

Université Pierre et Marie Curie



Master 2 mention : mathématiques et applications

Approximation diophantienne et géométrie
paramétrique des nombres

Anthony POELS

Directeur : Stéphane FISCHLER (Université Paris-Sud)

Date de soutenance : le 24 septembre 2015

Composition du jury :

Stéphane FISCHLER,
Patrice PHILIPPON,
Michel WALDSCHMIDT

Année 2014-2015

Table des matières

1	Introduction	3
2	Propriétés du déterminant, coordonnées de Grassmann et géométrie des nombres	6
2.1	Propriétés générales du déterminant, déterminants et matrices composés	6
2.2	Coordonnées de Grassmann et coordonnées de Grassmann duales	13
2.3	Déterminants généralisés	16
2.4	Géométrie des nombres - Corps convexe composé	19
3	Angles d'inclinaison entre deux sous-espaces	22
3.1	Produits scalaires successifs $\lambda_1, \dots, \lambda_f$	23
3.2	Quantités ν_1, \dots, ν_t	26
3.3	Angles d'inclinaison	27
4	Hauteur d'un sous-espace	30
4.1	Définition de la hauteur et premières propriétés	30
4.2	Seconde définition de la hauteur	33
4.3	Nombre de sous-espaces de hauteur $\leq H$	39
5	Approximations diophantiennes	40
5.1	Travail préliminaire	40
5.2	Théorèmes du Going-up et du Going-down	45
5.3	Approximations diophantiennes	52
5.4	De la qualité de ces approximations	54
6	Approximation diophantienne et géométrie paramétrique des nombres	55
6.1	Exposants d'approximation classiques	55
6.2	Approche du problème d'approximation simultanée par la géométrie paramétrique des nombres	56
7	Une conjecture résolue par Roy	57
7.1	Définitions et formulation du problème	57
7.2	Liens avec les exposants "classiques" d'approximation	59
7.3	La théorie de Schmidt et Summerer	62
7.3.1	Une famille générale de corps convexes	62
7.3.2	Trajectoires de points et graphes combinés	63
7.3.3	La famille principale de corps convexes	64
7.3.4	Familles de corps convexes pseudo-composés	64
7.3.5	Le théorème d'approximation de Schmidt et Summerer	65
7.4	Esquisse de la construction de Roy	66
7.4.1	Un premier résultat	67
7.4.2	Retour sur la hauteur de sous-espaces	69
7.4.3	Cas des systèmes rigides de grande maille	71
7.4.4	Systèmes réduits et approximations par des n -systèmes rigides	73
7.4.5	Preuve du théorème de Roy	74

8	Une application du théorème de Roy	75
8.1	Définitions et formulation du problème	76
8.2	Liens avec la géométrie paramétrique des nombres	78
8.3	La notion de n -systèmes généralisés	80
8.4	Une famille de n -systèmes généralisés	84
8.5	Preuve du théorème principal	86

1 Introduction

Ce mémoire se propose d'étudier deux aspects de la théorie de l'approximation diophantienne. Le premier est celui de la hauteur de sous-espaces vectoriels définis sur un corps de nombres K et des théorèmes d'approximation qui y sont liés. Schmidt a publié en 1967 un article fondateur sur ce sujet ([17]); l'objectif des parties 2 à 5 est d'étudier en détails la majeure partie de cet article de Schmidt. L'une des questions à la source de l'approximation diophantienne est la suivante (de manière informelle) : étant donné un nombre réel α , comment peut-on "bien" l'approcher par des rationnels p/q qui ne soient "pas trop compliqués", i.e. tels que p et q soient premiers entre eux et pas trop grands ? On peut voir ce problème d'une autre manière : étant donné une droite de \mathbb{R}^2 de coefficient directeur α , comment peut-on "bien" l'approcher par une droite de coefficient directeur p/q (on veut donc que l'angle entre ces deux droites soit le plus petit possible) qui ne soit pas trop "compliquée" ? De là on peut généraliser ce problème : étant donné un sous-espace vectoriel A de dimension d d'un espace vectoriel E de dimension n (euclidien ou hermitien), comment peut-on "bien" approcher A par des sous-espaces vectoriels B de dimension e définis sur un corps de nombres K (i.e. qui possèdent une base de vecteurs à coordonnées dans K) qui ne soient "pas trop compliqués" ? Pour définir cette notion de proximité, Schmidt utilise les angles d'inclinaison $\varphi(A, B)$ entre deux sous-espaces A et B ; cela fait l'objet de la partie 3. Deux sous-espaces sont "proches" si leurs angles d'inclinaison sont petits. Dans le cas où $d = 1$ (i.e. A est une droite) et $E = \mathbb{R}^n$, il n'y a qu'un seul angle d'inclinaison, qui correspond à l'angle aigu entre un vecteur directeur de A et son projeté orthogonal sur B (dans le cas où ce projeté est non nul, sinon l'angle vaut $\pi/2$). Pour caractériser la "complexité" d'un sous-espace B défini sur K , Schmidt utilise la notion de hauteur $H(B)$. Dans la partie 2 nous introduisons entre autres des outils utiles pour la définition de cette notion (comme les coordonnées de Grassmann). C'est dans la partie 4 que l'on définit explicitement la hauteur. Schmidt donne une deuxième définition équivalente utile et plus géométrique que nous étudions aussi et qui permet de l'exprimer en fonction du covolume d'un réseau d'un certain espace euclidien (Cf partie 4.2); c'est cette dernière définition qu'utilise notamment Roy dans son article [15]. Une fois tout cela posé, nous sommes en mesure d'énoncer les premiers résultats d'approximation diophantienne de Schmidt; c'est l'objet de la partie 5 (et plus particulièrement 5.3). L'objectif des parties 2–5 de ce mémoire est double : d'une part nous y avons développé (en proposant la plupart des preuves) les outils qu'utilise Schmidt sans démonstration en renvoyant à d'autres ouvrages (comme c'est le cas notamment pour les coordonnées de Grassmann), et d'autre part nous avons essayé de fournir un certain nombre de démonstrations que laisse partiellement ou totalement Schmidt au lecteur (par exemple le cas (b) des théorèmes 5.1.1 et 5.2.2). Notons également que les arguments de Schmidt pour le théorème 5.2.3 (going-down) pour le cas (b) ne fonctionnent pas à un endroit de la preuve; nous avons proposé une correction pour les rendre valables.

Le deuxième aspect abordé par ce mémoire (parties 6 à 8) est celui d'un domaine relativement nouveau fondé par Schmidt et Summerer : la géométrie paramétrique des nombres (Cf [18], [19] et les travaux plus récents de Roy [14], [15]). Dans ce cadre Schmidt et Summerer ont redémontré un certain nombre d'inégalités classiques d'approximation diophantienne et en ont découvert de nouvelles. L'idée de la géométrie paramétrique des nombres développée par Schmidt et Summerer dans [18] et [19] est d'étudier la suite des minima successifs d'une famille de corps convexes paramétrée par un réel $Q > 1$ pour un réseau fixé construit à partir d'un vecteur $u = (1, \xi_1, \dots, \xi_{n-1}) \in \mathbb{R}^n$ dont les coordonnées sont linéairement indépendantes sur \mathbb{Q} . Le logarithme en base Q

de ces minima est intimement lié à des exposants classiques d'approximation diophantienne. Dans la partie 6 nous parlons plus en détails de ces problèmes. Schmidt et Summerer ont pu, grâce à cette nouvelle manière de traiter les problèmes d'approximation rationnelle, redémontrer des inégalités classiques et en découvrir de nouvelles. Dans [19] les auteurs montrent que le n -uplet des minima successifs d'une famille de corps convexes symétriques par rapport à un certain réseau $\Lambda(\mathbf{u})$ (avec $\mathbf{u} \in \mathbb{R}^n$) peut être approximé par une fonction d'une certaine classe (les (n, γ) -systèmes, cf définition 7.3.7). Les auteurs conjecturent que réciproquement, à toute fonction de cette classe correspondrait un vecteur \mathbf{u} de sorte qu'elle approcherait les minima successifs associés au réseau $\Lambda(\mathbf{u})$ et à la famille de corps convexes paramétrés. Roy démontre dans un article récent [14] cette conjecture en simplifiant la classe de fonctions considérée (il le montre pour des n -systèmes rigides qui sont des $(n, 0)$ -systèmes particuliers, cf définition 7.1.2). La partie 7 est consacrée à l'article [14] et à la présentation du schéma de la preuve de Roy. Dans l'article [15], il donne une importante application de son théorème. Soit $n \geq 1$ un entier et $\mathbf{u} \in \mathbb{R}^n$ non nul. Pour tout $Q \geq 1$, l'auteur définit en suivant l'idée de Schmidt et Summerer le corps convexe suivant :

