Réseaux, sous-espaces et approximation diophantienne

Nicolas de Saxcé

17 mai 2024

Introduction

Le but de ce cours est de mettre en évidence quelques liens entre l'approximation diophantienne et l'étude des réseaux d'un espace euclidien.

L'ensemble $\mathbb{Q}$ des nombres rationnels est dense dans la droite réelle $\mathbb{R}$: tout élément θ dans $\mathbb{R}$ peut être approché par une suite de points rationnels $\frac{p}{q}$. Le domaine de l'*approximation diophantienne* est celui de l'étude de ces approximations rationnelles.

Chapitre 1

Approximation diophantienne dans $\mathbb{R}^n$

ch:rn

Commençons par le rappel de la théorie classique de l'approximation diophantienne en dimension 1.

1.1 La droite réelle

ss:r

Le théorème suivant est un exemple typique des résultats que nous chercherons à établir dans ce cours. On l'énonce parfois de la façon suivante : « Tout nombre réel est approchable à l'ordre 2 par des rationnels ».

ordre2

Théorème 1.1. *Pour tout $\theta \in \mathbb{R}$, il existe un rationnel $\frac{p}{q}$ arbitrairement proche de θ et tel que $\left|\theta - \frac{p}{q}\right| < \frac{1}{q^2}$.*

Démonstration. Nous donnons d'abord une démonstration de ce théorème qui repose sur la théorie des fractions continues, et donne en outre un algorithme pour construire une suite de rationnels approchant θ et vérifiant l'inégalité du théorème.

Sans perte de généralité, on peut supposer $\theta > 0$. Posons alors $\theta_0 = \theta$, puis pour tout $n \geq 0$,

$$a_n = \lfloor \theta_n \rfloor \quad \text{et} \quad \theta_{n+1} = \frac{1}{\theta_n - a_n}.$$

Notons $\begin{bmatrix} 1 \\ \theta \end{bmatrix}$ la droite engendrée par le vecteur $\begin{pmatrix} 1 \\ \theta \end{pmatrix}$ dans $\mathbb{R}^2$. Par récurrence, on observe que

$$\begin{bmatrix} 1 \\ \theta \end{bmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & a_0 \end{pmatrix} \cdots \begin{pmatrix} 0 & 1 \\ 1 & a_n \end{pmatrix} \begin{bmatrix} 1 \\ \theta_{n+1} \end{bmatrix} \tag{1.1}$$

thetan

et

$$\theta = a_0 + \cfrac{1}{a_1 + \cfrac{1}{\ddots + a_n + \frac{1}{\theta_{n+1}}}}. \tag{1.2}$$

theta

On peut définir deux suites d'entiers $(p_n)_{n\geq 0}$ et $(q_n)_{n\geq 0}$ par les égalités

$$\begin{pmatrix} q_n & q_{n+1} \\ p_n & p_{n+1} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & a_0 \end{pmatrix} \cdots \begin{pmatrix} 0 & 1 \\ 1 & a_{n+1} \end{pmatrix}.$$

En particulier, prenant $\theta_{n+1} = +\infty$ dans (**??**) ci-dessus, on trouve

$$\frac{p_n}{q_n} = a_0 + \cfrac{1}{a_1 + \cfrac{1}{\ddots + \frac{1}{a_n}}}.$$

De plus, (**??**) montre que θ est une fonction monotone de $\theta_{n+1} \in [a_{n+1}, +\infty)$ et donc appartient à l'intervalle $\left[\frac{p_n}{q_n}, \frac{p_{n+1}}{q_{n+1}}\right]$. (Les bornes ne sont pas nécessairement dans cet ordre.) On remarque alors que le déterminant de la matrice qui définit p_n et q_n est

$$\begin{vmatrix} q_n & q_{n+1} \\ p_n & p_{n+1} \end{vmatrix} = (-1)^n.$$

Cela implique $\frac{p_{n+1}}{q_{n+1}} - \frac{p_n}{q_n} = \frac{(-1)^n}{q_n q_{n+1}}$ et par suite

$$\left|\theta - \frac{p_n}{q_n}\right| \leq \frac{1}{q_n q_{n+1}} < \frac{1}{q_n^2}.$$

□

Exercice 1 (Théorème de Hurwitz)**.**

(a) Montrer qu'il existe n arbitrairement grand tel que $\frac{q_{n+1}}{q_n} > \phi := \frac{1+\sqrt{5}}{2}$.

(b) En déduire que $\left|\theta - \frac{p_n}{q_n}\right| \leq \frac{1}{q_n^2\sqrt{5}}$.

(c) Vérifier que si $c < 1/\sqrt{5}$, on a $\left|\phi - \frac{p}{q}\right| \geq \frac{c}{q^2}$ pour tout rationnel p/q.

Solution. (a) Les entiers q_n vérifient la relation de récurrence $q_{n+1} = a_{n+1}q_n + q_{n-1}$. Si $a_n \geq 2$ pour n arbitrairement grand, alors $q_{n+1} \geq 2q_n \geq \phi q_n$ et on a ce qu'on veut. Si $a_n = 1$ pour tout n suffisamment grand, et si $q_n \leq \phi q_{n-1}$, alors $q_{n+1} = q_n + q_{n-1} \leq (1 + 1/\phi)q_n = \phi q_n$.

(b) Remarquons que si $\frac{q_{n+1}}{q_n} > \phi$, alors $\frac{q_{n+1}}{q_n} + \frac{q_n}{q_{n+1}} > \phi + \frac{1}{\phi} = \sqrt{5}$, et donc

$$\left|\frac{p_{n+1}}{q_{n+1}} - \frac{p_n}{q_n}\right| = \frac{1}{q_n q_{n+1}} \leq \frac{1}{q_n q_{n+1}\sqrt{5}}\left(\frac{q_n}{q_{n+1}} + \frac{q_{n+1}}{q_n}\right) = \frac{1}{\sqrt{5}}\left(\frac{1}{q_{n+1}^2} + \frac{1}{q_n^2}\right).$$

Comme θ appartient à l'intervalle $\left[\frac{p_n}{q_n}, \frac{p_{n+1}}{q_{n+1}}\right]$, on doit avoir $\left|\theta - \frac{p_n}{q_n}\right| \leq \frac{1}{q_n^2\sqrt{5}}$ ou $\left|\theta - \frac{p_{n+1}}{q_{n+1}}\right| \leq \frac{1}{q_{n+1}^2\sqrt{5}}$.

(c) Notons $\phi' = \frac{1-\sqrt{5}}{2}$. Pour tout rationnel p/q qui approche ϕ, on peut écrire, pour $\varepsilon > 0$ arbitrairement proche de 0,

$$\begin{aligned}\frac{1}{q^2} &\leq \left|\phi' - \frac{p}{q}\right| \left|\phi - \frac{p}{q}\right| \\ &\leq (\phi' - \phi + \varepsilon) \left|\phi - \frac{p}{q}\right| \\ &= (\sqrt{5} + \varepsilon) \left|\phi - \frac{p}{q}\right|.\end{aligned}$$

Le théorème ?? ci-dessus peut aussi se démontrer très simplement à l'aide du principe des tiroirs de Dirichlet. On obtient même l'énoncé un peu plus précis suivant.

dirichletn1 **Théorème 1.2** (Dirichlet)**.** *Soit $\theta \in \mathbb{R}$. Pour tout $Q \in \mathbb{N}^*$, il existe $q \in \{0, \dots, Q\}$ et $p \in \mathbb{Z}$ tels que $\left|\theta - \frac{p}{q}\right| < \frac{1}{qQ}$. En particulier, l'inégalité $\left|\theta - \frac{p}{q}\right| < \frac{1}{q^2}$ admet une infinité de solutions $(p, q) \in \mathbb{Z}^2$.*

Démonstration. On découpe l'intervalle $[0, 1[$ en Q tiroirs $[\frac{k}{Q}, \frac{k+1}{Q}[$. Comme la famille $\{q\theta \mod 1\,;\ q = 0, \dots, Q\}$ contient $Q+1$ éléments, le principe des tiroirs montre que deux d'entre eux appartiennent au même tiroir. Par conséquent, il existe $q_1 > q_2$ dans $\{0, \dots, Q\}$ et $p_1, p_2 \in \mathbb{Z}$ tels que

$$|(q_1\theta - p_1) - (q_2\theta - p_2)| < \frac{1}{Q}.$$

Posant $q = q_1 - q_2$ et $p = p_1 - p_2$, on trouve bien $\left|\theta - \frac{p}{q}\right| < \frac{1}{qQ}$. □

Pour quantifier la qualité des approximations rationnelles à un réel u donné, on peut lui associer un exposant diophantien.

Définition 1.3. Étant donné $\theta \in \mathbb{R}$, on définit l'*exposant diophantien* $\beta(\theta) \in [2, +\infty]$ par

$$\beta(\theta) = \inf\left\{\beta > 0 \mid \exists c > 0 : \forall \frac{p}{q} \in \mathbb{Q},\ \left|\theta - \frac{p}{q}\right| \geq cq^{-\beta}\right\}.$$

Le lemme de Borel-Cantelli permet de calculer facilement l'exposant diophantien d'un point θ choisi aléatoirement dans $\mathbb{R}$ suivant la mesure de Lebesgue.

psn1 **Théorème 1.4** (Exposant presque sûr)**.** *Pour presque tout θ dans $\mathbb{R}$ au sens de la mesure de Lebesgue, $\beta(\theta) = 2$.*

Démonstration. D'après le principe de Dirichlet, $\beta(\theta) \geq 2$ pour tout θ, et il suffit donc de démontrer l'inégalité opposée. Si I est un intervalle borné de $\mathbb{R}$ et $\varepsilon > 0$, on pose, pour $q \in \mathbb{N}^*$,

$$A_q = \{\theta \in I \mid \exists p \in \mathbb{Z} :\ \left|\theta - \frac{p}{q}\right| \leq q^{-2-\varepsilon}\}.$$

Cet ensemble est réunion d'au plus $2q|I|$ intervalles de longueur $q^{-2-\varepsilon}$, donc

$$|A_q| \lesssim q^{-1-\varepsilon}.$$

En particulier, $\sum_{q\geq 1}|A_q| < +\infty$, et par le lemme de Borel-Cantelli, pour presque tout θ dans I, il existe q_0 tel que pour tout $q \geq q_0$, $\theta \notin A_q$. Cela montre que $\beta(\theta) \leq 2+\varepsilon$, et à la limite quand ε tend vers 0, on trouve bien $\beta(\theta) \leq 2$. □

Exercice 2 (Théorème de Khintchine)**.**
Étant donné une fonction $\psi\colon \mathbb{N} \to \mathbb{N}$, on considère l'équation

$$\left|\theta - \frac{p}{q}\right| \leq \psi(q) \qquad (E_\psi)$$ epsi

1. Montrer que si $\sum_{q\geq 1} q\psi(q) < +\infty$, alors, pour presque tout θ dans $\mathbb{R}$, l'inégalité (**??**) n'a qu'un nombre fini de solutions $\frac{p}{q}$ dans $\mathbb{Q}$.
2. *(Difficile)* Montrer que si ψ est décroissante et vérifie $\sum_{q\geq 1} q\psi(q) = +\infty$, alors, pour presque tout θ dans $\mathbb{R}$, l'inégalité (**??**) admet une infinité de solutions $\frac{p}{q}$.

Solution. 1.

2.

Le dernier résultat que nous voulons mentionner dans cette théorie de l'approximation des nombres réels par des rationnels est le célèbre théorème de Roth, qui montre que du point de vue de l'exposant diophantien, les points algébriques irrationnels se comportent comme les points génériques pour la mesure de Lebesgue.

Théorème 1.5 (Roth)**.** *Si $\theta \in \mathbb{R}\setminus\mathbb{Q}$ est algébrique, i.e. racine d'un polynôme à coefficients rationnels, alors $\beta(\theta) = 2$.*

La démonstration de ce théorème est difficile, et trop longue pour être incluse dans ces notes. Dans le cas particulier où θ est algébrique de degré 2, on peut en donner une démonstration élémentaire, cf. exercice ci-dessous.

Exercice 3 (Théorème de Liouville)**.**

1. Montrer qu'il existe $c > 0$ tel que pour tout rationnel $\frac{p}{q}$, $\left|\sqrt{2} - \frac{p}{q}\right| \geq \frac{c}{q^2}$. En déduire que $\beta(\sqrt{2}) = 2$.
2. Plus généralement, montrer que si θ est un nombre algébrique de degré d, alors $\beta(\theta) \leq d$. En déduire que $\theta = \sum_{n\geq 1} 10^{-n!}$ est transcendant.

Solution. 1.

2.

Remarque. S'il existe $c > 0$ tel que $\left|\theta - \frac{p}{q}\right| \geq \frac{c}{q^2}$ pour tout rationnel $\frac{p}{q}$, on dit que θ est mal approchable par les rationnels. L'argument de l'exercice ci-dessus montre que l'ensemble BA des réels mal approchables contient l'ensemble des irrationnels quadratiques. On peut par ailleurs montrer que BA est négligeable pour la mesure de Lebesgue, mais de dimension de Hausdorff égale à 1. On ne sait toujours pas si BA contient un nombre algébrique de degré strictement supérieur à 2.

1.2 La correspondance de Dani

Rappelons qu'un *réseau* Δ dans $\mathbb{R}^2$ est un sous-groupe discret à 2 générateurs. En d'autres termes, pour une certaine base (u_1, u_2) de $\mathbb{R}^2$,

$$\Delta = \mathbb{Z}u_1 \oplus \mathbb{Z}u_2.$$

L'action linéaire du groupe $\mathrm{GL}_2(\mathbb{R})$ des matrices inversibles à coefficients dans $\mathbb{R}$ induit une action transitive sur l'espace Ω_2 des réseaux de $\mathbb{R}^2$; le stabilisateur du réseau $\mathbb{Z}^2$ est égal au sous-groupe $\mathrm{GL}_2(\mathbb{Z})$ des matrices à coefficients entiers dont l'inverse est aussi à coefficients entiers. Cela permet d'identifier

$$\Omega_2 \simeq \mathrm{GL}_2(\mathbb{R})/\mathrm{GL}_2(\mathbb{Z}).$$

Pour décrire la position d'un élément Δ de l'espace Ω_2, nous utiliserons deux quantités :

— la *systole* $\lambda_1(\Delta) = \min\{\|v\| \ ; \ v \in \Delta \setminus \{0\}\}$

— le *covolume* $\mu_2(\Delta) = \|u_1 \wedge u_2\|$, si $\Delta = \mathbb{Z}u_1 \oplus \mathbb{Z}u_2$.

Un résultat fondamental de Minkowski montre que le covolume majore le carré de la systole.

Théorème 1.6 (Minkowski)**.** *Pour tout réseau Δ dans $\mathbb{R}^2$, $\lambda_1(\Delta)^2 \leq \frac{2}{\sqrt{3}}\mu_2(\Delta)$.*

Démonstration. Soit $u_1 \in \Delta \setminus \{0\}$ de norme minimale, et $u_2 \in \Delta$ minimal tel que (u_1, u_2) soit libre. Écrivons $u_2 = xu_1 + yu_1^\perp$, où $u_1^\perp$ est orthogonal à u_1, et de même norme. Alors, $\|u_2\|^2 = \|u_1\|^2(x^2 + y^2)$ et comme par choix de u_2, on doit avoir $\|u_2\| \leq \|u_2 \pm u_1\|$, on obtient

$$x^2 + y^2 \leq (x \pm 1)^2 + y^2$$

d'où $|x| \leq \frac{1}{2}$ puis $|y| \geq \frac{\sqrt{3}}{2}$. (Faire un dessin.) Cela implique $\mu_2(\Delta) \geq \|u_1 \wedge u_2\| = |y|\|u_1\|^2 \geq \frac{\sqrt{3}}{2}\lambda_1(\Delta)^2$. □

Exercice 4 (Sommes de deux carrés)**.**

1. Soit p un nombre premier impair. Montrer que -1 est un carré modulo p si et seulement si $p \equiv 1 \mod 4$.
2. Soit $p \equiv 1 \mod 4$, et a tel que $a^2 + 1 \equiv 0 \mod p$. À l'aide du réseau $\Delta = \mathbb{Z}\begin{pmatrix} a \\ 1 \end{pmatrix} + p\mathbb{Z}^2$, montrer qu'il existe deux entiers x, y tels que $p = x^2 + y^2$.
3. En déduire qu'un entier n est somme de deux carrés si et seulement si pour tout $p \equiv 3 \mod 4$, $v_p(n)$ est pair.

Solution.

1. Le groupe multiplicatif $(\mathbb{Z}/p\mathbb{Z})^*$ est cyclique, donc -1 est un carré si et seulement si $(-1)^{\frac{p-1}{2}} = 1$.
2. Tout élément $\begin{pmatrix} x \\ y \end{pmatrix}$ de Δ vérifie $x^2 + y^2 \equiv 0 \mod p$. Le réseau Δ est de covolume p, donc, d'après le premier théorème de Minkowski, il contient un élément tel que $x^2 + y^2 \leq \frac{2}{\sqrt{3}}p < 2p$. Comme $x^2 + y^2$ est un multiple de p, on doit avoir $x^2 + y^2 = p$.

3. L'égalité $(x^2+y^2)(z^2+t^2) = (xz+ty)^2 + (xt-yz)^2$ montre que l'ensemble des sommes de deux carrés est stable par produit, et avec la question précédente, cela implique qu'il contient tous les entiers n vérifiant $v_p(n) \equiv 0 \mod 2$ pour tout $p \equiv 3 \mod 4$. Réciproquement, supposons $n = x^2+y^2$. Le lemme chinois montre que pour tout p, l'équation $a^2 + b^2 \equiv 0 \mod p^{v_p(n)}$ admet une solution $(a,b) \not\equiv 0 \mod p^{v_p(n)}$. Écrivant $a = p^\alpha a_1$ et $b = p^\beta b_1$, on peut supposer $\alpha \leq \beta$ et alors $a_1^2 + p^{2(\beta-\alpha)} b_1^2 \equiv 0 \mod p^{v_p(n)-2\alpha}$. Si $v_p(n)$ est impair, alors $v_p(n) - 2\alpha \neq 0$, donc $x^2+y^2 \equiv 0$ a une solution non triviale dans $\mathbb{Z}/p^{v_p(n)-2\alpha}\mathbb{Z}$, et donc aussi dans $\mathbb{Z}/p\mathbb{Z}$. Donc -1 est un carré modulo p, et cela implique $p \equiv 1 \mod 4$.

On considère maintenant le sous-groupe à un paramètre $(a_t)_{t\in\mathbb{R}}$ dans $\mathrm{GL}_2(\mathbb{R})$ défini par

$$a_t = \begin{pmatrix} e^{-\frac{t}{2}} & 0 \\ 0 & e^{\frac{t}{2}} \end{pmatrix},$$

et on cherche à comprendre le comportement asymptotique d'une orbite $(a_t\Delta)_{t\in\mathbb{R}}$ dans l'espace Ω_2.

