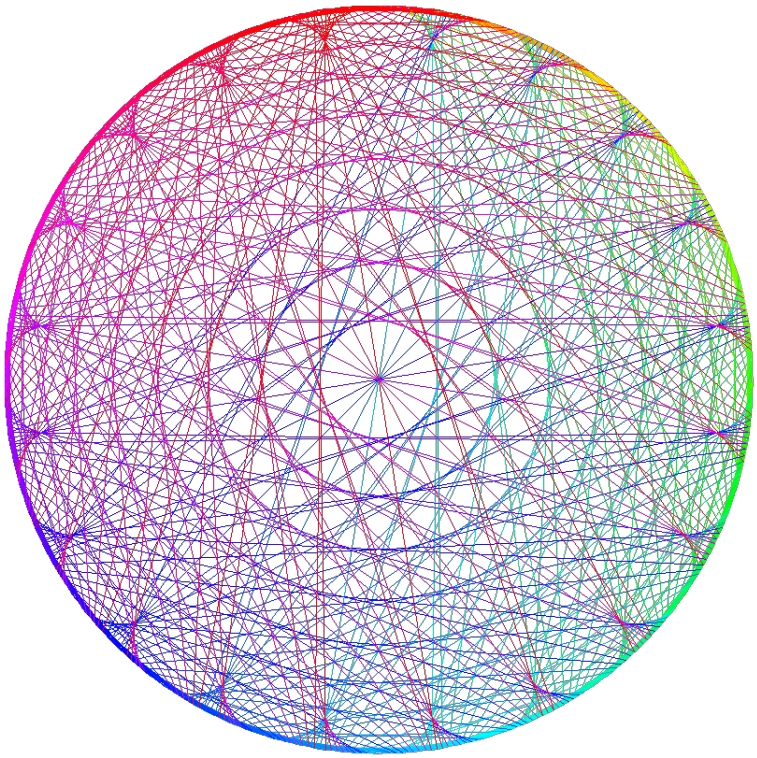
ÉCOLE POLYTECHNIQUE
UNIVERSITÉ PARIS-SACLAY

Projet Scientifique Collectif - Rapport final

Graphes expanseurs



DE LA MORINERIE Mathilde
GU Chenlin
HANGUIR Oussama
PAPIN Chloé
NGUYEN Manh Tien
VO Van Huy

Tuteur : FAVRE Charles
Coordinateur : BRUGALLÉ Erwan

Table des matières

Chapitre 1

Introduction

Au début du tronc commun, le professeur Kowalski a donné une conférence de mathématiques sur les graphes expanseurs. Ce type de graphes, très bien reliés et peu denses, est utilisé de façon surprenante dans de nombreux domaines de mathématiques et intervient naturellement en informatique théorique et en théorie de l'information. Un graphe est une notion abstraite de nature combinatoire qui permet d'étudier de nombreux modèles comme les réseaux de transport ou les comportements sociaux. Supposons que l'on ait besoin de construire un réseau de trains ; on modélise les stations par des sommets et les voies ferrées reliant les gares par des arêtes. La question de la robustesse du réseau se pose alors. Par exemple, si une station est fermée, on souhaite que le réseau ne soit pas paralysé. En outre, s'il y a trop de connexions au niveau d'une station, celle-ci risque d'être encombrée. La meilleure situation est alors que chaque station ait un petit nombre de liens, tout en gardant un réseau efficace. Dès lors, peut-on toujours construire des réseaux qui satisfont toutes ces propriétés ? Les graphes expanseurs sont un objet mathématique qui permet de construire des familles de graphes ayant ces deux propriétés.

Notre groupe s'est constitué autour d'une passion commune pour les mathématiques, et l'envie de monter un projet sur les graphes expanseurs. Nous formons un groupe diversifié, avec des élèves étrangers et français, issus de classes préparatoires ou de l'université, répartis dans plusieurs sections sportives, ce qui nous a permis d'adopter des approches différentes et complémentaires.

En choisissant les graphes expanseurs comme sujet de recherche et en travaillant en groupe sur un projet scientifique commun, nous avions identifié des objectifs à atteindre tout autant sur le plan scientifique que sur le plan des méthodes de travail. Tout d'abord, concernant les méthodes de travail, nous voulions nous enrichir de cette première expérience de travail scientifique en équipe. Nous avons appris à travailler ensemble dans un groupe hétérogène constitué de personnes aux origines et aux formations différentes. Notre objectif était d'apprendre à nous répartir équitablement le travail selon nos compétences propres, à nous organiser collectivement pour nous rencontrer et mettre en commun nos connaissances, afin de nous aider mutuellement à progresser. Ces compétences relationnelles nous serviront certainement dans notre vie professionnelle. Quant aux aspects scientifiques, nous voulions comprendre en détail plusieurs spécificités des graphes expanseurs, notion liée à de nombreux domaines des mathématiques fondamentales très différents, mais aussi à l'informatique théorique.

Concrètement nous avons suivi trois grands axes d'étude : la compréhension approfondie de la structure des graphes expanseurs et de leurs propriétés, la construction explicite de familles de graphes expanseurs par différentes approches, et leurs applications dans d'autres disciplines, en particulier l'informatique théorique. Les résultats auxquels nous sommes parvenus sont présentés dans la suite de ce rapport. Enfin nous conclurons par un retour sur notre expérience de projet scientifique collectif en analysant les points forts et les points perfectibles de notre travail.

Chapitre 2

Propriétés fondamentales et spectre des graphes expanseurs

2.1 Introduction

Dans cette partie nous allons d'abord présenter les graphes expanseurs qui sont des graphes bien connectés et pas trop denses. Ces graphes peuvent être caractérisés de deux façons différentes au travers de la constante de Cheeger et de la matrice d'adjacence, ces deux définitions étant équivalentes par l'inégalité de Cheeger. On étudiera ensuite les limites de la performance d'un graphe expanseur.

2.2 Caractérisation d'un graphe expanseur et constante de Cheeger

2.2.1 Rappels des définitions de base

Commençons par rappeler quelques définitions qui seront essentielles pour la suite de notre étude.

Définition 2.2.1. Graphe
Un **graphe** est un ensemble de points nommés **sommets** reliés par des **arêtes**.
Formellement ,un graphe $G = (V, E)$ est défini par la donnée de deux ensembles V et E tels que $E \subseteq V \times V$. Les éléments de V sont appelés sommets, tant que ceux de E sont nommés arêtes.
Les arêtes de la forme (v, v) sont appelées **boucles**. Si le graphe est non orienté et ne contient pas de boucle, on l'appelle **graphe simple**. Si $(v, w) \in E$, on dit que les sommets v et w sont **adjacents**, et que l'arête (v, w) est **incidente** à v et w.

Définition 2.2.2. Degré et régularité
L'**ordre** d'un graphe est le nombre de sommets de ce graphe.
Le **degré** d'un sommet (ou la valence d'un sommet) v est le nombre d'arêtes dans G incidents

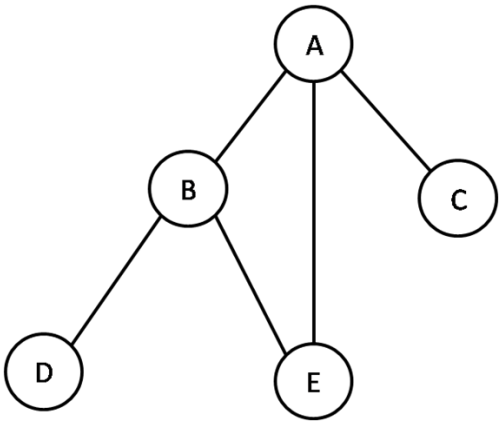


FIGURE 2.1 – Exemple d'un graphe non orienté

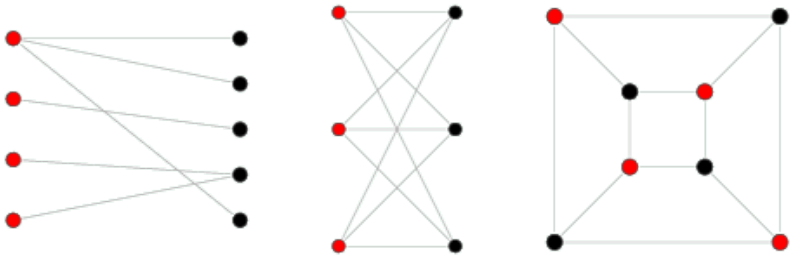


FIGURE 2.2 – Exemples de graphes bipartis

à v.
Si tous les sommets d'un graphe G ont le même degré k, on dit que G est k-régulier. Un graphe est **régulier** s'il est k-régulier pour un certain entier $k > 0$.

Définition 2.2.3. Connexité
Un graphe est **connexe** si on peut toujours relier deux sommets quelconques par un chemin (éventuellement réduit à une arête). Soit $G = (V, E)$ un graphe et $A \subset V$ un sous-ensemble de sommets. La **frontière** de A, notée ∂A, est l'ensemble des arêtes reliant un sommet de A à un sommet du complémentaire de A.

Définition 2.2.4. Graphe biparti
Un graphe est dit **biparti** s'il existe une partition de son ensemble de sommets en deux sous-ensembles L et R telle que chaque arête ait une extrémité dans L et l'autre dans R.

2.2.2 Une première caractérisation : la constante de Cheeger

Introduisons tout d'abord la constante de Cheeger qui traduit la robustesse d'un graphe, afin de donner une première définition des graphes expanseurs.

Définition 2.2.5. Constante de Cheeger
Dans un graphe $G(V, E)$, on définit la constante de Cheeger par :

$$h(G) = \inf_{F \subset E, |F| \leq \frac{1}{2}|E|} \frac{|\partial F|}{|F|}$$

où $|\partial F|$ est le cardinal de la frontière de F.

Intuitivement, la constante de Cheeger mesure la connexité et la robustesse du graphe. Il est clair que la constante de Cheeger est non nulle si et seulement si le graphe est connexe. La définition des graphes expanseurs généralise cette idée et mesure la connexité d'une famille de graphes.

Définition 2.2.6. Une famille de graphes ϵ-expanseurs
Une famille de graphes $\{G_n = (V_n, E_n)\}_n$ est une famille de graphes $\epsilon-$ expanseurs si elle satisfait les conditions suivantes :
(1) G_n est un graphe simple (pas de boucles et pas d'arêtes multiples) k-régulier
(2) $|V_n| \to \infty$
(3) $\exists \epsilon > 0$, tel que $\forall n : h(G_n) > \epsilon$

La constante de Cheeger induit une forte connexité du graphe, et la $k-$régularité impose que le graphe soit peu dense.

2.2.3 Une seconde caractérisation : la matrice d'adjacence

La constante de Cheeger est difficile à calculer en pratique, c'est pourquoi on introduit une seconde définition, plus facile à manipuler, au travers de la matrice d adjacence.

Définition 2.2.7. Matrice d'adjacence
On considère un graphe $G = (V, E)$ k-régulier à n sommets. On définit la matrice d'adjacence comme $(A)_{ij} = \mathbb{I}_{v_i \sim v_j}$. C'est-à-dire que A_{ij} vaut 1 si le i-ème sommet est relié au j-ème, et vaut 0 sinon.

C'est une matrice symétrique d'ordre n, elle a donc n valeurs propres réelles. On présente ici les liens entre ces valeurs propres et les propriétés du graphe.

Proposition 1.
(1) La matrice d'adjacence possède n valeurs propres réelles, on les note $k = \lambda_0 \geq \lambda_1 \geq \lambda_2 \cdots \geq \lambda_{n-1} \geq -k$
(2) $\lambda_0 > \lambda_1$ si et seulement si le graphe est connexe
(3) $\lambda_{n-1} = -k$ si et seulement si le graphe est biparti.

Démonstration.
(1) On cherche une base orthogonale $e_0, e_1, e_2, e_3 \cdots e_{n-1}$ telle que $Ae_i = \lambda_i e_i$. Soit $f : V \to \mathbb{R}$ une fonction sur les sommets, on peut considérer f comme un vecteur ${}^t(f(v_1), f(v_2), \cdots f(v_n))$. On définit le produit scalaire suivant :

$$\langle f, g \rangle_{l^2(V)} = \sum_{v \in V} f(v) g(v)$$

On établit que :

$$\left|\langle Af, f \rangle_{l^2(V)}\right| = \left| \sum_v \sum_{w \sim v} f(w) f(v) \right| \leq \frac{1}{2} \sum_v \sum_{w \sim v} \left(f^2(w) + f^2(v) \right) = k \sum_v f^2(v) \qquad (2.1)$$

Cela implique que $|\lambda_i| \leq k$ en prenant $f = e_i$. De plus, on sait que la fonction constante **1** est le vecteur propre associé à $\lambda_0 = k$.

(2) Supposons que le graphe est non-connexe. On peut décomposer sa matrice d'adjacence en deux blocs, qui correspondent respectivement à une composants connexe V_1 et au reste du graphe V_2. Les deux vecteurs $\mathbb{I}_{V_1}$ et $\mathbb{I}_{V_2}$ sont deux vecteurs propres orthogonaux associés à k, donc $\lambda_0 = \lambda_1$.
Réciproquement, on suppose que $\lambda_0 = \lambda_1$, on peut alors trouver deux vecteurs propres orthogonaux associés à k car A est symétrique. D'ailleurs, l'inégalité (2.1) implique que les éléments dans ces vecteurs valent 0 ou 1, ce qui permet de préciser les composantes connexes. Donc, le graphe n'est pas connexe.

(3) Pour le cas où $\lambda_{n-1} = -k$, le vecteur propre e_{n-1} vérifie $e_{n-1}(v) = -e_{n-1}(w)$ pour tout v relié à w d'après l'inégalité (2.1). On note V^+, V^- deux ensembles qu'on initialise par $V^+ = V^- = \emptyset$. On choisit arbitrairement un sommet v qu'on met dans V^+ et on met ses voisins dans V^-. Après on met les sommets restants w dans V^+ si $e_{n-1}(w) = e_{n-1}(v)$ et dans V^- si $e_{n-1}(w) = -e_{n-1}(v)$. En itérant ce processus on obtient $V = V^+ \sqcup V^-$ sans avoir de connexions entre V^+ et V^-.

Réciproquement, soit $G(V, E)$ un graphe biparti tel que $V = V^+ \cup V^-$ et tel qu'il n'y a pas d'arête reliant un point de V^+ et un autre de V^-. $\mathbb{I}_{V^+} - \mathbb{I}_{V^-}$ est un vecteur propre associée à la valeur $-k$. □

Nous avons maintenant les éléments pour indroduire la seconde définition d'une famille de graphes expanseurs.

Définition 2.2.8. Une famille de $\epsilon-$graphes expanseurs Une famille de graphes $\{G_n = (V_n, E_n)\}_n$ est une famille de graphes $\epsilon-$ expanseurs si elle satisfait les conditions suivantes :
(1) G_n est un graphe simple k-régulier
(2) $|V_n| \to \infty$
(3) $\exists \epsilon > 0$, tel que $\lambda_0(G_n) - \lambda_1(G_n) > k\epsilon$ uniformément pour tous les graphes G_n.

En réalité ces deux définitions sont équivalentes, comme nous allons le montrer dans ce qui suit.

Définition 2.2.9. Trou spectral
On appelle trou spectral la différence entre les deux plus grandes valeurs propres.

2.2.4 Lien entre les deux caractérisations

Dans cette sous-section, on présente un théorème qui décrit un graphe expanseur à l'aide de sa matrice d'adjacence, afin de prouver l'équivalence entre les deux définitions.

Théorème 2.2.1. ***L'inégalité de Cheeger***
Soit un graphe $G = (V, E)$ et A sa matrice d'adjacence. On note $k = \lambda_0 \geq \lambda_1 = k(1 - \epsilon) \geq \lambda_2 \cdots \geq \lambda_{n-1} \geq -k$ les valeurs propres de A. On a alors :

$$\frac{\epsilon}{2}k \leq h(G) \leq \sqrt{2\epsilon k}$$

Cette inégalité nous permet de voir directement que les deux définitions 2.2.6 et 2.2.4 sont équivalentes. Nous allons démontrer l'inégalité de gauche, la démonstration de celle de droite étant laissée en annexe.

Démonstration. D'abord, on montre que $\frac{\epsilon}{2}k \leq h(G)$. L'idée générale est de décomposer un vecteur dans selon la droite engendrée par vecteur $\mathbf{1}$ et selon son orthogonal.

$$\begin{aligned}
|\partial F| &= \langle A\mathbb{I}_F, \mathbb{I}_{F^c}\rangle_{l^2(V)} = \langle A\mathbb{I}_F, 1-\mathbb{I}_F\rangle_{l^2(V)} = \langle \mathbb{I}_F, A1\rangle_{l^2(V)} - \langle A\mathbb{I}_F, \mathbb{I}_F\rangle_{l^2(V)} \\
&= k\,|F| - \langle A(\mathbb{I}_F - \frac{|F|}{|E|}1 + \frac{|F|}{|E|}1), (\mathbb{I}_F - \frac{|F|}{|E|}1 + \frac{|F|}{|E|}1)\rangle_{l^2(V)} \\
&\geq k\,|F| - k(1-\epsilon)\langle(\mathbb{I}_F - \frac{|F|}{|E|}1), (\mathbb{I}_F - \frac{|F|}{|E|}1)\rangle_{l^2(V)} - \frac{|F|^2}{|E|}k \\
&= k\epsilon|F|(1-\frac{|F|}{|E|}) \geq \frac{1}{2}k\epsilon|F|
\end{aligned}$$

□

2.3 Limites de la performance du graphe expanseur

La constante de Cheeger et le trou spectral, c'est-à-dire la différence entre les deux plus grandes valeurs propres $\lambda_0 - \lambda_1$, caractérisent toutes deux la performance d'un graphe expanseur. Il est donc intéressant d'étudier les limites de cette performance dans une famille de graphes $k-$régulier en cherchant à borner le trou spectral.

2.3.1 Le théorème d'Alon et Boppana

Un premier théorème d'Alon et Boppana montre que le trou spectral ne peut pas être très grand.

Théorème 2.3.1 (Alon et Boppana). *Soit $\{G_n = (V_n, E_n)\}$ une famille de graphes connexes, k-réguliers et tels que $|V_n|$ tend vers infini. Alors les deuxièmes valeurs propres $\lambda_1(G_n)$ vérifient :*

$$\liminf_{n\to+\infty} \lambda_1(G_n) \geq 2\sqrt{k-1}$$

Le trou spectral a tendance à diminuer selon la formule :

$$\liminf_{n\to+\infty} trou(G_n) = \liminf_{n\to+\infty} \lambda_0(G_0) - \lambda_1(G_n) \leq k - 2\sqrt{k-1}$$

On va en fait démontrer un résultat plus fort affirmant qu'il y a une probabilité positive que $\lambda_1(G_n)$ appartienne à $[(2-\varepsilon)\sqrt{k-1}, k]$ pour $\epsilon > 0$.

Théorème 2.3.2. *Étant donné un $\varepsilon > 0$, il existe une constante C qui dépend de ε et k telle que pour chaque graphe G connexe, k-régulier et de n sommets, il y a au moins Cn valeurs propres dans $[(2-\varepsilon)\sqrt{k-1}, k]$.*

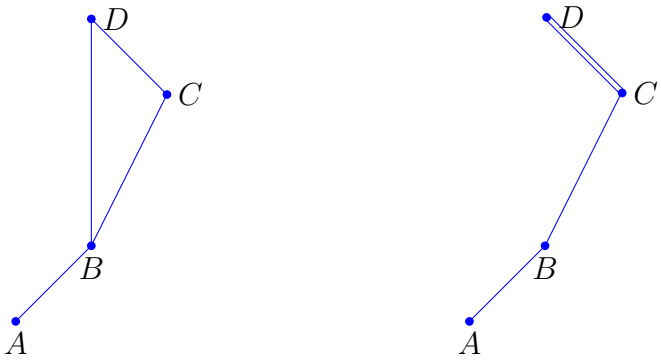


FIGURE 2.3 – Le chemin $A \to B \to C \to D$ à gauche est sans-retour, alors que $A \to B \to C \to D \to C$ (celui à droit) est interdit.

2.3.2 Démonstration du théorème

Présentons d'abord l'idée générale de la démonstration : l'idée, due à J.-P. Serre, est de trouver la fonction génératrice de la matrice d'adjacence A par des techniques combinatoires, pour arriver à l'égalité suivante :

$$\sum_{r \leq \frac{m}{2}} A_{m-2r} = (k-1)^{\frac{m}{2}} U_m\left(\frac{A}{2\sqrt{k-1}}\right)$$

où $(A_r)_{ij}$ est le nombre de chemins de longueur r partant du i-ème sommet au j-ème et U_m les polynômes de Chebyshev de deuxième espèce. En exploitant cette inégalité, on montre que le spectre des graphes G_n, vu comme une somme de distributions de Dirac réagissant sur U_m, est positif. On arrive au résultat attendu après une série d'applications du théorème de Banach-Alaoglu et d'autres outils d'analyse fonctionnelle.

Donnons d'abord quelques définitions :

Définition 2.3.1. Un chemin sans-retour sur un graphe non-orienté G est une suite de sommets $(x_0, x_1, \ldots, x_r)$ telle que $x_{i+1} \neq x_{i-1}\ \forall i = 2, \ldots, r-1$.

Soit A_k une matrice telle que $A_k(i,j)$ soit égal au nombre de chemins sans-retour de longueur k du sommet i au sommet j. On établit la relation suivante affirmant que A_k sont en fait des polynômes en A, la matrice d'adjacence. On pose $A_0 = I$.

Proposition 2. $A_0 = I, A_1 = A, A_2 = A^2 - kI, A_{r+1} = A_r A_1 - (k-1)A_{r-1}$.

En effet, il y a exactement k chemins avec retour de longueur 2 qui partent et finissent au même sommet et il n'y en a aucun entre deux sommets distincts. Pour construire un chemin sans-retour de longueur $r+1$ de i à j, il faut construire un chemin sans-retour de longueur r de i à h puis connecter h à j. Pour construire un chemin avec retour entre i et j, il faut que l'on parte de i à j en $r-1$ pas, puis suivre le chemin j, h, j, il y a donc $(k-1)A_{r-1}$ chemins avec retour entre i et j.

Lemme 1. *La fonction génératrice de la suite A_r est*

$$\sum_{r=0}^{\infty} A_r t^r = \frac{1-t^2}{1-At+(k-1)t^2}$$

Observons que cette fonction ressemble à la fonction génératrice des polynômes de Chebyshev de deuxième espèce. On rappelle la définition de polynômes de Chebyshev.