$$\mathcal{C}_{\mathbf{u}}(Q) = \{\mathbf{x} \in \mathbb{R}^{n+1} ; \|\mathbf{x}\| \leq 1, |\mathbf{x} \cdot \mathbf{u}| \leq Q^{-1}\}$$

et note $\lambda_1(\mathcal{C}_{\mathbf{u}}(Q)) \leq \dots \leq \lambda_{n+1}(\mathcal{C}_{\mathbf{u}}(Q))$ ses $n + 1$ minima successifs (pour le réseau \mathbb{Z}^n). Il pose aussi

$$L_{\mathbf{u},j}(q) = \log \lambda_j(\mathcal{C}_{\mathbf{u}}(e^q)) \quad (q \geq 0, 1 \leq j \leq n + 1),$$

et réunit ces applications en une unique application $\mathbf{L}_{\mathbf{u}} : [0, \infty[\rightarrow \mathbb{R}^{n+1}$ en posant

$$\mathbf{L}_{\mathbf{u}}(q) = (L_{\mathbf{u},1}(q), \dots, L_{\mathbf{u},n+1}(q)) \quad (q \geq 0).$$

Dans le but d'approximer la fonction $\mathbf{L}_{\mathbf{u}}$, Roy utilise une classe de fonctions déjà définie par Schmidt et Summerer dans le cas particulier suivant (ce sont les $(n, 0)$ -systèmes dans [19]). Soit $q_0 \geq 0$. Un n -système sur $[q_0, \infty[$ est une fonction continue affine par morceaux $\mathbf{P} = (P_1, \dots, P_n) : [q_0, \infty[\rightarrow \mathbb{R}^n$ vérifiant les conditions suivantes :

(S1) Pour tout $q \geq q_0$ on a $0 \leq P_1(q) \leq \dots \leq P_n(q)$ et $P_1(q) + \dots + P_n(q) = q$.

(S2) Si H est un sous-intervalle ouvert non vide de $[q_0, \infty[$ sur lequel \mathbf{P} est différentiable, alors il existe un entier r ($1 \leq r \leq n$) tel que P_r est de pente 1 sur H et P_j est constante sur H pour tout $j \neq r$.

(S3) Si $q > q_0$ est un point en lequel \mathbf{P} n'est pas différentiable et si les entiers r et s qui vérifient $P'_r(q^-) = P'_s(q^+) = 1$ sont tels que $r < s$, alors on a $P_r(q) = P_{r+1}(q) = \dots = P_s(q)$.

Roy prouve alors l'important résultat suivant :

Théorème 1.0.1. *Soit $\mathbf{u} \in \mathbb{R}^n$ non nul. Alors il existe $q_0 \geq 0$ et un n -système \mathbf{P} sur $[q_0, \infty[$ tel que $\mathbf{L}_{\mathbf{u}} - \mathbf{P}$ soit borné sur $[q_0, \infty[$. Réciproquement, pour tout n -système \mathbf{P} sur un intervalle $[q_0, \infty[$ avec $q_0 \geq 0$, il existe un vecteur $\mathbf{u} \in \mathbb{R}^n$ non nul tel que $\mathbf{L}_{\mathbf{u}} - \mathbf{P}$ soit borné sur $[q_0, \infty[$.*

En fait Roy prouve même ce résultat pour une classe de fonctions encore plus réduite dans [14] (les n -systèmes rigides ; cf la partie 7 de ce mémoire où nous présentons la construction de Roy). Dans [19], Schmidt et Summerer montrent la première assertion de ce théorème pour une classe de fonctions plus grande que nous passons sous silence ici (les (n, γ) -systèmes, cf [19] p. 59). Avant de parler de l'intéressante application que tire Roy de son théorème, rappelons les exposants classiques d'approximation diophantienne

que nous évoquions.

Soit $n \geq 1$ et $\mathbf{u} \in \mathbb{R}^{n+1} \setminus \{0\}$. Pour $j = 0, \dots, n-1$ on note $\omega_j(\mathbf{u})$ (resp. $\widehat{\omega}_j(\mathbf{u})$) la borne supérieure de l'ensemble des réels ω tels que pour des Q arbitrairement grands (resp. pour tout Q assez grand) il existe un sous-espace $S \subset \mathbb{R}^{n+1}$ défini sur \mathbb{Q} de dimension $j+1$ vérifiant

$$H(S) \leq Q \quad \text{et} \quad H(S)\text{dist}(\mathbf{u}, S) \leq Q^{-\omega},$$

où la distance (projective) $\text{dist}(\mathbf{u}, S)$ du vecteur \mathbf{u} au sous-espace S est égale, en notant proj_{S^\perp} la projection orthogonale sur S^\perp , à

$$\text{dist}(\mathbf{u}, S) = \frac{\|\text{proj}_{S^\perp}(\mathbf{u})\|}{\|\mathbf{u}\|}.$$

Géométriquement, cela représente le sinus de l'angle le plus petit entre la droite engendrée par \mathbf{u} et S ; cet angle est le premier angle d'inclinaison entre la droite engendrée par \mathbf{u} et S .

On peut établir un certain nombre de résultats sur ces exposants. D'abord, un théorème de Schmidt [17] (cf le théorème 5.3.3 de ce mémoire) assure que

$$\omega_j(\mathbf{u}) \geq \widehat{\omega}_j(\mathbf{u}) \geq \frac{j+1}{n-j} \quad (0 \leq j \leq n-1).$$

On a aussi les encadrements suivants :

Théorème 1.0.2 (Schmidt, Laurent). *Soit $n \geq 1$ un entier. Pour tout vecteur $\mathbf{u} \in \mathbb{R}^{n+1}$ non nul on a $\omega_0(\mathbf{u}) \geq 1/n$ et*

$$\frac{j\omega_j(\mathbf{u})}{\omega_j(\mathbf{u}) + j + 1} \leq \omega_{j-1}(\mathbf{u}) \leq \frac{(n-j)\omega_j(\mathbf{u}) - 1}{n-j+1} \quad (1 \leq j \leq n-1), \quad (1.1)$$

en définissant par convention le quotient de gauche comme étant égal à j et celui de droite à ∞ si $\omega_j(\mathbf{u}) = \infty$.

Les inégalités de droite peuvent se prouver à l'aide du going-up (théorème 5.8) et celles de gauche à l'aide du going-down (théorème 5.2.2) de [17] (Cf le théorème 8.1.3 de la partie 8.1 de ce mémoire pour le détail de la preuve). Ces inégalités ont été observées par Laurent dans [9] qui introduisit alors les exposants $\omega_j(\mathbf{u})$. Dans le même article il fait la remarque que chaque inégalité (1.1) prise individuellement est la meilleure possible car en les combinant on trouve les inégalités de transfert de Khinchine qui sont connues pour être optimales. On peut aussi montrer que l'ensemble des valeurs prises par $\omega_j(\mathbf{u})$ est tout l'intervalle $[(j+1)/(n-j), \infty[$.

En exprimant ces quantités $\omega_j(\mathbf{u}), \widehat{\omega}_j(\mathbf{u})$ dans le langage de la géométrie paramétrique des nombres et en utilisant le théorème 1.0.1, Roy montre

Théorème 1.0.3 (Roy, 2014). *Soit $n \geq 1$ un entier. Soient $\omega_0, \dots, \omega_{n-1} \in [0, \infty[$ vérifiant les inégalités $\omega_0 \geq 1/n$ et*

$$\frac{j\omega_j}{\omega_j + j + 1} \leq \omega_{j-1} \leq \frac{(n-j)\omega_j - 1}{n-j+1} \quad (1 \leq j \leq n-1).$$

Alors il existe un vecteur $\mathbf{u} \in \mathbb{R}^{n+1}$ dont les coordonnées sont \mathbb{Q} -linéairement indépendantes et tel que

$$\omega_j(\mathbf{u}) = \omega_j \quad \text{et} \quad \widehat{\omega}_j(\mathbf{u}) = \frac{j+1}{n-j} \quad (0 \leq j \leq n-1).$$

Ce théorème est le résultat principal de [15] (cf partie 8 de ce mémoire).

