quad \text{et} \quad \langle v_\varnothing, v_i\rangle = \frac{\lambda}{d}.$$

Pour $|\lambda| \leq d$, il existe $(\alpha, \beta) \in \mathbb{R}^2$ tels que pour tout $i \in \{1, \dots, d\}$, $a_i = \alpha v_\varnothing + \beta v_i$ et $b_i = \beta v_\varnothing + \alpha v_i$ vérifient $\|a_i\|_2 = \|b_i\|_2 = 1$ et $\langle a_i, b_i\rangle = 0$ (la paire (α, β) est unique à un signe près et (α, β) vérifie les égalités $(\alpha^2 + \beta^2)\lambda/d + 2\alpha\beta = 0$ et $\alpha^2 + \beta^2 + 2\alpha\beta\lambda/d = 1$). Pour $\lambda \in \sigma(T_d)$ nous avons la propriété d'orthogonalité suivante.

Lemme 6.2. — *Soit $\lambda \in \sigma(T_d)$. Il existe une paire $(p, q) \in [0, 1]^2$ telle que $p + q = 1$ et pour tout $f \in \mathbb{R}^d$,*

$$\|f\|_2^2 = \sum_{i=1}^{d} \Big(p\langle f, a_i\rangle^2 + q\langle f, b_i\rangle^2\Big).$$

De plus, p et q sont strictement positifs si $|\lambda| < 2\sqrt{d-1}$.

Démonstration. — On pose $c = \alpha + \lambda/d$. Nous avons $|c| < 1$ et, pour tout i, $\langle a_i, v_\varnothing\rangle = c$, $\sum_i a_i = cdv_\varnothing$ et, par symétrie, pour tous $i \neq j$, $\langle a_i, a_j\rangle = (dc^2 - 1)/(d-1) = c'$. En particulier, il existe un réel γ tel que $\langle a_i - \gamma v_\varnothing, a_j - \gamma v_\varnothing\rangle = 0$ pour tous $i \neq j$: en effet, l'équation $\gamma^2 - 2\gamma c + c' = 0$ a deux solutions réelles. Ainsi, la famille de vecteurs $((a_i - \gamma v_\varnothing)/\|a_i - \gamma v_\varnothing\|_2)_{i=1,\dots,d}$ forme une base orthonormée de $\mathbb{R}^d$. Aussi, la norme $\|a_i - \gamma v_\varnothing\|_2$ ne dépend pas de i. En particulier, d'après le théorème de Pythagore, pour tous $j \in \{1, \dots, d\}$ et $f \in \mathbb{R}^d$,

$$\begin{aligned}
\sum_{i=1}^{d} \langle f, a_i\rangle^2 &= \sum_{i=1}^{d} \langle f, a_i - \gamma v_\varnothing\rangle^2 + \sum_{i=1}^{d} 2\gamma\langle f, a_i - \gamma v_\varnothing\rangle\langle f, v_\varnothing\rangle + \sum_{i=1}^{d} \gamma^2\langle f, v_\varnothing\rangle^2 \\
&= \|a_j - \gamma v_\varnothing\|_2^2 \|f\|_2^2 + dc'\langle f, v_\varnothing\rangle^2,
\end{aligned}$$

où on a utilisé que $\sum_i a_i = cdv_\varnothing$ et $\gamma^2 - 2\gamma c + c' = 0$. L'identité ci-dessus appliquée à $f = v_\varnothing$ implique que $\|a_j - \gamma v_\varnothing\|_2^2 = d(1 - c^2)/(d-1)$. En remplaçant c et c' par leurs valeurs, on trouve que, pour tous $j \in \{1, \dots, d\}$ et $f \in \mathbb{R}^d$,

$$\sum_{i=1}^{d} \langle f, a_i\rangle^2 = \frac{d}{d-1}\Big((1 - \langle a_j, v_\varnothing\rangle^2)\|f\|_2^2 + (d\langle a_j, v_\varnothing\rangle^2 - 1)\langle f, v_\varnothing\rangle^2\Big).$$