Définition 2.3.2. Les polynômes de Chebyshev de deuxième espèce s'obtiennent en réécrivant $\frac{\sin(m+1)x}{\sin x}$ en fonction de $\cos x$ avec $m \in \mathbb{N}$ i.e

$$\frac{\sin(m+1)x}{\sin x} = U_m(\cos x)$$

Par exemple $U_0 = 1, U_1 = 2x, U_2 = 4x^2 - 1$. On a de plus la relation récurrente

$$U_{m+1} = 2xU_m - U_{m-1}$$

À partir de la relation de récurrence, on peut aussi exprimer la fonction génératrice .Celle-ci a une forme similaire à celle de A_m.

$$\sum_{m=0}^{\infty} U_m(x)t^m = \frac{1}{1-2xt+t^2}$$

Si on veut s'y ramener, il suffit de poser $T_m = \sum_{0\leq r\leq \frac{m}{2}} A_{m-2r}$ et on a

$$T_m = \sum_{0\leq r\leq \frac{m}{2}} A_{m-2r} = (k-1)^{\frac{m}{2}} U_m(\frac{A}{2\sqrt{k-1}}) \tag{2.2}$$

En prenant la trace du terme à gauche, on trouve un nombre positif qui est le nombre de chemins sans-retour. On obtient alors l'inégalité suivante :

Proposition 3. *Avec λ_i les valeurs propres de la matrice d'adjacence,*

$$\sum_i U_m(\frac{\lambda_i}{2\sqrt{k-1}}) \geq 0 \tag{2.3}$$

Pour les calculs qui vont suivre, posons

$$X_m(x) = U_m(x/2)$$

cela implique que :

$$X_m(2\cos\theta) = \frac{\sin(m+1)\theta}{\sin\theta}, \qquad X_{m+1} = xX_m - X_{m-1}$$

Après avoir établi l'inégalité 2.3, on applique un théorème d'analyse fonctionnelle pour conclure. On présentera ce théorème dans l'annexe.

Proposition 4. *Pour chaque mesure ν à support dans $[-L, L]$ satisfaisant $\int_{-L}^{L} X_m(x)d\nu(x) \geq 0$ pour tout $m = 1, 2, \ldots$, on a*

$$\nu([2-\varepsilon, L]) > 0$$

Théorème 2.3.3. *Soient $L \geq 2$ et $\varepsilon > 0$, il existe $C(\varepsilon, L) > 0$ tel que étant donné une mesure ν sur $[-L, L]$ satisfaisant $\int_{-L}^{L} U_m(x/2)d\nu(x) \geq 0$ pour tout $m = 1, 2, \ldots$, on a*

$$\nu([2-\varepsilon, L]) \geq C$$

De plus, une exploitation plus fine de (2.2) nous donne la convergence faible de la suite de mesures obtenues des spectres d'une famille de graphes.

Soit $\{G_n\}$ une famille de graphes k-réguliers connexes telle que la longueur des circuits les plus courts tende vers infini, et ν_n la mesure de probabilité correspondante au spectre de G_n divisé par $\sqrt{k-1}$ (c-à-d. $\nu_n = \frac{1}{|G_n|}\sum_{i=0}^{|G_n|-1} \delta_{\frac{\lambda_j(G_n)}{\sqrt{k-1}}}$). On a :

Théorème 2.3.4. *ν_n converge faiblement vers ν définie par $d\nu = \frac{\sqrt{4-x^2}}{2\pi}dx$*

Ceci donne un résultat similaire au théorème 2.3.2.

Corollaire 2.3.1. *Quelque soit $\varepsilon > 0$, il existe une constante $C(\varepsilon) > 0$ tel que pour toute famille $\{G_n\}$ vérifiant les conditions de (2.2.6), le nombre des valeurs propres de G_n qui sont dans $[-k, -(2-\varepsilon)\sqrt{k-1}]$ est au moins $C|G_n|$.*

La démonstration donne donc une limite supérieure du trou spectral.

Chapitre 3

Constructions des graphes expanseurs

3.1 Introduction

Dans cette partie, nous introduirons d'abord les graphes de Cayley et de Schreier, avant d'étudier la construction de Margulis, qui permet de trouver une famille explicite de graphes expanseurs. Cette construction repose sur la propriété (T) des groupes. Nous introduirons cette propriété grâce à la notion de constante de Kazhdan, puis nous procéderons à la démonstration de la construction. Enfin, nous présenterons deux autres constructions, l'une basée sur les graphe de Ramanujan, et l'autre reposant sur les graphes aléatoires.

3.2 Graphe de Cayley, graphe de Schreier

Avant de procéder à la construction de familles de graphes expanseurs, nous présenterons deux manières d'associer un graphe à un groupe. Le graphe de Cayley est celui que nous utiliserons le plus par la suite.

On notera xy le produit de deux éléments x et y d'un groupe.

Définition 3.2.1. Soit G un groupe et S une partie finie de G. On suppose que S est symétrique et ne contient pas l'élément neutre. On appelle *graphe de Cayley*, noté $Cay(G, S)$, le graphe dont :

- les sommets sont les éléments de G
- les arêtes sont les éléments de $\{(x, sx), x \in G, s \in S\}$

On remarque que $Cay(G, S)$ est $|S|$-régulier.

Remarque 1. *$Cay(G, S)$ est connexe si et seulement si S est une partie génératrice.*

La deuxième manière de construire un graphe est similaire, mais sera nécessaire pour la construction alternative. Dans cette construction, le groupe agit non pas sur lui-même mais sur un espace X. On notera sx l'action de l'élément s du groupe sur un élément x de l'espace.

Définition 3.2.2. Soit G un groupe et X un espace sur lequel l'action de G est définie. Soit S une partie symétrique finie de G telle que $sx \neq x$ pour tous $s \in S$ et $x \in X$, et $sx \neq s'x$ pour tous $s, s' \in S$ distincts et $x \in X$. On appelle *graphe de Schreier*, noté $Sch(G, S)$, le graphe dont :

- les sommets sont les éléments de X
- les arêtes sont les éléments de $\{(x, sx), x \in X, s \in S\}$

De nouveau, on obtient un graphe $|S|$-régulier.

3.3 La construction de Margulis

3.3.1 La propriété (T) de Kazhdan

Dans cette partie, on introduit la propriété (T) de Kazhdan qui est utilisée dans la construction de Margulis. Nous aurons besoin de la notion de représentation d'un groupe, que nous définissons ci-dessous.

3.3.1.1 Définitions de base

Commençons par définir la notion de représentation de groupe.

Définition 3.3.1. Représentation unitaire
Soit G un groupe localement compact, compactement engendré (c'est-à-dire qu'il est engendré par une partie compacte), à base dénombrable de voisinages. Une représentation unitaire de G est la donnée d'un espace de Hilbert séparable H et d'un homomorphisme continu $\rho : G \to U(H)$ où $U(H)$ est le groupe des transformations unitaires sur H. Par abus de notation, on considère souvent ρ comme la représentation de G (au lieu de $(\rho, U(H))$).

Définition 3.3.2. Représentation triviale
Si $\forall g \in G\ \rho(g) = Id_H$ avec H un espace de Hilbert, alors ρ est une représentation triviale de G.

Définition 3.3.3. Représentation régulière
Si G est un groupe discret, la représentation régulière de G est $\rho : G \to U(l^2(G))$ où $l^2(G)$ est l'espace des fonctions de carré sommable sur G et l'action de ρ sur G est donnée par :

$$\rho(g)f(x) = f(g^{-1}x)\ \ \forall (g, x) \in G^2$$

L'action d'un groupe G sur un espace X est dite *transitive* si pour tout $x \in X$ on a $\{gx,\ g \in G\} = X$.

Définition 3.3.4. Représentation quasi-régulière
Si (X, μ) est un espace mesuré (i.e. sur lequel il existe une mesure) sur lequel G agit de manière transitive en préservant la mesure, alors la représentation quasi-régulière $\pi_X : G \to U(L^2(X, \mu))$ est la représentation dont l'action sur l'espace de Hilbert $L^2(X, \mu)$ est donnée par la formule

$$\pi_X(g)f(x) = f(g^{-1}x)$$

Définition 3.3.5. Vecteur invariant
On dit que v est un vecteur G-invariant de la représentation $\rho : G \to U(H)$ si $\rho(g)v = v$ pour tout $g \in G$.

Définition 3.3.6. Vecteurs quasi-invariants
On dit qu'une représentation $\rho : G \to U(H)$ a des vecteurs quasi-invariants s'il existe une suite de vecteurs $v_n \in H$ telle que $\lim_{n\to\infty} \|\rho(g)v_n - v_n\| = 0$ pour tout $g \in G$.

Définition 3.3.7. Somme directe
Soient $\rho_1 : G \to U(H_1)$ et $\rho_2 : G \to U(H_2)$ deux représentations unitaires, leur somme directe $\rho_1 \oplus \rho_2 : G \to U(H_1 \oplus H_2)$ est aussi une représentation où $H_1 \oplus H_2$ est l'espace de Hilbert et la représentation est donnée par la formule

$$(\rho_1 \oplus \rho_2)(g)(v_1 \oplus v_2) = (\rho_1(g)v_1) \oplus (\rho_2(g)v_2)$$

La somme directe d'une suite dénombrable de représentations unitaires est définie de manière similaire.

Définition 3.3.8. La constante de Kazhdan et la propriété (T)
Soit $\rho : G \to U(H)$ une représentation unitaire d'un groupe localement compact G et soit S un sous-ensemble compact de G. La constante de Kazhdan $Kaz(G, S, \rho)$ est définie comme suit :

$$Kaz(G, S, \rho) = \inf_{v\in H\setminus\{0\}} \sup_{s\in S} \frac{\|\rho(s)v - v\|_H}{\|v\|_H}$$

Alors $Kaz(G, S, \rho)$ s'annule si ρ a un vecteur invariant non nul. La constante de Kazhdan $Kaz(G, S)$ est définie :

$$Kaz(G, S) = \inf_{\rho} Kaz(G, S, \rho)$$

pour tout ρ représentation unitaire de G n'ayant pas de vecteur invariant non nul.
On dit qu'un groupe G a la propriété (T) s'il existe un sous-ensemble compact S de G tel que $Kaz(G, S) > 0$.

3.3.1.2 Propriétés

Nous présentons certaines propriétés utiles de la constante de Kazhdan.

Théorème 3.3.1. *Soit G un groupe localement compact, soit $\rho : G \to U(H)$ une représentation, et soient S et S' deux sous-ensembles compacts de G. Alors :*

(i) Si $S \subset S'$, alors $Kaz(G, S, \rho) \leq Kaz(G, S', \rho)$ et $Kaz(G, S) \leq Kaz(G, S')$

(ii) $Kaz(G, S, \rho) = Kaz(G, S^{-1}, \rho) = Kaz(G, S \cup S^{-1}, \rho)$ et
$Kaz(G, S) = Kaz(G, S^{-1}) = Kaz(G, S \cup S^{-1})$

(iii) $Kaz(G, S, \rho) = Kaz(G, S \cup \{1\}, \rho)$ et $Kaz(G, S) = Kaz(G, S \cup \{1\})$

(iv) $Kaz(G, S^m, \rho) \leq mKaz(G, S, \rho)$ *et* $Kaz(G, S^m) \leq mKaz(G, S)$ *pour tout* $m \geq 1$.

(v) Si S *génère* G*, alors* G *a la propriété (T) si et seulement si* $Kaz(G, S) > 0$.

Démonstration. i) et (iii) sont triviaux. Comme pour tout $s \in S$ on a :

$$\|\rho(s)v - v\| = \|\rho(s)(v - \rho(s^{-1})v)\| = \|\rho(s^{-1})v - v\|$$

On en déduit :

$$\sup_{s\in S} \frac{\|\rho(s)v - v\|_H}{\|v\|_H} = \sup_{s\in S^{-1}} \frac{\|\rho(s)v - v\|_H}{\|v\|_H} = \sup_{s\in S\cup S^{-1}} \frac{\|\rho(s)v - v\|_H}{\|v\|_H}$$

pour tout $v \in H \setminus \{0\}$. Alors

$$Kaz(G, S, \rho) = Kaz(G, S^{-1}, \rho) = Kaz(G, S \cup S^{-1}, \rho)$$

et aussi

$$Kaz(G, S) = Kaz(G, S^{-1}) = Kaz(G, S \cup S^{-1})$$

On a donc démontré (ii). Pour démontrer (iv), on observe que tout élément s de S^m peut s'écrire sous forme $s = s_1 s_2 ... s_k$ où $k \leq m$ et $s_i \in S \ \ \forall i = \overline{1, k}$ et pour tout $v \in H \setminus \{0\}$:

$$\begin{aligned}
\frac{\|\rho(s)v - v\|_H}{\|v\|_H} &= \frac{\|\rho(s_1 s_2 ... s_k)v - v\|_H}{\|v\|_H} \\
&= \frac{\|\rho(s_1 s_2 ... s_k)v - \rho(s_1 s_2 ... s_{k-1})v + \rho(s_1 s_2 ... s_{k-1})v - ... + \rho(s_1) - v\|_H}{\|v\|_H} \\
&\leq \frac{\|\rho(s_1 s_2 ... s_k)v - \rho(s_1 s_2 ... s_{k-1})v\|_H}{\|v\|_H} + ... + \frac{\|\rho(s_1)v - v\|_H}{\|v\|_H} \\
&= \frac{\|\rho(s_1 s_2 ... s_{k-1})(\rho(s_k)v - v)\|_H}{\|v\|_H} + ... + \frac{\|\rho(s_1)v - v\|_H}{\|v\|_H} \\
&= \sum_{i=1}^{k} \frac{\|\rho(s_i)v - v\|_H}{\|v\|_H} \\
&\leq k \sup_{s\in S} \frac{\|\rho(s)v - v\|_H}{\|v\|_H} \\
&\leq m \sup_{s\in S} \frac{\|\rho(s)v - v\|_H}{\|v\|_H}
\end{aligned}$$

D'où :

$$\begin{aligned}
Kaz(G, S^m, \rho) &= \inf_{v\in H\setminus\{0\}} \sup_{s\in S^m} \frac{\|\rho(s)v - v\|_H}{\|v\|_H} \leq \inf_{v\in H\setminus\{0\}} m \sup_{s\in S} \frac{\|\rho(s)v - v\|_H}{\|v\|_H} \\
&= m \inf_{v\in H\setminus\{0\}} \sup_{s\in S} \frac{\|\rho(s)v - v\|_H}{\|v\|_H} = mKaz(G, S, \rho)
\end{aligned}$$

L'inégalité $Kaz(G, S^m) \leq mKaz(G, S)$ est déduite directement de ce résultat.
Maintenant, on va démontrer (v). Comme G a la propriété (T), il existe un sous-ensemble

compact S_1 de G tel que $Kaz(G, S_1) > 0$. Posons $S' = S \cup S^{-1} \cup \{1\}$, alors S' est encore un générateur de G. Comme G est un groupe localement compact, compactement engendré, à base dénombrable de voisinages, il existe un entier m tel que S'^m contient S_1 (Ce résultat est démontré en annexe). Alors d'après (i), (ii), (iii) et (iv) on a

$$Kaz(G, S_1) \leq Kaz(G, S'^m) \leq mKaz(G, S') = mKaz(G, S)$$

D'où $Kaz(G, S) > 0$. □

Théorème 3.3.2. *Soit G un groupe localement compact et S un sous-ensemble compact de G. Alors si $Kaz(G, S) = 0$, il existe une représentation unitaire $\rho : G \to U(H)$ n'ayant pas de vecteur invariant non nul tel que $Kaz(G, S, \rho) = 0$.*

Démonstration. Comme $Kaz(G, S) = \inf_\rho Kaz(G, S, \rho) = 0$, on peut trouver une suite de représentations unitaires $\rho_1, \rho_2, ...$ n'ayant pas de vecteur invariant non nul tel que

$$Kaz(G, S, \rho_i) < \frac{1}{i + \frac{1}{2}}$$

Considérons ensuite la représentation $\rho = \bigoplus_{i=1} \rho_i$ avec l'espace de Hilbert $H = \bigoplus_{i=1} H_i = \{v_1 \oplus v_2 \oplus ... \mid \sum_{i=1} \|v_i\|^2_{H_i} < \infty\}$. Puisque

$$Kaz(G, S, \rho_i) = \inf_{v \in H_i \setminus \{0\}} \sup_{s \in S} \frac{\|\rho_i(s)v - v\|_{H_i}}{\|v\|_{H_i}} = 0$$

on peut trouver un vecteur $v_i \in H_i \setminus \{0\}$ tel que $\|v_i\|_{H_i} = 1$ et

$$\sup_{s \in S} \|\rho_i(s)v - v\|_{H_i} = \sup_{s \in S} \frac{\|\rho_i(s)v - v\|_{H_i}}{\|v\|_{H_i}} < \frac{1}{i}$$

Puis posons $u_i = \bigoplus_{j=1}^i v_j \oplus 0 \in H$, alors $\|u_i\| = \sqrt{i}$. On a :

$$\begin{aligned}
\sup_{s \in S} \frac{\|\rho(s)u_i - u_i\|_H}{\|u_i\|_H} &= \sup_{s \in S} \frac{\|(\bigoplus_{k=1} \rho_k)(s)(\bigoplus_{j=1}^i v_j \oplus 0) - (\bigoplus_{j=1}^i v_j \oplus 0)\|_H}{\sqrt{i}} \\
&= \sup_{s \in S} \frac{\|\bigoplus_{j=1}^i (\rho_j)(s)v_j - \bigoplus_{j=1}^i v_j\|_H}{\sqrt{i}} \\
&= \sup_{s \in S} \frac{\|\bigoplus_{j=1}^i (\rho_j(s)v_j - v_j)\|_H}{\sqrt{i}} \\
&= \sup_{s \in S} \sqrt{\frac{\sum_{j=1}^i \|\rho_j(s)v_j - v_j\|^2_H}{i}} \\
&\leq \sqrt{\frac{\sum_{j=1}^i (\sup_{s \in S} \|\rho_j(s)v_j - v_j\|_H)^2}{i}} \\
&< \sqrt{\frac{\sum_{j=1}^i \left(\frac{1}{j}\right)^2}{i}} \\
&< \sqrt{\frac{\sum_{j=1} \left(\frac{1}{j}\right)^2}{i}}
\end{aligned}$$

On en déduit que :

$$\lim_{i\to\infty} \sup_{s\in S} \frac{\|\rho(s)u_i - u_i\|_H}{\|u_i\|_H} = 0$$

D'où

$$Kaz(G,S,\rho) = \inf_{v\in H\setminus\{0\}} \sup_{s\in S} \frac{\|\rho(s)v - v\|_H}{\|v\|_H} = 0$$

De plus, ρ n'a évidemment pas de vecteur invariant non nul. □

Le théorème ci-dessous nous donne une autre définition de la propriété (T).

Théorème 3.3.3. *Soit G un groupe localement compact tel que toute représentation ρ sur G qui a des vecteurs quasi-invariants possède un vecteur invariant non nul, alors G a la propriété (T).*

Démonstration. On raisonne par l'absurde. Supposons que G n'a pas la propriété (T), soit S un générateur compact de G, alors $Kaz(G,S) = 0$. D'après le théorème 2, il existe une représentation unitaire $\rho : G \to U(H)$ n'ayant pas de vecteur invariant non nul telle que $Kaz(G,S,\rho) = 0$. Par la définition de $Kaz(G,S,\rho)$, on peut trouver alors une suite $(v_n)_n$de vecteurs de module 1 dans H telle que

$$\lim_{n\to\infty} \sup_{s\in S} \|\rho(s)v_n - v_n\|_H = 0$$

Soit γ un élément de G. Puisque S génère G, γ peut s'écrire sous forme $\gamma = s_1 s_2 ... s_k$ où $k \in \mathbb{N}^*$ et $s_i \in S\ \forall i = \overline{1,k}$ et

$$\|\rho(\gamma)v_n - v_n\|_H = \|\rho(s_1 s_2 ... s_k)v_n - v_n\|_H \leq k \sup_{s\in S} \|\rho(s)v_n - v_n\|_H$$

Donc

$$\lim_{n\to\infty} \|\rho(\gamma)v_n - v_n\|_H \leq k \lim_{n\to\infty} \sup_{s\in S} \|\rho(s)v_n - v_n\|_H = 0$$

D'où $(v_n)_n$ est une suite des vecteurs quasi-invariants de ρ mais ρ n'a pas de vecteur invariant non nul. Cette contradiction nous donne le résultat voulu. □

Remarque : Nous verrons par la suite que cette définition est plus pratique pour démontrer qu'un groupe a la propriété (T), tandis que la définition se basant sur la constante de Kazhdan est plus utile pour voir le lien entre le constante de Kazhdan et le trou spectral d'un graphe de Cayley.

La proposition suivante est admise ici, la démonstration étant particulièrement ardue, mais elle est très importante.

Proposition 5. *Soit G un groupe localement compact, et soit Γ un réseau de G, alors Γ possède la propriété (T) si et seulement si G la possède.*

3.3.2 Un théorème central pour la construction

Le théorème suivant fournit une manière de construire une famille de graphes expanseurs à partir d'un groupe possédant la propriété (T).

Théorème 3.3.4. *Soit G un groupe discret finiment engendré et S un sous-ensemble symétrique de G qui génère G. Soit N_n une suite de sous-groupes distingués d'indice fini de G et soit $\pi_n : G \to G/N_n$ les surjections canoniques. Supposons que pour tout n suffisament grand, π_n est injective dans $S \cup \{1\}$. Alors si G a la propriété (T), la famille $Cay(G/N_n, \pi_n(S))$ pour n suffisament grand est une famille de graphes expanseurs.*

Démonstration. Pour démontrer ce théorème, nous avons besoin des deux lemmes suivants, dont la démonstration figure en annexe.

Lemme 2. *Lien entre la constante de Kazhdan et les graphes expanseurs*

Soit $Cay(G, S)$ un graphe de Cayley k-régulier, fini et soit $\epsilon > 0$. Si $Kaz(G, S) > \epsilon$, alors il existe $c > 0$ ne dépendant que de ϵ et k tel que $Cay(G, S)$ soit un graphe c-expanseur.

Ce lemme nous montre qu'il est possible d'utiliser la constante de Kazhdan pour travailler sur les graphes expanseurs. Cette constante est en effet plus pratique que la constante de Cheeger quand on travaille sur les groupes.

Lemme 3. *Soit G et G' deux groupes localement compacts tels qu'il existe un homomorphisme surjectif continu $\pi : G \to G'$. Soit S une partie compacte de G. Alors $Kaz(G', \pi(S)) \geq Kaz(G, S)$.*

Comme π_n est la projection canonique de G sur G/N_n, alors π_n est surjective et continue. Alors d'après le lemme (3), on a $Kaz(G/N_n, \pi_n(S)) \geq Kaz(G, S)$. Comme pour n suffisamment grand, π_n est injective dans $S \cup \{1\}$, alors $|\pi_n(S)| = |S|$ pour n suffisamment grand. En appliquant le lemme (2), on conclut que pour n suffisamment grand les graphes $Cay(G/N_n, \pi_n(S))$ sont des graphes c-expanseurs pour un c positif ne dépendant que de ϵ et k.

□

Conséquence : Ce théorème permet de construire la famille de graphes expanseurs de Margulis. On utilise pour cela le groupe $SL_d(\mathbb{Z})$, où $d \geq 3$, et la suite de quotients $SL_d(\mathbb{Z}/n\mathbb{Z})$.

3.3.3 La construction de Margulis

La construction de Margulis est donnée par le théorème suivant :

Théorème 3.3.5. *La famille de graphes $(Cay(SL_d(\mathbb{Z}/n\mathbb{Z})), \pi_n(S))$ pour $d \leq 3$ est une famille de graphes expanseurs.*

Si l'on remarque [1] que $SL_d(\mathbb{Z}/n\mathbb{Z})$ s'écrit comme un quotient de la forme $SL_d(\mathbb{Z})/N_n$, il suffit de montrer que $SL_d(\mathbb{Z})$ possède la propriété (T) pour $d \leq 3$: la section précédente permet alors de conclure.

Ici, nous traiterons seulement le cas $d = 3$, mais la démonstration peut se généraliser. Le raisonnement fait intervenir un certain nombre de résultats intermédiaires et s'appuie notamment sur la théorie de Fourier. Nous commencerons par montrer le théorème de Bochner pour établir un lien entre les espaces de Hilbert et les mesures de probabilité.