Définition 1.7. Le *taux de fuite* d'un réseau Δ sous l'action de (a_t) dans l'espace des réseaux est défini par

$$\gamma(\Delta) = \limsup_{t\to+\infty} \frac{-1}{t} \log \lambda_1(a_t\Delta).$$

Exercice 5.
1. Montrer que pour tout Δ dans Ω_2, $\gamma(\Delta) \in [0,1]$.
2. En utilisant l'ergodicité de l'action de $(a_t)_{t\in\mathbb{R}}$ sur Ω_2, montrer que pour presque tout Δ dans Ω_2, $\gamma(\Delta) = 0$.

Le lien entre l'approximation diophantienne et l'espace des réseaux se fait grâce à la correspondance de Dani, dont une forme est la proposition suivante.

Proposition 1.8 (Correspondance de Dani)**.** *Pour $\theta \in \mathbb{R}$, posons $u_\theta = \begin{pmatrix} 1 & 0 \\ -\theta & 1 \end{pmatrix}$ et $\Delta_\theta = u_\theta\mathbb{Z}^2$. Alors,*

$$\beta(\theta) = \frac{1}{\frac{1}{2} - \gamma(\Delta_\theta)}.$$

Démonstration. Supposons $\left|\theta - \frac{p}{q}\right| \leq q^{-\beta}$, i.e. $|p - q\theta| \leq q^{-\beta+1}$. Pour $t > 0$, on calcule

$$a_t u_\theta \begin{pmatrix} q \\ p \end{pmatrix} = \begin{pmatrix} e^{-\frac{t}{2}} q \\ e^{\frac{t}{2}}(p - q\theta) \end{pmatrix}$$

et choisissant $t > 0$ tel que $e^t = q^\beta$, on obtient donc, pour $v = \begin{pmatrix} q \\ p \end{pmatrix}$,

$$\|a_t u_\theta v\| \leq q^{-\frac{\beta}{2}+1} = e^{-t(\frac{1}{2}-\frac{1}{\beta})}.$$

Cela montre déjà que $\gamma \geq \frac{1}{2} - \frac{1}{\beta}$. Réciproquement, si $v = \begin{pmatrix} q \\ p \end{pmatrix}$ vérifie $\|a_t u_\theta v\| \leq e^{-\gamma t}$ pour $t > 0$, on en tire $q \leq e^{(\frac{1}{2}-\gamma)t}$ puis

$$\left|\theta - \frac{p}{q}\right| \leq \frac{1}{q} e^{-(\gamma+\frac{1}{2})t} \leq q^{-1+\frac{\gamma+\frac{1}{2}}{\frac{1}{2}-\gamma}} = q^{\frac{-1}{\frac{1}{2}-\gamma}}$$

d'où $\beta \geq \frac{1}{\frac{1}{2}-\gamma}$. □

Exercice 6. Soit Δ un réseau dans $\mathbb{R}^2$ admettant une base à coefficients algébriques. En admettant le théorème de Thue-Siegel-Roth, montrer que $\gamma(\Delta) = 0$ sauf si Δ contient un vecteur sur l'axe des abscisses, auquel cas $\gamma(\Delta) = 1$.

1.3 Approximation en dimension supérieure

sec:dimsup

Étant donné un point $\theta = (\theta_1, \ldots, \theta_n)$ dans $\mathbb{R}^n$, il existe traditionnellement deux problèmes d'approximation par des rationnels au point θ, qui généralisent tous deux le cadre de la droite réelle.

(A) Approximation simultanée

On définit l'exposant diophantien $\beta_1(\theta)$ pour l'approximation simultanée de la façon suivante :

$$\beta_1(\theta) = \sup\left\{\beta > 0 \;\middle|\; \exists(q, p_1, \ldots, p_n) : \begin{array}{c} (\frac{p_1}{q}, \ldots, \frac{p_n}{q}) \to \theta \\ \max_{1\leq i\leq n} \left|\theta_i - \frac{p_i}{q}\right| \leq q^{-\beta} \end{array}\right\}$$

Pour la suite, il sera approprié de voir cette approximation dans l'espace projectif $\mathbb{P}^n(\mathbb{R})$ des droites vectorielles dans $\mathbb{R}^{n+1}$. Pour x et y dans $\mathbb{P}^n(\mathbb{R})$, on note $\sphericalangle(x, y)$ l'angle entre les droites x et y et on définit la distance entre x et y par

$$d(x, y) = |\sin \sphericalangle(x, y)|.$$

La hauteur d'un point rationnel $v \in \mathbb{P}^n(\mathbb{R})$ — i.e. d'une droite dans $\mathbb{R}^{n+1}$ engendrée par un vecteur rationnel — est définie par

$$H(v) = \min\left\{\|\mathbf{v}\| \;;\; \mathbf{v} \in v \cap \mathbb{Z}^{n+1}\right\}.$$

ex:schanuel **Exercice 7** (Théorème de Schanuel)**.**
Montrer que le nombre le points rationnels v dans $\mathbb{P}^n$ tels que $H(v) \leq H$ est équivalent à $\frac{\mathrm{vol}(B_{\mathbb{R}^{n+1}}(0,1))}{\zeta(n+1)} \cdot H^{n+1}$ lorsque H tend vers $+\infty$.

L'exposant diophantien d'un point x dans $\mathbb{P}^n(\mathbb{R})$ pour l'approximation par des droites rationnelles est

$$\beta_1(x) = \sup\left\{\beta > 0 \;\middle|\; \exists v \in \mathbb{P}^n(\mathbb{Q}) : \begin{array}{c} v \to x \\ d(v, x) \leq H(v)^{-\beta} \end{array}\right\}.$$

Si x est la droite engendrée par le vecteur $(1, \theta_1, \ldots, \theta_n)$, on retrouve bien l'exposant $\beta_1(\theta)$ défini ci-dessus. Les résultats du paragraphe **??** se généralisent de la façon suivante. Ci-dessous, et dans la suite, on note $\overline{\mathbb{Q}}$ l'ensemble des nombres réels algébriques, i.e. qui sont racines d'un polynôme non nul à coefficients dans $\mathbb{Q}$. Par $\mathbb{P}^n(\overline{\mathbb{Q}})$ on désigne l'ensemble des droites dans $\mathbb{R}^{n+1}$ qui contiennent un vecteur dont toutes les coordonnées sont dans $\overline{\mathbb{Q}}$.

th:pnsimul **Théorème 1.9** (Propriétés de l'exposant diophantien β_1 sur $\mathbb{P}^n$)**.**

1. *(Dirichlet) Pour tout $x \in \mathbb{P}^n(\mathbb{R})$, $\beta_1(x) \geq \frac{n+1}{n}$.*

2. *(Borel-Cantelli) Pour presque tout* $x \in \mathbb{P}^n(\mathbb{R})$, $\beta_1(x) = \frac{n+1}{n}$.
3. *(Roth-Schmidt) Pour tout* $x \in \mathbb{P}^n(\overline{\mathbb{Q}})$ *hors de tout sous-espace rationnel strict,* $\beta_1(x) = \frac{n+1}{n}$.

Démonstration. (Dirichlet) Quitte à permuter les coordonnées, on peut supposer que la droite x est engendrée par le vecteur $(1, \theta_1, \dots, \theta_n)$. Montrons que pour tout entier $Q > 1$, il existe des entiers $q \in \{1, \dots, Q\}$ et $p_1, \dots, p_n \in \mathbb{Z}$ tels que pour chaque i, $\left|\theta_i - \frac{p_i}{q}\right| \leq \frac{1}{qQ^{\frac{1}{n}}}$.
La démonstration est essentiellement la même qu'en dimension $n = 1$: on considère les $Q+1$ points de $[0,1)^n$ obtenus par réduction modulo 1 des points $(q\theta_1, \dots, q\theta_n)$, $q = 0, \dots, Q$. Pour des raisons de volume, les cubes (modulo 1) de côté $Q^{-\frac{1}{n}}$ centrés en chacun de ces points ne sauraient être tous disjoints, donc il existe $q' < q''$ tels que pour certains entiers $p_1, \dots, p_n$,

$$\forall i \in \{1, \dots, n\}, \quad |q''\theta_i - q'\theta'_i - p_i| < Q^{-\frac{1}{n}}.$$

Cela donne ce qu'on veut, en posant $q = q'' - q'$.

(Borel-Cantelli) Le nombre de points rationnels de hauteur au plus H dans $\mathbb{P}^{n+1}$ satisfait

$$N_{\mathbb{P}^n}(H) = |\{v \in \mathbb{P}^n(\mathbb{Q}) \mid H(v) \leq H\}| \lesssim H^{n+1}.$$

Pour $\beta > 0$, on peut donc majorer la mesure de l'ensemble

$$A_H = \bigcup_{\substack{v \in \mathbb{P}^n(\mathbb{Q}): \\ H \leq H(v) < 2H}} B(v, H^{-\beta})$$

par $|A_H| \lesssim H^{n+1-n\beta}$. En particulier, si $\beta > \frac{n+1}{n}$, la somme $\sum_{H=2^k} |A_H|$ converge, et donc, d'après le lemme de Borel-Cantelli, pour presque tout x dans $\mathbb{P}^n(\mathbb{R})$, pour tout $H = 2^k$ suffisamment grand, $x \notin A_H$. Cela implique $\beta_1(x) \leq \frac{n+1}{n}$. Par le principe de Dirichlet, l'inégalité opposée est toujours vérifiée, donc $\beta_1(x) = \frac{n+1}{n}$ pour presque tout x dans $\mathbb{P}^1(\mathbb{R})$.

(Roth-Schmidt) Le calcul de l'exposant des points algébriques est plus subtil, il requiert l'introduction du théorème du sous-espace de Schmidt, et fait l'objet du reste de ce paragraphe. □

Nous admettrons le résultat fondamental suivant, qui a justement été démontré par Schmidt dans le but de généraliser le théorème de Roth en dimension supérieure.

Théorème 1.10 (Schmidt, théorème du sous-espace)**.** *Soit* $d \in \mathbb{N}^*$ *et* L *un élément de* $\mathrm{GL}_d(\overline{\mathbb{Q}})$*, dont on note* $L_1, \dots, L_d$ *les lignes. Pour tout* $\varepsilon > 0$*, l'ensemble des* $\mathbf{v} \in \mathbb{Z}^d$ *tels que*

$$|L_1(\mathbf{v}) \dots L_d(\mathbf{v})| \leq \|\mathbf{v}\|^{-\varepsilon}$$

est contenue dans une union finie d'hyperplans.

À partir de ce résultat, nous pouvons facilement calculer l'exposant d'un point x dans $\mathbb{P}^n(\overline{\mathbb{Q}})$.

Démonstration du théorème **??**, *3.* On peut supposer que la droite x est engendrée par un vecteur de la forme $(1, \theta_2, \dots, \theta_{n+1})$, où $\theta_i \in \overline{\mathbb{Q}}$. Notant $\mathbf{v} = (q, p_1, \dots, p_n$ un élément de $\mathbb{R}^{n+1}$, on définit $n+1$ formes linéaires sur $\mathbb{R}^{n+1}$ par

$$\begin{aligned} &L_1(\mathbf{v}) = q \\ &\forall i \geq 2,\ L_i(\mathbf{v}) = q\theta_i - p_i. \end{aligned}$$

Si v désigne la droite engendrée par le vecteur $\mathbf{v}$, on majore

$$\begin{aligned} &|L_1(\mathbf{v})| \leq \|\mathbf{v}\| \\ &\forall i \geq 2,\ |L_i(\mathbf{v})| \lesssim \|\mathbf{v}\| \cdot d(v, x). \end{aligned}$$

En particulier, si $d(v,x) \leq H(v)^{-\frac{n+1}{n}-\varepsilon}$, alors

$$|L_1(\mathbf{v}) \dots L_{n+1}(\mathbf{v})| \lesssim H(\mathbf{v})^{n+1} H(\mathbf{v})^{-n(\frac{n+1}{n}+\varepsilon)} \leq \|\mathbf{v}\|^{-n\varepsilon}$$

et d'après le théorème du sous-espace de Schmidt, les solutions $\mathbf{v}$ à cette inégalité sont contenues dans un nombre fini d'hyperplans rationnels. Si V est un tel hyperplan, alors $x \notin V$, par hypothèse, et donc x ne peut pas être approché par un élément $v = \mathbb{R}\mathbf{v}$ dans V Cela montre que l'inégalité $d(v,x) \leq H(v)^{-\frac{n+1}{n}-\varepsilon}$ n'a qu'un nombre fini de solutions $v \in \mathbb{P}^n(\mathbb{Q})$, d'où $\beta_1(x) \leq \frac{n+1}{n}$. Ici encore l'inégalité réciproque découle du principe de Dirichlet. □

(B) Approximation des formes linéaires

Le deuxième problème d'approximation classique associé à un point $\theta = (\theta_1, \dots, \theta_n)$ dans $\mathbb{R}^n$ est celui de la recherche de solutions entières $(q, p_1, \dots, p_n)$ à l'inégalité

$$|q + p_1\theta_1 + \dots + p_n\theta_n| \leq \left(\max_{1\leq i\leq n} |p_i|\right)^{-\beta+1}. \tag{1.3}$$

Cette inégalité s'interprète géométriquement comme un problème d'approximation dans l'espace projectif dual $\mathbb{P}^{*n}(\mathbb{R})$, que l'on identifie à l'ensemble des hyperplans de $\mathbb{R}^{n+1}$. Si $x^* = (1, \theta_1, \dots, \theta_n)^\perp$ désigne l'hyperplan orthogonal au vecteur $(1, \theta_1, \dots, \theta_n)$, et $v = \mathbb{R}\mathbf{v}$, où $\mathbf{v} = (q, p_1, \dots, p_n)$, l'inégalité ci-dessus peut se réécrire

$$d(v, x^*) \leq H(v)^{-\beta}$$

où cette fois $d(v, x^*)$ désigne la distance du point v à l'hyperplan x dans l'espace projectif $\mathbb{P}^n(\mathbb{R})$.

Comme ci-dessus, on définit l'exposant diophantien d'un élément x dans $\mathbb{P}^{*n}(\mathbb{R})$ pour l'approximation par des droites rationnelles est

$$\beta_1(x^*) = \sup\left\{\beta > 0 \;\middle|\; \exists v \in \mathbb{P}^n(\mathbb{Q}) : \begin{array}{c} v \to x^* \\ d(v, x^*) \leq H(v)^{-\beta} \end{array}\right\}.$$

À titre d'exercice, on laisse le soin au lecteur de vérifier que les résultats du paragraphe **??** se généralisent aussi à ce cadre.

Théorème 1.11 (Propriétés de l'exposant diophantien β_1 sur $\mathbb{P}^{*n}$)**.**

1. *(Dirichlet) Pour tout $x^* \in \mathbb{P}^{*n}(\mathbb{R})$, $\beta_1(x^*) \geq n+1$.*

2. *(Borel-Cantelli) Pour presque tout* $x^* \in \mathbb{P}^{*n}(\mathbb{R})$, $\beta_1(x^*) = n+1$.
3. *(Roth-Schmidt) Pour tout* $x^* \in \mathbb{P}^{*n}(\overline{\mathbb{Q}})$ *ne contenant aucun sous-espace rationnel non trivial,* $\beta_1(x^*) = n+1$.

Exercice 8 (Approximation par des formes linéaires)**.**
Expliquer comment on peut aussi interpréter l'inégalité (**??**) comme un problème d'approximation de la droite $x = \mathbb{R}(1, \theta_1, \dots, \theta_n)$ par un hyperplan rationnel. Si l'on note $\beta_n(x)$ l'exposant diophantien associé, justifier que $\beta_n(x) = \beta_1(x^\perp)$.

1.4 Exercices supplémentaires

Exercice 9 (Meilleures approximations)**.**

(a) Montrer que la suite $(|q_n\theta - p_n|)_{n\geq 0}$ est décroissante.

(b) Pour tout rationnel $\frac{p}{q}$ avec $q < q_n$, on a $|q\theta - p| \geq |q_{n-1}\theta - p_{n-1}|$.

(c) Réciproquement, montrer que si θ est irrationnel et

$$|q\theta - p| = \min_{q' \leq Q,\ p'} |q'\theta - p'|$$

pour un certain $Q \geq q$, alors $\frac{p}{q}$ apparaît dans la suite $(\frac{p_n}{q_n})_{n\geq 0}$.

Solution. (a) La matrice

$$\begin{pmatrix} 1 & 0 \\ -\theta & 1 \end{pmatrix} \begin{pmatrix} q_n & q_{n-1} \\ p_n & p_{n-1} \end{pmatrix} = \begin{pmatrix} q_n & q_{n-1} \\ p_n - q_n\theta & p_{n-1} - q_{n-1}\theta \end{pmatrix}$$

est de déterminant égal à 1 en valeur absolue. Par suite,

$$|p_{n-1} - q_{n-1}\theta| \geq \frac{1}{q_n} - \frac{q_{n-1}}{q_n}|p_n - q_n\theta|$$

et comme $1 \geq q_{n+1}|q_n\theta - p_n|$, on trouve

$$|p_{n-1} - q_{n-1}\theta| \geq |p_n - q_n\theta|(\frac{q_{n+1}}{q_n} - \frac{q_{n-1}}{q_n}) = a_{n+1}|p_n - q_n\theta| \geq |p_n - q_n\theta|.$$

(b) Le réseau unimodulaire

$$\begin{pmatrix} q_n^{-1} & 0 \\ 0 & q_n \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -\theta & 1 \end{pmatrix} \mathbb{Z}^2$$

admet une base donnée par

$$[u_1, u_2] = \begin{pmatrix} 1 & q_{n-1}/q_n \\ q_n(p_n - q_n\theta) & q_n(p_{n-1} - q_{n-1}\theta). \end{pmatrix}$$

Les vecteurs u_1 et u_2 sont tous deux dans $[0,1] \times [-1,1]$, avec des signes opposés sur la seconde coordonnée. Soit $v = \begin{pmatrix} q/q_n \\ q_n(p - q\theta) \end{pmatrix}$ un élément du réseau, avec $q < q_n$. En utilisant le fait que la première coordonnée de v appartient à $[0,1)$, on observe (faire un dessin) que la deuxième coordonnée de v est minorée en valeur absolue par la deuxième coordonnée de u_2, i.e.

$$q_n|p - q\theta| \geq q_n|p_{n-1} - q_{n-1}\theta|.$$

(c) Soit n tel que $q_{n-1} \leq Q < q_n$. La question précédente montre que $|q\theta - p| = \max_{q' \leq Q,\ p'} |q'\theta - p| = |q_{n-1}\theta - p_{n-1}|$, et cela implique $q = q_{n-1}$ car θ est irrationnel. (Si θ est rationnel, on peut aussi montrer $q = q_{n-1}$, mais c'est un peu plus subtil, cf. Khinchin, *Continued fractions*.)

Exercice 10 (Un théorème de Legendre)**.**
Montrer que si $\frac{p}{q}$ n'apparaît pas dans la suite $(\frac{p_n}{q_n})_{n\geq 0}$, alors $\left|\theta - \frac{p}{q}\right| \geq \frac{1}{2q^2}$.