Le même argument avec $(b_1, \dots b_d)$ donne pour tous $j \in \{1, \dots, d\}$ et $f \in \mathbb{R}^d$,

$$\sum_{i=1}^{d} \langle f, b_i\rangle^2 = \frac{d}{d-1}\Big((1 - \langle b_j, v_\varnothing\rangle^2)\|f\|_2^2 + (d\langle b_j, v_\varnothing\rangle^2 - 1)\langle f, v_\varnothing\rangle^2\Big).$$

Ainsi si $p + q = 1$, pour tous $j \in \{1, \dots, d\}$ et $f \in \mathbb{R}^d$,

$$\sum_{i=1}^{d} \Big(p\langle f, a_i\rangle^2 + q\langle f, b_i\rangle^2\Big) = \|f\|_2^2 + \Big(dp\langle a_j, v_\varnothing\rangle^2 + dq\langle b_j, v_\varnothing\rangle^2 - 1\Big)\frac{(d\langle f, v_\varnothing\rangle^2 - \|f\|_2^2)}{d-1}.$$

Un calcul immédiat donne $d\langle a_j, v_\varnothing\rangle^2$ et $d\langle b_j, v_\varnothing\rangle^2$ égaux à $(d \pm \sqrt{d^2 - \lambda^2})/2$. On en déduit que pour tout $\lambda \in \sigma(T_d)$, il existe une paire $(p, q) \in [0, 1]^2$ telle que $p + q = 1$ et

$dp\langle a_j, v_\varnothing\rangle^2 + dq\langle b_j, v_\varnothing\rangle^2 = d/2 + (2p-1)(\sqrt{d^2-\lambda^2})/2 = 1$ (pour $\lambda = 0$, on a $p = 1/d$, pour $|\lambda| = 2\sqrt{d-1}$, on a $p = 0$). Pour ce choix de (p, q) on obtient immédiatement l'énoncé du lemme. $\square$

Démonstration du théorème 6.1. — Nous démontrons d'abord que $\mathfrak{D}(X(t))' \geq 0$ pour tout $t \geq 0$. L'onde $X(t)$ étant invariante pour tout $t \geq 0$, il suffit de vérifier que $\mathfrak{D}(X(0))' \geq 0$. On rappelle que $Z(t) = \Sigma_0^{-1/2}(X_\varnothing(t), X_1(t), \ldots, X_{d-1}(t))$ est un vecteur aléatoire de $\mathbb{R}^d$ centré et de matrice de covariance identité. D'après (16), pour tout $t > 0$,

$$\mathfrak{D}(X(t)) = c + \mathfrak{H}(Z(t)) - (d/2)\mathfrak{H}(\langle a_1, Z(t)\rangle, \langle b_1, Z(t)\rangle),$$

où c est une constante indépendante de t. On pose $Z = Z(0)$. Par construction $Z(t) = Z + \sqrt{2t}N$ avec N vecteur gaussien standard dans $\mathbb{R}^d$. Ainsi, si f est la densité de Z et f_1 la densité $(\langle a_1, Z\rangle, \langle b_1, Z\rangle)$, la formule classique de Bruijn (voir [15, théorème 17.7.2]) affirme que

$$\mathfrak{D}(X(0))' = \int_{\mathbb{R}^d} \|\nabla f\|_2^2/f - (d/2)\int_{\mathbb{R}^2} \|\nabla f_1\|_2^2/f_1.$$

On a

$$\|\nabla f_1\|_2^2 = (\partial_{a_1} f_1)^2 + (\partial_{b_1} f_1)^2,$$

et d'après le lemme 6.2,

$$\|\nabla f\|_2^2 = \sum_{i=1}^d p(\partial_{a_i} f)^2 + q(\partial_{b_i} f)^2.$$