Théorème 3.3.6. *Théorème de Bochner : Soit $f : \mathbb{R}^d \to \mathbb{C}$ une fonction bornée, continue, définie semi-positive, i.e. $f(x) = \overline{f(-x)}$ pour tout $x \in \mathbb{R}^d$, et*

$$\int_{\mathbb{R}^d} \int_{\mathbb{R}^d} f(x-y)\, d\nu(x)\, d\overline{\nu}(y) \geq 0$$

pour toute mesure complexe finie ν. Alors, il existe une mesure non-négative finie μ sur $\mathbb{R}^d$ telle que f soit la transformée inverse de Fourier de μ, i.e.

$$f(x) = \int_{\mathbb{R}}^{d} e^{2\pi i x\cdot\xi} d\mu(\xi)$$

pour tout $x \in \mathbb{R}^d$.

Remarque 2. *Une démonstration partielle de ce théorème figure en annexe.*

À présent, soit $\rho : \mathbb{R}^d \to U(H)$ et $v \in H$. Nous construisons la fonction suivante, définie pour tout x dans $\mathbb{R}^d$ par

$$f_{v,v}(x) = \langle \rho(x)v, v\rangle_H$$

Il s'agit d'une fonction continue et bornée, vérifiant, par sesquilinéarité, $f_{v,v}(x) = \overline{f(-x)}$. En outre, utilisant la commutativité du groupe et le fait que $\rho(y)^{-1}$ et $\rho(y)$ soient adjoints puisque ρ est unitaire,

$$\begin{aligned}
\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} f_{v,v}(x-y)d\nu(x)d\overline{\nu}(y) &= \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \langle \rho(x-y)v, v\rangle_H d\nu(x)d\overline{\nu}(y) \\
&= \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \langle \rho(x)v, \rho(y)v\rangle_H d\nu(x)d\overline{\nu}(y) \\
&= \left\langle \int_{\mathbb{R}^d} \rho(x)v\, d\nu(x), \int_{\mathbb{R}^d} \rho(y)v\, d\nu(y) \right\rangle \\
&= \left\| \int_{\mathbb{R}^d} \rho(x)v\, d\nu(x) \right\|_H^2
\end{aligned}$$

Ceci garantit que $f_{v,v}$ est positive et semi-définie. Par application directe du théorème de Bochner, on obtient la transformée inverse de Fourier $\mu_{v,v}$, telle que

$$f_{v,v}(x) = \int_{\mathbb{R}}^{d} e^{2\pi i x\cdot\xi} d\mu_{v,v}(\xi)$$

1. Pour plus de détails voir en annexe.

Le calcul de $f_{v,v}(0)$ permet d'obtenir l'intégrale de $\mu_{v,v}$, qui vaut $\|v\|_H^2$. On peut alors construire une forme sesquilinéaire $\mu_{v,w}$ sur le principe de la polarisation, en posant

$$\mu_{v,w} := \frac{1}{4}(\mu_{v+w,v+w} - \mu_{v-w,v-w} + i\mu_{v+iw,v+iw} - i\mu_{v-iw,v-iw})$$

Un calcul simple permet de montrer que cette mesure est bien sesquilinéaire (cette propriété découle de la sesquilinéarité du produit scalaire hermitien).

Cette mesure nous servira par la suite.

Le lemme ci-dessous est un résultat qui sera utilisé dans la preuve de chacune des constructions présentées. La norme $\|\cdot\|_{TV}$ est la norme de la variation totale entre deux mesures.

Lemme 4. *Soit S un voisinage compact de l'identité dans $SL_2(\mathbb{R})$. Soit $\varepsilon > 0$. Enfin, soit μ une mesure de probabilité sur $\mathbb{R}^2$ vérifiant, pour tout $s \in S$,*

$$\|s_*\mu - \mu\|_{TV} \leq \varepsilon$$

où l'action de s sur $\mathbb{R}^2$ est l'action canonique, et où l'on définit $s_*\mu$ par $s_*\mu(E) = \mu(s^{-1}E)$ pour tout $E \subset \mathbb{R}^2$.

Alors, $\mu(\{0\}) = 1 - O(\varepsilon)$.

Démonstration. On utilise deux matrices particulières

$$a := \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$$

$$b := \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}$$

On définit également les sous-domaines suivants de $\mathbb{R}^2$: $A := \{(x,y) \in \mathbb{R}^2 / |x| < |y|\}$ et $B := \{(x,y) \in \mathbb{R}^2 / |x| > |y|\}$. On montre que $a^n A \subset B$ pour tout $n \geq 0$: en effet, $a^n = \begin{pmatrix} 1 & 2n \\ 0 & 1 \end{pmatrix}$. Soit $(x,y) \in A$. Alors $a^n(x,y) = (x + 2ny, y)$. Comme $|x| < |y|$, on a, par inégalité triangulaire,

$$\begin{aligned} |x + 2ny| &\geq 2n|y| - |x| \\ &> (2n-1)|y| + (|y| - |x|) \\ &\geq |y| \end{aligned}$$

d'où $a(x,y) \in B$.

Comme S est un voisinage de l'identité, il contient $a_n := \begin{pmatrix} 1 & 2/n \\ 0 & 1 \end{pmatrix}$ pour n entier suffisamment grand, fixé pour la suite. Alors, on a $a = a_n^n$. Par hypothèse, on a

$$\|(a_n)_*\mu - \mu\|_{TV} \leq \varepsilon$$

Or, par inégalité triangulaire,

$$\|a_*\mu - \mu\|_{TV} \leq \|(a_n^n)_*\mu - (a_n^{n-1})_*\mu\|_{TV} + \cdots + \|(a_n)_*\mu - \mu\|_{TV}$$

On vérifie que la mesure $(a_n^k)_*\mu$ possède la même propriété que μ, donc

$$\|a_*\mu - \mu\|_{TV} \leq n\varepsilon = O(\varepsilon)$$

Ainsi, comme $aA \subset B$, on a

$$\mu(B) \geq \mu(aA) = \mu(A) + O(\varepsilon)$$

De même, $bB \subset A$ donc

$$\mu(A) \geq \mu(bB) = \mu(B) + O(\varepsilon)$$

En utilisant ces deux inégalités, on déduit $\mu(B) \leq \mu(aA) \leq \mu(B) + O(\varepsilon)$, donc

$$\mu(B) = \mu(aA) + O(\varepsilon)$$

On procède de la même manière pour montrer que

$$\mu(B) = \mu(a^2A) + O(\varepsilon)$$

Enfin, en remarquant, d'après le paragraphe précédent, que tout élément (x, y) de a^2A vérifie $|x| > 3|y|$, on a

$$\{(x, y)/|y| < |x| < 3|y|\} \subset B \setminus a^2A$$

on déduit que l'ensemble $\{(x, y)/|y| < |x| < 3|y|\}$ a pour mesure $O(\varepsilon)$. Cet ensemble contient un secteur d'angle supérieur à un angle π/k où k est un entier donné. Alors, on peut recouvrir tout le plan excepté $\{0\}$ en réalisant un nombre fini (au maximum, $2k$) de rotations par des matrices de $SL_2(\mathbb{R})$. De même que pour a, on peut montrer que pour la rotation r d'angle π/k, on a

$$\|r_*\mu - \mu\|_{TV}\| = O(\varepsilon)$$

Alors, on peut conclure que $\mu(\mathbb{R}^2 \setminus \{0\}) = O(\varepsilon)$, d'où l'assertion du lemme. □

La proposition suivante offre un résultat sur l'existence de vecteurs invariants dans le groupe $SL_2(\mathbb{R}) \ltimes \mathbb{R}^2$, qui est isomorphe à un sous-groupe de $SL_3(\mathbb{R})$. Ce groupe peut être vu comme l'ensemble des transformations affines agissant sur $\mathbb{R}^2$: soit $(A, b) \in SL_2(\mathbb{R}) \ltimes \mathbb{R}^2$, alors pour tout $x \in \mathbb{R}^2$, on a $(A, b)x = Ax + b$. Ainsi, le produit de deux éléments (A, b) et (A', b') dans ce groupe est $(A', b')(A, b) = (A'A, A'b + b')$.

Proposition 6. *Soit S un voisinage compact de l'identité dans $SL_2(\mathbb{R}) \ltimes \mathbb{R}^2$, et $\rho : SL_2(\mathbb{R}) \ltimes \mathbb{R}^2 \to U(H)$ une représentation unitaire.*

Alors, si $Kaz(SL_2(\mathbb{R}) \ltimes \mathbb{R}^2, S, \rho)$ est arbitrairement petite selon le voisinage S choisi, alors H contient un vecteur $\mathbb{R}^2$-invariant non nul.

Avant de procéder à la démonstration, montrons le lemme suivant :

Lemme 5. *Soit $\rho : G \to U(H)$ une représentation unitaire d'un groupe localement compact G. Suppose qu'il existe un ensemble convexe fermé K dans H contenant un orbit $\{\rho(g)v_0 : g \in G\}$ d'un élément v_0 de H. Alors K contient un vecteur invariant.*

Démonstration. (lemme) Considérons l'ensemble $K' = \{v \in H : \rho(g)v \in K \ \ \forall g \in G\}$, alors K' est non vide. On va démontrer que K' est convexe et fermé.

Soit $v_1, v_2 \in K'$ alors $\rho(g)v_1, \rho(g)v_2 \in K \ \ \forall g \in G$, comme K est convexe, on déduit que $\forall\ \theta \in [0,1]\ \theta\rho(g)v_1 + (1-\theta)\rho(g)v_2 \in K \ \ \forall g \in G$ et donc $\theta v_1 + (1-\theta)v_2 \in K'$. La convexité de K' en découle.

Soit (v_n) une suite dans K' convergeant vers v dans H. Soit $g \in G$, comme $\rho(g)v_i \in K \ \ \forall i$, on déduit que $\rho(g)v \in K$. Comme g est choisi arbitrairment dans G, on a alors $\rho(g)v \in K \ \ \forall g \in G$ et donc $v \in K'$. Cela prouve que K' est fermé.

K' est convexe et fermé, la fonction $\| \bullet \|_H : K' \to \mathbb{R}$ atteint donc son minimum dans K'. Soit v_0 le vecteur de norme minimum dans K', et g un élément quelconque de G. Par définition on a, $\rho(g)v_0 \in K'$ et $(\rho(g)v_0 + v_0)/2$ l'est aussi. Alors par la convexité de K', on a $\|(\rho(g)v_0 + v_0)/2\| \leq \|v_0\|/2$, l'égalité ne se produit qu'au cas où $\rho(g)v_0 = v_0$. Comme v_0 est de norme minimum dans K', on déduit alors que l'égalité est vérifiée pour tout $g \in G$, ce qui veut dire que v_0 est un vecteur invariant.

On a donc prouvé que K contient un vecteur invariant. □

Démonstration. Soit $\varepsilon > 0$. Supposons $Kaz(SL_2(\mathbb{R}) \ltimes \mathbb{R}^2, S, H) < \varepsilon$, alors la définition de la constante de Kazhdan nous fournit un vecteur $v \in H$ tel que $\|\rho(s)v - v\|_H < \varepsilon$ pour tout $s \in S$.

Alors, soit $g \in SL_2(\mathbb{R}$. Pour tout $x \in \mathbb{R}^2$ on a

$$\langle \rho(x)\rho(g)v, \rho(g)v\rangle_H = \langle \rho(g)^*\rho(x)\rho(g)v, v\rangle_H$$

or $\rho(g)$ est unitaire, donc son adjoint est son inverse. Ainsi

$$\langle \rho(x)\rho(g)v, \rho(g)v\rangle_H = \langle \rho(g^{-1}xg)v, v\rangle$$

Le produit $g^{-1}xg$ peut se récrire pour davantage de clarté $(g^{-1}, 0)(I_2, x)(g, 0)$ et on remarque alors qu'il est égal à $(I_2, g^{-1}(x)$, d'où

$$\langle \rho(x)\rho(g)v, \rho(g)v\rangle_H = \langle \rho(g^{-1}(x))v, v\rangle_H$$

Alors, en reprenant les notations de la partie précédente, on a

$$\mu_{\rho(g)v,\rho(g)v}(x) = \mu_{v,v}(g(x))$$

pour tout $x \in \mathbb{R}^2$, or

$$\langle \rho(g^{-1}(x))v, v\rangle_H = \int_{\mathbb{R}^2} e^{2\pi i g^{-1}(x)\cdot\xi}\, d\mu_{v,v}(\xi)$$

et en passant à l'adjoint dans le produit scalaire, on obtient

$$\langle \rho(g^{-1}(x))v, v\rangle_H = \int_{\mathbb{R}^2} e^{2\pi i x\cdot(g^*)^{-1}\xi}\, d\mu_{v,v}(g^* \circ (g^*)^{-1}\xi)$$

et on conclut alors, par un changement de variables linéaire, que

$$\mu_{v,v}(g^{-1}(x)) = (g^*)_*\mu_{v,v}$$

Grâce à cette égalité et à la sesquilinéarité de μ, on peut assurer que

$$\|(g^*)_*\mu_{v,v} - \mu_{v,v}\|_{TV} = o(\|\rho(g)v - v\|_H$$

On applique alors le lemme pour trouver $\mu_{v,v}(\{0\}) \geq 1 - O(\varepsilon)$ et on conclut donc, d'après la définition de μ :

$$\langle \rho(x)v, v\rangle_H = 1 - O(\varepsilon)$$

quel que soit x dans $\mathbb{R}^2$. Plus précisément, $|\langle \rho(x)v, v\rangle_H - 1|$ est borné et on peut trouver ε suffisamment petit pour qu'il existe une boule fermée contenant l'orbite de v mais pas le vecteur nul. Cette boule étant fermée et convexe, on peut appliquer le lemme (5) : la boule contient un vecteur invariant. Ce vecteur est nécessairement non nul, d'où le résultat désiré. □

Grâce à la proposition précédente, nous pouvons exhiber un vecteur invariant par un sous-groupe isomorphe à $\mathbb{R}^2$ de $SL_3(\mathbb{R})$ (par exemple, celui engendré par $\left(\begin{smallmatrix} 1 & 0 & * \\ 0 & 1 & * \\ 0 & 0 & 1 \end{smallmatrix}\right)$). Nous souhaitons trouver un vecteur invariant par tout autre élément de $SL_3(\mathbb{R})$. Pour cela, montrons d'abord le résultat suivant, appelé phénomène de Mautner.

Il est d'abord nécessaire de définir les sous-groupes suivants de $SL_2(\mathbb{R})$.

$$U^+ := \{u_+(t), t \in \mathbb{R}\}$$

$$D := \{d(t), t \in \mathbb{R}\}$$

$$U^- := \{u_-(t), t \in \mathbb{R}\}$$

où l'on pose

$$u_+(t) := \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}$$

$$d(t) := \begin{pmatrix} e^t & 0 \\ 0 & e^{-t} \end{pmatrix}$$

$$u_-(t) := \begin{pmatrix} 1 & 0 \\ t & 1 \end{pmatrix}$$

Alors, on observe le phénomène suivant :

Proposition 7. *Phénomène de Mautner*
Soit $\rho : SL_2(\mathbb{R}) \to U(H)$ une représentation unitaire. Alors, tout vecteur de H qui est U^+-invariant est aussi $SL_2(\mathbb{R})$-invariant.

Ce résultat va nous permettre d'étendre la propriété d'invariance. Dans la preuve, nous travaillerons successivement avec des vecteurs des trois sous-groupes définis ci-dessus.

Démonstration. Soit v un vecteur U^+-invariant de H. En remarquant, pour tout t réel et tout ε strictement positif,

$$\begin{pmatrix} e^t & 0 \\ \varepsilon & e^{-t} \end{pmatrix} = u_+(\frac{e^t-1}{\varepsilon})u_-(\varepsilon)u_+(\frac{e^{-t}-1}{\varepsilon})$$

, on déduit

$$\left\langle \rho\left(\begin{pmatrix} e^t & 0 \\ \varepsilon & e^{-t} \end{pmatrix}\right) v, v \right\rangle_H = \langle \rho(u_+(\frac{e^t-1}{\varepsilon})u_-(\varepsilon)u_+(\frac{e^{-t}-1}{\varepsilon}))v, v\rangle$$

or ρ est unitaire, donc l'adjoint de $\rho(u_+(\frac{e^t-1}{\varepsilon}))$ est son propre inverse, c'est donc l'image par ρ d'un élément de U^+. En utilisant l'invariance de v par tout élément de U^+, on déduit

$$\left\langle \rho\left(\begin{pmatrix} e^t & 0 \\ \varepsilon & e^{-t} \end{pmatrix}\right) v, v \right\rangle_H = \langle \rho(u_-(\varepsilon))v, v\rangle$$

En faisant alors tendre ε vers 0, on peut conclure

$$\langle \rho(d(t))v, v\rangle_H = \langle v, v\rangle_H$$

Comme $\rho(d(t))$ est une isométrie, on calcule donc $\|d(t)v - v\|_H = 0$, et ceci pour tout t réel, ce qui permet de conclure que v est D-invariant. On applique une méthode similaire avec l'identité

$$d(t)u_-(s)d(-t) = u_-(e^{-t}s)$$

On utilise l'argument précédent, grâce à la D-invariance de v, pour établir

$$\langle u_-(s)v, v\rangle = \langle u_-(e^{-t}s)v, v\rangle$$

et on fait tendre s vers $+\infty$ d'où $\langle \rho(u_-(s))v, v\rangle_H = \langle v, v\rangle_H$. De même que précédemment, on conclut que v est U^--invariant. Enfin, comme U^+, D et U^- gérèrent $SL_2(\mathbb{R})$, nous pouvons conclure cette démonstration. □

À présent, par des manipulations astucieuses des résultats précédents, nous pouvons enfin prouver ce résultat final :

Proposition 8. *$SL_3(\mathbb{Z})$ possède la propriété (T).*

Démonstration. Tout d'abord, on admet que $SL_3(\mathbb{Z})$ est un réseau de $SL_3(\mathbb{R})$. La preuve est en effet assez difficile. Alors, d'après la proposition (5), il est équivalent de prouver que $SL_3(\mathbb{R})$ possède la propriété (T).

Soit S un voisinage compact générant $SL_3(\mathbb{R})$. Il nous faut prouver que la constante de Kazhdan $Kaz(SL_3(\mathbb{R}), S)$ est non nulle, ou de manière équivalente, que toute représentation admettant une suite de vecteurs quasi-invariants admet également un vecteur invariant.

Soit ρ une représentation admettant une suite de vecteurs quasi-invariants. Alors $Kaz(SL_3(\mathbb{R}), S, \rho)$ est nulle, or $SL_3(\mathbb{R})$ contient le sous-groupe

$$G_1 := \left\{ \left(\begin{array}{cc|c} & A & B \\ 0 & 0 & 1 \end{array}\right), A \in SL_2(\mathbb{R}), B \in \mathbb{R}^2 \right\}$$

qui est isomorphe à $SL_2(\mathbb{R}) \ltimes \mathbb{R}^2$. Alors $S \cap G_1 \neq \emptyset$, cette partie est un voisinage de l'identité dans G_1. On a $Kaz(SL_2(\mathbb{R}) \ltimes \mathbb{R}^2, G_1, \rho) \leq Kaz(\mathbb{R}), S, \rho)$, on peut donc appliquer la proposition (6) : il existe un vecteur $v \in H$ qui est K-invariant, où

$$K := \left\{ \left(\begin{array}{cc|c} & I_2 & M \\ 0 & 0 & 1 \end{array} \right), M \in \mathbb{R}^2 \right\}$$

qui est isomorphe à $\mathbb{R}^2$.

Considérons à présent les matrices de la forme

$$\begin{pmatrix} * & 0 & * \\ 0 & 1 & 0 \\ * & 0 & * \end{pmatrix}$$

où les astérisques sont à remplacer par des éléments quelconques de $\mathbb{R}^2$. Ces matrices forment un groupe noté G_2 isomorphe à $SL_2(\mathbb{R})$. On remarque alors que le sous-groupe isomorphe à U^+, engendré par les matrices de la forme

$$\begin{pmatrix} 1 & 0 & * \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

est inclus dans K. Le vecteur v est donc invariant par ce sous-groupe ; le phénomène de Mautner implique donc qu'il est G_2-invariant.

De même, en posant G_3 le sous-groupe engendré par les matrices de la forme

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & * & * \\ 0 & * & * \end{pmatrix}$$

on prouve que v est G_3-invariant.

G_2 et G_3 contiennent une partie des matrices de transvection $T_{i,j}(t) := I_3 + tE_{i,j}$ où t est réel et i, j sont deux entiers fixés dans $\{1, 2, 3\}$. Les autres matrices de transvection s'obtiennent par produit des premières, sur le principe des opérations sur les lignes et colonnes. Par exemple,

$$\begin{pmatrix} 1 & t & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & -t \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & t \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}$$

Lemme 6. *Les matrices de transvection engendrent $SL_3(\mathbb{R})$.*

Preuve du lemme : On vérifie tout d'abord que les matrices de permutation $P_{i,j}$ s'obtiennent comme le produit $T_{j,i}(-1)T_{i,j}(1)T_{j,i}(-1)$. Il sera donc légitime dans la suite de permuter les lignes et colonnes des matrices sur lesquelles on travaillera.

Nous aurons aussi besoin de matrices de la forme

$$C(\lambda) := \begin{pmatrix} \lambda & 0 & \dots & 0 \\ 0 & \lambda^{-1} & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & 1 \end{pmatrix}$$

où λ est un réel non nul. On peut vérifier l'identité $C(\lambda) = T_{2,1}(-\frac{1}{\lambda})T_{1,2}(\lambda-1)T_{2,1}(1)T_{1,2}(-\frac{\lambda-1}{\lambda})$.

Nous procédons par récurrence sur la taille de la matrice, bien que nous n'ayons besoin que du cas $d = 3$. Pour $d = 1$, le groupe spécial linéaire est trivial et est donc engendré par ses transvections.

Soit $d \in \mathbb{N}$. Supposons que le groupe $SL_d(\mathbb{R})$ soit engendré par ses transvections. Soit $A \in SL_d(\mathbb{R})$. Comme A est inversible, elle possède un coefficient non nul sur sa première ligne. Quitte à faire agir des matrices de permutation à droite de A, on obtient une matrice $A_0 = AP_{i_1,j_1}...P_{i_n,j_n}$ dont le coefficient supérieur gauche est non nul, notons-le λ_A. En faisant agir des matrices de transvection à droite, puis à gauche, on parvient à annuler les autres coefficients de la première ligne, puis de la première colonne. En multipliant par $C(\lambda_A^{-1})$, on obtient une matrice de la forme

$$\left(\begin{array}{c|ccc} 1 & 0 & \dots & 0 \\ \hline 0 & & & \\ \vdots & & A' & \\ 0 & & & \end{array}\right)$$

où A' est un élément de $SL_{d-1}(\mathbb{R})$. On peut, d'après l'hypothèse de récurrence, écrire A' comme un produit $M_1...M_k$ de matrices de transvection de $SL_{d-1}(\mathbb{R})$, donc la matrice ci-dessus peut s'écrire comme produit de tranvections $T_{i,j}$ de $SL_d(\mathbb{R})$ avec $i, j \geq 2$. Ainsi, en inversant les résultats ci-dessus, on trouve une écriture de A comme produit de matrices de transvection.