D'avantage que les détails techniques des preuves, dans les parties 6 à 8 nous avons essayé de retranscrire les structures et la plupart des constructions mises en jeu dans ces deux théorèmes.

2 Propriétés du déterminant, coordonnées de Grassmann et géométrie des nombres

2.1 Propriétés générales du déterminant, déterminants et matrices composés

Dans cette section sera présenté un certain nombre de résultats techniques sur le déterminant. La plupart sont issus de [1] et [5], les preuves proposées sont cependant parfois un peu différentes de celles des auteurs. Leur utilité première pour nous est de pouvoir établir les propriétés de base des coordonnées de Grassmann (et coordonnées de Grassmann duales) du paragraphe 2.2, ainsi que celles des corps convexes composés du paragraphe 2.4. La notion la plus importante est celle des matrices composées et adjointes composées p -èmes $A^{(p)}$ et $\text{Adj}^{(p)}A$. Les premiers résultats techniques s'exprimeront facilement en termes de propriétés de ces matrices. Par ailleurs, elles nous permettront de définir simplement ce qu'est un ensemble de coordonnées de Grassmann d'un sous-espace de \mathbb{K}^n .

Soit $1 \leq p \leq n$ deux entiers fixés. On pose $N = \binom{n}{p}$. Les conventions suivantes sont importantes et seront réutilisées dans les parties 2.2 et 2.4. Elles sont personnelles et inspirées de celles utilisées par Schmidt dans [17] pour introduire les coordonnées de Grassmann. On ordonne les p -uplets $(j_1 < \dots < j_p)$ d'éléments distincts de $\llbracket 1, n \rrbracket$ par l'ordre lexicographique (il y en a N); \underline{j} désignera le i -ème de ces p -uplets (attention, n et p sont sous-entendus dans cette définition, on s'assurera que le contexte dans lequel on l'utilisera ne laissera pas d'ambiguïté) et on notera $(i_1 < \dots < i_p)$ ses éléments. On notera également $\underline{j}' = (i_{p+1} < \dots < i_n)$ son complémentaire. Notons qu'il y a également N $(n-p)$ -uplets $(j_{p+1} < \dots < j_n)$ qu'on peut également ordonner par l'ordre lexicographique. Si \underline{j} est plus petit que \underline{j}' pour l'ordre lexicographique, alors \underline{j}' est plus grand que \underline{j} : on en déduit donc que \underline{j}' est le $(N - i + 1)$ -ème des $(n-p)$ -uplets. S'il n'y a pas d'ambiguïté, $\underline{j} = (j_1 < \dots < j_{n-p})$ pourra donc aussi désigner le j -ème des $(n-p)$ -uplets (avec nos notations, si \underline{j} est un p -uplet, alors $\underline{j}' = \underline{N - i + 1}$). Nous ne nous servirons pas tout de suite de ces notations.

$\mathcal{M}_{p,q}$ désignera dans cette section l'espace des matrices sur un anneau commutatif unitaire fixé (en pratique, ce sera souvent \mathbb{C} ou un sous-corps de \mathbb{C}) à p lignes et q colonnes.

Les notations suivantes (issues de [5]) seront commodes pour des raisons techniques.

Définition 2.1.1. Soient $i_1, \dots, i_n \in \{1, \dots, n\}$. Si $\sigma(k) := i_k$ définit une permutation de $\{1, \dots, n\}$ alors on note $\varepsilon^{i_1 \dots i_n}$ sa signature. Sinon on pose $\varepsilon^{i_1 \dots i_n} = 0$.

Définition 2.1.2. Soient (i_1, \dots, i_p) et (j_1, \dots, j_p) deux p -uplets d'entiers. Si $\{i_1, \dots, i_p\} = \{j_1, \dots, j_p\} =: I$ sont de cardinal p et inclus dans $\{1, \dots, n\}$ alors on définit $\delta_{\binom{n}{i_1 \dots i_p}}^{j_1 \dots j_p}$, ou plus simplement $\delta^{i_1 \dots i_p}$ s'il n'y a pas d'ambiguïté, comme étant le produit des signatures des permutations de I associées à (i_1, \dots, i_p) et (j_1, \dots, j_p) (en d'autres termes c'est la signature de la permutation qui envoie i_k sur j_k). Sinon, on pose $\delta_{\binom{n}{i_1 \dots i_p}}^{j_1 \dots j_p} = 0$.

Dans ce cadre, si $A = (a_{ij})_{i,j}$ est une matrice de taille n , alors on a

$$\det A = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) \prod_{i=1}^n a_{\sigma(i)i} \quad (2.1)$$

$$= \sum_{i_1, \dots, i_n=1}^n \varepsilon^{i_1 \dots i_n} \prod_{k=1}^n a_{i_k k}. \quad (2.2)$$

La deuxième formule peut paraître plus lourde que la première mais s'avère bien plus pratique pour faire des calculs avec des déterminants extraits. On a notamment la propriété utile suivante, qui vient directement de la définition de $\delta^{j_1 \dots j_n}_{i_1 \dots i_n}$ et de (2.2) :

Proposition 2.1.3. *Si $B = (b_{ij})_{i,j}$ est une matrice de taille $p \times q$ et A une sous-matrice de taille n dont les éléments sont ceux des lignes $i_1 < \dots < i_n$ et des colonnes $j_1 < \dots < j_n$ de B , alors*

$$\det A = \sum_{k_1, \dots, k_n=1}^p \delta^{k_1 \dots k_n}_{i_1 \dots i_n} \prod_{l=1}^n b_{k_l j_l}. \quad (2.3)$$

Le théorème suivant connu sous le nom de développement de Laplace est une généralisation des formules classiques de développement du déterminant par rapport à une ligne ou à une colonne.

Théorème 2.1.4 (Développement de Laplace). *Soit $A = (a_{ij})_{i,j}$ une matrice de taille n et (i_1, \dots, i_n) , (j_1, \dots, j_n) deux permutations de $\{1, \dots, n\}$. On fixe $1 \leq p \leq n$ un entier. Alors on a les deux formules suivantes :*

$$\det A = \sum_{\substack{k_1 < \dots < k_p \\ k_{p+1} < \dots < k_n}} \delta^{j_1 \dots j_n}_{k_1 \dots k_n} \begin{vmatrix} a_{k_1 j_1} & \dots & a_{k_1 j_p} \\ \vdots & & \vdots \\ a_{k_p j_1} & \dots & a_{k_p j_p} \end{vmatrix} \cdot \begin{vmatrix} a_{k_{p+1} j_{p+1}} & \dots & a_{k_{p+1} j_n} \\ \vdots & & \vdots \\ a_{k_n j_{p+1}} & \dots & a_{k_n j_n} \end{vmatrix}, \quad (2.4)$$

et

$$\det A = \sum_{\substack{k_1 < \dots < k_p \\ k_{p+1} < \dots < k_n}} \delta^{k_1 \dots k_n}_{i_1 \dots i_n} \begin{vmatrix} a_{i_1 k_1} & \dots & a_{i_1 k_p} \\ \vdots & & \vdots \\ a_{i_p k_1} & \dots & a_{i_p k_p} \end{vmatrix} \cdot \begin{vmatrix} a_{i_{p+1} k_{p+1}} & \dots & a_{i_{p+1} k_n} \\ \vdots & & \vdots \\ a_{i_n k_{p+1}} & \dots & a_{i_n k_n} \end{vmatrix}. \quad (2.5)$$

Preuve La démonstration présentée ici est tirée de [5] (chapitre II, section 8).

Prouvons (2.4) et fixons (j_1, \dots, j_n) .

On réarrange les facteurs a_{ij} dans (2.2) pour faire apparaître les j_k et on obtient en posant $k_l := i_{j_l}$:

$$\det A = \sum_{i_1, \dots, i_n=1}^n \varepsilon^{i_1 \dots i_n} \prod_{l=1}^n a_{k_l j_l} = \varepsilon^{j_1 \dots j_n} \times \sum_{k_1, \dots, k_n=1}^n \varepsilon^{k_1 \dots k_n} \prod_{l=1}^n a_{k_l j_l}, \quad (2.6)$$

en utilisant l'égalité $\varepsilon^{i_1 \dots i_n} = \varepsilon^{j_1 \dots j_n} \varepsilon^{k_1 \dots k_n}$.