L'onde X étant invariante, les lois de $(X_\varnothing, X_i)$ et donc de $(\langle a_i, Z\rangle, \langle b_i, Z\rangle)$ ne dépendent pas de i. On en déduit que

$$\mathfrak{D}(X(0))' = d\int_{\mathbb{R}^d} (p(\partial_{a_1} f)^2 + q(\partial_{b_1} f)^2)/f - (d/2)\int_{\mathbb{R}^2} ((\partial_{a_1} f_1)^2 + (\partial_{b_1} f_1)^2)/f_1. \tag{19}$$

Soient g, h des fonctions dans $L^1(\mathbb{R}^n)$. L'inégalité de Cauchy-Schwarz implique que si h est presque partout strictement positive sur le support de g,

$$\left(\int g\right)^2 \leq \left(\int g^2/h\right)\left(\int h\right). \tag{20}$$

On fixe une base orthonormale de $\mathbb{R}^d$ de la forme $(a_1, b_1, e_3, \ldots, e_d)$. Avec un abus de notation, on écrit la fonction f dans cette base orthonormale. Pour tout $y \in \mathbb{R}^2$, on applique l'inégalité (20) avec $n = d-2$ et, pour $z \in \mathbb{R}^{d-2}$, $h(z) = f(y, z)$, $g(z) = \partial_u f(y, z)$, $u \in \{a_1, b_1\}$. On trouve

$$\int_{\mathbb{R}^d} (\partial_{a_1} f)^2/f \geq \int_{\mathbb{R}^2} (\partial_{a_1} f_1)^2/f_1 \quad \text{et} \quad \int_{\mathbb{R}^d} (\partial_{b_1} f)^2/f \geq \int_{\mathbb{R}^2} (\partial_{b_1} f_1)^2/f_1.$$

On déduit donc de (19) que

$$\mathfrak{D}(X(0))' \geq (dp - d/2)\int_{\mathbb{R}^2} (\partial_{a_1} f_1)^2/f_1 + (dq - d/2)\int_{\mathbb{R}^2} (\partial_{b_1} f_1)^2/f_1.$$

Enfin l'onde X étant invariante, les lois de $(X_\varnothing, X_1)$ et $(X_1, X_\varnothing)$ sont égales. Ainsi, $f_1(y_1,y_2) = f_1(y_2,y_1)$ et, en particulier,

$$\int_{\mathbb{R}^2} (\partial_{a_1} f_1)^2/f_1 = \int_{\mathbb{R}^2} (\partial_{b_1} f_1)^2/f_1.$$

On obtient $\mathfrak{D}(X(0))' \geq 0$. C'est le premier énoncé du théorème.

Pour le second énoncé, il faut considérer le cas d'égalité dans l'inégalité de Cauchy-Schwarz (20) : nous avons égalité si et seulement si g/h est presque partout constant. On suppose sans perte de généralité que $|\lambda| < 2\sqrt{d-1}$. Dans ce cas p, q sont strictement positifs. Soit $g = -\log f$. En utilisant la continuité de f et de ses dérivées, on obtient $\mathfrak{D}(X(0))' = 0$ si et seulement si pour tous $y \in \mathbb{R}^2$ et $z \in \mathbb{R}^{d-2}$, $\partial_{a_1} g(y,z) = \partial_{a_1} g(y,0)$ et $\partial_{b_1} g(y,z) = \partial_{b_1} g(y,0)$. On en déduit que $g(y,z) = g(y,0) + k(z)$ pour une certaine fonction k dans $C^\infty(\mathbb{R}^{d-2})$. En d'autres termes, la variable $(\langle a_1, Z\rangle, \langle b_1, Z\rangle)$ est indépendante de $(\langle e_3, Z\rangle, \ldots, \langle e_d, Z\rangle)$. En utilisant les symétries de la densité f, on peut alors montrer que Z est nécessairement une loi gaussienne standard dans $\mathbb{R}^d$.