Ce lemme permet de conclure que G_2 et G_3 engendrent $SL_3(\mathbb{R})$, et le vecteur v est donc $SL_3(\mathbb{R})$-invariant. Cela permet de conclure notre démonstration : $SL_3(\mathbb{R})$ possède donc la propriété (T). Ainsi, la famille de Margulis est bien une famille de graphes expanseurs.

□

3.3.4 Une autre construction similaire

Nous présentons ici une construction qui n'utilise pas explicitement la propriété (T) de Kazhdan. On utilisera le produit semi-direct $SL_2(\mathbb{Z}) \ltimes \mathbb{Z}^2$. On note qu'il est isomorphe à un sous-groupe de $SL_3(\mathbb{Z})$, comme mentionné dans la parite précédente.

Théorème 3.3.7. *Soit S une partie génératrice finie symétrique de $SL_2(\mathbb{Z}) \ltimes \mathbb{Z}^2$, alors les graphes de Schreier $Sch((\mathbb{Z}/n\mathbb{Z})^2, \pi_n(S))$ forment une famille de graphes expanseurs où $\pi_n : SL_2(\mathbb{Z}) \ltimes \mathbb{Z}^2 \to SL_2(\mathbb{Z}/n\mathbb{Z}) \ltimes (\mathbb{Z}/n\mathbb{Z})^2$ est la projection canonique.*

Démonstration. On montre ce théorème par l'absurde. Supposons que les graphes de Schreier ne forment pas une famille de graphes expanseurs.

Soit ρ_n la représentation quasi-régulière de $SL_2(\mathbb{Z}) \ltimes \mathbb{Z}^2$ sur $(\mathbb{Z}/n\mathbb{Z})^2$. Par une preuve similaire à celle du lemme 1, on peut démontrer que s'il existe un ϵ positive tel que pour tout n et pour toute fonction $f \in l^2((\mathbb{Z}/n\mathbb{Z})^2)$:

$$\sup_{s\in S} \|\rho(s)f_n - f_n\|_{l^2((\mathbb{Z}/n\mathbb{Z})^2)} > \epsilon$$

alors les graphes de Schreier $Sch((\mathbb{Z}/n\mathbb{Z})^2, \pi_n(S))$ forment une famille de graphes c-expanseurs pour un c positif ne dépendant que de ϵ et $|S|$. Comme on a supposé que ces graphes ne forment pas une famille de graphes expanseurs, alors pour tout ϵ, on peut trouver un entier n et une fonction $f \in l^2(\mathbb{Z}/n\mathbb{Z})^2)$ tels que

$$\sup_{s\in S} \|\rho(s)f_n - f_n\|_{l^2(\mathbb{Z}/n\mathbb{Z})^2)} < \epsilon$$

Quitte à extraire une sous-suite, on peut conclure qu'il existe une suite de fonctions $(f_n) \in l^2(\mathbb{Z}/n\mathbb{Z})^2)$ telle que :

$$\sup_{s\in S)} \|\rho(s)f_n - f_n\|_{l^2(\mathbb{Z}/n\mathbb{Z})^2)} = o(1)$$

Soit e_1, e_2 les vecteurs générateurs du groupe Z^2, on a alors $\|\rho_n(e_j)f_n - f_n\|_{l^2(\mathbb{Z}/n\mathbb{Z})} = o(1)$ pour $j = 1, 2$. Maintenant, on considère la transformation de Fourier :

$$\mathcal{F}(f_n)(\xi_1, \xi_2) = \frac{1}{n} \sum_{x_1,x_2\in\mathbb{Z}/n\mathbb{Z}} f_n(x_1, x_2)e^{-2\pi i(x_1\xi_1+x_2\xi_2)/n}$$

On note que cette formule a été normalisée pour que la transformation de Fourier soit une isométrie dans $l^2(\mathbb{Z}/n\mathbb{Z})$. Ainsi :

On a :

$$|\mathcal{F}(f)(xi_1, \xi_2)|^2 = \frac{1}{n^2} \sum_{x_i,x_j\in(\mathbb{Z}/n\mathbb{Z})^2} f(x_{i_1}, x_{i_2})\overline{f(x_{j_1}, x_{j_2})}e^{2\pi i(\xi_1(x_{j_1}-x_{i_1}+\xi_2(x_{j_2}-x_{i_2}))/n}$$

Alors :

$$\begin{aligned}\|\mathcal{F}(f)\|_{l^2(\mathbb{Z}/n\mathbb{Z})^2} &= \frac{1}{n^2} \sum_{\xi_1,\xi_2\in\mathbb{Z}/n\mathbb{Z}} \sum_{x_i,x_j\in(\mathbb{Z}/n\mathbb{Z})^2} f(x_{i_1}, x_{i_2})\overline{f(x_{j_1}, x_{j_2})}e^{2\pi i(\xi_1(x_{j_1}-x_{i_1}+\xi_2(x_{j_2}-x_{i_2}))/n} \\ &= \frac{1}{n^2} \sum_{x_i,x_j\in(\mathbb{Z}/n\mathbb{Z})^2} f(x_{i_1}, x_{i_2})\overline{f(x_{j_1}, x_{j_2})} \sum_{\xi_1,\xi_2\in\mathbb{Z}/n\mathbb{Z}} e^{2\pi i(\xi_1(x_{j_1}-x_{i_1}+\xi_2(x_{j_2}-x_{i_2}))/n}\end{aligned}$$

Comme on a :

$$\sum_{\xi_1,\xi_2\in\mathbb{Z}/n\mathbb{Z}} e^{2\pi i(\xi_1(x_{j_1}-x_{i_1}+\xi_2(x_{j_2}-x_{i_2}))/n} = 0$$

pour tout x_i, x_j distincts dans $(\mathbb{Z}/n\mathbb{Z})^2$ et

$$\sum_{\xi_1,\xi_2\in\mathbb{Z}/n\mathbb{Z}} e^{2\pi i(\xi_1(x_{j_1}-x_{i_1}+\xi_2(x_{j_2}-x_{i_2}))/n} = n^2$$

si $x_i = x_j$. Alors

$$\|\mathcal{F}(f)\|_{l^2(\mathbb{Z}/n\mathbb{Z})^2} = \frac{1}{n^2} \sum_{x_i \in (\mathbb{Z}/n\mathbb{Z})^2} n^2 |f(x_{i_1}, x_{i_2})|^2 = \|f\|_{l^2(\mathbb{Z}/n\mathbb{Z})^2}$$

Maintenant, en remarquant que :

$$\mathcal{F}(\rho_n(e_j)f_n)(\xi_1, \xi_2) = \mathcal{F}(f(e_j^{-1} * \bullet))(\xi_1, \xi_2) = e^{-2\pi i \xi_j/n} \mathcal{F}(f_n)(\xi_1, \xi_2)$$

on a :

$$\|(e^{-2\pi i \xi_j/n} - 1)\mathcal{F}(f_n)\|_{l^2_{\xi_1,\xi_2}((\mathbb{Z}/n\mathbb{Z})^2)} = o(1)$$

Considérons la boule B_n de rayon $o(n)$ centrée à l'origine du groupe $(\mathbb{Z}/n\mathbb{Z})^2$ alors

$$\begin{aligned} \|\mathcal{F}(f_n)\|_{l^2_{\xi_1,\xi_2}((\mathbb{Z}/n\mathbb{Z})^2 \setminus B_n)} &\leq \frac{1}{\|e^{-2\pi i/n} - 1\|} \|(e^{-2\pi i \xi_j/n} - 1)\mathcal{F}(f_n)(\xi_1, \xi_2)\|_{l^2_{\xi_1,\xi_2}((\mathbb{Z}/n\mathbb{Z})^2 \setminus B_n)} \\ &\leq \frac{1}{\|e^{-2\pi i/n} - 1\|} \|(e^{-2\pi i \xi_j/n} - 1)\mathcal{F}(f_n)(\xi_1, \xi_2)\|_{l^2_{\xi_1,\xi_2}((\mathbb{Z}/n\mathbb{Z})^2)} = o(1) \end{aligned}$$

(car $\forall n \geq 2\ \forall (\xi_1, \xi_2) \in (\mathbb{Z}/n\mathbb{Z})^2 \setminus B_n : \|e^{-2\pi i \xi_j/n} - 1\| \geq \|e^{-2\pi i/n} - 1\|$)

Soit g_n la restriction de $\mathcal{F}(f_n)$ dans B_n, qui peut être identifiée avec un sous-ensemble de $\mathbb{Z}^2$(comme B_n est de rayon $o(n)$). Soit s un élément fixé quelconque de $SL_2(\mathbb{Z})$, s peut être écrit sous forme d'un produit fini des éléments de S : $s = s_1 s_2 .... s_m$, alors :

$$\|\rho_n(s)f_n - f_n\|_{l^2((\mathbb{Z}/n\mathbb{Z})^2)} \leq \sum_{i=1}^{m} \|\rho_n(s_i)f_n - f_n\|_{l^2((\mathbb{Z}/n\mathbb{Z})^2)} = o(1)$$

et par la transformation de Fourier :

$$\|\mathcal{F}(f_n(s^{-1}.\bullet)) - \mathcal{F}(f_n)\|_{l^2((\mathbb{Z}/n\mathbb{Z})^2)} = \|\mathcal{F}(f_n) \circ s^* - \mathcal{F}(f_n)\|_{l^2((\mathbb{Z}/n\mathbb{Z})^2)} = o(1)$$

Comme g_n est la restriction de f_n dans B_n et le rayon de B_n est plus petit que n, on en déduit :

$$\|g_n \circ s^* - g_n\|_{l^2(\mathbb{Z}^2)} = o(1)$$

En appliquant le lemme..., on conclut que

$$g_n(0) = 1 - o(1)$$

Mais comme f_n est de moyenne nulle, on doit avoir

$$\mathcal{F}(f_n)(0) = \frac{1}{n} \sum_{x_1, x_2 \in \mathbb{Z}/n\mathbb{Z}} f_n(x_1, x_2) = 0$$

et donc $g_n(0) = \mathcal{F}(f_n)(0) = 0$.

Cette contradiction entraine que les graphes de Schreier $Sch((\mathbb{Z}/n\mathbb{Z})^2, \pi(S))$ forment une famillle de graphes expanseurs. □

3.4 La construction des graphes de Ramanujan : une famille de graphes optimaux

Voici maintenant une seconde façon de construire une famille de graphes expanseurs, appelée construction de Ramanujan. Les graphes de Ramanujan permettent de se rapprocher de la borne de trou spectral défini en (2.3.2).

Définition 3.4.1. Graphe de Ramanujan Soit un graphe G $k-$ régulier simple (il n'y pas de boucles et d'arêtes multiples) et connexe. On note alors les valeurs propres de sa matrice d'adjacence $k = \lambda_0 > \lambda_1 \geq \lambda_2 \cdots \geq \lambda_{n-1} \geq -k$.

Si une famille de graphes $k-$réguliers vérifie l'inégalité

$$\lambda_1 \leq 2\sqrt{k-1} \tag{3.1}$$

on l'appelle une famille de graphes de Ramanujan.

Une famille de graphes de Ramanujan est optimale dans le sens où elle rapproche le trou spectral de sa limite théorique, mais son existence n'est pas évidente. C'est un résultat surprenant qui relie la théorie des nombres algébriques, la théorie des groupes et la théorie des graphes. Dans la sous-partie qui suit, nous allons rappeler les notations nécessaires à la construction des graphes de Ramanujan.

3.4.1 Préparation

Dans cette sous-partie, nous rappelons les définitions du groupe linéaire, des quaternions et de la décomposition d'un entier en somme de quatre carrés, qui nous seront utiles pour la construction des graphes de Ramanujan.

Définition 3.4.2. Groupe linéaire
Soit K un corps. On définit le groupe général linéaire d'ordre 2 par :

$$GL_2(K) = \{M \in M_2(K) | det(M) \neq 0\}$$

$$SL_2(K) = \{M \in M_2(K) | det(M) = 1\}$$

$$PGL_2(K) = GL_2(K)/\lambda Id$$

$$SGL_2(K) = SL_2(K)/\pm Id$$

On utilise aussi la notation $GL_2(q), SL_2(q), PGL_2(q), PSL_2(q)$ pour simplifier $GL_2(\mathbb{F}_q)$, $SL_2(\mathbb{F}_q)$, $PGL_2(\mathbb{F}_q)$, $PSL_2(\mathbb{F}_q)$.

Remarque 3. *Le cardinal des groupes susnommés est :*

$$\begin{aligned} |GL_2(q)| &= q(q-1)(q^2-1) \\ |SL_2(q)| = |PGL_2(q)| &= q(q^2-1) \\ |PSL_2(q)| &= \frac{1}{2}q(q^2-1)(q \neq 2) \\ |PSL_2(2)| &= 3 \end{aligned}$$

Définition 3.4.3. Quaternions
Soit K un corps, on définit un quaternion comme $a + a_1 i + a_2 j + a_3 k$ où $i^2 = j^2 = k^2 = -1, ij = k, jk = i, ki = j, a_0, a_1, a_2, a_3 \in K$. En plus, si $x, y \in K$ et satisfont $x^2 + y^2 + 1 = 0$, alors il existe un isomorphisme entre $H(K)$ et $M_2(K)$ donné par :

$$\psi(a_0 + a_1 i + a_2 j + a_3 k) = \begin{pmatrix} a_0 + a_1 x + a_3 y & -a_1 y + a_2 + a_3 x \\ -a_1 y - a_2 + a_3 x & a_0 - a_1 x - a_3 y \end{pmatrix} \tag{3.2}$$

Remarque 4. *L'isomorphisme entre $H(q)$ et $M_2(K)$ peut être considéré comme une généralisation de l'isomorphisme canonique des nombres complexes. Soit $a, b \in K'$*

$$\psi(a + bj) = \begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix} \tag{3.3}$$

Si on prend $K' = K + Ki$ alors

$$\bar{\psi}(a_0 + a_1 i + (a_2 + a_3 i)j) = \begin{pmatrix} a_0 + a_1 i & a_2 + a_3 i \\ -a_2 + a_3 i & a_0 - a_1 i \end{pmatrix} \tag{3.4}$$

Ceci est un isomorphisme puisque la conjugaison est aussi un isomorphisme. S'il existe $\bar{x}^2 + \bar{y}^2 = 1$, on écrit alors $i = \sqrt{\tilde{x}^2 + \tilde{y}^2}$ et définit

$$\psi(a_0 + a_1 i + (a_2 + a_3 i)j) = \begin{pmatrix} \tilde{x} & \tilde{y} \\ -\tilde{y} & \tilde{x} \end{pmatrix} \begin{pmatrix} a_0 + a_1\sqrt{\tilde{x}^2 + \tilde{y}^2} & a_2 + a_3\sqrt{\tilde{x}^2 + \tilde{y}^2} \\ -a_2 + a_3\sqrt{\tilde{x}^2 + \tilde{y}^2} & a_0 - a_1\sqrt{\tilde{x}^2 + \tilde{y}^2} \end{pmatrix} \begin{pmatrix} \tilde{x} & -\tilde{y} \\ \tilde{y} & \tilde{x} \end{pmatrix}$$

Cela donne l'isomorphisme comme dans l'expression 3.2 où $x = -\sqrt{\tilde{x}^2 + \tilde{y}^2}(\tilde{x}^2 - \tilde{y}^2)$, $y = 2\sqrt{\tilde{x}^2 + \tilde{y}^2}\tilde{x}\tilde{y}$.

Remarque 5. *Dans la construction du graphe de Ramanujan, on va appliquer cet isomorphisme dans le cas où $K = \mathbb{F}_q$, où q est premier. On remarque donc qu'il existe un couple (x, y) qui satisfait $x^2 + y^2 + 1 \equiv 0$. Comme l'ensemble $\{x^2, 0 \leq x \leq \frac{1+p}{2}\} \cup \{-1 - y^2, 0 \leq y \leq \frac{1+p}{2}\}$ contient $(p + 1)$ éléments, deux d'entre eux sont congrus par le principe des tiroirs. Mais les deux ne peuvent pas venir du même sous-ensemble. Donc, il existe $x^2 \equiv -1 - y^2 (\mod q)$.*

On définit aussi la norme carrée par

$$\mathcal{N}(a_0 + a_1 i + a_2 j + a_3 k) = a_0^2 + a_1^2 + a_2^2 + a_3^2$$

Inspiré par la décomposition en produit de facteurs premiers, on peut aussi faire la décomposition en produit de facteurs premiers dans $H(\mathbb{Z})$ par l'algorithme d'Euclide. Ici on remarque qu'une condition suffisant pour être premier est $\mathcal{N}(\alpha) = p$, où p est premier (en fait, elle est aussi nécessaire grâce au 3.4.1). Cette structure relie la décomposition en somme de quatre carrés avec la décomposition des quaternions.

$$a_0^2 + a_1^2 + a_2^2 + a_3^2 = n \Leftrightarrow \mathcal{N}(a_0 + a_1 i + a_2 j + a_3 k) = n$$

On énonce le théorème fondamental suivant sans démonstration. On trouvera plus de détails sur la structure du quaternion dans [8].

Théorème 3.4.1. *Soit p un nombre premier, alors le nombre de quaternions $H(\mathbb{Z})$ tels que $\mathcal{N}(\alpha) = p$ est $8(p + 1)$. Autrement dit, il existe $8(p + 1)$ différents $4 - tuples(a_0, a_1, a_2, a_3)$ vérifiant l'équation :*

$$a_0^2 + a_1^2 + a_2^2 + a_3^2 = p \tag{3.5}$$

3.4.2 Construction du graphe de Ramanujan

On commence la construction du graphe de Ramanujan de la même manière que pour les graphes de Cayley. On associe les sommets du graphe à une partie du groupe linéaire $PGL_2(q)$ où q est un premier. L'ensemble symétrique pour engendrer les arêtes est un peu compliqué : pour l'équation

$$a_0^2 + a_1^2 + a_2^2 + a_3^2 = p$$

d'après le théorème 3.5, il y a $8(p+1)$ solutions mais on ne prend que $(p+1)$ solutions comme représentants et on note leur ensemble S_p. Puis on les associe à une famille de matrices $S_{p,q}$ dans $PGL_2(q)$. On obtient une construction et on va expliquer l'ensemble symétrique $S_{p,q}$, la connexité et la répartition des spectres dans les sous-sections suivantes.

Théorème 3.4.2. ***Une construction du graphe de Ramanujan***
Soit p, q deux entiers premiers impairs distincts et $q > 2\sqrt{p}$. Si $(\frac{p}{q}) = 1$, c'est-à-dire s'il existe m tel que $p \equiv m^2 \mod q$. On définit

$$X^{p,q} = Cay(PSL_2(q), S_{p,q})$$

En revanche, si $(\frac{p}{q}) = -1$, c'est-à-dire $\forall m$ tel que $p \not\equiv m^2 \mod q$, on définit

$$X^{p,q} = Cay(PGL_2(q), S_{p,q})$$

Le graphe $X^{p,q}$ est connexe et c'est un graphe de Ramanujan. De plus, quand $q \to \infty$, c'est une famille d'expanseurs et en particulier une famille de graphes de Ramanujan.

Remarque 6. *La condition $q > 2\sqrt{p}$ assure que l'application de S_p à $S_{p,q}$ est bijective. Pour le cas où $(\frac{p}{q}) = 1$, $S_{p,q}$ engendre que $PSL_2(q)$ mais pas $PGL_2(q)$, donc il y a différent façon de le traiter.*

3.4.3 Ensemble symétrique $S_{p,q}$

Expliquons l'ensemble symétrique en détail. Soit $\alpha = (a_0, a_1, a_2, a_3)$ une solution tel que $a_0^2 + a_1^2 + a_2^2 + a_3^2 = p$, on sait qu'elle peut engendrer 8 différentes solutions par multiplier une unité $\pm 1, \pm i, \pm j, \pm k$.

Comme p est premier, il y a nécessairement un élément parmi a_0, a_1, a_2, a_3dont la parité est différente que celle des autres. On multiplie par $1, i, j, k$ de façon à ce que a_0 soit celui de parité différente.

Si a_0 n'est pas nul, on le multiplie par 1 ou -1 pour le rendre positif. Dans le cas où $a_0 = 0$, a_1 est impair donc non-nul. On multiplie α par 1 ou -1 pour que a_1 soit positif.

Définition 3.4.4. $S_p = \{\alpha \in H(\mathbb{Z}), \mathcal{N}(\alpha) = p, \alpha \equiv 1 \mod 2$, ou $\alpha \equiv i + j + k \mod 2$, en plus $a_0(\alpha) > 0$ ou $a_0(\alpha) = 0, a_1(\alpha) > 0\}$

Exemple 1. *Pour l'équation*

$$a_0^2 + a_1^2 + a_2^2 + a_3^2 = 3 \tag{3.6}$$

On sait que $3 = 1^2 + 1^2 + 1^2 + 0^2$*, mais à permutation et signe près, il y a* $C_4^1 * 2^3 = 32 = 8*(3+1)$ *solutions qui vérifient le théorème 3.5. On utilise alors la règle ci-dessus pour sélectionner le représentant dans* S_3 *:*

$$S_3 = \{(0,1,1,1),(0,1,1,-1),(0,1,-1,1),(0,1,-1,-1)\}$$

Voici un autre exemple pour l'équation

$$a_0^2 + a_1^2 + a_2^2 + a_3^2 = 23 \tag{3.7}$$

On sait que la solution unique, positive, et sans permutation est $23 = 3^2 + 3^2 + 2^2 + 1^2$*, mais à permutation et signe près, il y a* $C_4^2 * 2 * 2^4 = 192 = 8 * (23+1)$ *solutions qui vérifient aussi le théorème 3.5.* S_{23} *est donc*

$$S_{23} = \{(2,\pm1,\pm3,\pm3),(2,\pm3,\pm1,\pm3),(2,\pm3,\pm3,\pm1)\}$$

qui est un ensemble contenant $(23+1)$ *éléments.*

Afin d'associer S_p avec $PGL_2(q)$, on applique trois opérateurs.

Définition 3.4.5. Soit $\tau_q : H(\mathbb{Z}) \to H(\mathbb{F}_q)$ la surjection canonique de $H(\mathbb{Z})$ sur $H(\mathbb{F}_q)$. On définit :
$\psi_q : H(\mathbb{F}_q) \to GL_2(q)$ par l'isomorphisme entre $H(\mathbb{Z})$ et $M(\mathbb{Z})$ définie dans 3.2
$\phi : GL_2(q) \to PGL_2(q)$ par le morphisme canonique entre $GL_2(q)$ et $PGL_2(q)$
On définit $S_{p,q}$ l'ensemble symétrique pour engendrer les arêtes comme

$$S_{p,q} = (\phi \circ \psi_q \circ \tau_q)(S_p)$$

3.4.4 Connexité de $X^{p,q}$

La connexité de $X^{p,q}$ n'est pas évidente. On va construire un autre graphe $Y^{p,q}$ qui est isomorphe à $X^{p,q}$ et dont la connexité est plus claire.

On trouvera les détails de cette construction en annexe.