Solution. On raisonne par contraposée en supposant $|q(q\theta - p)| \leq \frac{1}{2}$. Le réseau unimodulaire

$$\begin{pmatrix} q^{-1} & 0 \\ 0 & q \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -\theta & 1 \end{pmatrix} \mathbb{Z}^2$$

contient le vecteur $u = \begin{pmatrix} 1 \\ q(q\theta - p) \end{pmatrix}$, qui appartient à $\{1\} \times [-1/2, 1/2]$. Tout autre vecteur $v = \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} q'/q \\ q(q'\theta - p') \end{pmatrix}$ du réseau appartenant à $[0,1] \times \mathbb{R}$ doit vérifier

$$|\det(u, v)| = |xq(q\theta - p) + y| \geq 1$$

et par conséquent

$$|q(q'\theta - p')| = |y| \geq 1 - \frac{x}{2} > \frac{1}{2} \geq |q(q\theta - p)|.$$

Cela montre que $|q\theta - p| = \min_{q' \leq q,\ p'} |q'\theta - p'|$, et d'après l'exercice précédent, $\frac{p}{q}$ doit apparaître dans la suite des quotients partiels $(\frac{p_n}{q_n})_{n\geq 0}$.

Exercice 11. Montrer que pour tout rationnel $\frac{p}{q}$ avec $q < q_n$ distinct de q_{n-1}, on a $\left|\theta - \frac{p}{q}\right| > \frac{1}{qq_n}$.

Solution. De l'égalité

$$\begin{vmatrix} 1 & q_{n-1}/q_n \\ q_n(q_n\theta - p_n) & q_n(q_{n-1}\theta - p_{n-1}) \end{vmatrix} = 1$$

on déduit

$$q_n|q_{n-1}\theta - p_{n-1}| \geq 1 - q_{n-1}|q_n\theta - p_n|.$$

Ensuite, on reprend le raisonnement de la deuxième question de l'exercice précédent : le vecteur $v = \begin{pmatrix} q/q_n \\ q_n(p - q\theta) \end{pmatrix}$ a sa première coordonnée dans $[0,1)$ et n'est pas égal à u_2, donc sa deuxième coordonnée est minorée par la somme des deuxièmes coordonnées de u_1 et u_2 :

$$\begin{aligned} q_n|q\theta - p| &\geq q_n|q_{n-1}\theta - p_{n-1}| + q_n|q_n\theta - p_n| \\ &\geq 1 - q_{n-1}|q_n\theta - p_n| + q_n|q_n\theta - p_n| \\ &> 1. \end{aligned}$$

Exercice 12. Montrer que la suite $((\cos n)^n)_{n\geq 1}$ ne tend pas vers 0 en l'infini.

Exercice 13 (Monotonie des quotients partiels)**.**
Montrer que les suites $(\frac{p_{2n}}{q_{2n}})_{n\geq 0}$ et $(\frac{p_{2n+1}}{q_{2n+1}})_{n\geq 0}$ sont adjacentes.

Solution. Les égalités

$$\begin{cases} q_{2n+2}p_{2n+1} - p_{2n+2}q_{2n+1} = -1 \\ q_{2n+1}p_{2n} - q_{2n}p_{2n+1} = 1 \end{cases}$$

impliquent que $q_{2n+1} \cdot \begin{vmatrix} q_{2n+2} & q_{2n} \\ p_{2n+2} & p_{2n} \end{vmatrix} = q_{2n+2} - q_{2n}$ et donc

$$\frac{p_{2n+2}}{q_{2n+2}} - \frac{p_{2n}}{q_{2n}} = \frac{q_{2n+2} - q_{2n}}{q_{2n+2}q_{2n+1}q_{2n}} > 0.$$

i.e. $(\frac{p_{2n}}{q_{2n}})_{n\geq 0}$ est croissante. On montre de même que $(\frac{p_{2n+1}}{q_{2n+1}})_{n\geq 0}$ est décroissante, et comme ces deux suites convergent toutes deux vers θ, elles sont adjacentes.

Exercice 14 (Borne inférieure sur l'erreur d'approximation)**.**
Montrer que pour tout n, $\left|\theta - \frac{p_n}{q_n}\right| > \frac{1}{q_n(q_{n+1}+q_n)}$.

Solution. L'intervalle $\left[\frac{p_n}{q_n}, \frac{p_{n+1}}{q_{n+1}}\right]$ contient θ. La suite de fractions médianes $\frac{p_n+p_{n+1}}{q_n+q_{n+1}}, \frac{p_n+2p_{n+1}}{q_n+2q_{n+1}}, \ldots, \frac{p_n+a_{n+1}p_{n+1}}{q_n+a_{n+1}q_{n+1}}$ est monotone. D'après l'exercice précédent, la fraction $\frac{p_{n+2}}{q_{n+2}} = \frac{p_n+a_{n+1}p_{n+1}}{q_n+a_{n+1}q_{n+1}}$ est située du même côté de θ que $\frac{p_n}{q_n}$, donc c'est aussi le cas de $\frac{p_n+p_{n+1}}{q_n+q_{n+1}}$. En particulier,

$$\left|\theta - \frac{p_n}{q_n}\right| \geq \left|\frac{p_n + p_{n+1}}{q_n + q_{n+1}} - \frac{p_n}{q_n}\right| = \frac{1}{q_n(q_n + q_{n+1})}.$$

Exercice 15. Le but de cet exercice est de justifier que la systole et le covolume permettent de décrire à une constante près la position d'un élément de l'espace des réseaux du plan euclidien.

1. Justifier qu'une métrique riemannienne invariante à droite sur $\mathrm{GL}_2(\mathbb{R})$ induit une distance riemannienne sur Ω_2. Vérifier que pour cette distance, une suite $(\Delta_n)_{n\geq 1}$ converge vers un élément $\Delta = \mathbb{Z}u_1 \oplus \mathbb{Z}u_2$ si, et seulement si, on peut écrire $\Delta_n = \mathbb{Z}u_1^{(n)} \oplus \mathbb{Z}u_2^{(n)}$, avec $\lim u_1^{(n)} = u_1$ et $\lim u_2^{(n)} = u_2$.
2. Montrer qu'à certaines constantes multiplicatives près, pour tous $\Delta, \Delta' \in \Omega_2$,
$$d(\Delta, \Delta') \asymp \max\left(\left|\log \frac{\mu_1(\Delta')}{\mu_1(\Delta)}\right|, \left|\log \frac{\mu_2(\Delta')}{\mu_2(\Delta)}\right|\right).$$
3. (Critère de Mahler) Montrer qu'une partie $A \subset \Omega_2$ est relativement compacte si et seulement si son image par l'application $\Delta \mapsto (\mu_1(\Delta), \mu_2(\Delta))$ est relativement compacte dans $(\mathbb{R}_+^*)^2$.

Chapitre 2

Réseaux dans $\mathbb{R}^d$ et sous-espaces rationnels

ch:sublattice

Pour présenter les différents énoncés d'approximation diophantienne dans $\mathbb{R}^n$ dans un cadre unifié, Schmidt a proposé en 1967 le problème suivant :

Fixons des entiers d, k et ℓ tels que $d \geq 2$ et $0 < k, \ell < d$. Étant donné un sous-espace x de dimension ℓ dans $\mathbb{R}^d$, étudier les sous-espaces rationnels v de dimension k proches de x.

Dans la suite, l'espace $\mathbb{R}^d$ est muni de sa structure euclidienne usuelle ; la norme est notée $\|\cdot\|$ et le produit scalaire $\langle \cdot, \cdot \rangle$. Rappelons que la distance entre un vecteur $\mathbf{u}$ et une partie fermée F est définie par $d(\mathbf{u}, F) = \min_{\mathbf{v} \in F} d(\mathbf{u}, \mathbf{v})$. Pour pouvoir évaluer la qualité d'une approximation rationnelle v de dimension k au sous-espace x de dimension ℓ, on définit aussi

$$d(v, x) = \begin{cases} \max_{\mathbf{u} \in v ; \|\mathbf{u}\|=1} d(\mathbf{u}, x) & \text{si } k \leq \ell \\ \max_{\mathbf{u} \in x ; \|\mathbf{u}\|=1} d(\mathbf{u}, v) & \text{si } \ell \leq k \end{cases}$$

Notons que $d(\cdot, \cdot)$ n'est pas à proprement parler une distance, puisque $d(v, x) = 0$ si et seulement si $v \subset x$ ou $x \subset v$.

Exercice 16. Vérifier qu'on a toujours $d(v, x) = d(v^\perp, x^\perp)$.

Solution. Si x est un sous-espace vectoriel de $\mathbb{R}^d$, alors pour tout $\mathbf{u} \in \mathbb{R}^d$,

$$d(\mathbf{u}, x) = \max\{\langle \mathbf{u}, \mathbf{w} \rangle \ ; \ \mathbf{w} \in x^\perp \text{ unitaire}\}.$$

Par conséquent,

$$\begin{aligned} d(v, x) &= \max\{\langle \mathbf{u}, \mathbf{w} \rangle \ ; \ \mathbf{u} \in v \text{ unitaire et } \mathbf{w} \in x^\perp \text{ unitaire}\} \\ &= d(x^\perp, v^\perp) = d(v^\perp, x^\perp), \end{aligned}$$

où la dernière égalité découle de la symétrie de la distance que nous avons définie.

On définit aussi la hauteur d'un sous-espace rationnel v comme le volume d'un domaine fondamental de v sous l'action de $v \cap \mathbb{Z}^d$:

$$H(v) = \mathrm{vol}\left(v/v \cap \mathbb{Z}^d\right).$$

Concrètement, si les vecteurs $u_1, \ldots, u_k$ forment une base de $v \cap \mathbb{Z}^d$, alors $H(v)$ est égale au volume du parallélépipède engendré par ces vecteurs. Nous verrons ci-dessous que pour tout $H \geq 0$, il n'y a qu'un nombre fini de sous-espaces rationnels v tels que $H(v) \leq H$. Dans la suite, nous noterons

$$X_{\ell,d}(\mathbb{R}) = \{x \leq \mathbb{R}^d \mid \dim x = \ell\}$$

la variété grassmannienne des sous-espaces de dimension ℓ dans $\mathbb{R}^d$, et

$$X_{\ell,d}(\mathbb{Q}) = \{x \leq \mathbb{R}^d \mid \dim x = \ell \text{ et } x \text{ est défini sur } \mathbb{Q}\}.$$

Les notions de distance et de hauteur ci-dessus permettent d'associer à un élément x dans $X_{\ell,d}(\mathbb{R})$ une famille d'exposants diophantiens $\beta_k(x)$, $k = 1, \ldots, d-1$, définis par

$$\beta_k(x) = \sup\left\{\beta > 0 \;\middle|\; \exists v \in X_{k,d}(\mathbb{Q}) : \begin{array}{l} d(v,x) \to 0 \\ d(v,x) \leq H(v)^{-\beta} \end{array}\right\} \qquad (2.1)$$

eq:betak

Le but de ce chapitre est d'obtenir les premières propriétés élémentaires de ces familles d'exposants, et d'énoncer les résultats qui seront démontrés dans la suite du cours. Nous commençons par quelques rappels sur les réseaux de l'espace euclidien $\mathbb{R}^d$.

2.1 Sous-groupes discrets de $\mathbb{R}^d$

pr:discret

Proposition 2.1. *Tout sous-groupe discret de $\mathbb{R}^d$ est de la forme $\Lambda = \mathbb{Z}\mathbf{v}_1 \oplus \cdots \oplus \mathbb{Z}\mathbf{v}_k$, où les vecteurs $\mathbf{v}_1, \ldots, \mathbf{v}_k$ sont linéairement indépendants sur $\mathbb{R}$.*

Démonstration. On procède par récurrence sur la dimension d. Pour $d = 0$, il n'y a rien à démontrer, supposons donc le résultat connu pour $d - 1 \geq 0$. Si Λ est un sous-groupe discret de $\mathbb{R}^d$, on choisit un élément $\mathbf{v}_1 \in \Lambda$ non nul de norme minimale. Soit $\pi\colon \mathbb{R}^d \to \mathbf{v}_1^\perp$ la projection orthogonale sur $\mathbf{v}_1^\perp$. Tout élément de $\mathbf{v}' \in \pi(\Lambda)$ admet une pré-image $\mathbf{v} \in \Lambda$ telle que $d(\mathbf{v}, \mathbf{v}_1^\perp) \leq \|\mathbf{v}_1\|$ et donc $\|\mathbf{v}\| \leq \|\mathbf{v}'\| + \|\mathbf{v}_1\|$. Comme Λ est discret, cela montre que $\pi(\Lambda)$ est un sous-groupe discret de $\mathbf{v}_1^\perp$. Par hypothèse de récurrence, on peut écrire,

$$\pi(\Lambda) = \mathbb{Z}\pi(\mathbf{v}_2) \oplus \cdots \oplus \mathbb{Z}\pi(\mathbf{v}_k)$$

pour certains vecteurs $\mathbf{v}_2, \ldots, \mathbf{v}_k$ dans Λ tels que $\pi(\mathbf{v}_2), \ldots, \pi(\mathbf{v}_k)$ soient linéairement indépendants. Si $\mathbf{v}$ est un élément de Λ, on peut écrire $\pi(\mathbf{v}) = \sum_{i\geq 2} n_i \pi(v_i)$, avec $n_i \in \mathbb{Z}$, donc

$$\pi\left(\mathbf{v} - \sum_{i\geq 2} n_i \mathbf{v}_i\right) = 0$$

d'où $\mathbf{v} - \sum_{i\geq 2} n_i \mathbf{v}_i = \lambda \mathbf{v}_1$ pour un certain $\lambda \in \mathbb{R}$. Mais par minimalité de $\|\mathbf{v}_1\|$, on doit avoir $\lambda \in \mathbb{Z}$, ce qui achève la démonstration. □

Définition 2.2 (Covolume d'un sous-groupe discret)**.** Si $\Lambda = \mathbb{Z}\mathbf{v}_1 \oplus \cdots \oplus \mathbb{Z}\mathbf{v}_k$ est un sous-groupe discret de $\mathbb{R}^d$, on note $|\Lambda|$ le volume du parallélépipède engendré par les éléments de la base $(\mathbf{v}_1, \ldots, \mathbf{v}_k)$.

Exercice 17.

1. Si Λ est un sous-groupe discret de $\mathbb{R}^d$, on définit le dual $\Lambda^* = \{\mathbf{w} \in \mathrm{Vect}_{\mathbb{R}} \Lambda \mid \forall \mathbf{v} \in \Lambda,\ \langle \mathbf{w}, \mathbf{v} \rangle \in \mathbb{Z}\}$. Montrer que $|\Lambda^*| = |\Lambda|^{-1}$.
2. Si $\Lambda_1 \leq \Lambda$ est un sous-groupe discret primitif, on pose $\Lambda/\Lambda_1 = \pi_{\Lambda_1^\perp}(\Lambda)$. Vérifier que $|\Lambda/\Lambda_1| = |\Lambda|/|\Lambda_1|$.
3. Montrer que pour tout v dans $X_{k,d}(\mathbb{Q})$, $H(v) = H(v^\perp)$.

Solution. 1.

2.

3. Étant donné un sous-espace rationnel v dans $\mathbb{R}^d$, notons $\Lambda_v = v \cap \mathbb{Z}^d$. On vérifie facilement que $\Lambda_v^* \supset \mathbb{Z}^d/\Lambda_{v^\perp}$, et donc

$$H(v)^{-1} = |\Lambda_v|^{-1} = |\Lambda_v^*| \leq \Lambda_{v^\perp}^{-1} = H(v^\perp)^{-1}$$

i.e.

$$H(v^\perp) \leq H(v).$$

En appliquant cette inégalité à $v^\perp$, on trouve aussi $H(v) = H((v^\perp)^\perp) \leq H(v^\perp)$, d'où l'égalité souhaitée.

2.2 Les théorèmes de Minkowski

Définition 2.3 (Réseau dans $\mathbb{R}^d$)**.** Un *réseau* Δ dans $\mathbb{R}^d$ est un sous-groupe discret de rang d. De façon équivalente, il existe une base $(\mathbf{v}_1, \ldots, \mathbf{v}_d)$ de $\mathbb{R}^d$ telle que

$$\Delta = \mathbb{Z}\mathbf{v}_1 \oplus \cdots \oplus \mathbb{Z}\mathbf{v}_d.$$

Le résultat le plus fondamental de la géométrie des nombres est le premier théorème de Minkowski, qui permet de majorer la norme de la systole $\lambda_1(\Delta)$ d'un réseau en fonction du covolume. Rappelons que $\lambda_1(\Delta)$ est par définition la norme minimale d'un élément non nul de Δ.

Théorème 2.4 (Premier théorème de Minkowski)**.** *Soit Δ un réseau dans $\mathbb{R}^d$ et C une partie convexe symétrique telle que $\mathrm{vol}\, C > 2^d|\Delta|$. Alors $C \cap \Delta \neq \{0\}$. En particulier,*

$$\lambda_1(\Delta)^d \leq \frac{2^d|\Delta|}{\mathrm{vol}\, B_{\mathbb{R}^d}(0,1)}.$$

Démonstration. Soit F un domaine fondamental pour Δ dans $\mathbb{R}^d$. Le convexe $C' = \frac{1}{2}C$ vérifie $\mathrm{vol}(C') = 2^{-d}\,\mathrm{vol}(C) > |\Delta| = |F|$. Par conséquent,

$$|C'| = \sum_{v \in \Delta} |C' \cap (F + v)| = \sum_{v \in \Delta} |(C' + v) \cap F| > |F|$$

et il doit donc exister $v_1 \neq v_2 \in \Delta$ tels que $C' + v_1 \cap C' + v_2 \neq \varnothing$. Cela donne, pour certains c_1, c_2 dans C', $c_1 + v_1 = c_2 + v_2$, et comme C est convexe et symétrique, le vecteur $v = v_2 - v_1 = c_1 - c_2$ est un élément non nul de $C \cap \Delta$.

Pour la deuxième assertion, il suffit d'observer que si $\lambda^d > \frac{2^d|\Delta|}{\operatorname{vol} B_{\mathbb{R}^d}(0,1)}$, alors le convexe $C = B_{\mathbb{R}^d}(0, \lambda)$ vérifie $\operatorname{vol} C > 2^d|\Delta|$. □

Exercice 18 (Théorème des quatre carrés)**.**

1. Montrer que pour tout nombre premier p, il existe une solution non triviale (a, b, c) à l'équation $a^2 + b^2 + c^2 \equiv 0 \mod p$.
2. À l'aide du réseau $\Delta = p\mathbb{Z}^4 + \mathbb{Z}(a, b, c, 0) + \mathbb{Z}(0, -c, b, a)$, montrer que p est somme de quatre carrés.
3. Conclure que tout entier positif est somme de quatre carrés.

Solution.