En effet, soit u_i la projection orthogonale de v_i sur $v_\varnothing^\perp$. On peut vérifier que pour tous i, j, $\langle u_i, u_j\rangle$ est non nul. Par hypothèse, pour tout $x \in \mathbb{R}^d$, $\partial_{u_i} g(x) = h(\langle x, u_i\rangle, \langle x, v_\varnothing\rangle)$ pour une certaine fonction h dans $C^\infty(\mathbb{R}^2)$. Donc

$$\partial_{u_j}\partial_{u_i} g(x) = \langle u_j, u_i\rangle \partial_1 h(\langle x, u_i\rangle, \langle x, v_\varnothing\rangle) = \partial_{u_i}\partial_{u_j} g(x) = \langle u_j, u_i\rangle \partial_1 h(\langle x, u_j\rangle, \langle x, v_\varnothing\rangle).$$

Pour $i \neq j$, l'espace vectoriel engendré par $(u_i, u_j, v_\varnothing)$ est de dimension 3. On en déduit que $\partial_1 h$ est nul et donc pour certaines fonctions h_0, h_1 dans $C^\infty(\mathbb{R})$,

$$\partial_{u_i} g(x) = \langle x, u_i\rangle h_0(\langle x, v_\varnothing\rangle) + h_1(\langle x, v_\varnothing\rangle).$$

En outre, $\sum_i v_i = \lambda v_\varnothing$. Donc $\sum_i u_i = 0$ et $\sum_i \partial_{u_i} g(x) = 0$. Cela implique que h_1 est nulle. L'espace vectoriel engendré par $(u_1, \ldots, u_d)$ est $v_\varnothing^\perp$. Il s'en suit que pour une certaine fonction h_2 dans $C^\infty(\mathbb{R})$,

$$g(x) = \|x - \langle x, v_\varnothing\rangle v_\varnothing\|_2^2 h_0(\langle x, v_\varnothing\rangle)/2 + h_2(\langle x, v_\varnothing\rangle).$$

On rappelle qu'en écrivant la fonction g dans la base $(a_1, b_1, e_3, \ldots, e_d)$, $g(y,z) = g(y,0) + k(z)$ pour tous $y \in \mathbb{R}^2$ et $z \in \mathbb{R}^{d-2}$. Le vecteur $v_\varnothing$ étant dans l'espace vectoriel engendré par (a_1, b_1), on obtient que la fonction h_0 est constante (car, avec $x = (y,z)$, $z \mapsto \|x - \langle x, v_\varnothing\rangle v_\varnothing\|_2^2$ n'est pas une fonction constante). Donc,

$$g(x) = \|x - \langle x, v_\varnothing\rangle v_\varnothing\|_2^2 h_0(0)/2 + h_2(\langle x, v_\varnothing\rangle) = \|x\|_2^2 h_2(0)/2 + h_3(\langle x, v_\varnothing\rangle).$$

Il reste à se souvenir que $\langle v_\varnothing, Z\rangle$ et $\langle v_1, Z\rangle$ ont même loi. Cela implique que la fonction h_3 est constante. Enfin, les relations $\mathbb{E} Z_i Z_j = \mathbf{1}_{i=j}$ et $\mathbb{E}1 = 1$ déterminent les valeurs de $h_2(0) = 1$ et $h_3(0) = (n/2)\log(2\pi)$. □

Remerciements. — Je remercie beaucoup Olivier Guichard et Emmanuel Kowalski pour leurs relectures attentives. Ce texte n'aurait pu aboutir sans la disponibilité et la patience d'Ágnes Backhausz et de Balász Szegedy pour m'expliquer certains points de leur travail remarquable. Je les en remercie très chaleureusement, *nagyon köszönöm*!

RÉFÉRENCES

[1] N. Alon, I. Benjamini, E. Lubetzky & S. Sodin. « Non-backtracking random walks mix faster ». *Commun. Contemp. Math.*, 9(4):585–603, 2007.

[2] N. Anantharaman. « Quantum ergodicity on regular graphs ». *Comm. Math. Phys.*, 353(2):633–690, 2017.

[3] N. Anantharaman & E. Le Masson. « Quantum ergodicity on large regular graphs ». *Duke Math. J.*, 164(4):723–765, 2015.