On fixe (k_1, \dots, k_n) avec $k_1 < \dots < k_p$ et $k_{p+1} < \dots < k_n$ (il y a $\binom{n}{p}$ tels n -uplets). On considère les permutations (k'_1, \dots, k'_n) de $\{1, \dots, n\}$ tels que (k'_1, \dots, k'_p) soit une permutation de $\{k_1, \dots, k_p\}$; (k'_{p+1}, \dots, k'_n) est alors une permutation de $\{k_{p+1}, \dots, k_n\}$. Il y a $p!(n-p)!$ telles permutations.

Comme $\varepsilon^{j_1 \dots j_n} \varepsilon^{k'_1 \dots k'_n} = \delta^{j_1 \dots j_n}_{k_1 \dots k_n} \delta^{k'_1 \dots k'_p}_{k_1 \dots k_p} \delta^{k'_{p+1} \dots k'_n}_{k_{p+1} \dots k_n}$, la contribution à la somme (2.6) de

ces permutations vaut

$$\begin{aligned} \delta^{j_1 \dots j_n}_{k_1 \dots k_n} \times \sum_{k'_1, \dots, k'_n=1}^n \delta^{k'_1 \dots k'_p}_{k_1 \dots k_p} \delta^{k'_{p+1} \dots k'_n}_{k_{p+1} \dots k_n} \prod_{l=1}^n a_{k'_l j_l} \\ = \delta^{j_1 \dots j_n}_{k_1 \dots k_n} \left| a_{k_i, j_l} \right|_{1 \leq i, l \leq p} \cdot \left| a_{k_{p+i}, j_{p+l}} \right|_{1 \leq i, l \leq n-p} \end{aligned}$$

par (2.3). D'où le résultat.

On obtient (2.5) par transposition. \square

Le théorème qui suit généralise la formule donnant l'expression d'un coefficient d'un produit de deux matrices en fonction des coefficients de chacun des facteurs. Ici, on exprime les mineurs d'un produit de deux matrices en fonction des mineurs de chacun des deux facteurs. Plus tard, on donnera une version matricielle de ce théorème.

Théorème 2.1.5. *Soit $A = (a_{ij})_{i,j} \in \mathcal{M}_{p,n}$, $B = (b_{ij})_{i,j} \in \mathcal{M}_{n,q}$ et $C = (c_{ij})_{i,j} = AB \in \mathcal{M}_{p,q}$. Soit $t \leq \min(p, q, n)$, $i_1 < \dots < i_t$ et $j_1 < \dots < j_t$. Alors*

$$\begin{vmatrix} c_{i_1, j_1} & \dots & c_{i_1, j_t} \\ \vdots & & \vdots \\ c_{i_t, j_1} & \dots & c_{i_t, j_t} \end{vmatrix} = \sum_{k_1 < \dots < k_t} \begin{vmatrix} a_{i_1, k_1} & \dots & a_{i_1, k_t} \\ \vdots & & \vdots \\ a_{i_t, k_1} & \dots & a_{i_t, k_t} \end{vmatrix} \cdot \begin{vmatrix} b_{k_1, j_1} & \dots & b_{k_1, j_t} \\ \vdots & & \vdots \\ b_{k_t, j_1} & \dots & b_{k_t, j_t} \end{vmatrix}. \quad (2.7)$$

Preuve La preuve est issue de [5] (chapitre II, section 8).

On note

$$\begin{aligned} D = \begin{vmatrix} c_{i_1, j_1} & \dots & c_{i_1, j_t} \\ \vdots & & \vdots \\ c_{i_t, j_1} & \dots & c_{i_t, j_t} \end{vmatrix} &= \sum_{\tau_1, \dots, \tau_t=1}^p \delta^{i_1 \dots i_t}_{\tau_1 \dots \tau_t} \prod_{l=1}^t c_{\tau_l, j_l} \quad (\text{par (2.3)}) \\ &= \sum_{h_1, \dots, h_t=1}^n \sum_{\tau_1, \dots, \tau_t=1}^p \delta^{i_1 \dots i_t}_{\tau_1 \dots \tau_t} \prod_{l=1}^t a_{\tau_l, h_l} \prod_{l=1}^t b_{h_l, j_l} \\ &= \sum_{h_1, \dots, h_t=1}^n \begin{vmatrix} a_{i_1, h_1} & \dots & a_{i_1, h_t} \\ \vdots & & \vdots \\ a_{i_t, h_1} & \dots & a_{i_t, h_t} \end{vmatrix} \prod_{l=1}^t b_{h_l, j_l}. \end{aligned}$$

Or, on peut écrire le déterminant dans la somme sous la forme

$$\begin{vmatrix} a_{i_1, h_1} & \dots & a_{i_1, h_t} \\ \vdots & & \vdots \\ a_{i_t, h_1} & \dots & a_{i_t, h_t} \end{vmatrix} = \sum_{k_1 < \dots < k_t} \delta^{k_1 \dots k_t}_{h_1 \dots h_t} \begin{vmatrix} a_{i_1, k_1} & \dots & a_{i_1, k_t} \\ \vdots & & \vdots \\ a_{i_t, k_1} & \dots & a_{i_t, k_t} \end{vmatrix},$$

donc finalement

$$\begin{aligned} D &= \sum_{k_1 < \dots < k_t} \left(\begin{vmatrix} a_{i_1, k_1} & \dots & a_{i_1, k_t} \\ \vdots & & \vdots \\ a_{i_t, k_1} & \dots & a_{i_t, k_t} \end{vmatrix} \sum_{h_1, \dots, h_t=1}^n \delta^{k_1 \dots k_t}_{h_1 \dots h_t} \prod_{l=1}^t b_{h_l, j_l} \right) \\ &= \sum_{k_1 < \dots < k_t} \begin{vmatrix} a_{i_1, k_1} & \dots & a_{i_1, k_t} \\ \vdots & & \vdots \\ a_{i_t, k_1} & \dots & a_{i_t, k_t} \end{vmatrix} \begin{vmatrix} b_{k_1, j_1} & \dots & b_{k_1, j_t} \\ \vdots & & \vdots \\ b_{k_t, j_1} & \dots & b_{k_t, j_t} \end{vmatrix}. \end{aligned}$$

\square

Remarque(s) 2.1.6. Dans le cas particulier où $p = q = t$, la formule précédente est appelée formule de Cauchy-Binet.

Rappelons-nous les notations introduites en début de paragraphe sur \underline{i} et \underline{j}' . Les définitions qui suivent sont personnelles.

Définition 2.1.7. Soit A une matrice de taille $r \times s$ avec $\min(r, s) \geq p$. Si $\underline{i} = (i_1 < \dots < i_p)$ est le i -ème p -uplet de $\{1, \dots, r\}$ et $\underline{j} = (j_1 < \dots < j_p)$ est le j -ème p -uplet de $\{1, \dots, s\}$, on note :

$$|A|_{\underline{i}, \underline{j}}^{(p)} = \begin{vmatrix} a_{i_1 j_1} & \dots & a_{i_1 j_p} \\ \vdots & & \vdots \\ a_{i_p j_1} & \dots & a_{i_p j_p} \end{vmatrix} \quad (2.8)$$

Si $r = s =: n$, on pose

$$|A|_{\underline{i}', \underline{j}'}^{(n-p)} = \begin{vmatrix} a_{i_{p+1} j_{p+1}} & \dots & a_{i_{p+1} j_n} \\ \vdots & & \vdots \\ a_{i_n j_{p+1}} & \dots & a_{i_n j_n} \end{vmatrix}.$$

Enfin (toujours dans le cas $n = r = s$), pour simplifier on notera

$$\varepsilon_{(n)}^{\underline{i}, \underline{j}} := \delta_{(n)}^{j_1 \dots j_n, i_1 \dots i_n},$$

ou simplement $\varepsilon^{\underline{i}, \underline{j}}$ s'il n'y a pas d'ambiguïté sur le n .