3.4.5 Spectres

On va montrer que notre construction est bien un graphe de Ramanujan, i.e. que $|\lambda_i| \leq 2\sqrt{p}$ sauf pour la valeur propre triviale. Pour estimer la répartition des valeurs propres, on rappelle la formule de la trace.

$$\sum_{0\leq r\leq \frac{m}{2}} f_{m-2r} = \frac{p^{\frac{m}{2}}}{n} \sum_{j=0}^{n-1} U_m(\frac{\lambda_i}{2\sqrt{p}})$$

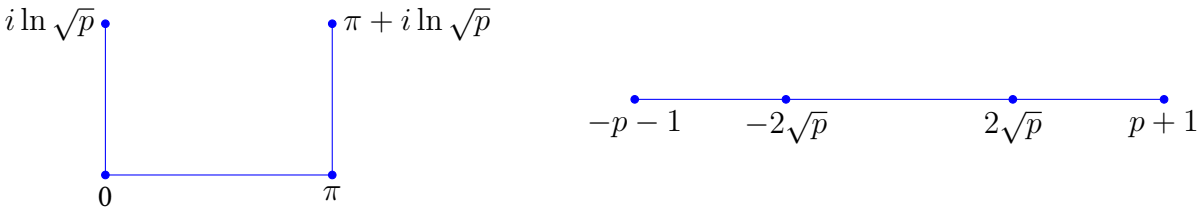


FIGURE 3.1 – L'intervalle $\mathcal{C}_R$ que l'on va utiliser pour paramétrer $[-(p+1), p+1]$. Notons que la fonction $z \longmapsto 2\sqrt{p}\cos z$ renvoie $\mathcal{C}_R$ à $[-p-1, p+1]$ avec $[i \ln \sqrt{p}, 0]$, $[0, \pi]$, $[\pi, \pi + i \ln \sqrt{p}]$ vers $[p+1, 2\sqrt{p}]$, $[2\sqrt{p}, -2\sqrt{p}]$ et $[-2\sqrt{p}, -p-1]$ respectivement.

où $U_m(\cos x) = \frac{\sin(m+1)x}{\sin x}$ et f_m signifie le nombre de cycles sans-retour de longueur m.

$$\sum_{0 \leq r \leq \frac{m}{2}} f_{m-2r} = \frac{p^{\frac{m}{2}}}{n} \sum_{j=0}^{n-1} U_m\left(\frac{\lambda_i}{2\sqrt{p}}\right)$$

où U_m est la famille de polynôme de Chebyshev, c-à-d. $U_m(\cos x) = \frac{\sin(m+1)x}{\sin x}$ et f_m signifie le nombre de cycles sans-retour de longueur m.

Quand on se place dans le cadre des graphes de Ramanujan $X^{p,q}$, il y a une autre interprétation de la somme à gauche :

Proposition 9.
Posons $Q(x_0, x_1, x_2, x_3) = x_0^2 + q^2(x_1^2 + x_2^2 + x_3^2)$. $\sum_{0 \leq r \leq \frac{m}{2}} f_{m-2r}$ vaut alors $\frac{1}{2} S_Q(p^m)$, avec $S_Q(p^m)$ le nombre de solutions dans $\mathbb{Z}^4$ de l'équation $Q(x_0, x_1, x_2, x_3) = p^m$ satisfaisant $x_1 \equiv x_2 \equiv x_3 \not\equiv x_0$ modulo 2.

La démonstration vient directement de la décomposition unique modulo p de $H(\mathbb{Z})$ (A.5.1). f_{m-2r} est le nombre d'éléments de $H(\mathbb{Z})$ de norme p^m dont la décomposition contient exactement r facteurs p. Le facteur 2 vient du choix de signe. La formule de trace devient ainsi :

$$S_Q(p^m) = \frac{2p^{\frac{m}{2}}}{n} \sum_{j=0}^{n-1} U_m\left(\frac{\lambda_i}{2\sqrt{p}}\right) \tag{3.8}$$

En utilisant des résultats de la géométrie algébrique, en particulier les conjectures de Weil, on arrive à une estimation plus fine.

Proposition 10.

$$S_Q(p^m) = \frac{4}{q(q^2-1)} \cdot \frac{p^{m+1}-1}{p-1} + \mathcal{O}_\epsilon(p^{m(\frac{1}{2}+\epsilon)})$$

À partir de cette proposition, on arrive facilement à la propriété de Ramanujan

Théorème 3.4.3. *La famille $\{X^{p,q}\}$ est de Ramanujan.*

Démonstration. Notons que la première partie, dite la série d'Eisenstein, est exactement la contribution des valeurs propres triviales (ou de la valeur propre triviale quand $(\frac{p}{q}) = 1$) dans la somme (3.8). Donc

$$\frac{1}{n} \sum_{\lambda_j \neq \pm(p+1)} \frac{\sin(m+1)\theta_j}{\sin\theta_j} = \mathcal{O}_\varepsilon(p^{\varepsilon m}/2) \qquad \forall m \in \mathbb{N}$$

où les θ_j varient dans $\mathcal{C}_R$, l'intervalle complexe est défini par la Figure 3.1. S'il existe $\theta_j \notin \mathbb{R}$, en notant ψ_j la partie imaginaire de θ_j et en ne considérant que le cas m pair, on a $\frac{\sin(m+1)\theta_j}{\sin\theta_j} = \frac{\sinh(m+1)\psi_j}{\sinh\psi_j} > 0$. L'estimation ci-dessus devient :

$$\frac{1}{n} \sum_{|\lambda_j| \leq 2\sqrt{p}} \frac{\sin(m+1)\theta_j}{\sin\theta_j} + \frac{1}{n} \sum_{|\lambda_j| > 2\sqrt{p}} \frac{\sinh(m+1)\psi_j}{\sinh\psi_j} = \mathcal{O}_\varepsilon(p^{\varepsilon m}/2) \qquad \forall m \in 2\mathbb{N}$$

Notons que $|\frac{\sin(m+1)\theta}{\sin\theta}| \leq m+1$, la première somme croît linéairement en m alors que la deuxième somme croît exponentiellement et avec tous ses termes positifs. Donc s'il existe $|\theta_j| > 2\sqrt{p}$, on prend ε plus petit que tous les ψ_j et on obtient une contradiction en laissant m tendre vers l'infini. Donc tous les θ_j sont réels et les $\{X^{p,q}\}$ sont des graphesde Ramanujan. □

Remarque 7. *Il existe aussi une estimation spectrale assez fine pour démontrer la propriété d'expansion (mais pas celle de Ramanujan) qui n'utilise que des notions élémentaires. Les raisonnements se basent sur deux points :*

(i) Chaque valeur propre non-triviale du spectre de $X^{p,q}$ est de multiplicité au moins $\frac{q-1}{2}$.

(ii) Le nombre de façons de décomposer un entier n en 3 carrés $r_3(n)$ est $\mathcal{O}_\varepsilon(n^{\varepsilon+\frac{1}{2}})$.

Le premier découle du fait que chaque espace propre associé à ces valeurs propres est naturellement une représentation de $PSL_2(q)$. Par la simplicité de $PSL_2(q)$, on peut déduire que toutes représentations non-triviales de $PSL_2(q)$ sont de degré au moins $\frac{q-1}{2}$. Le deuxième point n'est qu'une conséquence du fait que $r_2(n)$ est $\mathcal{O}_\varepsilon(n^\varepsilon)$.

Le premier point donnant une estimation de ψ_j en fonction de $S_Q(p^m)$ alors que le deuxième donne une estimation de ce dernier, on arrive à ;

$$|\lambda| \leq p^{5/6+\varepsilon} + p^{1/6-\varepsilon} \qquad \text{pour tout} \quad q > q_\varepsilon$$

d'où la propriété d'expansion.

Bien que la démonstration soit complexe et bien qu'elle utilise des connaissances dans de nombreux domaines, on a donc construit un graphe de Ramanujan, qui se rapproche des bornes limites indiquées précédemment.

3.5 La construction des graphes expanseurs en pratique : les graphes aléatoires

3.5.1 Modélisation

Une famille graphes aléatoires k-réguliers possède une très bonne chance d'être une famille de graphes expanseurs, c'est pourquoi dans la pratique, on va plutôt générer une famille

de graphes aléatoirement. Dans cette section, on va introduire ce qu'on entend par "une famille de graphes aléatoires k-réguliers"

En ne considérant que le cas où $k = 2l$ est pair, on peut modéliser un graphe G k-régulier de n sommets par l permutations $\{\pi_i\}_{i=1}^{l}$ sur $\{1, 2, .., n\}$. En effet, pour chaque sommet a de G, en le reliant aux sommets $\{\pi_i(a)\}_{i=1}^{l}$, on obtient un graphe $2l$-régulier si et seulement si les conditions suivantes sont satisfaites :

(i) $\pi_i(v) \neq {\pi_i}^{-1}(v)\ \forall i = 1, .., l$, i.e. chaque arête est comptée une fois dans chaque permutation.

(ii) $\pi_i(v) \neq \pi_j(v)\ \forall i \neq j$, i.e. chaque arête est comptée une fois dans toutes les permutations.

(iii) $\pi_i(v) \neq v\ \forall i = 1, .., l$, c'est-à-dire qu'il n'y a pas de boucle

3.5.2 Résultats

Le théorème suivant, démontré par Julius Petersen, assure que tous les graphes $2l$-réguliers admettent une telle représentation. L'idée est de construire, grâce à un circuit d'Euler à partir du graphe G, un graphe G' de $2n$ sommets, biparti et l-régulier, puis appliquer le théorème de Hall suivant.

Théorème 3.5.1 (Hall). *Soit $\mathcal{R}$ une relation, c'est-à-dire un sous-ensemble $A \times B$, telle que pour tout sous-ensemble C de A l'image de C dans B, i.e. $\{b \in B : \exists a \in C : (a, b) \in \mathcal{R}\}$ contient au moins autant d'éléments que C lui-même. Alors il existe un couplage parfait pour A, i.e. une injection f de A à B telle que $(a, f(a)) \in \mathcal{R}$ pour tous $a \in A$.*

Théorème 3.5.2 (2-facteur). *Soit G un graphe $2l$-régulier, G peut se décomposer en cycles disjoints, i.e. il existe l permutations $\pi_1, \ldots, \pi_l$ tel que l'ensemble des arêtes est exactement $\{(v, \pi_i(v)) :$ v sommet de $G\}$*

Démonstration. Notons qu'il ne faut que considérer le cas où G est connexe.

Comme les degrés de tous les sommets de G sont pairs, on a démontré dans la Proposition 11 (Annexe) que le graphe possède un circuit d'Euler, c'est-à-dire un circuit qui passe par toutes les arêtes exactement une fois (voir Figure 3.2).

On impose un sens (quelconque) à ce circuit d'Euler $\mathcal{C}$, et on construit un nouveau graphe G' de la façon suivante : pour chaque sommet v de G, on ajoute 2 sommets v^+, v^- dans G', et pour chaque arête (u, v) de G, on connecte u^- et v^+. Donc le graphe G' obtenu a $2n$ sommets et est l-régulier.

Notons maintenant que G' est biparti, il se compose des ensembles G^+ et G^- contenant les v^+ et v^-. Notons que grâce à la l-régularité de G',les conditions du théorème de Hall suivant sont satisfaites. En effet pour chaque sous-ensemble A de G^-, le nombre d'arêtes qui partent de A vaut $l|A|$. Or tous les sommets de G^+ sont de degré $\leq l$, donc l'image de A dans G^+ contient au moins autant d'éléments que A. En appliquant le théorème de Hall, on conclu qu'il existe une bijection de G^- à G^+ qui n'utilise que les arêtes de G'.

On a construit π_1, et en appliquant le même raisonnement $l - 1$ fois de plus, on arrive à la conclusion. □

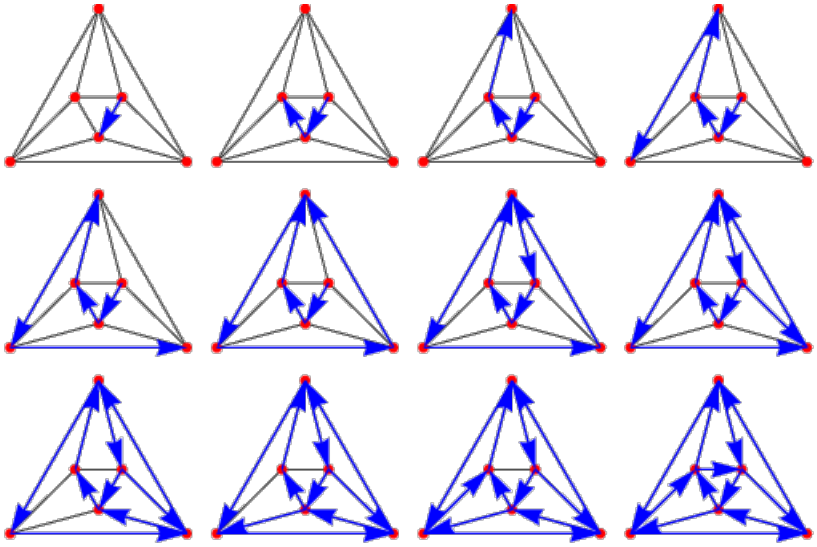


FIGURE 3.2 – Un circuit d'Euler existe si et seulement si l'on peut dessiner le graphe en un coup de crayon

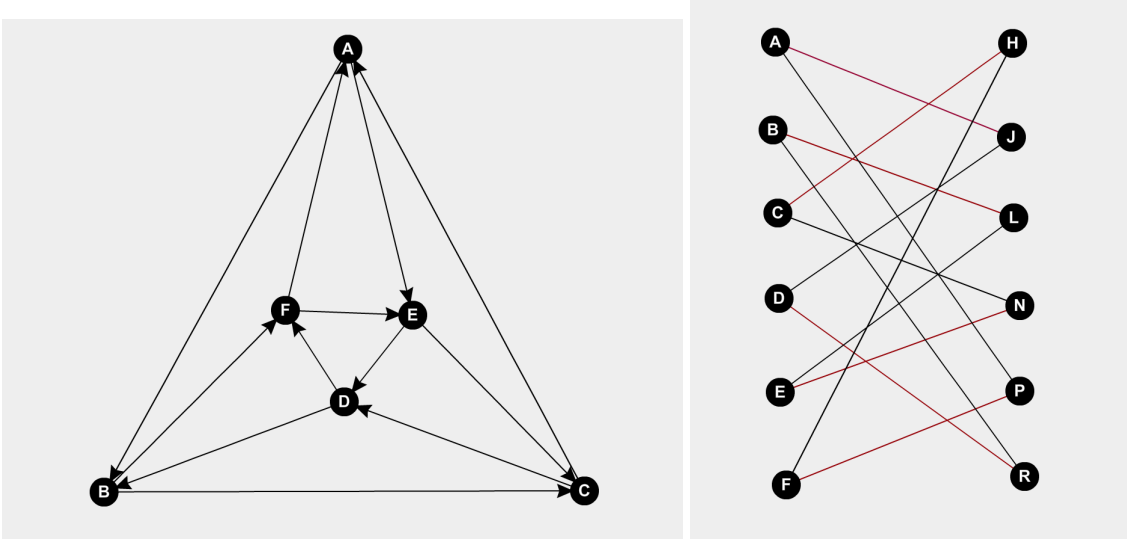


FIGURE 3.3 – Un nouveau graphe biparti G' de 12 sommets (à droite) à partir d'un graphe G de 6 sommets. Le théorème de Hall nous donne une permutation (arêtes en rouge)

On considère $\Omega_n = P_n^l$ comme univers muni de la probabilité uniforme et on prend $G \in \Omega_n$. On doit démontrer qu'il s'agit d'un bon espace, donc que :

- Tous les graphes k-réguliers font partie de Ω_n (Théorème 3.5.2)
- Les éléments de Ω_n sont "plutôt" les graphes k-régulier. En fait, on va démontrer qu'il y a un bonne probabilité indépendant de n qu'un élément de Ω_n soit un graphe k-régulier.
- La distribution engendrée par Ω_n sur l'ensemble de graphes k-réguliers "rassemble" la distribution uniforme

La deuxième propriété requiert quelques connaissances approfondies en probabilités et sera admise ici.

Théorème 3.5.3 (**G est $k-$régulier**). *Soit $G \in \Omega_n$. G est k-régulier avec une probabilité au moins $c > 0$ où c ne dépend que de k*

On a donc prouvé le fait qu'une famille prise au hasard ait une forte chance d'être une famille de graphes expanseurs. Ainsi en pratique les ingénieurs et chercheurs privilégient cette façon de générer des graphes expanseurs.

Avec $\varepsilon > 0$ et $k \in \mathbb{N}$ fixés, on va démontrer qu'en prenant G_n dans Ω_n, la probabilité que G_n soit k-régulier mais pas ε-expanseur tend vers 0 quand n tend vers infini. Rigoureusement, on a des résultats suivants ;

Théorème 3.5.4 (Les graphes aléatoires ont une bonne chance d'être expanseurs). *Il existe $\varepsilon > 0$ ne dépendant que de k tel que la probabilité que G soit k-régulier mais pas ε-expanseur tende vers 0 quand n tends vers infini et k fixé.*

Démonstration. Soit un graphe $G = (V, E)$ $2l-$régulier, $|V| = n$, et $A \subset V$. On définit $N(A) = \{v \in V | v \in A \text{ ou } \exists u \in A, u \sim v\}$. G n'est pas $c-$expanseur, c'est-à-dire :

$$\begin{aligned} & \forall c > 0 \quad , \exists A \subset V, \frac{|\partial A|}{|A|} \leq c \\ \Leftrightarrow \quad & \forall c > 0 \quad , \exists A \subset V, \frac{|N(A)|}{|A|} \leq 1 + c \\ \Leftrightarrow \quad & \forall c > 0 \quad , \exists A, B \subset V, |B| = (1+c)|A|, \forall i \leq l, \pi_i(A) \subset B \end{aligned}$$

On peut calculer la probabilité. En notant $S = \{$ Un graphes$2l-$régulier n'est pas $c-$expanseur$\}$ et on suppose que pour chaque permutation, il a probabilité uniforme. Donc

$$\begin{aligned} \mathbb{P}(S) \quad & \leq \quad \sum_{r < \frac{n}{2}} \sum_{A,B} \mathbb{P}(\pi_i(A) \subset B, \forall i < l) \\ & \leq \quad \sum_{r < \frac{n}{2}} \sum_{A,B} \Pi_{i=1}^{l} \mathbb{P}(\pi_i(A) \subset B) \\ & \leq \quad \sum_{1 \leq r \leq \frac{n}{2}} C_n^r C_{n-r}^{r'} \left(\frac{r + r'}{n} \right)^{rl} \end{aligned}$$

où $r' = \lfloor cr \rfloor + 1$. Après en utilisant la formule de Sterling, on fait une estimation très fine :

$$\mathbb{P}(S) \leq C \left[\left(\frac{1.7}{\sqrt{n}}^{l-(1+c)} \right) + (0.8)^{[l-(1+c)]\sqrt{n}} \right]$$

On prend c assez petit et quand $n \to \infty$, la probabilité tend vers 0, qui implique que l'on a pas mal de chance de trouver une famille de graphes expanseurs par hasard. □

On a donc prouvé le fait qu'une famille prise au hasard a une forte chance d'être une famille de graphes expanseurs. Ainsi en pratique les ingénieurs et chercheurs privilégient cette façon de générer des graphes expanseurs.

Chapitre 4

Graphes expanseurs et codes correcteurs d'erreurs

4.1 Principe des codes correcteurs

4.1.1 Principe général

Les graphes expanseurs peuvent être utilisés dans la pratique pour obtenir des codes correcteurs d'erreurs performants. Ces codes sont utilisés aussi bien dans le stockage d'information, comme dans les CD ou les DVD pour les protéger en cas de rayures par exemple, que dans les communications avec les satellites, les téléphones ou les modems. Nous présenterons dans un premier temps le principe des codes correcteurs d'erreurs, puis présenterons plus en détail la façon d'obtenir un code correcteur performant à partir d'un graphe expanseur.

Pour modéliser un échange d'information entre deux personnes, on appelle émetteur celui qui envoie le message, et récepteur celui qui le reçoit. Le message est transmis au travers d'un canal, mais l'imperfection de ce canal, appelée bruit, peut parfois modifier le message en chemin. Pour modéliser ce bruit, on part du principe que l'information est envoyée à l'émetteur sous forme de mots de même longueur k constitués de bits 0 ou 1. On appelle *canal binaire symétrique* un canal où chaque bit a une probabilité p d'être modifié en cours de route. Par ailleurs, l'apparition d'erreurs est supposée indépendante pour chaque bit transmis.

Pour pallier le problème de l'apparition d'erreurs, et donc de la réception de mauvais messages, on peut coder le message de l'émetteur avant de l'envoyer au travers du canal, en lui ajoutant de l'information redondante. Ainsi, au lieu d'envoyer un mot de longueur k, on envoie un mot de longueur $n > k$, appelé mot du code, tel que la redondance ajoutée permette de corriger et de détecter les éventuelles erreurs par le décodage du message. L'ensemble des mots de longueur n après codage constitue le code, et les processus de codage et décodage du message permettent la détection et la correction d'erreurs.

4.1.2 Définitions

Définition 4.1.1. Un code C est un ensemble de chaînes binaires de longueur n. Ses éléments sont appelés les mots du code.

Remarque 8. *Par la suite nous noterons indifféremment F_2^n ou $\{0,1\}^n$ pour désigner $\{0,1\}^n$ vu comme espace vectoriel.*

La distance utilisée sur le code est appelée la distance de Hamming.

Définition 4.1.2. La **distance de Hamming** entre x et $y \in F_2^n$, qu'on notera $d(x,y)$, est le nombre de coordonnées où x et y diffèrent.

$$d(x,y) = |\{i : x_i \neq y_i\}|$$

Exemple : $d(0110, 1101) = 3$

Définition 4.1.3. La **distance d'un code** $C \subset F_2^n$, notée $dist(C)$, est la distance de Hamming minimale entre deux mots différents du code.

$$dist(C) = \min_{x \neq y \in C} d(x,y)$$

Remarque 9. *Plus la distance d'un code est grande, plus il est résistant aux erreurs, car il faut d'avantage pour passer d'un mot du code à un autre.*

Définition 4.1.4. Taux d'un code
Soit $C \subset F_2^n$ un code transformant les mots de longueur k en des mots de longueur n. Le **taux** de C est défini par :

$$rate(C) = \frac{k}{n} = \frac{log_2(|C|)}{n}$$

Remarque 10. *Le taux d'un code est toujours inférieur à 1. Par ailleurs, plus il est petit, plus on a ajouté d'information redondante, et plus les mots qu'on envoie sont longs.*

Ces définitions vont nous permettre d'énoncer trois propriétés essentielles caractérisant un code correcteur performant :

- La **fiabilité** : Le code est fiable s'il minimise la distance de Hamming entre le message de l'émetteur et le message reçu par le récepteur, c'est-à-dire le nombre d'erreurs entre les deux messages.
- L'**efficacité** : Plus on introduit de redondance dans un mot du code, plus le message est long à transmettre, et moins on peut transmettre de mots. Ainsi il est important d'optimiser l'utilisation du canal en maximisant le taux $r = \frac{k}{n}$, i.e. le rapport entre la taille du message initial et la taille du message après ajout de redondance, tout en conservant assez de redondance pour pouvoir corriger les erreurs.
- La **faisabilité** : pour qu'un code correcteur soit performant, il est essentiel que le codage et le décodage puissent s'effectuer en temps polynomial, ce qui s'avère primordial lorsque la taille des données à traiter est très grande.