1. L'ensemble des carrés modulo p est de cardinal $\frac{p+1}{2}$, de même que l'ensemble des $-1 - b^2$, donc on peut trouver a et b tels que $a^2 = -1 - b^2$. Avec $c = 1$, on trouve bien $a^2 + b^2 + c^2 \equiv 0 \mod p$.
2. L'image du réseau Δ dans $(\mathbb{Z}/p\mathbb{Z})^4$ est un sous-groupe d'ordre p^2, donc Δ est un réseau de covolume p^2 dans $\mathbb{R}^4$. D'après le théorème de Minkowski, il existe un élément (x, y, z, t) dans Δ tel que $x^2+y^2+z^2+t^2 \leq p \cdot \frac{4}{\operatorname{vol} B_{\mathbb{R}^4}(0,1)}$. Or, on calcule facilement

$$\begin{aligned}\operatorname{vol} B_{\mathbb{R}^4}(0,1) &= \int \mathbb{1}_{x^2+y^2+z^2+t^2\leq 1}\, dx\, dy\, dz\, dt \\ &= \pi \int (1 - x^2 - y^2)\mathbb{1}_{x^2+y^2\leq 1}\, dx\, dy \\ &= 2\pi^2 \int_0^1 (1 - r^2) r\, dr \\ &= \tfrac{\pi^2}{3}.\end{aligned}$$

   Par suite, $x^2+y^2+z^2+t^2 \leq p \cdot \frac{12}{\pi^2} < 2p$. Comme tout élément de Δ vérifie $x^2 + y^2 + z^2 + t^2 \equiv 0 \mod p$, on doit avoir $x^2 + y^2 + z^2 + t^2 = p$.
3. La multiplicativité de la norme sur l'algèbre des quaternions montre que le produit de deux sommes de quatre carrés est encore une somme de quatre carrés. Comme l'ensemble des sommes de quatre carrés contient tous les nombres premiers, il est égal à l'ensemble des entiers naturels.

Pour décrire plus précisément la forme d'un réseau Δ dans $\mathbb{R}^d$, on pose la définition suivante.

Définition 2.5 (Minima successifs)**.** Les *minima successifs* d'un réseau Δ dans $\mathbb{R}^d$ sont les nombres réels positifs $\lambda_1(\Delta) \leq \lambda_2(\Delta) \leq \cdots \leq \lambda_d(\Delta)$ définis par

$$\lambda_i(\Delta) = \inf\{\lambda > 0 \mid B(0, \lambda) \cap \Delta \text{ contient } i \text{ vecteurs linéairement indépendants}\}.$$

Le second théorème de Minkowski exprime le fait que les vecteurs qui réalisent les minima successifs sont toujours presque orthogonaux. Ci-dessous, l'espace $\mathbb{R}^d$ est muni de sa norme euclidienne.

Théorème 2.6 (Second théorème de Minkowski)**.** *Pour tout réseau* Δ *dans* $\mathbb{R}^d$*,*

$$|\Delta| \leq \lambda_1(\Delta)\dots\lambda_d(\Delta) \leq \frac{2^d|\Delta|}{\operatorname{vol} B_{\mathbb{R}^d}(0,1)}.$$

Démonstration. Notons $(v_i)_{1\leq i\leq d}$ une famille linéairement indépendante d'éléments de Δ tels que pour chaque i, $\|v_i\| = \lambda_i(\Delta)$. Comme le réseau $\Delta_1 = \mathbb{Z}v_1 \oplus \dots \oplus \mathbb{Z}v_d$ est inclus dans Δ, on peut majorer

$$|\Delta| \leq |\Delta_1| \leq \prod_{i=1}^{d} \|v_i\| = \lambda_1(\Delta)\dots\lambda_d(\Delta).$$

Pour l'autre inégalité du théorème, on note $(u_i)_{1\leq i\leq d}$ une base orthonormée de $\mathbb{R}^d$ (obtenue par le procédé de Gramm-Schmidt) telle que pour tout i,

$$V_i := \operatorname{Vect}(u_1,\dots,u_i) = \operatorname{Vect}(v_1,\dots,v_i).$$

Soit $T\colon \mathbb{R}^d \to \mathbb{R}^d$ l'application linéaire telle que pour tout $i = 1,\dots,d$, $Tu_i = \lambda_i(\Delta)^{-1}u_i$, et $\Delta' = T\Delta$. Montrons que $\lambda_1(\Delta') \geq 1$. Pour cela, si $v \in \Delta$, on écrit

$$v = \sum_{i=1}^{j} \alpha_i v_i, \quad \text{avec} \quad \alpha_j \neq 0.$$

Comme v est linéairement indépendant de $(v_1,\dots,v_{j-1})$, on a $\|v\| \geq \lambda_j(\Delta)$. Par ailleurs, comme (u_i) est orthonormée, $\|T^{-1}|_{V_j}\| = \lambda_j(\Delta)$, et par suite

$$\|Tv\| \geq \frac{1}{\|T^{-1}|_{V_i}\|}\|v\| \geq 1,$$

d'où $\lambda_1(\Delta') \geq 1$. Pour conclure, on applique le premier théorème de Minkowski dans le réseau Δ' :

$$1 \leq \lambda(\Delta')^d \leq \frac{2^d|\Delta'|}{\operatorname{vol} B_{\mathbb{R}^d}(0,1)} = \frac{1}{\lambda_1(\Delta)\dots\lambda_d(\Delta)} \cdot \frac{2^d|\Delta|}{\operatorname{vol} B_{\mathbb{R}^d}(0,1)}.$$

□

Exercice 19 (Norme et théorèmes de Minkowski)**.**

1. Montrer que si l'espace $\mathbb{R}^d$ est muni d'une norme arbitraire, on a toujours

$$\frac{1}{d!} \cdot |\Delta| \leq \lambda_1(\Delta)\dots\lambda_d(\Delta) \leq \frac{2^d}{\operatorname{vol} B_{\mathbb{R}^d}(0,1)} \cdot |\Delta|.$$

2. Vérifier que ces inégalités sont optimales.

Solution. 1.

2.

2.3 Sous-espaces rationnels de hauteur bornée

Comme application du second théorème de Minkowski, nous allons montrer un encadrement asymptotique du nombre de sous-espaces rationnels de dimension k et de hauteur inférieure à H.

Proposition 2.7. *Pour tout entier $d \geq 1$, pour tout $k = 1, \dots, d-1$, le nombre*

$$N_{k,d}(H) = |\{v \in X_{k,d}(\mathbb{Q}) \mid H(v) \leq H\}|$$

vérifie

$$N_{k,d}(H) \asymp_d H^d.$$

Démonstration. Montrons d'abord par récurrence sur k que $N_{k,d}(H) \lesssim H^d$. Pour $k = 1$, le résultat est clair puisque $N_{k,d}(H)$ est égal au nombre de points primitifs dans $\mathbb{Z}^d$ de norme au plus H. (Voir exercice **??** sur le théorème de Schanuel.) Supposons qu'on ait montré que pour tout H, $N_{k-1,d}(H) \lesssim_d H^d$. Par le second théorème de Minkowski, si $v \in X_k(\mathbb{Q})$ vérifie $H(v) \leq H$, alors il existe $v' \in X_{k-1}(\mathbb{Q})$ tel que $v' \leq v$ et $H(v') \lesssim H^{\frac{k-1}{k}}$. Si $v' \in X_{k-1,d}(\mathbb{Q})$, les éléments $v \in X_{k,d}(\mathbb{Q})$ contenant v' correspondent aux vecteurs primitifs du réseau $\mathbb{Z}^d/(v' \cap \mathbb{Z}^d)$, de covolume $H(v')^{-1}$ et de dimension $d-k+1$; de plus, la norme du vecteur de $\mathbb{Z}^d/(v' \cap \mathbb{Z}^d)$ correspondant à v est égale à $H(v)H(v')^{-1} \leq HH(v')^{-1}$. Par le cas $k = 1$, le nombre de tels sous-espaces est donc majoré par $\lesssim \left(HH(v')^{-1}\right)^{d-k+1} H(v')$ et par conséquent,

$$\begin{aligned}
N_{k,d}(H) &\leq \sum_{\substack{v' : \\ H(v') \lesssim H^{\frac{k-1}{k}}}} \left(HH(v')^{-1}\right)^{d-k+1} H(v') \\
&\lesssim \sum_{\substack{n : \\ 2^n \lesssim H^{\frac{k-1}{k}}}} \sum_{\substack{v' : \\ 2^n \leq H(v') < 2^{n+1}}} \left(H2^{-n}\right)^{d-k+1} 2^n \\
&\lesssim \sum_{n \,:\, 2^n \lesssim H^{\frac{k-1}{k}}} 2^{dn} \left(H2^{-n}\right)^{d-k+1} 2^n \\
&\lesssim H^{d-k+1} \cdot \left(H^{\frac{k-1}{k}}\right)^k \\
&\lesssim H^d.
\end{aligned}$$

La démonstration de l'inégalité réciproque est analogue, par induction rétrograde sur k. Remarquons d'abord qu'il existe $c > 0$ tel que $N_{d-1,d}(H) = N_{1,d}(H) \geq cH^d$. Supposons $N_{k,d}(H) \geq cH$ pour tout H. Le même raisonnement que ci-dessus donne pour une constante $c_0 = c_0(d)$,

$$\sum_{\substack{v' : \\ H(v') \leq H}} \left(H^{\frac{k}{k-1}} H(v')^{-1}\right)^{d-k+1} H(v') \geq N_{k,d}(cH^{\frac{k}{k-1}}) \geq cc_0^{\frac{dk}{k-1}} H^{\frac{dk}{k-1}}$$

Le calcul ci-dessus, avec la majoration $N_{k-1,d}(H) \lesssim H^d$, montre qu'on peut choisir $c_1 = c_1(d) > 0$ tel que

$$\sum_{\substack{v' : \\ H(v') \leq c_1 H}} \left(H^{\frac{k}{k-1}} H(v')^{-1}\right)^{d-k+1} H(v') \leq \tfrac{1}{2} cc_0^{\frac{dk}{k-1}} H^{\frac{dk}{k-1}}.$$

Cela permet de conclure

$$\begin{aligned} N_{k-1,d}(H) c_1^{-d+k} H^{\frac{d}{k-1}} &\geq \sum_{\substack{v' : \\ c_1 H \leq H(v') \leq H}} \left(H H(v')^{-1}\right)^{d-k+1} H(v') \\ &\geq \tfrac{1}{2} c c_0^{\frac{dk}{k-1}} H^{\frac{dk}{k-1}} \end{aligned}$$

et donc $N_{k-1,d}(H) \gtrsim H^d$. □

Application : Valeur heuristique de l'exposant diophantien

Supposons pour simplifier $1 \leq k \leq \ell < d$. Pour v dans $X_{k,d}(\mathbb{Q})$, l'inégalité $d(x,v) \leq \varepsilon$ définit dans $X_{\ell,d}(\mathbb{R})$ un voisinage de taille ε de la sous-variété

$$E_v = \{x \in X_{\ell,d}(\mathbb{R}) \mid x \geq v\} \subseteq X_{\ell,d}(\mathbb{R}).$$

Cette sous-variété est de dimension $(\ell-k)(d-\ell)$, et donc de codimension $k(d-\ell)$. Par conséquent, pour une mesure riemannienne sur $X_{\ell,d}(\mathbb{R})$, on peut évaluer

$$|\{x \in X_{\ell,d}(\mathbb{R}) \mid d(x,v) \leq \varepsilon\}| \asymp \varepsilon^{k(d-\ell)}.$$

Comme il y a dans $X_{k,d}(\mathbb{Q})$ à peu près H^d points de hauteur au plus H, on en déduit que la somme

$$\sum_{v \in X_{k,d}(\mathbb{Q})} |\{x \in X_{\ell,d}(\mathbb{R}) \mid d(x,v) \leq H(v)^{-\beta}\}|$$

converge si et seulement si $\beta > \frac{d}{k(d-\ell)}$. Cela suggère que $\frac{d}{k(d-\ell)}$ est une valeur critique pour l'exposant diophantien $\beta_k(x)$ défini au début de cette partie, par la formule (**??**).

eq:betak

Le but de la suite de ce cours sera de démontrer le résultat suivant. On rappelle qu'un *pinceau* dans $X_{\ell,d}$ est une sous-variété de la forme

$$\mathcal{P}_{W,r} = \{x \in X_{\ell,d}(\mathbb{R}) \mid \dim x \cap W \geq r\}$$

où $W \leq \mathbb{R}^d$ est un sous-espace vectoriel, et r un entier positif. Un tel pinceau est dit *rationnel* si le sous-espace W est rationnel, et *contraignant* si $\frac{r}{\dim W} > \frac{\ell}{d}$.

Théorème 2.8 (Approximation diophantienne dans la grassmannienne)**.** *Soient $1 \leq k \leq \ell < d$ des entiers fixés.*

(1) Pour tout x dans $X_{\ell,d}(\mathbb{R})$, $\beta_k(x) \geq \frac{d}{k(d-\ell)}$.

(2) Pour presque tout x dans $X_{\ell,d}(\mathbb{R})$, $\beta_k(x) = \frac{d}{k(d-\ell)}$.

(3) Pour tout x dans $X_{\ell,d}(\overline{\mathbb{Q}})$ non contenu dans un pinceau rationnel contraignant, $\beta_k(x) = \frac{d}{k(d-\ell)}$.

Pour la démonstration de ce théorème, nous établissons dans la partie suivante une correspondance entre les exposants $\beta_k(x)$ et l'existence de certains petits vecteurs le long d'une orbite diagonale dans un espace de réseau bien choisi.

Exercice 20. Vérifier que sans la restriction $k \leq \ell$, l'exposant critique pour l'approximation d'un élément $x \in X_{\ell,d}(\mathbb{R})$ par des sous-espaces rationnels de dimension k est égal à $\frac{d}{\min(k,\ell)(d-\max(k,\ell))}$. Expliquer pourquoi on peut déduire le cas général du cas où $k \leq \ell$.

2.4 Exercices supplémentaires

Exercice 21. Le but de cet exercice est de donner une démonstration un peu différente du premier théorème de Minkowski. Nous montrerons même le résultat un peu plus précis suivant : si Δ est un réseau de $\mathbb{R}^d$ et C un convexe symétrique, le cardinal de $\Delta \cap C$ est minoré par $\frac{|C|}{2^d \mu_d(\Delta)}$.

1. Justifier qu'on peut supposer sans perte de généralité que $\Delta = \mathbb{Z}^d$.
2. Pour $n \geq 1$, notons $\Delta_n = \frac{1}{n}\mathbb{Z}^d$. Donner un équivalent du cardinal de $\frac{C}{2} \cap \Delta_n$ lorsque n tend vers l'infini.
3. En déduire que pour n assez grand, l'application $\frac{C}{2} \cap \Delta_n \to \Delta_n/\Delta$ admet une fibre de cardinal supérieur à $\frac{|C|}{2^d}$, et conclure.
4. Retrouver le point premier **??** du théorème ci-dessus.

Exercice 22.

1. Construire un réseau Δ dans $\mathbb{R}^3$ tel que le sous-réseau qui réalise $\mu_2(\Delta)$ ne contienne pas les deux vecteurs qui réalisent les deux premiers minima $\lambda_1(\Delta)$ et $\lambda_2(\Delta)$.
2. Justifier que si Δ est un réseau dans $\mathbb{R}^d$ et $v_1, \ldots, v_d$ des vecteurs tels que pour chaque i, $\|v_i\| = \lambda_i(\Delta)$, alors le sous-groupe $\Delta_0 = \mathbb{Z}v_1 \oplus \ldots \mathbb{Z}v_d$ vérifie $[\Delta : \Delta_0] \leq \frac{2^d}{|B(0,1)|}$.
3. Construire un réseau Δ dans $\mathbb{R}^d$, $d \geq 4$ et des vecteurs $v_1, \ldots, v_d$ tels que pour chaque i, $\|v_i\| = \lambda_i(\Delta)$, mais néanmoins $\Delta > \mathbb{Z}v_1 \oplus \cdots \oplus \mathbb{Z}v_d$.

Chapitre 3

La correspondance de Dani

ch:dani

Cette partie est dédiée à la correspondance qui relie les exposants diophantiens $\beta_k(x)$, $k = 1, \dots, d-1$ d'un élément x dans $X_{\ell,d}(\mathbb{R})$ à certaines orbites dans l'espace des réseaux. Nous commençons par le cas plus simple où l'on approche x par des droites, i.e. $k = 1$.

3.1 Approximation par des droites

Le groupe $G = \mathrm{SL}_d(\mathbb{R})$ agit transitivement sur la variété $X_{\ell,d}(\mathbb{R})$, et le stabilisateur du point base $x_0 = \mathrm{Vect}(e_1, \dots, e_\ell)$ est égal au sous-groupe parabolique

$$P = \left\{ g \in G \mid g = \begin{pmatrix} A & B \\ 0 & C \end{pmatrix}, \ \dots \right\}.$$

Nous utiliserons dans la suite l'identification

$$\begin{array}{ccc} P\backslash G & \to & X_{\ell,d}(\mathbb{R}) \\ Pg & \mapsto & g^{-1}x_0 \end{array}$$

et le sous-groupe diagonal à un paramètre

$$a_t = \mathrm{diag}(e^{-\frac{(d-\ell)t}{d}}, \dots, e^{-\frac{(d-\ell)t}{d}}, e^{-\frac{(d-\ell)t}{d}}, \dots, e^{-\frac{(d-\ell)t}{d}}).$$

Proposition 3.1 (Correspondance de Dani, $k = 1$)**.** *Soit $x \in X_{\ell,d}(\mathbb{R})$ et $u_x \in G$ tel que $x = Pu_x$. L'exposant diophantien pour l'approximation de x par des droites rationnelles est donné par*

$$\beta_1(x) = \frac{1}{\frac{d-\ell}{d} - \gamma_1(x)},$$

où $\gamma_1(x) = \limsup_{t\to+\infty} \frac{-1}{t} \log \lambda_1(a_t u_x \mathbb{Z}^d)$.

Démonstration. Soit $\beta < \beta_1(x)$. Par définition de $\beta_1(x)$, il existe $v \in \mathbb{P}^1(\mathbb{R})$ arbitrairement proche de x tel que $d(x, v) \leq H(v)^{-\beta}$. Soit $\mathbf{v} \in \mathbb{Z}^d$ un vecteur primitif tel que $v = \mathbb{R}\mathbf{v}$. On décompose alors $u_x\mathbf{v}$ suivant les espaces propres de a_t :

$$u_x\mathbf{v} = \mathbf{v}_x^{(0)} + \mathbf{v}_x^{(1)}$$

de sorte que

$$a_t u_x \mathbf{v} = e^{-\frac{(d-\ell)t}{d}} \mathbf{v}_x^{(0)} + e^{\frac{\ell t}{d}} \mathbf{v}_x^{(1)}.$$

À certaines constantes près dépendant du choix de u_x, grâce au fait que v est proche de x, on a

$$H(v) = \|\mathbf{v}\| \asymp \|u_x \mathbf{v}\| \asymp \|\mathbf{v}_x^{(0)}\|$$

et

$$d(v,x) = d(v, u_x^{-1} x_0) \asymp d(u_x v, x_0) \asymp H(v)^{-1} \|\mathbf{v}_x^{(1)}\|.$$

Par conséquent,

$$\|a_t u_x \mathbf{v}\| \asymp H(v) \max\left(e^{-\frac{(d-\ell)t}{d}}, e^{\frac{\ell t}{d}} d(x,v)\right).$$

On choisit alors $t > 0$ tel que $e^t = H(v)^\beta$, ce qui donne

$$\|a_t u_x \mathbf{v}\| \lesssim e^{-t\left(\frac{d-\ell}{d} - \frac{1}{\beta}\right)}$$

puis $\gamma_1(x) \geq \frac{d-\ell}{d} - \frac{1}{\beta}$, i.e. $\beta_1(x) \leq \frac{1}{\frac{d-\ell}{d} - \gamma_1(x)}$.