On peut réécrire (2.4), (2.5) et (2.7) avec ces nouvelles notations :

Théorème 2.1.8. Soit $A = (a_{ij})_{i,j}$ une matrice de taille n . On fixe $1 \leq p \leq n$ entier. Alors pour tous i et j compris entre 1 et $N = \binom{n}{p}$ on a les formules suivantes :

$$\det A = \sum_{k=1}^N \varepsilon^{k, j} |A|_{\underline{k}, \underline{j}}^{(p)} |A|_{\underline{k}', \underline{j}'}^{(n-p)}, \quad (2.9)$$

et

$$\det A = \sum_{k=1}^N \varepsilon^{\underline{i}, \underline{k}} |A|_{\underline{i}, \underline{k}}^{(p)} |A|_{\underline{i}', \underline{k}'}^{(n-p)}. \quad (2.10)$$

Théorème 2.1.9. Soit $A = (a_{ij})_{i,j} \in \mathcal{M}_{p,n}$, $B = (b_{ij})_{i,j} \in \mathcal{M}_{n,q}$ et $C = (c_{ij})_{i,j} = AB \in \mathcal{M}_{p,q}$. On pose $n' := \min(p, q, n)$, on fixe $t \leq n'$ et on pose $N := \binom{n'}{t}$. Alors pour tout i compris entre 1 et $\binom{p}{t}$ et tout j compris entre 1 et $\binom{q}{t}$ on a

$$|C|_{\underline{i}, \underline{j}}^{(t)} = \sum_{k=1}^N |A|_{\underline{i}, \underline{k}}^{(t)} |B|_{\underline{k}, \underline{j}}^{(t)}. \quad (2.11)$$

On peut définir alors les matrices composée et adjointe composée p -èmes d'une matrice (cf [1] ou [5] p.291). Notons que Hodge et Pedoe les définissent après avoir défini les ensemble de coordonnées de Grassmann, alors que nous allons au contraire définir les coordonnées de Grassmann comme étant une certaine matrice composée p -ème (cf le début du paragraphe 2.2).

Définition 2.1.10. Soit A une matrice de taille $r \times s$ et $1 \leq p \leq \min(r, s)$ fixé. On note $R := \binom{r}{p}$ et $S := \binom{s}{p}$. On définit $A^{(p)}$ - appelée matrice composée p -ème de A - comme étant la matrice de taille $R \times S$ dont le coefficient (i, j) est

$$A_{i,j}^{(p)} = |A|_{\underline{i}, \underline{j}}^{(p)}. \quad (2.12)$$

Si $r = s =: n$ (i.e. que A est carrée), chaque mineur d'ordre p de A est associé dans le développement de Laplace (2.9) à un cofacteur d'ordre $n - p$ (partir de (2.10) donne le même cofacteur). On note $\text{Adj}^{(p)}A$ et on appelle matrice composée p -ème adjointe la matrice transposée de la matrice dont les coefficients sont ces cofacteurs. Ainsi, le coefficient d'indice (i, j) de $\text{Adj}^{(p)}A$ est

$$\left(\text{Adj}^{(p)}A\right)_{i,j} = \varepsilon^{j,i} |A|_{\underline{j'}, \underline{i'}}^{(n-p)}. \quad (2.13)$$

Cette égalité permettra dans la suite d'établir (2.20) qui fournit une relation entre $\text{Adj}^{(p)}A$ et $A^{(n-p)}$. Un petit calcul et (2.11) permettent d'établir l'agréable propriété suivante :

Proposition 2.1.11. *Soit $A \in \mathcal{M}_{p,n}$ et $B \in \mathcal{M}_{n,q}$ deux matrices, et $1 \leq p \leq \min(n, p, q)$ un entier. On pose $N = \binom{n}{p}$. Alors*

$$I_n^{(p)} = I_N \quad \text{et} \quad (AB)^{(p)} = A^{(p)}B^{(p)}. \quad (2.14)$$

Théorème 2.1.12. *Si A est une matrice carrée de taille n et $1 \leq p \leq n$, on a la formule*

$$A^{(p)} \cdot \text{Adj}^{(p)}A = \text{Adj}^{(p)}A \cdot A^{(p)} = (\det A)I_N, \quad (2.15)$$

où I_N désigne la matrice identité de \mathcal{M}_N . En particulier, on en déduit que

$$\det A^{(p)} = (\det A)^{\binom{n-1}{p-1}}, \quad (2.16)$$

et

$$\det \text{Adj}^{(p)}A = (\det A)^{\binom{n-1}{p}}. \quad (2.17)$$

Preuve Les arguments sont issus de [1] (chapitre 2 et 5).

Montrons d'abord (2.16) et (2.17) à partir de (2.15). On suppose (provisoirement) ici que nos matrices sont à coefficients dans le corps \mathbb{C} . On considère alors $\det A$ comme un polynôme en n^2 variables, de degré n . Par (2.15) on a l'identité polynomiale

$$(\det A^{(p)}) \cdot (\det \text{Adj}^{(p)}A) = (\det A)^{\binom{n}{p}}.$$

Remarquons que $\det A$ est un polynôme premier. L'argument est tiré de [1] (chapitre 2, section 18) : en effet, sinon $\det A$ s'écrirait PQ avec P et Q non constants. P contient au moins un certain a_{ij} et donc par (2.2), Q ne contient aucun a_{it} ni a_{tj} (pour tout t). De même, Q contient un a_{rs} et le même argument assure que P ne peut contenir aucun a_{rt} ni a_{ts} . Mais alors ni a_{is} ni a_{rj} n'apparaissent dans P et Q donc dans $PQ = \det A$, ce qui est visiblement faux. Donc $\det A$ est premier.

On en déduit l'existence de constantes non nulles λ et μ et d'entiers α et β tels que $\det A^{(p)} = \lambda(\det A)^\alpha$ et $\det \text{Adj}^{(p)}A = \mu(\det A)^\beta$. En évaluant en $A = xI_n$ on a $A^{(p)} = x^p I_N$. On en déduit immédiatement que $\lambda = 1$, $x^{p \binom{n}{p}} = x^{n\alpha}$, donc $\alpha = \binom{n-1}{p-1}$, et par suite que $\mu = 1$ et $\beta = \binom{n-1}{p}$.

Reste donc à prouver (2.15).

Par (2.12) et (2.13), le coefficient d'indice (i, j) de $A^{(p)}\text{Adj}^{(p)}A$ est alors :

$$\left(A^{(p)}\text{Adj}^{(p)}A\right)_{i,j} = \sum_k \varepsilon^{j,k} |A|_{i,k}^{(p)} |A|_{j',k'}^{(n-p)}$$

Par (2.10) utilisé avec (j, j') , cette somme est égale au déterminant de la matrice obtenue à partir de A en remplaçant la j_l -ème ligne par la i_l -ème ligne, $l = 1 \dots p$. En effet, si on appelle B la matrice ainsi obtenue, le déterminant $|B|_{j,k}^{(p)}$ sera alors égal

à $|A|_{\underline{i}, \underline{k}}^{(p)}$ et $|B|_{\underline{j}', \underline{k}'}^{(n-p)}$ restera inchangé et égal à $|A|_{\underline{j}', \underline{k}'}^{(n-p)}$. Puisque le déterminant est n -linéaire alterné par rapport aux lignes, B est de déterminant nul si $i \neq j$ (car cela signifie que deux lignes au moins de B sont identiques), de déterminant égal à $\det A$ si $i = j$. D'où

$$A^{(p)} \cdot \text{Adj}^{(p)} A = (\det A) I_N.$$

Un raisonnement similaire permet d'aboutir à

$$\left(\text{Adj}^{(p)} A \cdot A^{(p)} \right)_{\underline{i}, \underline{j}} = \sum_k \varepsilon^{k, \underline{i}} |A|_{\underline{k}, \underline{j}}^{(p)} |A|_{\underline{k}', \underline{i}'}^{(n-p)} = \delta^{j_1 \dots j_p}_{i_1 \dots i_p} \det A,$$

ce qui donne

$$\text{Adj}^{(p)} A \cdot A^{(p)} = (\det A) I_N.$$

On peut enfin remarquer que (2.15), (2.16) et (2.17) sont des identités polynomiales en leurs coordonnées à coefficients entiers. En particulier, elles restent vraies sur n'importe quel anneau commutatif unitaire. \square

Remarque : $\text{Adj}^{(1)} A$ est simplement la transposée de la comatrice de A et les formules précédentes sont plus classiques. Le théorème suivant permet de relier les mineurs d'ordre p d'une matrice inversible à ceux d'ordre $(n-p)$ de son inverse, et réciproquement.