Remarque 11. *En réalité, il est plus intéressant d'utiliser des familles de codes correcteurs afin de pouvoir coder des mots de longueurs différentes, potentiellement très grands, et d'étudier plus facilement les propriétés asymptotiques de ces codes.*

4.1.3 Exemples

Donnons maintenant quelques exemples pour clarifier ces notions :

4.1.3.1 Code par répétition

Une façon naïve de corriger et de détecter des erreurs de transmission est d'envoyer plusieurs fois la même information. Par exemple, pour envoyer le mot 1011, on peut répéter chaque lettre un certain nombre de fois.

$$1011 \underset{Codage}{\rightarrow} 11111\ 00000\ 11111\ 11111 \underset{Transmission}{\rightarrow} 11100\ 00010\ 11010\ 11101 \underset{Decodage}{\rightarrow} 1011$$

On peut alors reconstituer le mot d'origine en gardant la lettre qui apparaît le plus souvent. On remarque que si plus de 3 erreurs surviennent dans une même chaîne de 5 bits, cela conduirait à mot décodé différent.

Bien sûr, plus le nombre de répétitions est grand, plus le code est fiable. Cependant ce code est n'est pas efficace puisque son taux est de 1/5. Le canal de transmission est donc utilisé très longtemps pour envoyer un petit message, ce qui est loin d'être idéal.

4.1.3.2 Code de parité

Étudions un deuxième exemple de code correcteur appelé code de parité, qui permet de bien comprendre le principe de codage de l'information.

Un code de parité consiste à ajouter à la suite de chaque mot un bit de plus correspondant à la parité de ce mot, c'est à dire 0 si le mot contient un nombre pair de bits valant 1, et 1 sinon. Par exemple, le mot 1010110 sera codé par 10101100, alors que le mot 1111100 sera codé par 11111001.

À la réception, le dernier bit permet de contrôler l'absence d'erreur. En effet, si le mot reçu est 11111000, le récepteur sait qu'une erreur s'est introduite dans le message.

Bien que le taux du code soit bien plus satisfaisant que l'exemple ci-dessus, ce code n'est pas optimal. Tout d'abord, il ne permet pas de savoir où se trouve l'erreur dans le message. De plus, il ne permet pas de détecter les erreurs si elles sont en nombre pair. Par exemple, si 10101010 devient 11111010, le bit de parité reste correct. Enfin, si c'est le bit de parité lui-même qui est modifié pendant la transmission, le décodage peut détecter une erreur alors qu'il n'y en a pas.

Intéressons-nous donc à la probabilité de détecter une erreur : les erreurs sont indépendantes les unes des autres, avec une probabilité p pour chaque bit d'être modifié. Cette situation correspond donc à un schéma de Bernoulli, et la probabilité d'avoir k erreurs dans le message suit une loi binomiale de paramètres n (la longueur du mot) et p. Notons X le nombre d'erreurs détectées. Ainsi :

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$$

La probabilité qu'une anomalie soit détectée est, d'après les considérations ci-dessus :

$$p_{det} = P(X = 1) + P(X = 3) + P(X = 5) + P(X = 7)$$

De plus, la probabilité que le message contienne au moins une erreur est de :

$$p_{err} = 1 - P(X = 0) = 1 - (1 - p)^8$$

Pour $p = 0,1$ par exemple, on obtient que $p_{det}/p_{err} = 0,74$ des messages erronés sont détectés. Comme plus d'un quart des erreurs qui surviennent ne sont pas détectées, cet exemple montre l'importance de diminuer le nombre de messages erronés non reconnus.

4.1.3.3 Code de parités croisées

Une façon d'améliorer le codage ci-dessus est d'utiliser un code appelé code de parités croisées. Il s'agit d'un double codage par bit de parité : on range les L mots transmis dans un tableau de L lignes, puis on complète chaque ligne et chaque colonne par un bit de parité, comme sur le schéma ci-dessous, avec $L = 3$:

$$\begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix}$$

A la réception, l'étude de la parité de chaque ligne et chaque colonne permet de détecter certaines erreurs. L'intérêt de ce code est que tous les messages ne comprenant qu'une seule erreur sont détectés et corrigés. Par exemple :

$$\begin{pmatrix} 0 & 0 & 1 & \underline{0} \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & \underline{1} & 0 & 1 \end{pmatrix}$$

Cependant, quand plusieurs erreurs apparaissent, elles ne peuvent pas toujours être corrigées ni même détectées, comme par exemple si elles apparaissent sur la même ligne comme ci-dessous :

$$\begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & \underline{1} & \underline{0} & 1 \end{pmatrix}$$

Dans ce cas, l'erreur peut être détectée mais pas localisée.
De même lorsque les erreurs sont disposées en carré comme ceci :

$$\begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix}$$

Dans ce cas, on ne peut même pas détecter l'existence d'erreurs. Si X est la variable aléatoire comptant le nombre de bits mal transmis dans un message, la probabilité d'exactitude p_{exa} d'un message après décodage est telle que :

$$p_{exa} \geq P(X = 0) + P(X = 1)$$

Avec les mêmes hypothèses d'indépendance et d'équiprobabilité que dans l'exemple précédent, X est une variable aléatoire de loi binomiale pour les paramètres : $n = 16$ et $p = 0,1$, et l'on a :

$$P(X = 0) + P(X = 1) = (1 - p)^n + np(1 - p)^{n-1} \simeq 0.51$$

Comparons sur cet exemple la probabilité d'exactitude d'un message décodé à celle à la réception exacte du mot d'information non codé. Ce dernier, qui est de longueur $k = 9$, est alors transmis directement avec la probabilité de transmission sans erreur :

$$p(0) = (1 - p)^k \simeq 0.38$$

On obtient :

$$\frac{p_{exa} - p(0)}{p(0)} \geq 0.35$$

soit une amélioration d'au moins 35%.

Remarquons cependant que le codage de parités croisées, qui nécessite un certain nombre de bits de redondance, a l'inconvénient d'augmenter le temps nécessaire aux opérations de codage et de décodage.

Ces exemples, dont la performance est limitée, permettent tout de même de comprendre le fonctionnement d'un code correcteur d'erreur. Après ces considérations générales, étudions l'intérêt de l'utilisation de codes linéaires.

4.2 Codes linéaires

4.2.1 Motivations

Puisque les mots sont des suites de k bits, tout mot peut s'identifier à un vecteur appelé vecteur binaire, dans l'espace vectoriel $\{0,1\}^k$ muni de l'addition bit par bit (modulo 2) et de la multiplication par un scalaire dans $\{0,1\}$. Par exemple $10110 + 00010 = 10100$ et $1 \times 00010 = 00010$. Notons que l'addition et la soustraction ne sont qu'une seule et même opération dans $\{0,1\}^k$.

On appelle *fonction de codage* la fonction qui à un mot associe son homologue codé. Comme on se trouve dans un espace vectoriel $\{0,1\}^k$, il est intéressant de choisir une application linéaire pour cette fonction de codage. Ainsi tous les mots du code sont connus facilement dès qu'on connaît le codage d'une base de $\{0,1\}^k$, et l'image de cette base est une base du code.

Définition 4.2.1. Code linéaire
Un code $C_n \subset F_2^n$ est dit linéaire de dimension k et de longueur n s'il a une structure d'espace vectoriel et a 2^k éléments.

4.2.2 Code linéaire systématique

Définition 4.2.2. Code systématique
On appelle code systématique un code pour lequel tout mot du code se construit en prolongeant le mot d'information par de l'information redondante, appelée clé de contrôle.

Dans l'exemple du code de parité ci-dessus, le bit de parité constitue la clé de contrôle du mot d'information. Comme la fonction de codage est linéaire, la clé de la somme de deux mots est la somme des clés de ces deux mots. Par exemple $010\mathbf{1} + 011\mathbf{0} = 001\mathbf{1}$, où la clé est indiquée en gras.

Pour un codage systématique, on peut représenter la fonction de codage sous forme d'une matrice :

$$\begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & 0 & \ddots & 1 \\ c_{11} & c_{21} & \dots & c_{k1} \\ \vdots & \ddots & \ddots & \vdots \\ c_{1r} & c_{2r} & \dots & c_{kr} \end{pmatrix}$$

La partie inférieure est appelée matrice des clés de la base canonique, et $r = n - k$. Prenons un exemple. Soit le code linéaire de matrice génératrice :

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}$$

Ici : $k = 3, n = 5$ et $r = 2$. À chaque mot de longueur 3 on ajoute une clé de longueur 2.

On cherche le codage du mot d'information (011). Il s'obtient en multipliant ce mot par la matrice génératrice. Ainsi le mot codé est (01111).

Maintenant que nous savons comment coder les mots, se pose ensuite le problème du contrôle des messages. Étant donné un message reçu m, comment savoir s'il appartient ou non au code ? Nous allons introduire pour cela la matrice de contrôle qui permet de vérifier simplement l'appartenance d'un mot au code.

On considère dans ce qui suit que les espaces $\{0,1\}^n$ sont munis du produit scalaire canonique.

Définition 4.2.3. Code orthogonal
On appelle code orthogonal le code formé par l'ensemble des vecteurs de $\{0,1\}^n$ orthogonaux aux vecteurs du code. Ce code est linéaire, de longueur n et de dimension $(n - r)$.
En d'autres termes, si f est la fonction de codage linéaire, le code orthogonal, qu'on note souvent par $C^\perp$ est tel que : $C^\perp = (Im(f))^\perp$

Grâce au code orthogonal, on peut caractériser l'appartenance d'un mot au code de la façon suivante :

Propriété 1. *Soit G' une matrice génératrice du code orthogonal. Soit c un mot de longueur n. Alors c appartient au code si et seulement si ${}^tG'c = 0$.*

Définition 4.2.4. Matrice de contrôle
On appelle matrice de contrôle une matrice dont le noyau est constitué des mots du code.

D'après ce qui précède, tout code linéaire admet pour matrice de contrôle la transposée d'une matrice génératrice de son code orthogonal. La matrice de contrôle permet donc de vérifier de manière efficace l'appartenance d'un mot au code.

Dans notre cas, la matrice du code orthogonal est relativement simple à trouver. En effet la matrice obtenue par juxtaposition de la matrice des clés à gauche et de la matrice identité de taille $n-k$ à droite convient.

$$\begin{pmatrix} c_{11} & c_{21} & \dots & c_{k1} & 1 & \dots & 0 \\ \vdots & \ddots & \ddots & \vdots & 0 & \ddots & \vdots \\ c_{1r} & c_{2r} & \dots & c_{kr} & 0 & \dots & 1 \end{pmatrix}$$

Le codage systématique donne une première idée du fonctionnement d'un code linéaire pour lequel les clés de contrôle sont clairement identifiées. Cependant cette idée peut être généralisée.

4.2.3 Cas général

L'intérêt du codage systématique est de placer l'information en évidence dans le mot du code. Mais en réalité toute application linéaire injective f de $\{0,1\}^k$ dans un sous-espace vectoriel de $\{0,1\}^n$ de dimension k définit un code linéaire. Un code linéaire de longueur n et dimension k se construit donc à l'aide d'une matrice G, appelée matrice génératrice, dont les k vecteurs colonnes forment un système libre dans $\{0,1\}^n$, donc une base du code. Un code ayant plusieurs bases, il peut donc avoir plusieurs matrices génératrices.

Prenons un exemple. Si on cherche à construire un code de longueur 5 et de dimension 3, il faut choisir trois vecteurs de $\{0,1\}^5$ linéairement indépendants, comme 11100, 01111 et 10110. Le code admet donc pour matrice génératrice :

$$\begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

Dans ce cas, (100) est codé par (111000), (010) par (01111) et (001) par (10110). Plus généralement, tout mot d'information $i \in \{0,1\}^3$ est codé par Gi, le produit de la matrice G par le vecteur i.

De même que pour le codage systématique, la matrice de contrôle permet de vérifier facilement qu'un mot appartient bien au code. Il est intéressant de noter que le code ayant plusieurs matrices génératrices, il a aussi plusieurs matrices de contrôle.

4.2.4 Algorithme de décodage

Intéressons-nous maintenant au décodage d'un message. Que se passe-t-il si le produit de la matrice de contrôle et du mot reçu n'est pas nul, c'est-à-dire si le mot n'appartient pas au code ? Comment corrige-t-on l'erreur ?

On utilise un algorithme de décodage qui va trouver le mot x le plus proche : après avoir reçu la chaîne binaire y de longueur n, et tant qu'il y a une variable dont les contraintes des voisins ne sont pas satisfaites, on change la valeur de cette variable. En d'autres termes, étant donné un mot y n'appartenant pas au code, on change sa i^{eme} coordonnée tant que $w(A(y + e_i)) < w(Ax)$ (où w désigne le poids), où A est la matrice de contrôle de $C(G)$. Cet algorithme est connu sous le nom de *Belief Propagation*.

Maintenant que nous comprenons mieux le fonctionnement et l'utilité des codes linéaires, intéressons-nous au lien entre les codes et les graphes.

4.3 Lien entre les codes correcteurs et les graphes expanseurs

4.3.1 Définitions

Après avoir exposé le principe des codes correcteurs d'erreurs et montré l'avantage que présente l'utilisation des codes linéaires, nous sommes maintenant en mesure de faire le lien entre les codes et les graphes.

Commençons par rappeler quelques définitions :

Définition 4.3.1. Un graphe $G = (V, E)$ est défini par la donnée d'un ensemble de sommets V et d'un ensemble d'arêtes E, chaque arc étant une paire de sommets. Un graphe est dit non-orienté si les arêtes peuvent être parcourues indifféremment dans un sens ou dans l'autre. Un graphe est dit régulier si chaque sommet a le même nombre d'arêtes.

Remarque 12. *On ne considère pas ici les multigraphes, c'est-à-dire qu'il n'y a pas de boucles sur un sommet ou d'arêtes multiples entre deux sommets.*

On trouvera sur la figure (4.1) un exemple de graphe non-orienté 3-régulier. On remarque qu'en numérotant les points du graphe on peut lui associer une matrice $M = (a_{ij})$ où le coefficient a_{ij} vaut 1 si les sommets i et j sont reliés par une arête, et 0 sinon.

Ainsi sur l'exemple ci-dessus la matrice associée au graphe est :

$$\begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix}$$

Cette matrice est carrée, symétrique (puisque le graphe est non-orienté), avec des 0 sur la diagonale (car un sommet n'est jamais relié à lui-même).

Étudions maintenant un type de graphe particulier appelé graphe biparti .

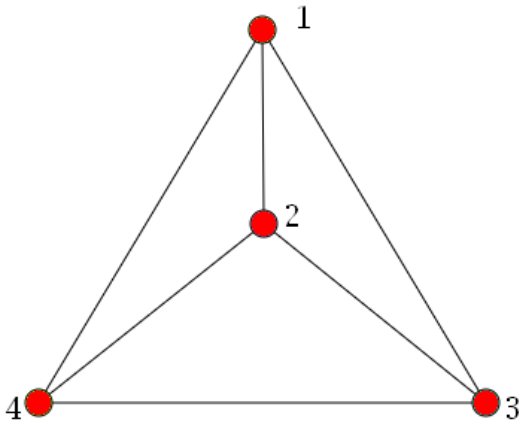


FIGURE 4.1 – Graphe 3-régulier

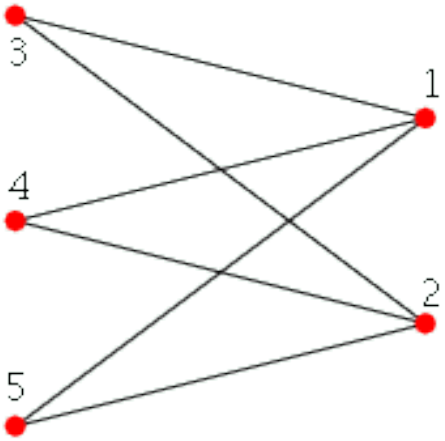


FIGURE 4.2 – Graphe biparti

Définition 4.3.2. Graphe biparti
Un graphe $G = (V_L, V_R, E)$ est dit biparti s'il existe une partition de son ensemble de sommets en deux sous-ensembles V_L et V_R telle que chaque arête ait une extrémité dans V_L et l'autre dans V_R.

Dans la figure (4.2) on trouvera un graphe biparti, sa matrice d'adjacence est :

$$\begin{pmatrix} 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \end{pmatrix}$$

La matrice de ce graphe est très particulière. Elle est définie par blocs, avec deux blocs nuls, et les deux autres blocs sont les transposés l'un de l'autre. C'est ce bloc non nul en bleu sur l'exemple qui va être intéressant pour faire le lien avec les codes : en effet ce bloc issu

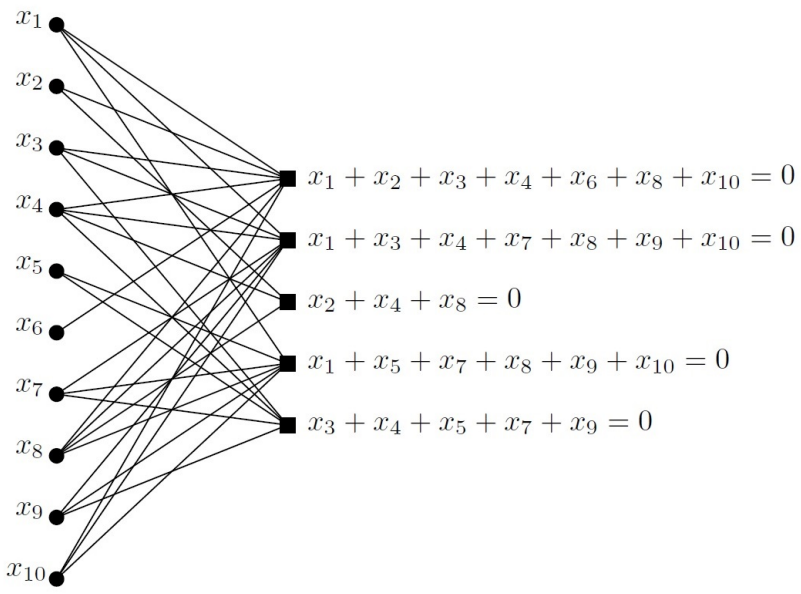


FIGURE 4.3 – Graphe de contraintes d'un code

d'une matrice d'un graphe biparti est utilisé comme matrice de contrôle d'un code linéaire, c'est-à-dire qu'un mot appartiendra au code si et seulement si le produit de ce mot par le bloc est nul. C'est ce que nous allons expliciter dans ce qui suit.

4.3.2 Codes correcteurs d'erreurs et graphes

Expliquons formellement comment obtenir des codes linéaires à l'aide de graphes bipartis.

Supposons que G est un graphe biparti avec n nœuds à gauche (appelés nœuds messages) et r nœuds à droite (appelés nœuds de contrôle). Ce graphe peut générer un code linéaire de longueur n et de dimension au moins $n - r$ de la façon suivante : les n nœuds messages (à gauche) correspondent aux n coordonnées d'un mot du code. Les mots du code sont alors les vecteurs $(x_1, \ldots, x_n)$ tels que pour chaque nœud de contrôle, la somme des coordonnées correspondantes à ses voisins est nulle modulo 2, comme représenté sur la figure (4.3).

De manière plus pratique, et en regardant la matrice d'adjacence du graphe $A = (a_{ij})$ de taille $r \times n$, le code linéaire défini par le graphe est alors l'ensemble des vecteurs $x = (x_1, \ldots, x_n)$ tels que $A.x^t = 0$. La matrice A est ainsi la matrice de contrôle du code créé. Inversement, chaque matrice binaire de taille $r \times n$ peut aussi générer un graphe biparti entre n messages et r nœuds de contrôle, et le code défini comme le noyau de cette matrice sera le code associé à ce graphe.

4.3.3 La performance apportée par les graphes expanseurs

Les codes correcteurs d'erreurs performants allient deux propriétés contradictoires : ils contiennent beaucoup de mots, mais ces mots doivent être très distants les uns des autres,

au sens de la distance de Hamming. En effet, plus les mots sont distants, plus le code est robuste car le passage d'un mot à un autre demande beaucoup de modifications. Or les graphes expanseurs possèdent eux aussi deux propriétés contradictoires puisqu'un graphe expanseur est bien connecté mais contient peu d'arêtes (on dit qu'il est peu dense).

C'est l'équilibre apporté par les graphes expanseurs entre deux propriétés antagonistes qui va permettre d'obtenir des codes correcteurs performants.

Considérons un graphe biparti $G = (V_L, V_R, E)$ dont la partie gauche notée $V_L$est K-régulière (c'est-à-dire que chaque sommet a k arêtes sortantes), et tel que $|V_L| = n$ et $|V_R| = m$ avec $m < n$.

Remarque 13. *On notera $C(G)$ le code issu du graphe G.*

On définit alors le taux d'expansion des points à gauche :

Définition 4.3.3. Taux d'expansion à gauche
Le taux d'expansion des points à gauche noté $L(G, d)$ (pour *left vertex expansion ratio* en anglais) est le minimum de $\frac{|\Gamma(S)|}{|S|}$ calculé pour tout sous-ensemble non vide $S \subset V_L$ vérifiant $|S| \leqslant d$, où $\Gamma(S)$ désigne l'ensemble des voisins des points de S.
Autrement dit, un tel ensemble vérifie $|\Gamma(S)| \geqslant L(G, d)|S|$.

$$L(G, d) = \min_{S \subset V_L,\, |S| \leqslant d} \frac{|\Gamma(S)|}{|S|}$$

On note que $L(G, d)$ ne peut dépasser k pour aucun graphe k-régulier à gauche et aucun d. On présentera comment de tels graphes avec un "grand" $L(G, d)$ génèrent des codes asymptotiquement bons et comment ils permettront un décodage efficace.

Remarque 14. *On remarque que ce taux d'expansion à gauche n'est autre que la constante de Cheeger pour la partie gauche du graphe.*

On aura besoin, dans ce qui suit, d'utiliser la notion de poids.

Définition 4.3.4. Poids
Le poids d'un mot $x \in F_2^n$ est égal au nombre de coordonnées de x qui sont non nulles. On le note $w(x)$ et on a : $w(x) = d(x, 0)$

Remarque 15. *Par linéarité, la distance d'un code est égale au poids minimal de ses mots non nuls.*

Théorème 4.3.1. *Le code $C(G)$ a une grande distance*
Si $L(G, d) > \frac{k}{2}$ alors $dist(C(G)) \geqslant d$

Démonstration. Montrons que pour chaque sous-ensemble $S \subset V_L$ tel que $|S| \leqslant d$ on peut trouver un "unique voisin" dans les sommets à droite, c'est-à-dire qu'il existe $v \in V_R$ tel que $[\Gamma(v) \cap S| = 1$. En effet, si $E(S, \Gamma(S))$ désigne l'ensemble des arêtes entre S et $\Gamma(S)$, alors la K-régularité à gauche fait que $|E(S, \Gamma(S))| = K|S|$, mais en même temps $|\Gamma(S)| > \frac{K|S|}{2}$ par hypothèse car $S \subset V_L$ et $|S| \leqslant d$.

Ainsi pour un nœud dans $\Gamma(S)$, le nombre moyen d'arêtes partant de ce point vers S est inférieur strictement à 2, et comme tous les nœuds dans $\Gamma(S)$ ont déjà au moins un voisin dans S, alors il existe $v \in \Gamma(S)$ tel que $[\Gamma(v) \cap S| = 1$.

Utilisons cette propriété pour montrer que tout mot du code non nul $x \in C(G)$ a un poids supérieur à d. Soit $S \subset V_L$ le support de x, i.e. l'ensemble des coordonnées i tels que $x_i = 1$. Si $|S| \leq d$, par ce qui précède $\Gamma(S)$ contient un nœud v qui a un seul voisin dans S, mais alors la v-ème coordonnée de $A.x$ n'est pas nulle. Ceci contredit le fait que $x \in C(G)$. □

Montrons maintenant que le décodage se fait efficacement en temps linéaire.