Réciproquement, supposons $\gamma \in (0, \frac{d-\ell}{d})$ et pour $t > 0$ arbitrairement grand, $\mathbf{v} \in \mathbb{Z}^d$ vérifie $\|a_t u_x \mathbf{v}\| \leq e^{-\gamma t}$. Si l'on note v la droite engendrée par le vecteur $\mathbf{v}$, alors

$$\|a_t u_x \mathbf{v}\| \asymp H(v) \max\left(e^{-\frac{(d-\ell)t}{d}}, e^{\frac{\ell t}{d}} d(x,v)\right) \leq e^{-\gamma t},$$

donc $H(v) \leq e^{(\frac{d-\ell}{d} - \gamma)t}$ et

$$d(x,v) \leq H(v)^{-1} e^{-t(\gamma + \frac{\ell}{d})} \leq H(v)^{-1 - \frac{\gamma + \frac{\ell}{d}}{\frac{d-\ell}{d} - \gamma}} = H(v)^{-\frac{1}{\frac{d-\ell}{d} - \gamma}}.$$

□

Comme première application de cette correspondance, nous pouvons retrouver le théorème de Dirichlet, qui donne la valeur minimale de l'exposant diophantien $\beta_1(x)$.

Corollaire 3.2 (Dirichlet, Minoration de l'exposant diophantien)**.** *Pour tout x dans $X_{\ell,d}(\mathbb{R})$, $\beta_1(x) \geq \frac{d}{d-\ell}$.*

Démonstration. Comme l'action du sous-groupe $(a_t)_{t\in\mathbb{R}}$ préserve le volume, le réseau $a_t u_x \mathbb{Z}^d$ est de covolume constant, et d'après le premier théorème de Minkowski, cela implique $\lambda_1(a_t u_x \mathbb{Z}^d) \lesssim 1$. Par conséquent, $\gamma_1(x) \geq 0$, puis $\beta_1(x) \geq \frac{d}{d-\ell}$. □

Nous verrons au paragraphe **??** que la correspondance de Dani permet aussi d'obtenir la valeur presque sûre de l'exposant diophantien. Mais dans le cas particulier $k = 1$, il est aussi simple de montrer directement le résultat.

Exercice 23. Vérifier que pour presque tout x dans $X_{\ell,d}(\mathbb{R})$, $\beta_1(x) = \frac{d}{d-\ell}$.

Solution. Pour v dans $\mathbb{P}^1(\mathbb{Q})$, l'ensemble

$$\{x \in X_{\ell,d}(\mathbb{R}) \mid d(v,x) \leq \varepsilon\}$$

est un voisinage de taille ε de la sous-variété $E_v = \{x \in X_{\ell,d}(\mathbb{R}) \mid x \supseteq v\}$. La variété E_v est de dimension $(\ell-1)(d-\ell)$ (donc de codimension $d-\ell$ dans $X_{\ell,d}(\mathbb{R})$) donc

$$\big|\{x \in X_{\ell,d}(\mathbb{R}) \mid d(v,x) \leq H(v)^{-\beta}\}\big| \asymp H(v)^{-\beta(d-\ell)}.$$

Si $\beta > \frac{d}{d-\ell}$, en utilisant le fait que $\big|\{v \in \mathbb{P}^1(\mathbb{Q}) \mid H(v) \leq H\}\big| \lesssim H^d$, on trouve

$$\sum_{v\in\mathbb{P}^1(\mathbb{Q})} \big|\{x \in X_{\ell,d}(\mathbb{R}) \mid d(v,x) \leq H(v)^{-\beta}\}\big| < +\infty,$$

et avec le lemme de Borel-Cantelli, cela montre que pour presque tout x, l'inégalité $d(v,x) < H(v)^{-\beta}$ n'a qu'un nombre fini de solutions, i.e. $\beta_1(x) \leq \frac{d}{d-\ell}$.

3.2 Approximation par des sous-espaces

Avec quelques modifications, on peut généraliser la correspondance de Dani pour comprendre l'exposant $\beta_k(x)$ à partir d'une orbite diagonale dans un espace de réseaux. Ici encore, on note $G = \mathrm{SL}_d(\mathbb{R})$, P le stabilisateur du sous-espace $x_0 = \mathrm{Vect}(e_1,\dots,e_\ell)$, et on utilise l'identification

$$\begin{array}{ccc} P\backslash G & \to & X_{\ell,d}(\mathbb{R}) \\ Pg & \mapsto & g^{-1}x_0. \end{array}$$

Rappelons que la puissance extérieure $\wedge^k\mathbb{R}^d$ est un espace vectoriel engendrée par une base (e_I), où I décrit l'ensemble des parties de $\{1,\dots,d\}$ à k éléments. Il existe une unique application k-linéaire alternée

$$\begin{array}{ccc} \mathbb{R}^d \times \cdots \times \mathbb{R}^d & \to & \wedge^k\mathbb{R}^d \\ (v_1,\dots,v_k) & \mapsto & v_1 \wedge \cdots \wedge v_k \end{array}$$

telle que pour tout $I = \{i_1 < i_2 < \cdots < i_k\}$, $e_{i_1} \wedge \cdots \wedge e_{i_k} = e_I$. Un élément de $\wedge^k\mathbb{R}^d$ de la forme $v_1 \wedge \dots v_k$ est dit *décomposable* ; notons que dès que $k \notin \{1, d-1\}$, il existe des vecteurs non décomposables, par exemple $e_1\wedge e_2 + e_3 \wedge e_4$ dans $\wedge^2\mathbb{R}^4$.

Le groupe $G = \mathrm{SL}_d(\mathbb{R})$ agit linéairement sur $\wedge^k\mathbb{R}^d$ via la formule

$$g \cdot (v_1 \wedge \cdots \wedge v_k) = gv_1 \wedge \cdots \wedge gv_k$$

étendue par linéarité à $\wedge^k\mathbb{R}^d$ tout entier.

À un sous-groupe discret $\Lambda = \mathbb{Z}v_1 \oplus \cdots \oplus v_k$ dans $\mathbb{R}^d$, on associe le vecteur $\mathbf{w}_\Lambda = v_1 \wedge \cdots \wedge v_k$, qui, au signe près, ne dépend du choix de la base $v_1,\dots,v_k$. Si l'on munit $\wedge^k\mathbb{R}^d$ de la structure euclidienne pour laquelle la base (e_I) est orthonormée, on peut relier covolume et norme dans $\wedge^k\mathbb{R}^d$.

Proposition 3.3. *Si Λ est un sous-groupe discret de rang k dans $\mathbb{R}^d$, alors $|\Lambda| = \|\mathbf{w}_\Lambda\|$.*

Démonstration. Montrons d'abord que $K = \mathrm{SO}_d(\mathbb{R})$ agit par isométries sur $\wedge^k\mathbb{R}^d$. Si $J = \{j_1 < j_2 < \cdots < j_k\}$, on calcule

$$ge_J = ge_{j_1} \wedge \cdots \wedge ge_{j_k} = \sum_{i_1,\ldots,i_k}^{d} g_{i_1 j_1} \ldots g_{i_k j_k} e_{i_1} \wedge \cdots \wedge e_{i_k}.$$

et de même, pour $J = \{j'_1 < \cdots < j'_k\}$,

$$ge_{J'} = \sum_{i_1,\ldots,i_k}^{d} g_{i_1 j'_1} \ldots g_{i_k j'_k} e_{i_1} \wedge \cdots \wedge e_{i_k}.$$

Par conséquent,

$$\langle ge_J, ge_{J'} \rangle = \sum_{i_1,\ldots,i_k} g_{i_1 j_1} \ldots g_{i_k j_k} g_{i_1 j'_1} \ldots g_{i_k j'_k} = ({}^t g g)_{j_1 j'_1} \ldots ({}^t g g)_{j_k j'_k},$$

et si $g \in \mathrm{SO}_d(\mathbb{R})$, ${}^t g g = 1$ donc $\langle ge_J, ge_{J'} \rangle = \delta_{JJ'}$. L'image par g de la base orthonormée (e_I) est orthonormée, donc g induit une isométrie de $\wedge^k\mathbb{R}^d$.

Soit maintenant Λ un sous-groupe discret de rang k dans $\mathbb{R}^d$. Quitte à multiplier Λ par un élément $r \in \mathrm{SO}_d(\mathbb{R})$ (qui préserve à la fois le covolume et la norme sur $\wedge^k\mathbb{R}^d$), on peut supposer que l'espace vectoriel engendré par Λ est égal à $\mathrm{Vect}(e_1, \ldots, e_k)$. Soit $g_1 \in \mathrm{GL}_k(\mathbb{R})$ tel que $\Lambda = g_1\mathbb{Z}^k$, et $g \in \mathrm{GL}_d(\mathbb{R})$ tel que $g = g_1$ sur $\mathrm{Vect}(e_1, \ldots, e_k)$ et $g = 1$ sur $\mathrm{Vect}(e_{k+1}, \ldots, e_d)$. Alors,

$$\|\mathbf{w}_\Lambda\| = \|g \cdot (e_1 \wedge \cdots \wedge e_k)\| = |\det g_1|$$

tandis que par les propriétés de la mesure de Lebesgue sur $\mathbb{R}^k$,

$$|\Lambda| = |\det g_1|.$$

Cela donne l'égalité souhaitée. □

Rappelons que le sous-groupe diagonal $(a_t)_{t\in\mathbb{R}}$ est défini par

$$a_t = \mathrm{diag}(e^{-\frac{(d-\ell)t}{d}}, \ldots, e^{-\frac{(d-\ell)t}{d}}, e^{\frac{(d-\ell)t}{d}}, \ldots, e^{\frac{(d-\ell)t}{d}}).$$

Si $k \leq \ell$, l'élément a_t agit sur $\wedge^k\mathbb{R}^d$ avec pour valeurs propres

$$e^{-\frac{k(d-\ell)}{d}t}, e^{-\left(\frac{k(d-\ell)}{d}-1\right)t}, e^{-\left(\frac{k(d-\ell)}{d}-2\right)t}, \ldots$$

On note $\pi^+ \colon \wedge^k\mathbb{R}^d \to \wedge^k\mathbb{R}^d$ le projecteur spectral de a_t associé à la valeur propre $e^{-\frac{k(d-\ell)}{d}t}$; en d'autres termes, π^+ est la projection orthogonale sur l'espace $\mathrm{Vect}(e_I \ ; \ I \subset \{1, \ldots, \ell\})$.

Si Λ est un sous-groupe additif de $\mathbb{R}^d$, on note $\wedge^k\Lambda$ le sous-groupe de $\wedge^k\mathbb{R}^d$ engendré par les vecteurs de la forme $v_1 \wedge \cdots \wedge v_k$, où les v_i sont des éléments de Λ. On laisse le soin au lecteur de vérifier que si Λ est un sous-groupe discret (resp. un réseau) de $\mathbb{R}^d$, alors $\wedge^k\Lambda$ est un sous-groupe discret (resp. un réseau) de $\wedge^k\mathbb{R}^d$.

Exercice 24. Vérifier que si Λ est un sous-groupe discret de rang ℓ dans $\mathbb{R}^d$, alors, pour tout $k \leq \ell$, $\wedge^k\Lambda$ est un sous-groupe discret de $\wedge^k\mathbb{R}^d$ de covolume $|\wedge^k\Lambda| = |\Lambda|^{\frac{k}{\ell}\binom{\ell}{k}}$.

Solution. Écrivons $\Lambda = g\mathbb{Z}^\ell$, où $\mathbb{Z}^\ell = \mathbb{Z}e_1 \oplus \cdots \oplus \mathbb{Z}e_\ell$ et $g \in \mathrm{GL}_d(\mathbb{R})$. L'élément g peut s'écrire $g = k_1 a k_2$, avec $k_1, k_2 \in O_d(\mathbb{R})$ et $a = \mathrm{diag}(a_1, \dots, a_d)$, $a_i > 0$. Comme k_1 et k_2 agissent par isométries sur $\mathbb{R}^d$ et toutes ses puissances extérieures, il suffit de vérifier le résultat lorsque $g = a$. Or,

$$|a\mathbb{Z}^\ell| = \prod_{i=1}^{\ell} a_i$$

tandis que le covolume de $\wedge^k(a\mathbb{Z}^\ell)$ est égal au produit des valeurs propres de a sur $\wedge^k\mathbb{Z}^\ell$, i.e.

$$|\wedge^k(a\mathbb{Z}^\ell)| = \prod_{1\le i_1<\cdots<i_k\le\ell} a_{i_1}\dots a_{i_k} = \left(\prod_{i=1}^{\ell} a_i\right)^{\frac{k}{\ell}\binom{\ell}{k}}.$$

Cela montre l'égalité souhaitée.

La généralisation de la correspondance de Dani à l'approximation par des sous-espaces de dimension k fait intervenir les petits vecteurs dans des réseaux de $\wedge^k\mathbb{R}^d$. Il convient de noter qu'outre la norme de ces petits vecteurs, on doit aussi contrôler leur direction.

pr:danigen **Proposition 3.4** (Correspondance de Dani généralisée)**.** *Fixons des entiers $1 \le k \le \ell < d$. Soit x dans $X_{\ell,d}(\mathbb{R})$ et $u_x \in G$ tel que $x = Pu_x$. Alors,*

$$\beta_k(x) = \frac{1}{\frac{k(d-\ell)}{d} - \gamma_k(x)},$$

où

$$\gamma_k(x) = \sup\left\{\gamma \in \mathbb{R} \;\middle|\; \exists t \to +\infty :\ \exists \mathbf{w} \in a_t u_x \wedge^k \mathbb{Z}^d :\ \begin{array}{c}\|\mathbf{w}\| \le e^{-\gamma t}\\ \text{et } \|\pi^+(\mathbf{w})\| \ge \frac{1}{2}\|\mathbf{w}\|\end{array}\right\}.$$

Démonstration. Soit $\beta < \beta_k(x)$. Il existe $v \in X_{k,d}(\mathbb{Q})$ proche de x tel que $d(v,x) \le H(v)^{-\beta}$. Soit $\mathbf{v} \in \wedge^k\mathbb{Z}^d$ le vecteur associé à v, de sorte que $H(v) = \|\mathbf{v}\|$. On décompose $u_x\mathbf{v}$ suivant les espaces propres de a_t :

$$u_x\mathbf{v} = \mathbf{v}_x^{(0)} + \mathbf{v}_x^{(1)} + \dots$$

et donc

$$a_t u_x \mathbf{v} = e^{-\frac{k(d-\ell)}{d}t}\mathbf{v}_x^{(0)} + e^{-\left(\frac{k(d-\ell)}{d}-1\right)t}\mathbf{v}_x^{(1)} + \dots$$

Pour contrôler la norme de $a_t u_x \mathbf{v}$, nous aurons besoin du lemme suivant.

lm:vr **Lemme 3.5.** *Pour tout $v \in X_{k,d}(\mathbb{Q})$ proche de x, à certaines constantes multiplicatives près dépendant du choix de u_x, on a :*

(i) $\|\mathbf{v}_x^{(0)}\| \asymp H(v)$;

(ii) $\|\mathbf{v}_x^{(1)}\| \asymp H(v)d(v,x)$;

(iii) $\forall r \ge 2, \quad \|\mathbf{v}_x^{(r)}\| \lesssim H(v)d(v,x)^r$.

Démonstration. Tout d'abord,

$$H(v) = \|\mathbf{v}\| \asymp \|u_x \mathbf{v}\| \asymp \max_{r\geq 0} \|\mathbf{v}_x^{(r)}\|.$$

Par ailleurs, si $E_x = \{y \in X_{k,d}(\mathbb{R}) \mid y \subseteq x\}$, alors $d(v,x) \asymp d(v, E_x)$. Notons $x_0 = \mathrm{Vect}(e_1, \dots, e_\ell)$ et $V^+ = \wedge^k x_0 \subseteq \wedge^k \mathbb{R}^d$. Dans le plongement $X_{k,d}(\mathbb{R}) \hookrightarrow \mathbb{P}(\wedge^k \mathbb{R}^d)$, la sous-variété E_x s'envoie sur $u_x^{-1} V^+$, donc

$$\begin{aligned} d(v,x) &\asymp \tfrac{1}{H(v)} d(u_x v, V^+) \\ &\asymp \tfrac{1}{H(v)} \max_{r\geq 1} \|\mathbf{v}_x^{(r)}\|. \end{aligned}$$

Si v est assez proche de x (i.e. $d(v,x)$ assez petit), cela implique que $\max_{r\geq 1}\|\mathbf{v}_x^{(r)}\|$ est petit devant $H(v) \asymp \max_{r\geq 0}\|\mathbf{v}_x^{(r)}\|$, et donc

$$\|\mathbf{v}_x^{(0)}\| = \max_{r\geq 0} \|\mathbf{v}_x^{(r)}\| \asymp H(v).$$

Ensuite, quitte à permuter les vecteurs de la base canonique, on peut écrire

$$\frac{1}{\|\mathbf{v}_x^{(0)}\|} u_x \mathbf{v} = \begin{pmatrix} I_k & 0 & 0 \\ 0 & I_{\ell-k} & 0 \\ (u_{ij}) & 0 & I_{d-\ell} \end{pmatrix} e_{\{1,\dots,k\}}.$$

Alors, $d(v,x) \asymp \max_{i,j} |u_{ij}|$, tandis que

$$\frac{\mathbf{v}_x^{(1)}}{\|\mathbf{v}_x^{(0)}\|} = \sum_{\substack{1\leq j\leq k \\ \ell < i \leq d}} \pm u_{ij} e_{(\{1,\dots,k\}\setminus\{j\})\cup\{i\}}$$

donc

$$\|\mathbf{v}_x^{(1)}\| \asymp \|\mathbf{v}_x^{(0)}\| d(v,x) \asymp H(v) d(v,x).$$

Enfin, pour $r \geq 2$, les coordonnées de $\frac{\mathbf{v}_x^{(r)}}{\|\mathbf{v}_x^{(0)}\|}$ sont des polynômes homogènes de degré r en les variables u_{ij}, donc

$$\frac{\|\mathbf{v}_x^{(r)}\|}{\|\mathbf{v}_x^{(0)}\|} \lesssim \left(\max_{i,j} |u_{ij}|\right)^r$$

puis

$$\|\mathbf{v}_x^{(r)}\| \lesssim H(v) d(v,x)^r.$$

□

Avec ce lemme, on majore

$$\begin{aligned} \|a_t u_x \mathbf{v}\| &\lesssim \max\left(e^{-\frac{k(d-\ell)}{d}t}\|\mathbf{v}_x^{(0)}\|, e^{-\left(\frac{k(d-\ell)}{d}-1\right)t}\|\mathbf{v}_x^{(1)}\|, \dots\right) \\ &\lesssim H(v) e^{-\frac{k(d-\ell)}{d}t} \max\left(1, e^{-t} d(v,x), e^{-2t} d(v,x)^2, \dots\right) \end{aligned}$$

et choisissant t tel que $e^t = H(v)^\beta$, on obtient

$$\|a_t u_x \mathbf{v}\| \lesssim H(v) e^{-\frac{k(d-\ell)}{d}t} = e^{-\left(\frac{k(d-\ell)}{d} - \frac{1}{\beta}\right)t}.$$