Théorème 2.1.13 (de Jacobi). *Soit A une matrice de taille n et $1 \leq p \leq n$ un entier fixé. Soit i et j compris entre 1 et $N = \binom{n}{p}$. Alors :*

$$\left(\text{Adj}^{(1)} A \right)^{(p)} = (\det A)^{p-1} \text{Adj}^{(p)} A.$$

En particulier, si A est inversible et puisque $A^{-1} = (\det A)^{-1} \text{Adj}^{(1)} A$, on a

$$(A^{-1})^{(p)} = (\det A)^{-1} \text{Adj}^{(p)} A. \quad (2.18)$$

Preuve Hodge et Pedoe ne donnent pas d'énoncé matriciel de ce théorème dans [5] mais directement une égalité sur les coefficients des matrices mises en jeu (cf Thm VIII section 8 chapitre II). La preuve proposée ici est également différente de la leur : plus matricielle comme le suggère son énoncé, et utilisant des arguments plus algébriques qui reposent sur la topologie de \mathbb{C} pour éviter les distinctions de cas, et qui s'étendent ensuite à un anneau commutatif unitaire quelconque. Supposons pour commencer que A est inversible et que nos matrices sont à coefficients dans le corps \mathbb{C} . Alors, d'après (2.15)

$$A^{-1} = (\det A)^{-1} \text{Adj}^{(1)} A.$$

D'autre part, en appliquant (2.14) à (2.15) avec $p = 1$, on a

$$A^{(p)} \left(\text{Adj}^{(1)} A \right)^{(p)} = \left(\text{Adj}^{(1)} A \right)^{(p)} A^{(p)} = (\det A)^p I_N,$$

donc

$$A^{-1} = (\det A)^{-p} \left(\text{Adj}^{(1)} A \right)^{(p)}.$$

En identifiant ces deux expressions et en multipliant par $(\det A)^p$ on trouve le résultat annoncé. On remarque ensuite que la formule est polynomiale en ses coordonnées,

valable sur l'ouvert dense de \mathbb{C}^{n^2} correspondant à $\det A \neq 0$. Elle est donc vraie même dans le cas $\det A = 0$. Pour passer à un anneau commutatif unitaire quelconque, il suffit encore une fois de remarquer que c'est une formule polynomiale en ses coordonnées et à coefficients entiers. \square

Terminons cette section avec un dernier théorème qui permet de lier $\text{Adj}^{(p)} A$ et $A^{(n-p)}$. Ce résultat servira directement pour les coordonnées de Grassmann duales.

Théorème 2.1.14. *Soit n en entier et $1 \leq p \leq n$ fixé. On pose $N := \binom{n}{p}$ et pour tout i compris entre 1 et N , $\varepsilon_i := \varepsilon^{i_1, \dots, i_n}$. On définit*

$$M := \begin{pmatrix} 0 & \cdots & 0 & \varepsilon_N \\ \vdots & & \ddots & 0 \\ 0 & \ddots & & \vdots \\ \varepsilon_1 & 0 & \cdots & 0 \end{pmatrix}. \quad (2.19)$$

Soit A une matrice de taille n . Alors

$$M \left(\text{Adj}^{(p)} A \right) M^{-1} = A^{(n-p)}. \quad (2.20)$$

En particulier, si A est inversible on a les formules

$${}^t(A^{-1})^{(p)} = (\det A)^{-1} M^{-1} A^{(n-p)} M \quad (2.21)$$

$${}^t(A^{-1})^{(n-p)} = (\det A)^{-1} M A^{(p)} M^{-1}. \quad (2.22)$$

Preuve Aitken laisse ces résultats en exercice dans son livre.

On note $\delta_{i,j}$ le symbole de Kronecker qui vaut 1 si $i = j$, 0 sinon. Si i et j sont compris entre 1 et N , les coefficients d'indices (i, j) de M et M^{-1} sont respectivement $M_{i,j} = \delta_{j, N-i+1} \varepsilon_{N-i+1}$ et $(M^{-1})_{i,j} = \delta_{j, N-i+1} \varepsilon_i$. On a alors

$$\begin{aligned} \left(M \left(\text{Adj}^{(p)} A \right) M^{-1} \right)_{i,j} &= \sum_{k,l} M_{i,k} \left(\text{Adj}^{(p)} A \right)_{k,l} (M^{-1})_{l,j} \\ &= \varepsilon_{N-i+1} \varepsilon_{N-j+1} \left(\text{Adj}^{(p)} A \right)_{N-i+1, N-j+1} \\ &= |A|_{\underline{(N-i+1)'}, \underline{(N-j+1)'}}^{(n-p)}, \end{aligned}$$

la dernière égalité étant obtenue à partir de (2.13) en utilisant le fait que $\varepsilon_{N-i+1} \varepsilon_{N-j+1} \varepsilon_{\underline{N-i+1}, \underline{N-j+1}} = 1$. On conclut en se rappelant que $\underline{(N-i+1)'}$ et $\underline{(N-j+1)'}$ (dans le contexte des $(n-p)$ -uplets de $\{1, \dots, n\}$).

La deuxième assertion du théorème s'obtient aisément en se rappelant que $A^{-1} = (\det A)^{-1} \times \text{Adj}^{(1)} A$ pour la première formule, et en remplaçant A par A^{-1} pour la deuxième. \square

Remarque(s) 2.1.15.

La transposition commute avec la composée p -ème d'une matrice et $M^{-1} = {}^t M$.

2.2 Coordonnées de Grassmann et coordonnées de Grassmann duales

On peut naturellement adopter un point de vue projectif pour traiter des coordonnées de Grassmann (c'est le point de vue de Hodge et Pedoe dans [5], mais ce n'est pas celui que l'on adoptera ici). Remarquons également qu'un certain nombre de résultats de ce paragraphe restent vrais en considérant un module libre de rang fini sur un anneau commutatif unitaire quelconque. Nous nous restreignons au cas d'un corps \mathbb{K} puisque c'est le cadre qui nous intéresse. Les coordonnées de Grassmann permettent d'associer à un sous-espace de \mathbb{K}^n de dimension p une droite de $\mathbb{K}^{\binom{n}{p}}$. C'est grâce à cette association que Schmidt définit la hauteur d'un sous-espace de dimension $p > 1$ dans [17]. En fin de section nous verrons que l'on peut définir les coordonnées de Grassmann de manière intrinsèque pour un sous-espace d'un espace vectoriel E de dimension finie à l'aide de l'algèbre extérieure p -ème : si x_1, \dots, x_p sont des vecteurs libres de E , le vecteur $x_1 \wedge \dots \wedge x_p$ est un ensemble de coordonnées de Grassmann de l'espace vectoriel qu'ils engendrent. Si cette définition est élégante et offre une notation agréable, elle reste cependant peu pratique pour effectuer des calculs explicites, c'est pourquoi nous commencerons par en donner une différente.

Le résultat principal de cette sous-partie est le théorème 2.2.6 qui relie les coordonnées de Grassmann et les les coordonnées de Grassmann duales d'un sous-espace.

Définition 2.2.1. On note φ la forme bilinéaire symétrique non dégénérée sur \mathbb{K}^n définie par

$$\varphi(X, Y) = \sum_{i=1}^n x_i y_i.$$

Si G est un sous-espace vectoriel de \mathbb{K}^n , on notera $G^{\perp, \varphi}$ son orthogonal pour φ (sous-espace vectoriel de codimension $\dim G$).

Soit F un sous-espace de \mathbb{K}^n de dimension p . On note $N := \binom{n}{p}$. On choisit X_1, \dots, X_p une base de F et on forme $A \in \mathcal{M}_{n,p}$ la matrice dont les colonnes sont les X_i .

Définition 2.2.2. On appelle ensemble de coordonnées de Grassmann de F un vecteur de la forme $A^{(p)} (\in \mathcal{M}_{N,1})$ avec A comme ci-dessus.

La proposition suivante établit les premières propriétés des coordonnées de Grassmann. On trouvera une preuve du point (i) écrite d'une manière différente dans [5] (p.289 – 290).

Proposition 2.2.3. Soit X (resp. X') un ensemble de coordonnées de Grassmann d'un espace vectoriel F (resp. F') de dimension p , et u un automorphisme de \mathbb{K}^n de matrice U . Alors

- (i) $F = F' \iff (\exists \alpha \neq 0 \text{ tel que } X = \alpha X')$.
- (ii) $U^{(p)} X$ est un ensemble de coordonnées de Grassmann du sous espace $u(F)$.