Théorème 4.3.2. *Décodage efficace*
Soit G un graphe biparti K-régulier à gauche vérifiant $L(G, d) > \frac{3}{4}K$, et y un mot de longueur n vérifiant $d(x, y) \leqslant \frac{d}{2}$. Alors l'itération de l'algorithme précédent sur y retournera x après un nombre d'itérations au plus linéaire en n.

Démonstration. On rappelle que A est la matrice de contrôle du code. On indexera parfois par des sommets v du graphe, il faut comprendre que l'indice est dans ce cas le numéro du nœud dans son ensemble de droite ou de gauche.

Soit $y^{(i)}$ le vecteur (ou le mot de longueur n) généré par l'algorithme après i itérations, avec $y^{(0)} = y$. Et notons D_i l'ensemble des erreurs à la i-ème itération, i.e. $D_i = \{v | y^{(i)} \neq x_v\}$.

On doit alors prouver que D_t est vide pour $t = O(n)$. Considérons l'ensemble D_i, et supposons que D_i n'est pas vide et que $|D| \leq d$. On partitionne $\Gamma(D_i)$ en deux ensembles : l'ensemble des voisins satisfaits S et l'ensemble des voisins non satisfaits U. C'est-à-dire que U est le support du vecteur $A.y$. On a alors :

$$|S| + |U| = |\Gamma(D_i)| > \tfrac{3}{4}K|D_i|$$

Comptons maintenant le nombre d'arêtes entre D_i et $\Gamma(D_i) = S \cup U$. Il y a au moins $|U|$ arêtes sortantes de U et au moins $2|S|$ sortantes de S. Pour U c'est immédiat par connexité du graphe. Pour S on prouve qu'il doit y avoir un nombre pair d'arêtes de S vers $|D_i|$. Notons que pour un nœud $v \in S$ on a $(A.y^{(i)})_v = 0$ car v fait partie des voisins satisfaits. S'il n'y a qu'une seule arête de v vers D_i alors $(A.y^{(i)})_v = 1 + (A.x)_v = 1$ et on a donc une contradiction. Le résultat est le même si on suppose un nombre impair d'arêtes. On conclut donc que chaque nœud v de S a un nombre pair de voisins dans D_i. Ainsi :

$$|U| + 2|S| \leq K|D_i|$$

Une combinaison linéaire des deux dernières inégalités donne :

$$|U| > \tfrac{1}{2}K|D_i|$$

Par conséquence, il existe une variable dans D avec plus de $\frac{1}{2}K$ voisins non satisfaits. Ceci implique que tant qu'il y a des erreurs ($y^{(i)} \neq x$) et que $|D_i| \leq d$, une variable sera changée par l'algorithme. Ce changement transmet tous les nombre de $|U|$ vers $|S|$ et vice versa. Mais comme ce changement s'opère sur un noeud avec plus de voisins non satisfaits que de voisins satisfaits ($|U| > \frac{1}{2}K|D_i| \geq |S|$), on obtient que $w(A(y^{(i)} + e_j)) < w(Ay^{(i)})$ où j correspond au numéro du nœud v.

Finalement, ceci implique que $|U|$ décroit avec chaque itération de l'algorithme. Ainsi, si la distance entre $y^{(i)}$ et x ne dépasse jamais d, alors l'algorithme s'arrête en donnant x comme résultat après un nombre linéaire d'itérations.

Pour terminer la preuve du théorème, il faut démontrer qu'on a toujours $|D_i| \leq d$. Rappelons que par hypothèse $|D_0| \leq \frac{d}{2}$, ce qui implique que $|U_0| \leq |\Gamma(D_0)| \leq K\frac{d}{2}$. En effet, après chaque itération, $|D_i|$ change de ± 1. Si après une itération $|D_i|$ dépasse d, il existe un indice j tel que $|D_j| = d$ et donc vérifiant $|U_j| > \frac{1}{2}K|D_j| = \frac{1}{2}Kd$. Mais ceci contredit le fait que $|U_0| \leq K\frac{d}{2}$ et que $|U_i|$ décroit avec chaque itération. □

4.3.4 Existence et construction des graphes bipartis expanseurs : les "lossless"

Considérons un graphe biparti G, K-régulier à gauche et tel que $|V_L| = N$ et $|V_R| = M$. On dira que G est un (D, γ)-expanseur si $L(G, D) > \gamma$. On avait déjà établi que l'expansion à gauche sera toujours inférieure à K, ainsi le mieux qu'on puisse espérer est d'avoir une constante γ très proche de K. Lorsque $\gamma = (1 - \epsilon)K$ pour un ϵ petit on qualifie le graphe expanseur de "lossless". On ne peut atteindre une telle expansion que lorsque $D \simeq \frac{M}{K}$

Un code correcteur d'erreurs (de longueur N) issu de ces graphes expanseurs "lossless" vérifiera les hypothèses des théorèmes (4.3.1) et (4.3.2), aura donc une distance supérieure à D et effectuera le décodage en un temps au plus linéaire. De plus, ce code aura un taux supérieur à $1 - \frac{M}{N}$, qu'on peut rendre arbitrairement proche de 1 à degré K constant, et qui pour $\kappa = \frac{M}{N}$ suffisamment petit, s'approchera de la borne de Gilbert-Varshamov. Cette borne restreint les valeurs possible du taux (voir corollaire (C.1.1) en annexe). Le théorème qui suit assure l'existence de ces graphes :

Théorème 4.3.3. *Pour chaque triplets de suites ϵ_n, $M_n \leq N_n$, il existe une famille explicite de graphes G_n bipartis K_n-réguliers à gauche qui sont des $(D_n, 1 - \epsilon_n)$-expanseurs "lossless", où $K_n \leq (N_n/\epsilon M_n)^c$ pour une certaine constante c et avec limsup$_n$ $K_n/(\epsilon_n M_n/D_n) > 0$.*

Commentaire : Le cas qui est intéressant pour notre application est de choisir $\epsilon_n = \epsilon$ petit fixé, et de fixer aussi $N_n = N$ et $M_n = M = \kappa N$ avec δ petit pour avoir un taux proche de 1. Dans ce cas le théorème nous assure l'existence d'une suite de graphes bipartis avec $N \longrightarrow \infty$ et $M <= \kappa N$ K-réguliers (pour un K fixe) et $(D, 1-\epsilon)$-expanseurs avec $D \geq c'N$ pour une constante $c' > 0$

Construire des graphes expanseurs "lossless" reste un défi majeur. Cependant, Capalbo, Reingold, Vadhan, et Wigderson ont pu exhiber une construction qui fait un pas important dans cette direction. Pour tout $\delta > 0$ et pour un d suffisamment grand, c'est une construction explicite de familles de graphes bipartis expanseurs dont le degré à gauche est d et dont l'expansion à gauche vaut $(1 - \delta)d$ pour les sous ensembles petits. La construction est basée sur une généralisation du concept du "produit Zig-Zag", qui est une opération sur les graphes à des objets mathématiques appelés "Conductors".

Chapitre 5

Retour sur l'expérience

5.1 Contexte

Le projet scientifique collectif est un projet proposé à tous les élèves de deuxième année, visant à travailler collectivement par groupe de cinq à sept élèves afin d'identifier, de poser, et de travailler sur un problème d'envergure. Les sujets sont choisis au mois de juin, puis des échéances sont répartis tout au long de l'année sous forme de rapports écrits ou de présentations orales, jusqu'au rapport final et à la soutenance au mois d'avril et mai.

Chaque groupe est en outre rattaché à un tuteur pouvant apporter un éclairage ou des conseils sur le sujet choisi, à un coordinateur chargé de valider le projet, et à un cadre militaire référant pouvant apporter des conseils sur la gestion humaine.

5.2 Mode d'organisation

Pour élaborer notre projet, nous avons institué, à l'initiative de notre tuteur, un fonctionnement en binôme. Les trois équipes ainsi constituées se sont partagés les domaines étudiés. Régulièrement, chaque sous-groupe a fait part de son avancée au reste de ses camarades. La formation de ces binômes a été réalisée collectivement dans une perspective de mixité des origines et formations précédentes. Nous avons ensuite choisi les sujets selon nos préférences.

Chenlin GU et Manh Tien NGUYEN ont étudié les définitions et propriétés approfondies des graphes expanseurs, ainsi que la construction des graphes de Ramanujan. Van Huy VO et Chloé PAPIN ont étudié la construction de Margulis d'une famille de graphes expanseurs, basée sur la notion de propriété (T) de Kazhdan. Oussama HANGUIR et Mathilde DE LA MORINERIE se sont concentrés sur les applications des graphes expanseurs à l'informatique théorique.

L'organisation du travail était basée sur une mise au point systématique le lundi. À cette occasion, nous déjeunions ensemble pour discuter de nos progrès et difficultés, ainsi que des orientations à prendre. C'était également un moyen pour nous d'identifier de nouveaux problèmes avant d'en faire part à notre tuteur. Nous consignions le contenu de ces réunions afin de s'y référer plus tard et de planifier notre investissement dans la durée. Nous profitions ensuite de l'après-midi pour approfondir nos sujets respectifs par binôme. Chaque binôme

évoluait de manière autonome durant le restant de la semaine. Les deux membres se répartissaient le travail et en discutaient. Comme nous sommes issus de formations et de pays variés, cette organisation est très profitable : nous avons des méthodes de travail et des connaissances différentes, dont la mise en commun est bénéfique. De ce fait, nous apprenons en outre à développer nos aptitudes relationnelles. Il était prévu que chaque binôme rencontre notre tuteur, Monsieur Charles FAVRE, de façon hebdomadaire. Celui-ci travaille dans le laboratoire Laurent Schwartz, et a généreusement accepté de nous encadrer sur ce sujet exigeant. Le créneau du lundi après-midi était privilégié, toutefois nous pouvions également le rencontrer d'autres jours, ou bien le contacter via Skype. Nous avons travaillé sur un projet mathématique essentiellement théorique. Nous souhaitions comprendre les résultats existants et explorer en profondeur ce sujet passionnant. Pour cela, nous avons étudié des documents que nous a proposés Monsieur Charles FAVRE, ou que nous avons trouvés nous-mêmes. Les documents étaient principalement sous forme électronique et ont été partagés par l'intermédiaire du site des Projets Scientifiques Collectifs.

5.3 Objectifs atteints

5.4 Analyse des points positifs et négatifs

Le projet scientifique collectif était pour plusieurs d'entre nous une première expérience de travail en groupe sur un sujet d'envergure, et nous avons beaucoup appris tant sur le plan organisationnel que sur le plan du management. A l'heure du bilan, voici les points positifs et les points perfectibles que nous avons dégagés.

5.4.1 Enseignements positifs

Pour commencer, cette expérience de travail en équipe nous a permis d'améliorer nos capacités organisation et de management. Nous nous sommes appliqués à respecter les échéances imposées, à répartir notre travail sur la durée pour profiter pleinement des six mois de projet, et à respecter notre calendrier de travail.

Par ailleurs, nous nous sommes efforcés de nous répartir les tâches équitablement en fonction des compétences et des affinités de chacun, ce qui a été facilité par notre très grande complémentarité au sein du groupe. En effet nous avons essayé de mettre à profit les capacités mathématiques, informatiques, relationnelles, organisationnelles et fédératrices de chacun, afin de faire fructifier nos talents respectifs. C'est certainement un acquis positif et un réel enrichissement mutuel.

De plus, le PSC a été l'occasion de découvrir le monde de la recherche au travers de l'étude de publications scientifiques, mais également grâce aux rencontres avec des chercheurs des laboratoires de l'école. En particulier, les discussions avec Monsieur Charles FAVRE, notre tuteur, Monsieur Alain Couvreur, chercheur au LIX, ou lors des séminaires de mathématiques, nous ont permis de réaliser l'importance de discuter, échanger, et demander conseil sur un sujet afin d'avoir une approche différente et un éclairage nouveau.

Enfin, le travail de synthèse et de vulgarisation de nos recherches nécessaire à l'élaboration des différents rapports et soutenances nous a appris à prendre du recul sur nos connaissances et à les expliquer le plus clairement possible, auprès d'auditeurs différents, dans des temps impartis variables. Ainsi les trois séminaires de mathématiques des élèves auxquels nous avons participé ont été un excellent entraînement, et nous ont permis de nous familiariser avec la présentation orale.

5.4.2 Points à améliorer

Tout au long de nos recherches, deux types de difficultés se sont présentées à nous.

Tout d'abord, une difficulté technique : l'étude des graphes expanseurs est un sujet très stimulant mais également complexe, et il nous arrivait fréquemment de buter sur une démonstration. Pour surmonter cela, nous discutions régulièrement en groupe afin de trouver à plusieurs la solution d'un problème rencontré par un sous-groupe, et, si cela ne suffisait pas, nous sollicitions notre tuteur. Par ailleurs, les documents sur lesquels nous nous appuyions étaient le plus souvent en anglais, et nous avions parfois des difficultés pour traduire les termes techniques en français. Pour cela nous avons trouvé quelques documents francophones qui nous ont permis de trouver les mots français correspondants lorsqu'ils existaient, sinon nous conservions le terme anglo-saxon. La seconde difficulté majeure que nous avons rencontrée porte sur la coordination. Nos travaux par sous-groupe étant de complexité croissante et avançant dans des directions très différentes, il n'était pas toujours facile de suivre l'avancée des autres sous-groupes. C'est pourquoi nous tachions de nous retrouver régulièrement pour qu'un sous-groupe explique au reste des membres de l'équipe l'avancée de ses travaux sous forme d'un petit exposé. Malheureusement, il était parfois ardu de trouver un créneau libre où tous les membres du groupe pouvaient se retrouver, particulièrement quand les conférences étaient programmées sur la plage horaire habituellement réservée au PSC. La troisième difficulté était une difficulté d'harmonisation lors de la mise en commun de nos travaux pour les différents rapports demandés. En effet nous avions sous-estimé le temps nécessaire pour harmoniser nos recherches, et la rédaction de ces rapports, souvent très chrono-phages, nous stoppait dans nos recherches. C'est pourquoi nous en avons tenu-compte pour l'élaboration du rapport final pour lequel nous avons commencé l'harmonisation très en amont.

Jusqu'ici, les difficultés techniques, de coordination et d'harmonisation que nous avons rencontrées ont donc toujours pu être surmontées grâce au travail de l'ensemble des membres du groupe.

Annexe A

A.1 La suite de démonstration de l'inégalité de Cheeger

Cette inégalité est une version discrète d'une inégalité géométrique. On définit le Laplacian sur un graphe comme la façon suivante :

$$\triangle = \frac{A}{k} - 1;$$

Ceci implique que $sp(\triangle) = \{0 = \tilde{\lambda}_0 \leq \tilde{\lambda}_1 \cdots \tilde{\lambda}_{n-1} \leq 2\}$

En fait, l'inégalité de Cheeger est plutôt geométrique ; Pour le montrer, considérons une mesure m sur les sommets et une autre mesure μ pour les arêtes telles que ;

$$m(v) = \sum_w \mu_{vu}, \forall v \in V$$

Dans un graphe k-régulier, $\mu_{vu} = \mu_{uv} = \frac{1}{k}\mathbf{I}_{v\sim u}$. L'additivité de cette mesure permet d'écrire la mesure d'un ensemble et de sa frontière.

$$m(F) = |F| = \sum_v \mathbf{I}_{v\in F} \tag{A.1}$$

$$\mu(\partial F) = \sum_e \mu_e \mathbf{I}_{e\in\partial F} = \frac{1}{2}\sum_{v\in V}\sum_{w\in V} \mu_{vw}\mathbf{I}_{vw\in\partial F} \tag{A.2}$$

On définit ensuite la constante de Cheeger correspondante à ces mesure, et on établit l'inégalité de Cheeger :

$$\tilde{h}(G) = \inf_{F\subset E, m(F)\leq\frac{1}{2}m(E)} \frac{\mu(\partial F)}{m(F)} = \frac{1}{k}h(G) \tag{A.3}$$

$$\frac{\tilde{\lambda}_1}{2} \leq \tilde{h}(G) \leq \sqrt{2\tilde{\lambda}_1} \tag{A.4}$$

On montre plusieurs lemmes pour la deuxième partie de la démonstration.

Lemme 7. ***Formule de Green***

$$\langle -\triangle f, f\rangle_{l^2(V)} = \langle \nabla f, \nabla f\rangle_{l^2(E)}$$

Où

$$\langle \nabla f, \nabla f \rangle_{l^2(E)} = \frac{1}{2} \sum_v \sum_w (f(v) - f(w))^2 \mu_{vw}$$

Démonstration.

$$\begin{aligned}
\langle -\triangle f, f \rangle_{l^2(V)} &= \sum_v ((1 - \frac{A}{k}) f)(v) f(v) \\
&= \sum_v f^2(v) - \sum_v \sum_w f(w) f(v) \mu_{vw} \\
&= \sum_v \sum_w \left(\frac{(f^2(v) + f^2(w))}{2} - f(w) f(v) \right) \mu_{vw} \\
&= \frac{1}{2} \sum_v \sum_w (f(v) - f(w))^2 \mu_{vw} \\
&= \langle \nabla f, \nabla f \rangle_{l^2(E)}
\end{aligned}$$

□

Lemme 8. ***L'inégalité de Poincaré*** *Soit* $f \perp 1$

$$\langle \nabla f, \nabla f \rangle_{l^2(E)} = \langle -\triangle f, f \rangle_{l^2(V)} \geq \epsilon \langle f, f \rangle_{l^2(V)}$$

Lemme 9. ***Formule de Coarea***

$$\sum_{e \in E} |\nabla f| \mu_e = \int_\infty^\infty \mu(\partial \Omega_t) dt \text{ où } \Omega_t = \{v | f(v) > t\}$$

Démonstration.

$$\begin{aligned}
\int_\infty^\infty \mu(\partial \Omega_t) dt &= \int_\infty^\infty \sum_{e \in E} \mathbf{I}_{\{f(x) > t, f(y) \leq t\}} \mu_{xy} dt \\
&= \sum_{e \in E} \int_\infty^\infty \mathbf{I}_{\{f(x) > t, f(y) \leq t\}} \mu_{xy} dt \\
&= \sum_{e \in E} |f(x) - f(y)| \mu_{xy} = \sum_{e \in E} |\nabla f| \mu_e
\end{aligned}$$

□

Corollaire A.1.1.
Soit f *une fonction sur les sommets telle que* $m(\{f \geq 0\}) \leq \frac{1}{2} m(E)$*, on a :*

$$\sum_{e \in E} |\nabla f| \mu_e \geq \tilde{h}(G) \sum_{v \in V} f(v) m(v)$$

Démonstration.

$$\begin{aligned}\sum_{e\in E}|\nabla f|\mu_e &= \int_{\infty}^{\infty}\mu(\partial\Omega_t)\geq\int_0^{\infty}\mu(\partial\Omega_t)dt\\ &\geq \int_0^{\infty}\tilde{h}(G)m(\Omega_t)dt\\ &= \tilde{h}(G)\sum_{v\in V}\int_0^{\infty}\mathbf{I}_{\{f(v)>t\}}dt\\ &= \tilde{h}(G)\sum_{v\in V}f(v)m(v)\end{aligned}$$

□

Démontrons maintenant la deuxième partie de l'inégalité :

Démonstration. On note f le vectuer propre associé à $\tilde{\lambda}_1$. Le but est de monter que

$$\tilde{\lambda}_1\geq\frac{\tilde{h}^2(G)}{2}$$

On suppose aussi que $m(supp(f^+))\leq\frac{1}{2}m(E)$. Alors, d'après la formule de Green

$$\tilde{\lambda}_1\langle f,f^+\rangle_{l^2(V)}=\langle-\triangle f,f^+\rangle_{l^2(V)}=\langle\nabla f,\nabla f^+\rangle_{l^2(E)}$$

On observe que

$$\langle f,f^+\rangle_{l^2(V)}=\langle f^+,f^+\rangle_{l^2(V)}$$

$$(f(x)-f(y))(f^+(x)-f^+(y))\geq(f^+(x)-f^+(y))^2$$

Donc on a

$$\tilde{\lambda}_1\geq\frac{\langle\nabla f^+,\nabla f^+\rangle_{l^2(E)}}{\langle f^+,f^+\rangle_{l^2(V)}}$$

On remarque que l'inégalité de Poincaré ne marche pas directement ici, mais on va utiliser le corollaire (62 $A.1.1$) pour une estimation plus fine. D'après le corollaire :

$$\begin{aligned}\tilde{h}(G)\sum_{v\in V}(f^+(v))^2m(v) &\leq \sum_{e\in E}\nabla(f^+)^2(e)\mu_e\\ &= \frac{1}{2}\sum_v\sum_w(f^+(w)-f^+(v))(f^+(w)+f^+(v))\mu_{wv}\\ &\leq \frac{1}{2}(\sum_v\sum_w(f^+(w)-f^+(v))^2\mu_{vw})^{\frac{1}{2}}(\sum_v\sum_w(f^+(w)+f^+(v))^2\mu_{wv})^{\frac{1}{2}}\\ &\leq \frac{1}{2}(2\langle\nabla f^+,\nabla f^+\rangle_{l^2(E)})^{\frac{1}{2}}(\sum_v\sum_w(2(f^+(w))^2+2(f^+(v))^2\mu_{wv})^{\frac{1}{2}}\\ &= (2\langle\nabla f^+,\nabla f^+\rangle_{l^2(E)})^{\frac{1}{2}}\langle f^+,f^+\rangle_{l^2(V)}\end{aligned}$$

Cette estimation implique $\frac{\langle\nabla f^+,\nabla f^+\rangle_{l^2(E)}}{\langle f^+,f^+\rangle_{l^2(V)}}\geq\frac{\tilde{h}^2(G)}{2}$, donc $\tilde{\lambda}_1\geq\frac{\tilde{h}^2(G)}{2}$ □

A.2 La théorie des graphes

Proposition 11. *Soit G un graphe dont tous les sommets sont de degré pair, G possède alors un circuit d'Euler*

Démonstration. Prenons le chemin $\mathcal{C}$ le plus long qui ne passe qu'une fois par chaque arête, comme le degré des points de départ et d'arrivée est pair, il s'agit en fait d'un circuit.

On montre ensuite que ce circuit utilise tous les arêtes en partant d'un de ses sommets. En fait, s'il existe un point A dans ce circuit tel que le composant connexe de $G \setminus \mathcal{C}$ qui contient A n'est pas ce point seul, comme tous les sommets sont de degré pair, il existe un circuit $\mathcal{C}'$ (le chemin le plus long par exemple) dans $G \setminus \mathcal{C}$ passant par A. En joignant $\mathcal{C}$ et $\mathcal{C}'$, on arrive à un chemin plus long que $\mathcal{C}$.

Donc $\mathcal{C}$ contient tous les arêtes de sa composante connexe, on déduit que $\mathcal{C}$ passe par tous les sommets de G. □

A.3 Les polynômes de Chebyshev de deuxième espèce

Cette partie vise à démontrer la Proposition 4. On va d'abord commencer par les définitions et les propriétés des polynômes de Chebyshev de deuxième espèce.