Comme la plus grande coordonnée de $a_t u_x \mathbf{v}$ est atteinte le long de V^+, on a aussi (quitte à diminuer t d'une constante) $\|\pi^+(a_t u_x \mathbf{v})\| \geq \frac{1}{2}\|a_t u_x \mathbf{v}\|$ et donc $\gamma_k(x) \geq \frac{k(d-\ell)}{d} - \frac{1}{\beta}$ i.e.

$$\beta_k(x) \leq \frac{1}{\frac{k(d-\ell)}{d} - \gamma_k(x)}.$$

Réciproquement, supposons que pour $t > 0$ arbitrairement grand, on puisse trouver $\mathbf{v} \in \wedge^k \mathbb{Z}^d$ tel que

$$\|a_t u_x \mathbf{v}\| \leq e^{-\gamma t} \quad \text{et} \quad \|\pi^+(a_t u_x \mathbf{v})\| \geq \tfrac{1}{2}\|a_t u_x \mathbf{v}\|.$$

D'après le lemme de Mahler ci-dessous, on peut supposer que $\mathbf{v}$ est un élément décomposable de $\wedge^k \mathbb{Z}^d$. Soit $v \in X_{k,d}(\mathbb{Q})$ l'élément correspondant à $\mathbf{v}$. Avec les notations de la première partie de la démonstration, on a alors

$$\max\left(e^{-\frac{k(d-\ell)}{d}t}\|\mathbf{v}_x^{(0)}\|, e^{-\left(\frac{k(d-\ell)}{d}-1\right)t}\|\mathbf{v}_x^{(1)}\|, \dots\right) \lesssim \|a_t u_x \mathbf{v}\| \leq e^{-\gamma t} \qquad (3.1)$$

eq:gamma

et le maximum doit être essentiellement atteint sur la première coordonnée, i.e.

$$\|\mathbf{v}_x^{(0)}\| \gtrsim \max_{r \geq 1} e^{rt}\|\mathbf{v}_x^{(r)}\|.$$

Comme $t > 0$ est arbitrairement grand, cela implique que $d(v,x) \leq \frac{\max_{r\geq 1}\|\mathbf{v}_x^{(r)}\|}{\|\mathbf{v}_x^{(0)}\|}$ est arbitrairement petit. On peut donc appliquer le lemme **??** : $\|\mathbf{v}_x^{(0)}\| \asymp H(v)$ et $\|\mathbf{v}_x^{(1)}\| \asymp H(v)d(v,x)$. L'inégalité (**??**) donne alors

lm:vr

eq:gamma

$$H(v) \asymp \|\mathbf{v}_x^{(0)}\| \lesssim e^{\left(\frac{k(d-\ell)}{d}-\gamma\right)t}$$

tandis que $\|\pi^+(a_t u_x \mathbf{v})\| \geq \frac{1}{2}\|a_t u_x \mathbf{v}\|$ implique

$$d(v,x) \lesssim e^{-t} \lesssim H(v)^{-\frac{1}{\frac{k(d-\ell)}{d}-\gamma}}.$$

□

Nons démontrons maintenant le lemme de Mahler que nous avons utilisé dans la démonstration ci-dessus.

lm:mahler

Lemme 3.6 (Mahler)**.** *Soit Δ un réseau dans $\mathbb{R}^d$. Les minima successifs du réseau $\wedge^k\Delta$ dans $\wedge^k\mathbb{R}^d$ sont essentiellement égaux aux nombres*

$$\lambda_I(\Delta) = \lambda_{i_1}(\Delta)\dots\lambda_{i_k}(\Delta); \qquad I = \{i_1 < i_2 < \dots < i_k\} \subseteq \{1,\dots,d\}.$$

De plus, $\wedge^k\Delta$ contient une famille de vecteurs décomposables qui réalisent ces minima successifs à une constante multiplicative près ne dépendant que de d.

Démonstration. Soit $(v_i)_{1\leq i\leq d}$ une famille de vecteurs linéairement indépendants dans Δ tels que pour chaque i, $\|v_i\| = \lambda_i(\Delta)$. Pour $I = \{i_1 < \dots < i_k\}$, le vecteur $\mathbf{v}_I = v_{i_1} \wedge \dots \wedge v_{i_k}$ vérifie

$$\|\mathbf{v}_I\| \leq \lambda_{i_1}(\Delta)\dots\lambda_{i_k}(\Delta) = \lambda_I(\Delta).$$

Mais par ailleurs, d'après le second théorème de Minkowski appliqué au réseau Δ,

$$\prod_I \lambda_I(\Delta) = \left(\prod_{i=1}^d \lambda_i(\Delta)\right)^{\frac{k}{d}\binom{d}{k}} \asymp |\Delta|^{\frac{k}{d}\binom{d}{k}} = |\wedge^k \Delta|.$$

Les vecteurs $\mathbf{v}_I \in \wedge^k\Delta$ sont linéairement indépendants et vérifient $\prod_I \|\mathbf{v}_I\| \asymp |\wedge^k\Delta|$; le second théorème de Minkowski (dans $\wedge^k\Delta$) implique qu'ils réalisent les minima successifs de $\wedge^k\Delta$ à une constante près. □

Lorsque $k \geq 2$, la condition $\|\pi^+(\mathbf{w})\| \geq \frac{1}{2}\|\mathbf{w}\|$ est indispensable pour définir $\gamma_k(x)$, et on ne peut donc pas appliquer le premier théorème de Minkowski pour conclure que $\gamma_k(x) \geq 0$ pour tout x. Cependant, comme nous l'expliquons dans le paragraphe suivant, la correspondance permet déjà de calculer la valeur de $\gamma_k(x)$ pour presque tout x dans $X_{\ell,d}(\mathbb{R})$.

3.3 Valeur presque sûre de $\beta_k(x)$

ss:betaps

Commençons par un corollaire important de la correspondance établie au paragraphe précédent.

cor:nulextremal **Corollaire 3.7** (Taux de fuite nul ⇒ extrémalité)**.** *Si $x \in X_{\ell,d}(\mathbb{R})$ vérifie*

$$\lim_{t\to+\infty} \tfrac{1}{t}\log\lambda_1(a_t u_x \mathbb{Z}^d) = 0,$$

alors $\beta_k(x) = \frac{d}{k(d-\ell)}$ pour tout $k = 1, \ldots, \ell$.

Démonstration. Soit $x \in X_\ell(\mathbb{R})$ tel que $\lim_{t\to\infty}\frac{1}{t}\log\lambda_1(a_t s_x\mathbb{Z}^d) = 0$. Pour tout $\varepsilon > 0$ on a, pour tout $t > 0$ suffisamment grand, $\lambda_1(a_t s_x\mathbb{Z}^d) \geq e^{-\varepsilon t}$. Par le second théorème de Minkowski, cela implique

$$e^{-\varepsilon t} \leq \lambda_1(a_t s_x\mathbb{Z}^d) \leq \cdots \leq \lambda_d(a_t s_x\mathbb{Z}^d) \leq e^{d\varepsilon t}.$$

Soient $u_1, \ldots, u_d$ dans $a_t s_x\mathbb{Z}^d$ des vecteurs qui réalisent ces minima successifs. D'après le lemme **??** ci-dessus, les vecteurs

$$\mathbf{u}_\tau = u_{\tau_1} \wedge \cdots \wedge u_{\tau_k}, \qquad \tau = \{\tau_1 < \cdots < \tau_k\} \subset \{1, \ldots, d\},\ \operatorname{card}\tau = k$$

réalisent les minima successifs de $\wedge^k a_t s_x\mathbb{Z}^d$ à une constante multiplicative près qui ne dépend que de d. En particulier, en faisant tendre ε vers 0, on trouve $\lim_{t\to\infty}\frac{1}{t}\log\lambda_1(\wedge^k a_t s_x\mathbb{Z}^d) = 0$ ce qui implique $\gamma_k(x) \leq 0$ i.e. $\beta_k(x) \leq \frac{d}{k(d-\ell)}$.

Pour l'inégalité réciproque, notons que les vecteurs $\mathbf{u}_\tau$ engendrent un sous-réseau d'indice borné dans $\wedge^k a_t s_x\mathbb{Z}^d$, et forment une famille essentiellement orthogonale. Par conséquent, il existe τ tel que le vecteur $\mathbf{u} = \mathbf{u}_\tau$ vérifie $\|\pi^+(\mathbf{u})\| \gtrsim \|\mathbf{u}\|$. Comme on a aussi $\|\mathbf{u}_\tau\| \lesssim e^{kd\varepsilon}$, cela donne

$$\gamma_k(x) \gtrsim kd\varepsilon$$

puis, en faisant tendre ε vers zéro, $\gamma_k(x) \geq 0$. Par la proposition **??**, cela implique $\beta_k(x) \geq \frac{d}{k(d-\ell)}$. □

Avec quelques propriétés élémentaires de la mesure de Haar sur l'espace des réseaux, cette observation permet de démontrer le théorème suivant, qui est le but de ce paragraphe. Ci-dessous, et dans toute la suite, la variété $X_{\ell,d}(\mathbb{R})$ est munie d'une « mesure de Lebesgue », i.e. de n'importe quelle mesure sur $X_{\ell,d}(\mathbb{R})$ équivalente à la mesure de Hausdorff en dimension $\dim X_{\ell,d} = \ell(d-\ell)$ pour une métrique riemannienne.

th:beta1ps **Théorème 3.8** (Valeur presque sûre de l'exposant diophantien)**.** *Soit des entiers $1 \leq k \leq \ell < d$. Pour presque tout x dans $X_{\ell,d}(\mathbb{R})$, $\beta_k(x) = \frac{d}{k(d-\ell)}$.*

Comme ci-dessus, on note $G = \mathrm{SL}_d(\mathbb{R})$ et $\Gamma = \mathrm{SL}_d(\mathbb{Z})$. L'espace Ω des réseaux unimodulaires dans $\mathbb{R}^d$ s'identifie à G/Γ et supporte une unique mesure m_Ω invariante à gauche par G telle que

$$\forall f \in C_c(G), \qquad \int_G f = \int_{G/\Gamma} \left(\sum_{\gamma \in \Gamma} f(g\gamma) \right) dm_\Omega(g\Gamma).$$

Le théorème ?? découlera facilement du lemme suivant.

lm:voispointe **Lemme 3.9** (Mesure d'un voisinage de la pointe)**.** *Étant donné un entier $d \geq 1$, à certaines constantes multiplicatives près ne dépendant que de d, pour tout $\varepsilon > 0$,*

$$m_\Omega(\{\Delta \mid \lambda_1(\Delta) < \varepsilon\}) \lesssim \varepsilon^d.$$

Admettons momentanément ce lemme, et voyons comment en déduire le théorème ??.

Démonstration du théorème ??. Comme la mesure m_Ω est invariante par G, la borne du lemme ci-dessus donne, pour tout $t > 0$,

$$m_\Omega(\{\Delta \mid \lambda_1(a_t\Delta) < e^{-\varepsilon t}\}) \lesssim e^{-d\varepsilon t}.$$

Avec le lemme de Borel-Cantelli, cela implique, pour presque tout Δ dans Ω,

$$\lim_{t \to +\infty} \frac{1}{t} \log \lambda_1(a_t\Delta) = 0.$$

Or, cette propriété est invariante par translation de Δ par un élément du sous-groupe parabolique $P = \mathrm{Stab}_G \mathrm{Vect}(e_1, \ldots, e_\ell)$, puisque pour tout p dans P, l'élément $a_t p a_{-t}$ converge lorsque t tend vers $+\infty$ et qu'on peut écrire $a_t p \Delta = (a_t p a_{-t}) a_t \Delta$. Par conséquent, on a aussi, pour presque tout x dans $X_{\ell,d}(\mathbb{R}) \simeq P\backslash G$,

$$\lim_{t \to +\infty} \frac{1}{t} \log \lambda_1(a_t u_x \mathbb{Z}^d) = 0$$

et d'après le corollaire ??, cela implique $\beta_k(x) = \frac{d}{k(d-\ell)}$ pour tout $k = 1, \ldots, \ell$. □

Le dernier paragraphe de ce chapitre a pour but de démontrer l'encadrement asymptotique donné par le lemme ??; nous y verrons au passage une importante formule de Siegel, qui permet de mieux comprendre la mesure m_Ω sur l'espace des réseaux.

3.4 Formule de Siegel

Si l'on note L le stabilisateur du vecteur e_1 dans la représentation standard, l'espace quotient $G/L \simeq \mathbb{R}^d \setminus \{0\}$ supporte une unique mesure invariante $m_{G/L}$ telle que

$$\forall f \in C_c(G), \qquad \int_G f = \int_{G/L} \left(\int_L f(g\ell)\, d\ell \right) dm_{G/L}(gL).$$

À un facteur près, $m_{G/L}$ coïncide avec la mesure de Lebesgue sur $\mathbb{R}^d$. **Exercice.** Calculer le facteur de proportionnalité si m_Ω est une mesure de probabilité. Les mesures m_Ω et $m_{G/L}$ sont reliées par le théorème suivant.

Théorème 3.10 (Formule de Siegel)**.** *Si $f \in C_c(\mathbb{R}^d)$, la transformée de Siegel $\tilde{f}$ de f, définie sur Ω par l'expression*

$$\tilde{f}(\Delta) = \sum_{\mathbf{v} \in \Delta \text{ primitif}} f(\mathbf{v})$$

vérifie

$$\int_\Omega \tilde{f} = \int_{\mathbb{R}^d} f.$$

Démonstration. Notons $\Gamma_L = \Gamma \cap L$, et montrons que si $\phi \in C_c(G/\Gamma_L)$, alors

$$\int_{G/\Gamma_L} = \int_{G/L} \left(\int_{L/\Gamma_L} \phi(gu\Gamma_L)\, du \right) d(gL).$$

Comme tout élément ϕ dans $C_c(G/\Gamma_L)$ peut s'obtenir comme projection d'un élément de $C_c(G)$ (cf. Raghunathan, Lemma 1.1), il suffit de vérifier l'égalité lorsque ϕ est de la forme

$$\phi(g\Gamma_L) = \sum_{\gamma \in \Gamma_L} \psi(g\gamma), \quad \text{avec } \psi \in C_c(G).$$

Dans ce cas, la formule découle des définitions des mesures de Haar sur G/L et L/Γ_L.

Le même raisonnement, en échangeant les rôles de Γ et L, montre que

$$\int_{G/\Gamma_L} \phi = \int_{G/\Gamma} \left(\sum_{\Gamma/\Gamma_L} \phi(g\gamma\Gamma_L) \right) d(g\Gamma).$$

Choisissant $\phi(g\Gamma_L) = f(ge_1)$, on obtient

$$\int_{L/\Gamma_L} f(gue_1)\, du = \mathrm{vol}(L/\Gamma_L) \cdot f(ge_1) = \mathrm{vol}(L/\Gamma_L) \cdot f(gU)$$

tandis que

$$\sum_{\gamma \in \Gamma/\Gamma_L} f(g\gamma e_1) = \sum_{v \in \Delta \text{ primitif}} f(v) = \tilde{f}(g\Gamma).$$

Si l'on sait que $\mathrm{vol}(L/\Gamma_L)$ est fini, cela donne la formule souhaitée. Or, le sous-groupe L est isomorphe au produit semi-direct $\mathrm{SL}_{d-1}(\mathbb{R}) \ltimes \mathbb{R}^{d-1}$, donc le quotient L/Γ_L est une extension compacte de $\mathrm{SL}_{d-1}(\mathbb{R})/\mathrm{SL}_{d-1}(\mathbb{Z})$, dont on peut supposer par récurrence qu'il est de volume fini. Cela achève notre démonstration. □

À l'aide de la formule de Siegel, nous pouvons facilement borner la mesure de l'ensemble des réseaux contenant un petit vecteur.

lm:voispointe

Démonstration du lemme **??**. On applique la formule à la fonction $f = \mathbb{1}_{B(0,\varepsilon)}$. Dans ce cas,

$$\begin{aligned}\tilde{f}(\Delta) &= |\{\mathbf{v} \in \Delta \text{ primitif} \mid \|\mathbf{v}\| < \varepsilon\}| \\ &\geq \mathbb{1}_{\{\lambda_1(\Delta)<\varepsilon\}}\end{aligned}$$

et donc

$$m_\Omega(\{\lambda_1 < \varepsilon\}) \leq \int_\Omega \tilde{f} = \int_{\mathbb{R}^d} f = c_d \varepsilon^d.$$

□

Exercice 25 (Une généralisation de la formule de Siegel)**.**
Une famille de vecteurs $(\mathbf{v}_1, \ldots, \mathbf{v}_k)$ d'un réseau Λ est dite *primitive* si elle se complète en une base de Λ. Étant donné une fonction f sur $(\mathbb{R}^d)^k$, on définit la k-ième transformée de Siegel $\tilde{f}^k$ par

$$\tilde{f}^k = \sum_{(\mathbf{v}_1,\ldots,\mathbf{v}_k) \text{ primitive}} f(\mathbf{v}_1, \ldots, \mathbf{v}_k).$$

1. Montrer que $\int_\Omega \tilde{f}^k = c_{k,d} \int_{\mathbb{R}^{dk}} f$, avec $c_{d,k} = \frac{1}{\zeta(d)\zeta(d-1)\ldots\zeta(d-k+1)}$.
2. En déduire que $m_\Omega(\{\lambda_1 < \varepsilon\}) \gtrsim_d \varepsilon^d$.

Chapitre 4

Approximation des points algébriques

ch:algebrique

Le théorème de Roth montre que si θ est un nombre algébrique réel irrationnel, alors, pour tout $\varepsilon > 0$, l'inégalité $\left|\theta - \frac{p}{q}\right| < \frac{1}{q^{2+\varepsilon}}$ n'a qu'un nombre fini de solutions $\frac{p}{q}$ dans $\mathbb{Q}$. Nous avons aussi vu au paragraphe ?? que pour les deux formes d'approximation — simultanée ou par formes linéaires — dans $\mathbb{R}^n$, Schmidt a généralisé ce théorème en calculant l'exposant $\beta_1(x)$ d'un point x dans $\mathbb{P}^n(\overline{\mathbb{Q}})$ ou l'exposant $\beta_1(x^*)$ d'un hyperplan x^* dans $\mathbb{P}^{*n}(\overline{\mathbb{Q}})$. Dans le présent chapitre, nous montrons que ces résultats sont des cas particuliers d'une formule générale pour l'exposant $\beta_k(x)$ d'un sous-espace x de dimension ℓ dans $\mathbb{R}^d$ qui admet une base constituée de vecteurs à coordonnées dans $\overline{\mathbb{Q}}$.

sec:rn

La démonstration de cette formule repose sur la correspondance établie au chapitre précédent, et sur l'interprétation du théorème du sous-espace de Schmidt en termes d'orbites diagonales dans l'espace des réseaux. Nous commençons par développer le formalisme des polygones de Grayson, qui sera commode pour décrire la géométrie de l'espace des réseaux.