Preuve (ii) provient directement de (2.14) et de la définition d'un ensemble de coordonnées de Grassmann.

Pour (i)(\Rightarrow). Remarquons que si $X = A^{(p)}$ et que $Y = B^{(p)}$ sont deux ensembles de coordonnées de Grassmann de F , avec $A, B \in \mathcal{M}_{n,p}$, alors il existe P , matrice de taille p inversible telle que $A = BP$. On en déduit par (2.14) et le fait que $P^{(p)} = \det P$ que $X = \alpha Y$ avec $\alpha = \det P \neq 0$.

Pour (i)(\Leftarrow). Par (2.14) et en composant avec un automorphisme idoine u , on peut se ramener au cas où F est l'espace vectoriel engendré par les p premiers vecteurs de la base canonique de \mathbb{K}^n et où $X = (1, 0, \dots, 0)$ provient de la matrice dont les colonnes sont les p premiers vecteurs de la base canonique. Le résultat devient alors plus immédiat : si X' provient de A' , alors tout mineur formé à partir de $p - 1$ lignes de A' choisies parmi les p premières et d'une i -ème ligne ($i > p$) est nul. Donc nécessairement cette i -ème ligne est nulle et on a bien $F' = F$. \square

Définition 2.2.4. Si F est un espace vectoriel de dimension p de \mathbb{K}^n , on note F^* la droite de \mathbb{K}^N engendrée par un ensemble de coordonnées de Grassmann de F .

Par ce qui précède, cela ne dépend pas du choix de l'ensemble des coordonnées de Grassmann (et cela caractérise F parmi les sous-espaces de dimension p).

Définition 2.2.5. Un ensemble de coordonnées de Grassmann duales de F est un ensemble de coordonnées de Grassmann de $F^{\perp, \varphi}$.

Remarquons que c'est un vecteur de \mathbb{K}^N puisque $\dim F^{\perp, \varphi} = n - p$. Un lien très fort existe entre les coordonnées de Grassmann et les coordonnées de Grassmann duales. Si on note τ l'automorphisme de \mathbb{K}^n associé à la matrice M (définie par (2.19) - de manière explicite on a $\tau(\eta_1, \dots, \eta_N) = (\varepsilon_N \eta_N, \dots, \varepsilon_1 \eta_1)$) - alors on a le

Théorème 2.2.6.

$$(F^{\perp, \varphi})^* = \tau(F^*). \quad (2.23)$$

Preuve Les arguments qui suivent sont ceux utilisés par Hodge et Pedoe dans [5] mais écrits dans un langage un peu différent (plus matriciel et explicite).

Soit X un ensemble de coordonnées de Grassmann de F qui provient de $A \in \mathcal{M}_{n,p}$. Soit $B \in \mathcal{M}_{n,n-p}$ dont les colonnes sont une base de $F^{\perp, \varphi}$. Par définition $Y := B^{(n-p)}$ est un ensemble de coordonnées de Grassmann duales de F . De plus, A et B sont reliés par l'équation

$${}^t B A = 0.$$

Maintenant, si u est un automorphisme de \mathbb{K}^n de matrice U , si on change F en $u(F)$, alors A est changé en UA et B est changé en ${}^t U^{-1} B$. Par (2.14), on en déduit que X est changé en $U^{(p)} X$, et donc MX est changé en $MU^{(p)} M^{-1} (MX)$. Par ailleurs Y est changé en $({}^t U^{-1})^{(n-p)} Y$. Or, par (2.22), $({}^t U^{-1})^{(n-p)} = (\det U)^{-1} MU^{(p)} M^{-1}$. Finalement, $\tau(u(F)^*)$ et $(u(F)^{\perp, \varphi})^*$ sont les images respectives de $\tau(F^*)$ et $(F^{\perp, \varphi})^*$ par un même automorphisme (l'automorphisme de matrice $MU^{(p)} M^{-1}$ dans la base canonique), donc il suffit de montrer le théorème pour $u(F)$.

On se ramène ainsi au cas où les colonnes de A sont les p premiers vecteurs de la base canonique et celles de B les $(n - p)$ derniers vecteurs de la base canonique.

Dans ce cas-là, $X = (1, 0, \dots, 0)$ et $Y = (0, \dots, 0, 1)$ et donc MX et Y sont proportionnels, d'où (2.23). \square

On peut adopter un autre point de vue sur les coordonnées de Grassmann - qui a l'avantage de permettre une définition intrinsèque des coordonnées de Grassmann d'un sous-espace F d'un espace vectoriel E de dimension finie n - en les reliant à

l'algèbre extérieure. En contrepartie, il faudra passer par une identification dès qu'on voudra les manipuler de manière explicite. Pour cela, il suffit de remarquer qu'on a un isomorphisme canonique entre $\bigwedge^p(\mathbb{K}^n)$ et \mathbb{K}^N : en notant e_j le j -ème vecteur de la base canonique de \mathbb{K}^n et E_i le i -ème de la base canonique de \mathbb{K}^N , on sait que $(e_{i_1} \wedge \cdots \wedge e_{i_p})_{\underline{i}}$ constitue une base de $\bigwedge^p(\mathbb{K}^n)$ (où \underline{i} désigne toujours le i -ème p -uplet $(i_1 < \cdots < i_p)$ lorsqu'on a ordonné les p -uplets par l'ordre lexicographique) ;

$$e_{i_1} \wedge \cdots \wedge e_{i_p} \longrightarrow E_i \quad (2.24)$$

est alors l'isomorphisme recherché. C'est toujours cette identification qu'on considérera dans la suite, sauf mention explicite du contraire.

Remarquons de plus que si $x_1, \dots, x_p \in \mathbb{K}^n$, et si on définit A comme étant la matrice dont la i -ème colonne est x_i , alors avec l'identification précédente on a :

$$x_1 \wedge \cdots \wedge x_p = A^{(p)}.$$

Ainsi, si les x_i sont libres et que F est l'espace vectoriel qu'ils engendrent, alors $x_1 \wedge \cdots \wedge x_p$ est un ensemble de coordonnées de Grassmann de F .

Nous allons maintenant introduire une généralisation des coordonnées de Grassmann : les espaces composés p -ème. Les résultats qui suivent sont personnels et seront utilisés dans certaines démonstrations du paragraphe 2.3.

Définition 2.2.7. Soit F un sous-espace de \mathbb{K}^n de dimension k . On pose $N := \binom{n}{p}$. On définit $F^{(p)}$, le sous-espace composé p -ème de F comme étant l'espace vectoriel de \mathbb{K}^N engendré par les $x_1 \wedge \cdots \wedge x_p, x_1, \dots, x_p \in F$.

Proposition 2.2.8. Si $k \geq p$ et A est une matrice de taille $n \times k$ dont les colonnes forment une base de F , alors les colonnes de $A^{(p)}$ forment une base de $F^{(p)}$. En particulier, $\dim F^{(p)} = \binom{k}{p}$.

Preuve Soient (X_1, \dots, X_k) et (X'_1, \dots, X'_k) deux bases de F . On forme A et A' dont les colonnes sont respectivement les X_i et les X'_i . Alors il existe une matrice U inversible telle que $A' = AU$. En particulier, par (2.14) on a $(A')^{(p)} = A^{(p)}U^{(p)}$ avec $U^{(p)}$ inversible. On en déduit que les colonnes de $(A')^{(p)}$ et $A^{(p)}$ forment deux bases d'un même espace vectoriel qu'on note provisoirement G .

Par ailleurs, si $\underline{j} = (j_1 < \cdots < j_p)$, alors la j -ème colonne de $A^{(p)}$ est égale à $X_{j_1} \wedge \cdots \wedge X_{j_p}$. Donc $G \subset F^{(p)}$. Réciproquement, si $X_1, \dots, X_p \in F$ sont libres, alors on peut les compléter en X_1, \dots, X_k une base de F et si B est la matrice dont la j -ème colonne est X_j , alors la première colonne de $B^{(p)}$ sera égale à $X_1 \wedge \cdots \wedge X_p$. Donc $F^{(p)} = G$.