Lemme 10. *Pour la famille $\{X_m\}$, les plus grandes racines sont $\alpha_m = 2\cos\frac{\pi}{m+1}$ et on a :*

$$X_k X_l = \sum_{m=0}^{k} X_{k+l-2m} \tag{A.5}$$

$$\frac{X_m}{x-\alpha_m} = \sum_{i=0}^{m-1} X_{m-1-i}(\alpha_m) X_i \tag{A.6}$$

$$\frac{X_m^2}{x-\alpha_m} = \sum_{i=0}^{2m-1} y_i X_i, \quad y_i \geq 0 \tag{A.7}$$

Démonstration. (A.5) peut s'obtenir par récurrence sur k. (A.7) s'obtient de (A.5) et (A.6) et du fait que $X_i(\alpha_m) \leq 0$ quelque soit $i \leq m$. Pour (A.6) :

$$\begin{aligned}
\sum_{i=0}^{m-1} X_{m-1-i}(\alpha_m) X_i &= (x-\alpha_m)\left[X_{m-1}(\alpha_m)X_0 + X_{m-2}(\alpha_m)X_1 + \cdots + X_0(\alpha_m)X_{m-1}\right] \\
&= (x-\alpha_m)X_{m-1}(\alpha_m)X_0 - \alpha_m\left[X_{m-2}(\alpha_m)X_1 + \ldots X_0(\alpha_m)X_{m-1}\right] \\
&+ X_{m-2}(\alpha_m)(X_0+X_2) + \cdots + X_0(\alpha_m)(X_{m-2}+X_m) \\
&= \left[(x-\alpha_m)X_{m-1}(\alpha_m) + X_{m-2}(\alpha_m)\right]X_0 - X_{m-1}(\alpha_m)X_1 + X_m \\
&= -X_m(\alpha_m) + X_m = X_m
\end{aligned}$$

□

On va considérer d'abord le cas où la constante C peut dépendre de ν, c'est-à-dire que :

Proposition 12.
Pour chaque ν à support dans $[-L, L]$ satisfaisant $\int_{-L}^{L} X_m(x)d\nu(x) \geq 0$ et pour tout $m = 1, 2, \ldots$, on a :

$$\nu([2-\varepsilon, L]) > 0$$

Démonstration. Comme $\frac{X_m^2}{x-\alpha_m} < 0$ pour $x < \alpha_m$, on trouve $\nu([-L, \alpha_m]) = 0$. Supposons que $\nu([2-\varepsilon, L]) = 0$, pour m assez grand, $\alpha_m \in [2-\varepsilon, L]$, donc $\nu = 0$ (absurde !). □

A.4 Le théorème de Banach-Alaoglu et la conclusion

Dans cette partie, on cherche à démontrer le Théorème 2.3.3 à partir de la Proposition 4. En utilisant le théorème de Banach-Alaoglu, on va démontrer une borne uniforme pour toute mesure positive ν sur $[-L, L]$. On rappelle d'abord quelques outils d'analyse fonctionnelle.

Théorème A.4.1 (*de représentation de Riesz-Markov*). *L'espace dual de $C([-L, L])$ est l'espace de mesures complexes, boréliennes et régulières sur $[-L, L]$ avec la variation total comme norme.*

On remarque aussi que l'ensemble de mesure de probabilité est un sous-ensemble fermé dans la boule unitaire et donc faiblement* compact dans la topologie, d'après le théorème de Banach-Alaoglu.

Théorème A.4.2. *Soient $L \geq 2$ et $\varepsilon > 0$, il existe $C(\varepsilon, L) > 0$ et une mesure ν sur $[-L, L]$ tels que $\int_{-L}^{L} U_m(x/2)d\nu(x) \geq 0$ pour tout $m = 1, 2, \ldots$, on a*

$$\nu([2-\varepsilon, L]) \geq C$$

Démonstration. Considérons la fonction χ qui s'annule sur $x < L - \varepsilon$, qui vaut 1 sur $x > L$ et linéaire par morceaux. Comme ν est positive, on a :

$$\langle \nu, 1_{[2-\varepsilon, L]} \rangle \geq \langle \nu, \chi \rangle \geq \langle \nu, 1_{[2-\varepsilon/2, L]} \rangle.$$

En interprétant χ comme fonction continue sur l'ensemble B de mesures de probabilité dans $[-L, L]$, par compacité, le minimum de χ sur B est atteint, d'où l'existence de la constante $C(\varepsilon, L)$ uniforme. □

A.5 Connexité de $X^{p,q}$

Dans cette, on va se concentrer sur le diagramme commutatif suivant et expliquer les notations,

$$\begin{array}{ccccc} S_p \subset \Lambda' & \xrightarrow{\tau_q} & H(\mathbb{F}_q) & \xrightarrow{\psi_q} & GL_2(q) \\ {\scriptstyle /\ p^r}\downarrow{\scriptstyle Q} & & \downarrow & & \downarrow{\scriptstyle \phi} \\ \Lambda & \xrightarrow{\pi_q} & H(\mathbb{F}_q)/Z(q) & \longrightarrow & PGL_2(q) \end{array}$$

puis conclure l'existence de $Y^{p,q}$ dans certaines conditions. La première étape est la définition de Λ'.

Définition A.5.1. On considère une famille d'éléments dans $H(\mathbb{Z})$.

$$\Lambda' = \{\alpha \in H(Z), \mathcal{N}(\alpha) = p^r, \alpha \equiv 1 \mod 2, \textit{ ou } \alpha \equiv i+j+k \mod 2, = p^r\}$$

Soit $\alpha, \beta \in \Lambda'$, on définit la relation d'équivalence comme $\alpha \sim \beta \Leftrightarrow \alpha p^m = \pm \beta p^n$. Donc, une classe d'équivalence est définie comme

$$Q : \Lambda' \to \Lambda = \Lambda'/\sim \tag{A.8}$$

On rappelle la décomposition du quaternion en éléments premiers. On utilise l'algorithme d'Euclide par division à droite. En observant

$$\alpha = \beta\gamma \Rightarrow \mathcal{N}(\alpha) = \mathcal{N}(\beta)\mathcal{N}(\gamma)$$

et 3.4.1, tous les éléments dans S_p sont premiers, donc ils peuvent engendrer les quaternions dont la norme au carré est puissance de p.

Théorème A.5.1. *Soit $\alpha \in H(\mathcal{Z}), \mathcal{N}(\alpha) = p^k$, alors $\alpha = \epsilon p^r w_m$ où ϵ est unité, w_m est un mot réduit formé par S_p, cette écriture est unique. En particulier, soit $\alpha \in \Lambda' \Rightarrow \alpha = \pm p^r w_m$.*

On applique la relation d'équivalence A.8 et on réduit le graphe.

Corollaire A.5.1. *Λ est un groupe engendré par S_p, en plus, le graphe $Cay(\Lambda, Q(S_p))$ est un arbre $(p+1)$- régulier.*

Démonstration. On utilise le théorème A.5.1 pour le montrer. On remarque que dans Λ, un cycle est un mot réduit $\alpha_1 \cdot \alpha_2 \cdots \alpha_l = 1$. Mais A.5.1nous dit l'expression d'un mot est unique. Donc $Cay(\Lambda, Q(S_p))$ un arbre. □

Définition A.5.2.

$$Z_q = \{\alpha \in H(\mathbb{F}_{||})^*, \alpha = \bar{\alpha}\}$$

On déduit le groupe morphisme par module

$$\pi_q : \Lambda \to H(\mathbb{F}_q)^*/Z_q$$

et on note le noyau $\Lambda(q)$.

On démontre le diagramme commutatif suivant. Comme p, q sont tous premiers, $\forall t \in Z, \exists r$ t.q $p^r \equiv t$ (mod q), donc on peut changer l'ordre de module et la relation d'équivalence. En plus, $\Lambda/\Lambda(q) \backsimeq Im(\pi_q)$. En suite, on déduit la bijection β entre $H(\mathbb{F}_q)/Z(q)$ et $PGL_2(q)$. On peut aussi montrer que la bijection et la relation d'équivalence est aussi commutative.

Finalement, on obtient une autre construction. On sait que $Y^{p,q}$ est une composante connexe de $X^{p,q}$, mais [8] précise le cas où $X^{p,q}$ et $Y^{p,q}$ sont identiques.

Théorème A.5.2. *On note $T_{p,q} = (\pi_q \circ Q)(S_p), \Lambda/\Lambda(q) = (\pi_q \circ Q)(\Lambda')$ et $Y^{p,q} = Cay(\Lambda/\Lambda(q), T_{p,q})$, dans le cas où $p > 5$ et $q > p^8$, le graphe $X^{p,q}$ isomorphe au graphe $Y^{p,q}$ et donc connexe.*

On va expliquer la construction de $Y^{p,q}$ en détail. C'est aussi un graphe de Cayley. Tout d'abord on prend une partie Λ' du quaternion $H(\mathbb{Z})$ comme sommets et S_p comme l'ensemble symétrique afin d'engendrer un arbre infini $p+1$ régulier. Après on applique la relation d'équivalence, c'est-à-dire on fusionne les sommets et finalement on obtient $Y^{p,q}$.

Corollaire A.5.2. *Le graphe $X^{p,q}$ est connexe.*

Annexe B

B.1 Démonstration du lemme (2)

Démonstration. Soit $\rho = \tau|_{l^2(G)_0}$ la restriction de la représentation régulière de G sur le groupe de fonction de moyenne nulle. Comme $Kaz(G,S) > \epsilon$, par définition de la constante de Kazhdan, on a $Kaz(G,S,\rho) > \epsilon$.

Soit F un sous ensemble de G tel que $F \neq \emptyset$ et $|F| \leq \frac{|G|}{2}$. On considère la fonction $f \in l^2(G)_0$ telle que $f_{|F} = a$ et $f_{|G\setminus F} = b$ ($a, b \in \mathbb{R}$), $\|f\|_{l^2(G)_0} = 1$ et on pose $|F| = m$, $|\bar{F}| = n$. Alors on a $ma + nb = 0$ (comme f est de moyenne nulle) et $ma^2 + nb^2 = 1$ (la fonction est de norme 1).

Comme $Kaz(G,S,\rho) \geq \epsilon$, on a :

$$\sup_{s\in S} \|\rho(s)f - f\| = \sup_{s\in S} \|f(s^{-1} * \bullet) - f(\bullet)\| \geq Kaz(G,S,\rho) > \epsilon$$

On en déduit qu'il existe $s \in S$ tel que $\| f(s^{-1} * \bullet) - f(\bullet) \| \geq \epsilon$.

Considérons maintenant $F' = \{g \in F | s^{-1}g \in F\}$ et pose $|F'| = l$ alors

$$\| f(s^{-1} * \bullet) - f(\bullet) \|^2 = 2l(a-b)^2 \geq \epsilon^2$$

Mais comme $ma + nb = 0$ et $ma^2 + nb^2 = 1$, on a $a^2 = \frac{n}{m(n+m)}$ et $b^2 = \frac{m}{n(n+m)}$. On en déduit que :

$$2m(a-b)^2 \leq 4m(a^2+b^2) = \frac{4m}{m+n}(\frac{m}{n} + \frac{n}{m}) \leq 4(\frac{m^2}{n^2} + 1) \leq 8$$

Les deux inégalités ci-dessus entrainent que

$$\frac{l}{m} \geq \frac{\epsilon^2}{8}$$

D'ailleurs par la définition, on a

$$\frac{l}{m} = \frac{\partial F}{F} \leq h(G)$$

et donc

$$\frac{\epsilon^2}{8} \leq h(G)$$

D'après l'inégalité de Cheeger :

$$h(G) \leq \sqrt{2ck}$$

où c est la meilleure constante telle que $Cay(G, S)$ est un graphe c-expanseur, on déduit alors :

$$c \geq \frac{\epsilon^4}{128k}$$

Le lemme est prouvé. □

B.2 Démonstration du lemme (3)

Démonstration. Ce lemme découle en fait directement de la définition de la constante de Kazhdan. Comme π est continue, $\pi(S)$ est compact. Soit $\rho' : G' \to U(H)$ une représentation de G', alors par la définition $Kaz(G', \pi(S), \rho') = \inf_{v\in H, \|v\|=1} \sup_{s'\in\pi(S)} \|\rho'(s')v - v\|_H = \inf_{v\in H, \|v\|=1} \sup_{s\in S} \|\rho'(\pi(s))v - v\|_H = Kaz(G, S, \rho' \circ \pi)$. Donc

$$Kaz(G', \pi(S)) = \inf_{\rho'_{G'}} Kaz(G', \pi(S), \rho') = \inf_{\rho'_{G'}} Kaz(G, S, \rho' \circ \pi) \geq \inf_{\rho_G} Kaz(G, S, \rho) = Kaz(G, S)$$

ce qui prouve le lemme. □

B.3 Lemme pour la démonstration de 3.3.1

Proposition 13. *Soit G un groupe localement compact, compactement engendré, à base dénombrable de voisinages. Soit S une partie génératrice compacte, symétrique de G qui contient l'identité et S' une partie compacte quelconque de G. Alors, il existe un entier m tel que S^m contient S'.*

Démonstration. Considérons les ensembles S, S^2,.... Comme S est symétrique et contient l'identité, on déduit que $S^i \subset S^j$ pour tout $i < j$ ainsi que S^i est symétrique pour tout i. Maintenant, on démontre qu'il existe un n tel que S^n est de l'intérieure non vide. Ainsi, on a :

$$\cup_{k=1}^{\infty} S^i = G$$

qui est de l'intérieur non vide, tandis que si S^i est de l'intérieur vide pour tout i, d'après le théorème de Baire, on déduit que G est de l'intérieur vide. Il faut donc qu'il existe un entier n tel que S^n est de l'intérieur non vide, c'est à dire que S^n contient un ouvert O. Comme G est un groupe topologique et S^n est symétrique, O^{-1} est aussi un ouvert. De plus, S^{2n} contient $O' = OO^{-1}$ qui est un ouvert contenant l'identité.

Maintenant, pour tout $s \in S'$, on a que sO' est un ouvert et les ensembles sO' couvrent S'. Par la compacité de S', il existe un nombre fini d'éléments $\{s_1, s_2, ..., s_k\}$ de S' tels que $\cup_{i=1}^{k} s_i O'$ couvrent S'. Mais comme S est une partie génératrice de G, il existe un entier p tel que S^p couvre $\{s_1, s_2, ..., s_k\}$ et donc S^{2n+p} couvrent $\cup_{i=1}^{k} s_i O'$ ainsi que S'. On peut alors conclure la proposition.

□

B.4 Le groupe $SL_d(\mathbb{Z}/n\mathbb{Z})$

Pour construire le groupe $SL_d(\mathbb{Z}/n\mathbb{Z})$, on passe au quotient dans $SL_d(\mathbb{Z})$ par le sous-groupe H_n défini comme le sous-groupe engendré par les éléments de la forme

$$I_d + nE_{ij}, i \neq j$$

où E_{ij} est la matrice élémentaire dont seul le coefficient le la i-ème ligne et j-ième colonne est non nul, et vaut 1.

On montre que ce sous-groupe est distingué. Soit un générateur h de H_n et $x \in SL_d(\mathbb{Z})$, il suffit pour cela de montrer qu'on a bien $xhx^{-1} \in H_n$. Le générateur h peut s'écrire $I_d + nE_{ij}$. On peut montrer, par récurrence et en utilisant l'algorithme d'Euclide, que H_n est l'ensemble des matrices dont les coefficients diagonaux sont congrus à 1 modulo n et dont les autres coefficients sont multiples de n. Alors :

$$\begin{aligned} x^{-1}hx &= x^{-1}I_d x + nx^{-1}E_{ij}x \\ &= I_d + nx^{-1}\sum_{k=1}^{d} x_{jk}E_{jk} \\ &= I_d + n\sum_{a=1}^{d}\sum_{k=1}^{d}\sum_{l=1}^{d}(x^{-1})_{al}(x_{jk}(E_{ik})_{lk})E_{ak} \\ &= I_d + n\sum_{a=1}^{d}\sum_{k=1}^{d}(x^{-1})_{ai}x_{jk}E_{ak} \end{aligned}$$

Ainsi, $x^{-1}hx$ appartient bien à H_n, qui est donc distingué. Les classes de conjugaison associées sont composées des matrices dont les coefficients correspondants sont congrus entre eux deux à deux. Elles forment bien un groupe isomorphe à $SL_d(\mathbb{Z}/n\mathbb{Z})$.

B.5 Démonstration du théorème de Bochner

Démonstration. Dans le cas où f est dans L^2, par application du théorème de Plancherel, il existe une transformée de Fourier $\hat{f}$ dans L^2, pour laquelle on a

$$f(x) = \int_{\mathbb{R}}^{d} e^{2\pi i x\cdot\xi}\hat{f}(\xi)d\xi$$

au sens des distributions tempérées. En appliquant l'inégalité ci-dessus avec $d\nu = g(x)dx$, on a

$$\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} f(x-y)\, g(x)dx\, \overline{g(y)}dy \geq 0$$

Par changement de variable, en posant $u = x - y$ et $h(t) = \overline{g(-t)}$ pour tout $t \in \mathbb{R}^d$, on obtient

$$\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} f(u)\, g(x)h(u-x)\, dx\, du \geq 0$$

On reconnait alors un produit de convolution, d'où

$$\int_{\mathbb{R}^d} f(u)\, g \star h(u)du \geq 0$$

On applique alors le théorème de Plancherel :

$$\int_{\mathbb{R}^d} \hat{f}\hat{g}\hat{h} \geq 0$$

Or $\hat{h} = \overline{\hat{g}}$ d'où

$$\int_{\mathbb{R}^d} \hat{f}(\xi)|\hat{g}(\xi)|^2 d\xi \geq 0$$

En choisissant $\hat{g}$ de manière adéquate, on montre que $\hat{f}$ est positive presque partout.

Soit ϕ une fonction continue à support compact égale à 1 au voisinage de 0 et d'intégrale 1. Alors, on a, par théorème de Plancherel,

$$\int_{\mathbb{R}} f(x)\, R^d\phi(Rx)\, dx = \int_{\mathbb{R}} \hat{f}(\xi)\hat{\phi}(\xi/R)\, d\xi$$

On fait tendre R vers $+\infty$: alors, la distribution $R^d\phi(Rx)$ converge vers la masse de Dirac, donc le membre de gauche tend vers $f(0)$. Ainsi

$$f(0) = \lim_{R\to+\infty} \int_{\mathbb{R}} \hat{f}(\xi)\hat{\phi}(\xi/R)\, d\xi$$

Par continuité séquentielle de la transformée de Fourier, $\hat{\phi} \to 1$ d'où

$$f(0) = \int_{\mathbb{R}^d} \hat{f}(\xi)d\xi$$

Ainsi, $\hat{f}$ est intégrable. On pose alors $d\mu(\xi) = \hat{f}(\xi)d\xi$ pour conclure.

On peut étendre la démonstration au cas général, mais nous l'admettons ici. La preuve repose sur une approximation par des fonctions L^2.

□

Annexe C

C.1 Bornes asymptotiques

Les deux propriétés paradoxales, portant sur la distance et le taux d'un code, dépendent fortement l'une de l'autre. On présente ici les bornes supérieures et inférieures basiques du compromis entre la distance d'un code et sa taille (et donc son taux). Introduisons d'abord quelques définitions.

Définition C.1.1. La boule de Hamming de centre $x \in F_2^n$ et de rayon r est l'ensemble

$$B(x,r) = \{y \in F_2^n : d(x,y) \leqslant r\}$$

On note $v(n,r)$ le volume de la boule de rayon r dans la boule de Hamming. Formellement :

$$v(n,r) = \sum_{i=0}^{r} \binom{n}{i}$$

Théorème C.1.1. *Pour $n \in N$ et $d \geqslant 0$, il existe un code avec une distance supérieure ou égale à d et une taille supérieure ou égale à $\frac{2^n}{v(n,d)}$.*

Démonstration. Ceci est un résultat de l'algorithme glouton suivant (de temps d'exécution exponentiel) qui construit un code de distance d. On initialise $S = \{0,1\}^n$, $C = \emptyset$. À chaque étape, on choisit un point quelconque $x \in S$ et on l'ajoute à C. On supprime ensuite de S tous les points dont la distance par rapport à x est inférieure à d. On retrouve la borne du théorème car la taille initiale de S est 2^n, et à chaque itération la taille de S est réduite d'au plus $v(n,r)$. □

Comme on a déjà précisé, un code linéaire peut être connu par la donnée d'une base du sous-espace $C \subset F_2^n$ ou par la donnée d'une matrice de contrôle A d'ordre $m \times n$ telle que $C = \{x | Ax = 0\}$. On peut démontrer que la distance du code C est le plus petit nombre de colonnes de A dont la somme est nulle.

On peut construire A colonne par colonne, tout en s'assurant de ne pas créer une famille liée de moins de d colonnes. On peut construire la j-ème colonne sous cette condition, étant donné que :

$$\sum_{r=0}^{d-1} \binom{j-1}{r} < 2^m$$

car les colonnes restantes ne doivent pas coïncider avec la somme d'un ensemble de moins de $d-1$ colonnes dans A.

Soit $\delta = d/n$ la distance normalisée du code pour $\delta \leq 1/2$. La somme $v(n,r) = \sum_{i=0}^{r} \binom{n}{i}$ est dominée par le dernier terme $\binom{n}{\delta n}$. Le taux du code résultant est donc est donc supérieur à $log(2^n/\binom{n}{\delta n})/n$. Comme $log(\binom{n}{\delta n})/n$ peut être approximé par la fonction d'entropie binaire $H(\delta) = -\delta log\delta - (1-\delta)log(1-\delta)$. On obtient ainsi la version asymptotique de la borne de Gilbert-Varshamov.

Corollaire C.1.1. ***Borne de Gilbert-Varshamov***

Pour tout $\delta \leq \frac{1}{2}$ et pour n assez grand, il existe un code (linéaire) de distance normalisée δ et de taux $r \geq 1 - H(\delta)$.

Théorème C.1.2. *Tout code C de longueur n, de distance d et de distance normalisée δ vérifie $|C| \leqslant \frac{2^n}{v(n,d/2)}$ et $r \leqslant 1 - H(\frac{\delta}{2})$.*

Démonstration. Pour un code C de distance d, les boules de rayon $d/2$ et de centre les points de C doivent être disjointes. On en déduit le résultat en divisant la taille de l'espace 2^n par le volume de chaque boule. □

Corollaire C.1.2. *Un code de distance normalisée δ et de taux r vérifie :*

$$r \leq 1 - H(\delta/2)$$

Bibliographie

[1] A. Lubotzky, *Discrete Groups, Expanding Graphs and Invariant Measures*. Basel : Birkhäuser Basel, 1994.

[2] S. Hoory, N. Linial, and A. Wigderson, *"Expander Graphs and their applications,"* Bulletin of the American Mathematical Society, vol. 43, no. 4, pp. 439–561, Aug. 2006.

[3] A. Lubotzky, *"Expander Graphs in Pure and Applied Mathematics,"* arXiv :1105.2389 [math], May 2011.

[4] T. Tao, *Expansion in finite simple groups of Lie type,* vol. 164. American Mathematical Soc., 2015.

[5] B. Bekka, P. de la Harpe, and A. Valette, *Kazhdan's Property (T)*. 2007.

[6] J.Urquidi, *Expander Graphs and Error Correcting Codes*. Master Thesis. Bordeaux. Université de Bordeaux1. October 2012.

[7] G.Gardam, *Expander Graphs and Kazhdan's Property (T)* . Master Thesis. University of Sydney October 2012.

[8] Elementary Number Theory, Group Theory and Ramanujan Graphs, 1 edition. Cambridge University Press, 2003.

[9] A. Poli, L. Huguet, et G. Cullmann, Codes correcteurs : théorie et applications. Paris, France, Italie, Espagne, 1989.