4.1 Sous-modularité et polygone de Grayson

Les vecteurs qui réalisent les minima successifs d'un réseau ne sont pas uniquement définis, ce qui peut être gênant. Dans ce paragraphe, nous associons à chaque réseau un drapeau partiel uniquement défini, et étroitement relié aux vecteurs qui réalisent les minima successifs. Les idées de la construction seront encore utiles au paragraphe, pour interpréter le théorème du sous-espace de Schmidt.

Définition 4.1. Les *covolumes successifs* $\mu_1(\Delta), \dots, \mu_d(\Delta)$ d'un réseau Δ dans $\mathbb{R}^d$ sont les réels strictement positifs définis pour $i = 1, \dots, d$ par

$$\mu_i(\Delta) = \min\{|V| \ ; \ V \text{ sous-groupe de rang } i \text{ dans } \Delta\}.$$

Exercice 26. Montrer qu'à certaines constantes multiplicatives près ne dépendant que de d, la donnée des covolumes successifs est équivalente à celle des minima successifs : $\mu_i(\Delta) \asymp \lambda_1(\Delta) \dots \lambda_i(\Delta)$.

pr:hn **Proposition 4.2** (Filtration de Harder-Narasimhan d'un réseau)**.** *Soit Δ un réseau dans $\mathbb{R}^d$ et $c_\Delta\colon \{0,\dots,d\} \to \mathbb{R}$ la plus grande fonction convexe telle que $c_\Delta(0) = 0$ et pour chaque $i \geq 1$, $c_\Delta(i) \leq \log \mu_i(\Delta)$. Si i est un point angulaire de c_Δ, il existe un unique sous-groupe V_i de rang i dans Δ tel que $|V_i| = \mu_i(\Delta)$. De plus, si $I = \{i_1,\dots,i_k\}$ est l'ensemble des points angulaires de c_Δ, alors les sous-espaces V_{i_s}, $s = 1,\dots,k$ forment un drapeau partiel de Δ :*

$$\{0\} < V_{i_1} < \cdots < V_{i_k} < \Delta.$$

La proposition ci-dessus découlera d'un résultat général sur les applications sous-modulaires sur la variété grassmannienne, le théorème **??** ci-dessous. Rappelons que si K est un corps quelconque et $d \in \mathbb{N}^*$, la variété grassmannienne $\mathrm{Grass}(K^d)$ est par définition l'ensemble des sous-espaces vectoriels de K^d.

Définition 4.3 (Sous-modularité)**.** Soit K un corps quelconque. Une application $\tau : \mathrm{Grass}(K^d) \to \mathbb{R}$ est dite *sous-modulaire* si elle vérifie

$$\forall V, W \in \mathrm{Grass}(K^d), \quad \tau(V \cap W) + \tau(V + W) \leq \tau(V) + \tau(W).$$

La fonction covolume sur les sous-groupes d'un réseau de $\mathbb{R}^d$ fournit l'exemple fondamental d'application sous-modulaire.

pr:smcovol **Proposition 4.4.** *Étant donné un réseau Δ dans $\mathbb{R}^d$, on peut identifier l'ensemble des sous-groupes primitifs de Δ à la variété grassmannienne $\mathrm{Grass}(\mathbb{Q}^d)$. Alors, la fonction définie par $\tau(W) = \log|W|$ est sous-modulaire.*

Démonstration. Choisissons des éléments décomposables $\mathbf{u}$, $\mathbf{v}$ et $\mathbf{w}$ dans $\wedge^*\Delta$ tels qu'avec l'identification d'un sous-groupe primitif avec son représentant dans $\wedge^*\Delta$,

$$V \cap W = \mathbf{u}, \qquad V = \mathbf{u} \wedge \mathbf{v}, \quad \text{et} \quad W = \mathbf{u} \wedge \mathbf{w}.$$

Il s'agit de voir que les volumes des parallélépipèdes correspondants vérifient

$$\|\mathbf{u}\|\|\mathbf{u} \wedge \mathbf{v} \wedge \mathbf{w}\| \leq \|\mathbf{u} \wedge \mathbf{v}\|\|\mathbf{u} \wedge \mathbf{w}\|.$$

Pour cela, on remarque que $\|\mathbf{u} \wedge \mathbf{v} \wedge \mathbf{w}\| = \|\mathbf{u} \wedge \mathbf{v}\| \cdot \|p_{U^\perp}(\mathbf{w})\|$, où $p_{U^\perp}$ est la projection orthogonale sur $U^\perp$, tandis que $\|\mathbf{u}\wedge\mathbf{w}\| = \|\mathbf{u}\|\cdot\|p_{V^\perp}(\mathbf{w})\|$. L'inégalité souhaitée découle alors du fait que $\|p_{V^\perp}(\mathbf{w})\| \leq \|p_{U^\perp}(\mathbf{w})\|$. □

L'importance de la notion de sous-modularité provient de la construction donnée par le théorème suivant, qui généralise la proposition **??**.

th:hn **Théorème 4.5.** *Soit K un corps de caractéristique nulle et $\tau : \mathrm{Grass}(K^d) \to \mathbb{R}$ une application sous-modulaire. Soit $c : [0,d] \to \mathbb{R}$ la plus grande fonction convexe dont le graphe soit situé en-dessous de tous les points $(\dim W, \tau(W))$, $W \leq K^d$. Si $I = \{i_1 < \cdots < i_k\}$ désigne l'ensemble des points angulaires de c, il existe un unique drapeau partiel*

$$F : \quad \{0\} < V_{i_1} < V_{i_2} < \cdots < V_{i_k} < K^d$$

tel que pour chaque s, $\dim V_{i_s} = i_s$ et $c(i_s) = \tau(V_{i_s})$. De plus, tout sous-espace W tel que $\tau(W) = c(\dim W)$ est compatible avec le drapeau F.

Définition 4.6. La fonction $c : [0, d] \to \mathbb{R}$ est appelée *polygone de Grayson*, et le drapeau partiel F est la *filtration de Harder-Narasimhan* associée à la fonction sous-modulaire τ.

La démonstration du théorème hn **??** se fonde sur l'observation géométrique suivante, dite « règle du parallélogramme » : si V, W sont deux sous-espaces de K^d, et si on place les points $(\dim V \cap W, \tau(V \cap W))$, $(\dim V, \tau(V))$ et $(\dim V + W, \tau(V + W))$, alors le point $(\dim W, \tau(W))$ est situé sur la demi-droite verticale au-dessus du quatrième point du parallélogramme.

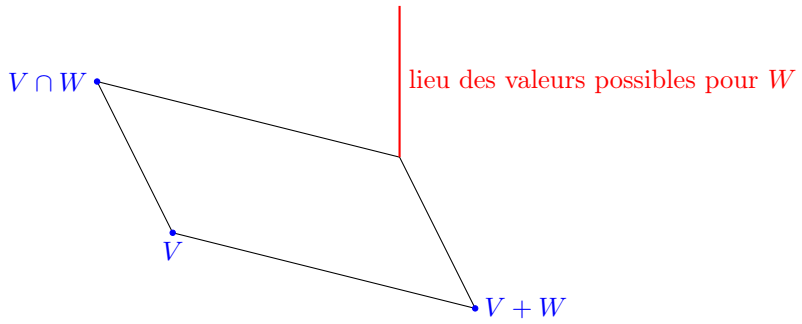

FIGURE 4.1 – La règle du parallélogramme

Lemme 4.7 (Lemme de sous-modularité)**.** *Soit $\tau : \mathrm{Grass}(K^d) \to \mathbb{R}$ une fonction sous-modulaire telle que $\tau(0) = 0$. L'ensemble des sous-espaces V tels que*

$$\frac{\tau(V)}{\dim V} = \inf_W \frac{\tau(W)}{\dim W}$$

admet un unique plus grand élément.

Démonstration. Posons

$$a := \inf_V \frac{\tau(V)}{\dim V}.$$

Si V et W sont deux sous-espaces tels que $\frac{\tau(V)}{\dim V} = \frac{\tau(W)}{\dim W} = a$, la sous-modularité et la définition de a comme valeur minimale permettent de majorer

$$\begin{aligned}\tau(V + W) &\leq \tau(V) + \tau(W) - \tau(V \cap W) \\ &\leq \tau(V) + \tau(W) - a \dim(V \cap W) \\ &= a \dim V + a \dim W - a \dim(V \cap W) = a \dim(V + W).\end{aligned}$$

L'ensemble des sous-espaces qui réalisent la borne inférieure $\inf_W \frac{\tau(W)}{\dim W}$ est stable par addition, donc la somme de tous ces sous-espaces est l'unique plus grand élément de cet ensemble. □

Exercice 27. Le lecteur attentif aura noté une erreur dans la démonstration ci-dessus : on ne sait pas a priori que la borne inférieure $a = \inf_W \frac{\tau(W)}{\dim W}$ est atteinte.

1. Soit V_n une éventuelle suite de sous-espaces de dimension maximale telle que $\tau(V_n) \to -\infty$. Montrer que si D est une droite fixée quelconque, alors $\tau(V_n + D) \to -\infty$. En déduire que τ est nécessairement minorée.
2. Soit k maximal tel qu'il existe V de dimension k approchant la borne inférieure $a = \inf_W \frac{\tau(W)}{\dim W}$. Justifier qu'il existe $\varepsilon > 0$ tel que pour tout V' de dimension strictement supérieure à k, $\frac{\tau(V')}{\dim V'} \geq a + \varepsilon$.
3. Expliquer comment corriger la démonstration du lemme de sous-modularité.

Démonstration du théorème ??. Le lemme de sous-modularité donne exactement l'existence du sous-espace V_{i_1}. On procède ensuite par récurrence. Supposons définis des sous-espaces $V_{i_1} < \cdots < V_{i_s}$ et une fonction $c^{(s)} : [0, d] \to \mathbb{R}$ tels que

1. $c^{(s)}$ est convexe, affine par morceaux, avec pour points angulaires $\{i_1, \ldots, i_{s-1}\}$;
2. $c^{(s)}$ est située au-dessous de tout point de la forme $(\dim V, \tau(V))$, $V \leq K^d$;
3. Si $\tau(V) = c^{(s)}(\dim V)$, alors le sous-espace V est compatible avec le drapeau partiel $V_{i_1} < \cdots < V_{i_s}$.

Le lemme de sous-modularité appliqué dans l'espace quotient K^d/V_{i_s} montre qu'il existe un unique sous-espace $V_{i_{s+1}}$ contenant V_{i_s}, et qui minimise le rapport $\frac{\tau(V)-\tau(V_{i_s})}{\dim V - \dim V_{i_s}}$ parmi tous les sous-espaces contenant V_{i_s}. Notons

$$a = \frac{\tau(V_{i_{s+1}}) - \tau(V_{i_s})}{\dim V_{i_{s+1}} - \dim V_{i_s}}.$$

Pour définir la fonction $c^{(s+1)}$, on ajoute le point $(i_s, \tau(V_{i_s}))$ et on impose une pente égale à a sur le segment $[i_s, d]$. Nous voulons vérifier que $c^{(s+1)}$ vérifie encore les trois propriétés ci-dessus. La première est claire, par construction. Comme $c^{(s+1)}$ et $c^{(s)}$ coïncident sur l'intervalle $[0, i_s]$, les deuxième et troisième propriétés sont satisfaites si $\dim V \leq i_s$. Si $\dim V > i_s$, on a par construction de $c^{(s+1)}$ (noter que $V + V_{i_s}$ contient V_{i_s})

$$\tau(V + V_{i_s}) \geq c^{(s+1)}(\dim(V + V_{i_s}))$$

et

$$\tau(V \cap V_{i_s}) \geq c^{(s+1)}(\dim(V \cap V_{i_s})).$$

Les points correspondants à V_{i_s}, $V \cap V_{i_s}$ et $V + V_{i_s}$ sont tous au-dessus du graphe de $c^{(s+1)}$. Comme $c^{(s+1)}$ est convexe, la règle du parallélogramme montre que le point $(\dim V, c^{(s+1)}(\dim V))$ est aussi au-dessus du graphe de $c^{(s+1)}$. Formellement, les pentes de $c^{(s+1)}$ sont croissantes donc

$$c^{(s+1)}(\dim(V + V_{i_s})) - c^{(s+1)}(\dim V_{i_s}) \geq c^{(s+1)}(\dim V) - c^{(s+1)}(\dim(V \cap V_{i_s}))$$

puis

$$\begin{aligned}\tau(V) &\geq \tau(V \cap V_{i_s}) + \tau(V + V_{i_s}) - \tau(V_{i_s}) \\ &\geq c^{(s+1)}(\dim(V \cap V_{i_s})) + c^{(s+1)}(\dim(V + V_{i_s})) - c^{(s+1)}(\dim V_{i_s}) \\ &\geq c^{(s+1)}(\dim V).\end{aligned}$$

De plus, en cas d'égalité, on doit avoir $c^{(s+1)}(\dim(V+V_{i_s}))-c^{(s+1)}(\dim V_{i_s}) = c^{(s+1)}(\dim V)-c^{(s+1)}(\dim(V\cap V_{i_s}))$, et comme $c^{(s+1)}$ admet un angle strict en i_s, cela force $\dim(V\cap V_{i_s})=i_s$ i.e. $V\supset V_{i_s}$. Mais alors, par définition de $V_{i_{s+1}}$, on a $V\subset V_{i_{s+1}}$. □

Remarque. On peut aussi comprendre les minima successifs d'un réseau Δ via la théorie de la réduction : toute matrice g dans $\mathrm{GL}_d(\mathbb{R})$ admet une décomposition de Siegel

$$g = kan\gamma,$$

où $k\in O_d(\mathbb{R})$, $\gamma\in\mathrm{GL}_d(\mathbb{Z})$, n est unipotente triangulaire supérieure et vérifie $|n_{ij}|\leq\frac{1}{2}$ si $i<j$, et $a=\mathrm{diag}(a_1,\dots,a_d)$, avec $\forall i,\ a_{i+1}\geq\frac{\sqrt{3}}{2}a_i$. Alors, pour chaque i, $\log\lambda_i(\Delta)\asymp a_i$. Cela permet notamment de montrer facilement que si l'espace Ω des réseaux dans $\mathbb{R}^d$ est muni d'une distance riemannienne qui provient d'une métrique riemannienne invariante à droite sur $\mathrm{GL}_d(\mathbb{R})$, alors, pour tous réseaux Δ,Δ',

$$d(\Delta,\Delta') = \|c_\Delta - c_{\Delta'}\| + O_d(1).$$

Exercice 28. Soit Δ un réseau dans $\mathbb{R}^d$ et $F_\Delta:\ \{0\}<V_{i_1}<\cdots<V_{i_k}<\Delta$ sa filtration de Harder-Narasimhan.

1. Justifier que pour tout s, $\log\mu_{i_s}(\Delta)=c_\Delta(i_s)$.
2. Montrer que pour tout $i\in\{1,\dots,d\}$, $\log\mu_i(\Delta)=c_\Delta(i)+O_d(1)$.
3. Justifier que la donnée du polygone de Grayson d'un réseau Δ est essentiellement équivalente à celle des minima successifs, ou des covolumes successifs de Δ.

4.2 Le théorème du sous-espace paramétrique

Comme auparavant, nous noterons dans ce paragraphe $\overline{\mathbb{Q}}\subset\mathbb{R}$ l'ensemble des nombres réels algébriques. Nous voulons interpréter théorème du sous-espace de Schmidt en termes d'orbites diagonales de réseaux dans $\mathbb{R}^d$ qui admettent une base constituée de vecteurs à coordonnées dans $\overline{\mathbb{Q}}$. Le but de cette partie est le résultat suivant qui décrit le comportement asymptotique d'une telle orbite au premier ordre.

th:sep **Théorème 4.8** (Théorème du sous-espace paramétrique)**.** *Soit Δ un réseau algébrique, i.e. $\Delta=L\mathbb{Z}^d$ avec $L\in\mathrm{GL}_d(\overline{\mathbb{Q}})$, et $(a_t)_{t\in\mathbb{R}}$ un sous-groupe diagonal à un paramètre. Alors,*

1. *le diagramme de Grayson renormalisé $\frac{1}{t}c_{a_t\Delta}$ converge en $+\infty$ vers une limite c_∞ ;*
2. *si $i_1<\cdots<i_k$ désignent les points angulaires de c_∞, il existe un drapeau partiel $\{0\}<V_{i_1}<\cdots<V_{i_k}<\mathbb{Z}^d$ tel que pour tout $t>0$ assez grand, pour $s=1,\dots,k$, les i_s premiers minima successifs de $a_t\Delta$ dont atteints dans $a_tLV_{i_s}$.*

Les observations de la partie précédente sur les applications sous-modulaires vont nous permettre de construire facilement la limite c_∞ et le drapeau partiel associé, ce qui sera à la base de la démonstration du théorème.

Définition 4.9. Le *taux d'expansion* d'un sous-groupe discret $V \leq \mathbb{R}^d$ par le flot diagonal $(a_t L)_{t\in\mathbb{R}}$ est la quantité

$$\tau_L(V) = \lim_{t\to+\infty} \frac{1}{t} \log|a_t L V|.$$

Remarque. Cette limite est bien définie, c'est le logarithme de la plus grande valeur propre de a_1 apparaissant dans la décomposition de LV suivant les espaces propres de a_t dans $\wedge^{\dim V}\mathbb{R}^d$. Elle ne dépend que du sous-espace vectoriel engendré par V.

Nous déduirons le théorème **??** du théorème du sous-espace de Schmidt, que nous avons déjà utilisé au paragraphe **??**, et que nous rappelons ci-dessous.

Théorème 4.10 (Théorème du sous-espace de Schmidt)**.** *Soit $L \in \mathrm{GL}_d(\overline{\mathbb{Q}})$ et $L_1, \dots, L_d$ les formes linéaires sur $\mathbb{R}^d$ données par les lignes de L. Pour tout $\varepsilon > 0$, l'ensemble des solutions $v \in \mathbb{Z}^d$ à l'inégalité*

$$|L_1(v) \dots L_d(v)| \leq \|v\|^{-\varepsilon}$$

est inclus dans une union finie d'hyperplans.

Exercice 29. Démontrer le théorème de Schmidt dans le cas particulier $L \in \mathrm{GL}_d(\mathbb{Q})$.

Dans un cours plus complet, on démontrerait le théorème de Schmidt, et alors le théorème **??** pourrait apparaître comme un résultat intermédiaire dans la démonstration. Quoiqu'il en soit, il est bon de savoir que les deux énoncés sont équivalents.

Démonstration du théorème **??**. Il découle de la proposition **??** que le taux de contraction par $(a_t L)_{t\in\mathbb{R}}$ définit une application sous-modulaire sur $\mathrm{Grass}(\mathbb{Q}^d)$. Notons c_∞ le polygone de Grayson et $\{0\} < V_{i_1} < \dots < V_{i_k} < \mathbb{Q}^d$ la filtration de Harder-Narasimhan associés à τ_L, et montrons que les conclusions du théorème sont alors satisfaites.