[4] Á. Backhausz & B. Szegedy. « On the almost eigenvectors of random regular graphs ». arXiv:1607.04785, 2016.

[5] Á. Backhausz & B. Szegedy. « On large girth regular graphs and random processes on trees ». *Random Structures Algorithms*, 2018. À paraître.

[6] R. Bauerschmidt, J. Huang, A. Knowles & H.-T. Yau. « Bulk eigenvalue statistics for random regular graphs ». *Ann. Probab.*, 45(6A):3626–3663, 2017.

[7] R. Bauerschmidt, A. Knowles & H.-T. Yau. « Local semicircle law for random regular graphs ». *Comm. Pure Appl. Math.*, 70(10):1898–1960, 2017.

[8] E. A. Bender & E. Canfield. « The asymptotic number of labeled graphs with given degree sequences ». *Journal of Combinatorial Theory, Series A*, 24(3):296 – 307, 1978.

[9] I. Benjamini & O. Schramm. « Recurrence of distributional limits of finite planar graphs ». *Electron. J. Probab.*, 6 :no. 23, 13, 2001.

[10] B. Bollobás. « A probabilistic proof of an asymptotic formula for the number of labelled regular graphs ». *European J. Combin.*, 1(4):311–316, 1980.

[11] C. Bordenave. « A new proof of Friedman's second eigenvalue theorem and its extension to random lifts ». arXiv:1502.04482, 2015.

[12] C. Bordenave. « Spectrum of random graphs ». In *Advanced topics in random matrices*, volume 53 of *Panor. Synthèses*, pages 91–150. Soc. Math. France, Paris, 2017.

[13] L. Bowen. « The ergodic theory of free group actions : entropy and the f-invariant ». *Groups Geom. Dyn.*, 4(3):419–432, 2010.

[14] S. Brooks & E. Lindenstrauss. « Non-localization of eigenfunctions on large regular graphs ». *Israel J. Math.*, 193(1):1–14, 2013.

[15] T. M. Cover & J. A. Thomas. *Elements of information theory*. Wiley-Interscience [John Wiley & Sons], Hoboken, NJ, second edition, 2006.

[16] I. Dumitriu, T. Johnson, S. Pal & E. Paquette. « Functional limit theorems for random regular graphs ». *Probab. Theory Related Fields*, 156(3-4):921–975, 2013.

[17] I. Dumitriu & S. Pal. « Sparse regular random graphs : spectral density and eigenvectors ». *Ann. Probab.*, 40(5):2197–2235, 2012.

[18] Y. Elon. « Gaussian waves on the regular tree ». arXiv:0907.5065, 2009.

[19] J. Friedman. « A proof of Alon's second eigenvalue conjecture and related problems ». *Mem. Amer. Math. Soc.*, 195(910):viii+100, 2008.

[20] V. Harangi & B. Virág. « Independence ratio and random eigenvectors in transitive graphs ». *Ann. Probab.*, 43(5):2810–2840, 2015.

[21] W.-C. W. Li & P. Solé. « Spectra of regular graphs and hypergraphs and orthogonal polynomials ». *European J. Combin.*, 17(5):461–477, 1996.

[22] R. Roland Bauerschmidt, J. Huang & H.-T. Yau. « Local Kesten–McKay law for random regular graphs ». arXiv:1609.09052, 2016.

[23] S. Sodin. « Random matrices, nonbacktracking walks, and orthogonal polynomials ». *J. Math. Phys.*, 48(12):123503, 21 p., 2007.

[24] N. C. Wormald. « Models of random regular graphs ». In *Surveys in combinatorics, 1999 (Canterbury)*, volume 267 of *London Math. Soc. Lecture Note Ser.*, pages 239–298. Cambridge Univ. Press, Cambridge, 1999.

Charles BORDENAVE

CNRS
Université Aix-Marseille
Institut de Mathématiques de Marseille
39 rue Frédéric Joliot Curie
13453 Marseille - France
E-mail : charles.bordenave@univ-amu.fr