Remarquons que pour $s = 1, \dots, k$, par définition du taux de contraction, $\frac{1}{t}\log|a_t L V_{i_s}| = \tau_L(V_{i_s}) + o(1)$. Cela implique que

$$\limsup \frac{1}{t} \log \mu_{i_s}(a_t L\mathbb{Z}^d) \leq \tau_L(V_{i_s}) = c_\infty(i_s).$$

En d'autres termes, pour $\varepsilon > 0$, pour $t > 0$ grand, la fonction convexe $\frac{1}{t}c_{a_t\Delta}$ est située au-dessous de $c_\infty + \varepsilon$ en tout point angulaire de c_∞, donc $\frac{1}{t}c_{a_t\Delta} \leq c_\infty + \varepsilon$. Ainsi,

$$\limsup \frac{1}{t} c_{a_t\Delta} \leq c_\infty.$$

Pour montrer la limite souhaitée, il suffit de minorer les pentes de $\frac{1}{t}c_{a_t\Delta}$ à droite de chaque point angulaire en montrant que pour $s = 1, \dots, k$,

$$\liminf \frac{1}{t} \log \lambda_{i_s+1}(a_t L\mathbb{Z}^d) \geq \frac{\tau_L(V_{i_{s+1}}) - \tau_L(V_{i_s})}{i_{s+1} - i_s}.$$

Pour fixer les idées, écrivons $a_t = \mathrm{diag}(e^{A_1 t}, \dots, e^{A_d t})$, avec $A_1 \geq \dots \geq A_d$, ce qui est toujours possible, quitte à permuter les éléments de la base canonique

de $\mathbb{R}^d$. Commençons par le cas $s = 0$; on veut donc montrer que pour $\varepsilon > 0$, pour tout $t > 0$ assez grand, $\lambda_1(a_t L\mathbb{Z}^d) \geq e^{(\frac{\tau_L(V_{i_1})}{i_1} - \varepsilon)t}$. Soit $V \leq \mathbb{Z}^d$ de dimension minimale tel que pour $t > 0$ arbitrairement grand, il existe $v \in V$ tel que $\|a_t L v\| \leq e^{(\frac{\tau_L(V_{i_1})}{i_1} - \varepsilon)t}$. On associe à V un ensemble $J_V \subset \{1, \dots, d\}$ de cardinal $\dim V$ de la façon suivante

$$\begin{cases} j_1 \text{ est minimal tel que } L_{j_1}|_V \neq 0 \\ j_2 \text{ est minimal tel que } (L_{j_1}|_V, L_{j_2}|_V) \text{ est libre} \\ \dots \text{ etc.} \end{cases}$$

Alors,

$$\tau_L(V) = \sum_{j \in J_V} A_j.$$

Donc, pour tout $v \in V$ vérifiant $\|a_t L v\| \leq e^{(\frac{\tau_L(V_{i_1})}{i_1} - \varepsilon)t}$, on a

$$\begin{aligned} \prod_{j \in J_V} |L_j(v)| &= e^{-\tau_L(V)t} \prod_{j \in J_V} e^{A_j t} |L_j(v)| \\ &\leq e^{-\tau_L(V)t} e^{(\dim V)(\frac{\tau_L(V_{i_1})}{i_1} - \varepsilon)t} \\ &\leq e^{-(\dim V)\varepsilon t}. \end{aligned}$$

Mais $\|v\| \lesssim e^{-A_1 t} \|a_t L v\| \leq e^{At}$, où $A = |A_1| + |A_d|$, et donc, pour $\varepsilon' = \frac{\varepsilon \dim V}{A}$,

$$\prod_{j \in J_V} |L_j(v)| \leq \|v\|^{-\varepsilon'}.$$

D'après le théorème du sous-espace de Schmidt, les solutions de cette inégalité sont contenues dans un nombre fini de sous-espaces stricts $W < V$, mais par minimalité de V, pour $t > 0$ assez grand, chaque tel W ne contient pas de solution à $\|a_t L v\| \leq e^{(\frac{\tau_L(V_{i_1})}{i_1} - \varepsilon)t}$. Cela montre ce qu'on voulait : $\liminf \frac{1}{t} \log \lambda_1(a_t L\mathbb{Z}^d) \geq \frac{\tau_L(V_{i_1})}{i_1}$.

Pour $s \geq 1$, on procède par récurrence, en supposant le résultat connu pour $s-1$. Soit $V \leq \mathbb{Z}^d$ contenant $V_{i_{s-1}}$ et de dimension minimale tel que pour $t > 0$ arbitrairement grand, il existe $v \in V$ tel que

$$\|a_t L v\| \leq e^{t(\frac{\tau_L(V_{i_s}) - \tau_L(V_{i_{s-1}})}{i_s - i_{s-1}} - \varepsilon)}.$$

Pour $j \in J_V \setminus J_{V_{i_{s-1}}}$, on peut choisir successivement des éléments $\alpha_{j\ell} \in \overline{\mathbb{Q}}$ tels que

$$M_j = L_j - \sum_{\ell \in J_{V_{i_{s-1}}}} \alpha_{j\ell} L_\ell \equiv 0 \quad \text{sur } V_{i_{s-1}}.$$

Alors, $\|a_t M v\| \lesssim \|a_t L v\|$ et donc

$$\begin{aligned} \prod_{j \in J_V \setminus J_{V_{i_{s-1}}}} |M_j(v)| &\leq e^{-t(\tau_L(V) - \tau_L(V_{i_{s-1}}))} e^{t(\dim V - i_{s-1})(\frac{\tau_L(V_{i_s}) - \tau_L(V_{i_{s-1}})}{i_s - i_{s-1}} - \varepsilon)} \\ &\leq e^{-t\varepsilon(\dim V - i_{s-1})} \leq \|v\|^{-\varepsilon'}. \end{aligned}$$

Le théorème du sous-espace de Schmidt dans l'espace quotient $\mathbb{Z}^d / V_{i_{s-1}}$ permet donc de conclure comme dans le cas $s = 0$ déjà traité ci-dessus. □

4.3 Application aux variétés grassmanniennes

Comme au chapitre **??**, on fixe des entiers $1 \leq k \leq \ell < d$, et on s'intéresse aux approximations d'un élément x dans $X_{\ell,d}(\mathbb{R})$ par des points rationnels dans $X_{k,d}(\mathbb{Q})$. Rappelons que l'exposant diophantien $\beta_k(x)$ est défini par

$$\beta_k(x) = \sup\left\{\beta > 0 \;\middle|\; \exists v \in X_{k,d}(\mathbb{Q}) : \begin{array}{l} d(v,x) \to 0 \\ d(v,x) \leq H(v)^{-\beta} \end{array}\right\}.$$

De plus, si s_x est un élément de $\mathrm{SL}_d(\mathbb{R})$ tel que $x = s_x^{-1} \cdot \mathrm{Vect}(e_1, \dots, e_\ell)$, et

$$a_t = \mathrm{diag}(e^{-\frac{(d-\ell)t}{d}}, \dots, e^{-\frac{(d-\ell)t}{d}}, e^{\frac{(d-\ell)t}{d}}, \dots, e^{\frac{(d-\ell)t}{d}}),$$

la proposition **??** relie l'exposant $\beta_k(x)$ à l'orbite du réseau $a_t s_x \mathbb{Z}^d$ par la formule

$$\beta_k(x) = \frac{1}{\frac{k(d-\ell)}{d} - \gamma_k(x)},$$

où

$$\gamma_k(x) = \sup\left\{\gamma \in \mathbb{R} \;\middle|\; \exists t \to +\infty : \; \exists \mathbf{w} \in a_t u_x \wedge^k \mathbb{Z}^d : \begin{array}{c} \|\mathbf{w}\| \leq e^{-\gamma t} \\ \text{et } \|\pi^+(\mathbf{w})\| \geq \frac{1}{2}\|\mathbf{w}\| \end{array}\right\}.$$

Les résultats du paragraphe précédent vont nous permettre de calculer explicitement la quantité $\gamma_k(x)$ lorsque x est un sous-espace défini sur $\overline{\mathbb{Q}}$, i.e. admettant une base de vecteurs à coordonnées dans $\overline{\mathbb{Q}}$.

Pour cela, commençons par remarquer qu'avec les notations ci-dessus, le taux de contraction d'un sous-espace $W \leq \mathbb{R}^d$ sous l'action de $a_t u_x$ est donné par

$$\tau_x(W) = -\frac{1}{d}\left((d-\ell)\dim x \cap W - \ell(\dim W - \dim x \cap W)\right),$$

donc

$$\frac{\tau_x(W)}{\dim W} = \frac{\ell}{d} - \frac{\dim x \cap W}{\dim W}$$

est minimal si et seulement si $\frac{\dim x \cap W}{\dim W}$ est maximal.

Par conséquent, le drapeau $\{0\} = V_0 < V_{i_1} < \cdots < V_{i_k} < V_d = \mathbb{Q}^d$ du théorème **??** s'obtient par récurrence de la façon suivante : V_{i_1} est l'unique sous-espace rationnel de dimension maximale qui maximise le quotient $\frac{\dim x \cap V_{i_1}}{\dim V_{i_1}}$, et par récurrence, V_{i_s} est l'unique sous-espace rationnel contenant $V_{i_{s-1}}$ de dimension maximale et qui maximise $\frac{\dim x \cap V_{i_s} - \dim x \cap V_{i_{s-1}}}{i_s - i_{s-1}}$. De plus, les pentes

$$\Lambda_i := \lim_{t \to +\infty} \tfrac{1}{t} \log \lambda_i(a_t s_x \mathbb{Z}^d)$$

du polygone limite c_∞ sont données par

$$\Lambda_i = \frac{\ell}{d} - \frac{\dim x \cap V_{i_s} - \dim x \cap V_{i_{s-1}}}{i_s - i_{s-1}} \quad \text{si } i_s < i \leq i_{s+1}.$$

Ces observations permettent déjà d'obtenir l'exposant d'un sous-espace algébrique hors de certaines contraintes rationnelles. On rappelle qu'un *pinceau* dans $X_{\ell,d}$ est une sous-variété de la forme

$$\mathcal{P}_{W,r} = \{x \in X_{\ell,d}(\mathbb{R}) \mid \dim x \cap W \geq r\}$$

où $W \leq \mathbb{R}^d$ est un sous-espace vectoriel, et r un entier positif. Un tel pinceau est dit *rationnel* si le sous-espace W est rationnel, et *contraignant* si $\frac{r}{\dim W} > \frac{\ell}{d}$.

Théorème 4.11 (Extrémalité d'un sous-espace algébrique non dégénéré)**.** *Si $x \in X_{\ell,d}(\overline{\mathbb{Q}})$ n'est inclus dans aucun pinceau rationnel contraignant, alors, pour tout $k \in \{1, \dots, \ell\}$, $\beta_k(x) = \frac{d}{k(d-\ell)}$.*

Démonstration. Si x n'est inclus dans aucun pinceau rationnel contraignant, le drapeau ci-dessus est réduit à $\{0\} = V_0 < V_d = \mathbb{Z}^d$, et pour tout i, $\Lambda_i = 0$. En particulier, $\lim_{t\to+\infty} \frac{1}{t} \log \lambda_1(a_t s_x \mathbb{Z}^d) = 0$, et d'après le corollaire **??**, cela implique $\beta_k(x) = \frac{d}{k(d-\ell)}$ pour tout $k = 1, \dots, \ell$. □

cor:nulextremal

Plus généralement, nous allons montrer le théorème suivant, qui donne une formule explicite pour $\gamma_k(x)$ lorsque x appartient à $X_{\ell,d}(\overline{\mathbb{Q}})$.

expalg

Théorème 4.12 (Exposants d'un sous-espace algébrique)**.** *Soit $x \in X_\ell(\overline{\mathbb{Q}})$ et $\{0\} < V_{i_1} < \cdots < V_{i_k} < \mathbb{Z}^d$ le drapeau partiel défini ci-dessus. Pour $s = 1, \dots, k$, posons*

$$j_s = \dim x \cap V_{i_s}.$$

Alors, pour $k = 1, \dots, \ell$, notant $k_s = \min(j_s, k)$,

$$\gamma_k(x) = -\sum_{s=0}^{r} (k_{s+1} - k_s)\Lambda_{i_s+1} = \frac{-k\ell}{d} + \sum_{s=0}^{r} \frac{(k_{s+1} - k_s)(j_{s+1} - j_s)}{i_{s+1} - i_s}.$$

Démonstration. Considérons le sous-réseau de $\wedge^k \mathbb{Z}^d$ défini par

$$S = \wedge^{k_1} V_{i_1} \wedge (\wedge^{k_2 - k_1} V_{i_2}) \wedge \cdots \wedge (\wedge^{k_{r+1} - k_r} V_{i_{r+1}}).$$

th:sep

Le théorème **??** implique que pour tout $\varepsilon > 0$, pour tout $t > 0$ suffisamment grand, le réseau $a_t s_x S$ admet une base de vecteurs dont la norme est majorée par

$$\exp[t(\varepsilon + \sum_{s=0}^{r} (k_{s+1} - k_s)\Lambda_{i_s+1})].$$

De plus, par définition des entiers k_s, ce sous-espace contient un vecteur pur $\mathbf{v}_x \in \wedge^k x$. L'élément $s_x \mathbf{v}_x$ appartient à l'image $\mathrm{Vect}(e_I;\ I \subset \{1, \dots, \ell\})$, donc il existe dans $a_t s_x S$ un vecteur $\mathbf{v}$ tel que $\|\pi^+(\mathbf{v})\| \gtrsim \|\mathbf{v}\|$ et $\|\mathbf{v}\| \leq e^{t(\sum_{s=0}^{r} (k_{s+1} - k_s)\Lambda_{i_s+1} + O(\varepsilon))}$. Comme $\varepsilon > 0$ est arbitrairement petit, cela montre déjà l'inégalité

$$\gamma_k(x) \geq -\sum_{s=0}^{r} (k_{s+1} - k_s)\Lambda_{i_s+1}.$$

th:sep

Réciproquement, le théorème **??** montre aussi que pour $\varepsilon > 0$, pour tout $t > 0$ suffisamment grand, tout vecteur $\mathbf{v}$ dans $\wedge^k \mathbb{Z}^d$ tel que

$$\|a_t s_x \mathbf{v}\| \leq e^{t(-\varepsilon + \sum_{s=0}^{r} (k_{s+1} - k_s)\Lambda_{i_s+1})}$$

appartient à un sous-espace

$$S' = \wedge^{k'_1} V_{i_1} \wedge (\wedge^{k'_2 - k'_1} V_{i_2}) \wedge \cdots \wedge (\wedge^{k'_{r+1} - k'_r} V_{i_{r+1}})$$

avec pour un certain u, $k'_u > k \geq k_u$. Par définition des entiers k_s, $s = 1, \dots, r$, le sous-espace S' ne contient aucun vecteur pur non nul de $\wedge^k x$, et donc il existe $c > 0$ tel que pour tout vecteur pur $\mathbf{v} \in S'$, $\|s_x \mathbf{v} - \pi^+(s_x \mathbf{v})\| \geq c\|s_x \mathbf{v}\|$. Cela implique

$$\begin{aligned}\|a_t s_x \mathbf{v}\| &\geq c e^{-t(\frac{k}{\ell} - \frac{1}{\ell} - \frac{1}{d-\ell})} \|s_x \mathbf{v}\| \\ &\gtrsim e^{t(\frac{1}{\ell} + \frac{1}{d-\ell})} \|\pi^+(a_t s_x \mathbf{v})\|\end{aligned}$$

et montre que le vecteur $a_t s_x \mathbf{v}$ ne saurait satisfaire $\|\pi^+(a_t s_x \mathbf{v})\| \geq \frac{1}{2}\|a_t s_x \mathbf{v}\|$. Ainsi, $\gamma_k(x) \leq \varepsilon - \sum_{s=0}^{r}(k_{s+1} - k_s)\Lambda_{i_s+1}$. Lorsque ε tend vers zéro, on obtient le résultat souhaité. □

Cette formule permet déjà de montrer que l'exposant diophantien $\beta_k(x)$ d'un sous-espace défini sur $\overline{\mathbb{Q}}$ est toujours supérieur à l'exposant générique $\frac{d}{k(d-\ell)}$, et donne une condition nécessaire et suffisante pour qu'il y ait une égalité.

Corollaire 4.13. *Pour tout $x \in X_\ell(\overline{\mathbb{Q}})$ et tout $k \leq \ell$, on a $\beta_k(x) \geq \frac{d}{k(d-\ell)}$, avec égalité si et seulement si x n'est inclus dans aucun pinceau rationnel contraignant.*

Démonstration. D'après la proposition ??, l'exposant diophantien $\beta_k(x)$ est donné par $\beta_k(x) = \frac{1}{\frac{k(d-\ell)}{d} - \gamma_k(x)}$, et il suffit donc de montrer que $\gamma_k(x) \geq 0$, avec égalité si et seulement si x n'est inclus dans aucun pinceau rationnel contraignant. Lorsque $k = \ell$, on a $k_s = j_s$ pour chaque s, et donc

$$\gamma_\ell(x) = \frac{-k\ell}{d} + \sum_{s=0}^{r} \frac{(j_{s+1} - j_s)^2}{i_{s+1} - i_s}.$$

On écrit alors

$$\begin{aligned}d\sum_{s=0}^{r} \frac{(j_{s+1} - j_s)^2}{i_{s+1} - i_s} &= \Big(\sum_{s=0}^{r} \frac{(i_{s+1} - i_s)^2}{i_{s+1} - i_s}\Big)\Big(\sum_{s=0}^{r} \frac{(j_{s+1} - j_s)^2}{i_{s+1} - i_s}\Big) \\ &\geq \Big(\sum_{s=0}^{r} j_{s+1} - j_s\Big)^2 = \ell^2\end{aligned}$$

et cela montre que $\gamma_\ell(x) \geq 0$. L'inégalité $\gamma_k(x) \geq 0$ pour $k \leq \ell$ découle du cas particulier $k = \ell$, car comme la fonction $s \mapsto \frac{j_{s+1} - j_s}{i_{s+1} - i_s}$ est décroissante,

$$\sum_{s=0}^{r} \frac{(k_{s+1} - k_s)(j_{s+1} - j_s)}{i_{s+1} - i_s} \geq \frac{k}{\ell} \sum_{s=0}^{r} \frac{(j_{s+1} - j_s)^2}{i_{s+1} - i_s}.$$

En effet, la fonction constante par morceaux $f : x \mapsto \frac{j_{s+1} - j_s}{i_{s+1} - i_s}$ si $j_s < x \leq j_{s+1}$ est décroissante, et l'égalité ci-dessus s'écrit simplement $\int_0^k f(x)\mathrm{d}x \geq \frac{k}{\ell}\int_0^\ell f(x)\mathrm{d}x$.

Si $\gamma_k(x) = 0$, les calculs ci-dessus montrent que $\gamma_1(x) = 0$, et cela implique $\frac{j_1}{i_1} = \frac{\ell}{d}$ ce qui par définition de V_{i_1} n'est possible que si $V_{i_1} = \mathbb{Q}^d$, i.e. x n'est inclus dans aucun pinceau rationnel contraignant. Réciproquement, si x n'est inclus dans aucun pinceau rationnel contraignant, on doit avoir $V_{i_1} = \mathbb{Q}^d$, et donc $\gamma_k(x) = 0$ pour tout $k = 1, \dots, \ell$. □

Remarque. Nous verrons au chapitre suivant que cette minoration de l'exposant diophantien est encore valable pour tout x dans $X_{\ell,d}(\mathbb{R})$.