Pour la dimension, on remarque que la dimension de $F^{(p)}$ ne dépend que de celle de F puisque si u est un automorphisme de \mathbb{K}^n de matrice U , alors $u(F)^{(p)} = U^{(p)}F^{(p)}$. On est ramené au cas où $X_i = e_i$. Le résultat est alors immédiat. \square

Corollaire 2.2.9. Soit u un endomorphisme de \mathbb{K}^n dont on note $u^{(p)}$ l'endomorphisme composé p -ème de \mathbb{K}^N associé. On a

$$\text{Im} \left(u^{(p)} \right) = (\text{Im } u)^{(p)}. \quad (2.25)$$

Preuve Si le rang de u est $< p$, les deux membres de l'égalité sont nuls. Si le rang de u est plus grand que p , si on note U sa matrice dans la base canonique, alors on sait que $\text{Im } u$ est engendré par les colonnes de U , et de même, $\text{Im } u^{(p)}$ est engendré par les colonnes de $U^{(p)}$. La proposition précédente permet de conclure.

□

Proposition 2.2.10. *Soit F et F' deux sous-espaces de E de dimension $\geq p$. Alors*

$$F \subset F' \Leftrightarrow F^{(p)} \subset (F')^{(p)}.$$

En particulier si (x_1, \dots, x_p) est libre,

$$x_1 \wedge \dots \wedge x_p \in F^{(p)} \Leftrightarrow \text{pour tout } i \ x_i \in F. \quad (2.26)$$

Preuve \Rightarrow est trivial.

\Leftarrow Comme $F^{(p)}$ est engendré par les $x_1 \wedge \dots \wedge x_p$ ($x_i \in F$), on est ramené au cas où $\dim F = p$ et $F = \text{Vect}(x_1, \dots, x_p)$, l'espace vectoriel engendré par x_1, \dots, x_p . On pose $k := \dim F'$. Par ailleurs, quitte à composer par un automorphisme idoine, on peut supposer que $F' = \text{Vect}(e_1, \dots, e_k)$. Le sous-espace vectoriel $(F')^{(p)}$ a alors pour base les $e_{i_1} \wedge \dots \wedge e_{i_p}$ ($i_1 < \dots < i_p \leq k$).

Soit $j_0 > k$ et $\underline{j}' = \{j_1 < \dots < j_{p-1}\} \cup \{j_0\}$ avec $j_{p-1} \leq k$. Alors le j' -ème coefficient de $x_1 \wedge \dots \wedge x_p$ (i.e. la composante selon $E_{j'}$ avec les notations de (2.24)) est nul car $E_{j'} = e_{j'_1} \wedge \dots \wedge e_{j'_p} \notin (F')^{(p)}$.

Si on note A la matrice dont les colonnes sont les x_i , cela signifie que les p lignes j_0, \dots, j_{p-1} sont liées. Or $x_1 \wedge \dots \wedge x_p \in (F')^{(p)}$ est non nul, donc il existe $j_1 < \dots < j_p \leq k$ tels que $|A|_{\underline{j}, \underline{j}_1}^{(p)} \neq 0$, i.e que les lignes j_1, \dots, j_p de A soient libres. La ligne j_0 sera alors combinaison linéaire de n'importe quelles $j-1$ lignes choisies parmi j_1, \dots, j_p , donc ne peut être que nulle. On en déduit immédiatement que pour tout $j_0 > k$, pour tout l , la composante selon e_{j_0} de x_l est nulle et on a donc bien $x_1, \dots, x_p \in \text{Vect}(e_1, \dots, e_k) = F'$.

□

2.3 Déterminants généralisés

La majeure partie de cette section est tirée de [17], à l'exception des preuves que Schmidt laisse au lecteur dans son article et qui sont ici personnelles.

Dans cette section, E désigne un espace euclidien ou hermitien de dimension n dont on note $\langle \cdot, \cdot \rangle$ le produit scalaire, $\mathbb{K} = \mathbb{R}$ ou \mathbb{C} , et m un entier strictement positif.

Si X_1, \dots, X_m sont m vecteurs de E , on note $\text{Vect}(X_1, \dots, X_m)$ le sous-espace vectoriel qu'ils engendrent. De même, si A, B, \dots sont des sous-espaces vectoriels de E , on note encore $\text{Vect}(A, B, \dots)$ le sous-espace qu'ils engendrent. Etant donnés X_1, \dots, X_m , on peut former la matrice

$$\left(\langle X_i, X_j \rangle \right)_{1 \leq i, j \leq m}.$$

Si (Z_1, \dots, Z_l) est une famille orthonormée de E telle que $\text{Vect}(Z_1, \dots, Z_l)$ contienne $\text{Vect}(X_1, \dots, X_m)$, en notant M la transposée de la matrice des X_i dans la base $(Z_i)_i$ (i.e. si on écrit $X_i = \sum c_{ij} Z_j$ alors $M = (c_{ij})_{i,j}$) et M^* sa transposée (resp. sa transconjugée) si E est euclidien (resp. hermitien), alors on a la formule utile (que n'utilise pas explicitement Schmidt) :

$$\left(\langle X_i, X_j \rangle \right)_{\substack{1 \leq i \leq m \\ 1 \leq j \leq m}} = MM^*. \quad (2.27)$$

En particulier, en prenant $(Z_i)_i$ de cardinal m , il est immédiat que la matrice précédente est de déterminant ≥ 0 avec égalité si et seulement si la famille (X_1, \dots, X_m) est liée. On peut alors définir

Définition 2.3.1. On note

$$D(X_1, \dots, X_m) = \sqrt{\det \left(\langle X_i, X_j \rangle \right)_{i,j}}, \quad (2.28)$$

appelé déterminant généralisé de la famille (X_1, \dots, X_m) .

Ce déterminant généralisé D possède les propriétés suivantes :

Proposition 2.3.2.

- (i) $D(X_1, \dots, X_m) \geq 0$ (et $D(X_1, \dots, X_m) = 0 \Leftrightarrow (X_1, \dots, X_m)$ liée)
- (ii) $D(X_{\sigma(1)}, \dots, X_{\sigma(m)}) = D(X_1, \dots, X_m)$ pour tout $\sigma \in \mathfrak{S}_n$.
- (iii) $D(X_1, \dots, tX_k, \dots, X_m) = |t|D(X_1, \dots, X_k, \dots, X_m)$.
- (iv) $D(X_1, \dots, X_k, \dots, X_l + tX_k, \dots, X_m) = D(X_1, \dots, X_k, \dots, X_l, \dots, X_m)$.
- (v) $D(\tau(X_1), \dots, \tau(X_m)) = D(X_1, \dots, X_m)$ si τ est une transformation unitaire.

Preuve (v) vient de (2.28) et de la définition d'une transformation unitaire (qui préserve par définition le produit scalaire). Les autres assertions découlent immédiatement de (2.27). □

Remarque(s) 2.3.3. On fixe $\mathfrak{B} = (Z_1, \dots, Z_n)$ une base orthonormée de E et on note M la transposée de la matrice des $(X_i)_i$ dans la base \mathfrak{B} . Soit τ un automorphisme de E dont on note P la matrice dans \mathfrak{B} . Enfin on note M_τ la transposée de la matrice des $(\tau(X_i))_i$ dans la base \mathfrak{B} . Alors on a la relation $M_\tau = M {}^t P$, où ${}^t P$ désigne la transposée de P . Notant alors \bar{P} la conjuguée (au sens complexe) de P (remarque : $\bar{P} = P$ dans le cas euclidien), on en déduit :

$$M_\tau M_\tau^* = M ({}^t P \bar{P}) M^*. \quad (2.29)$$

En particulier cela prouve de nouveau (v).

Rappelons-nous la définition (2.8). La proposition qui suit est une traduction du lemme 1 section 2 de [17] dans un langage différent ; on y exploite notamment les notations introduites dans les paragraphes précédents, en particulier le point de vue produit extérieur qui fournit une définition élégante du déterminant généralisé et qui rend plus apparentes certaines de ses propriétés.

Proposition 2.3.4. Soit $X_1, \dots, X_m \in E$. On note M la matrice dont la i -ème ligne est X_i écrit dans une base \mathfrak{B} orthonormée de E . On pose $N := \binom{n}{m}$. Alors

$$D^2(X_1, \dots, X_m) = \sum_{i=1}^N \left| |M|_{\underline{1}, i}^{(m)} \right|^2.$$

Autrement dit :

$$D(X_1, \dots, X_m) = \|X_1 \wedge \dots \wedge X_m\|_2, \quad (2.30)$$

où $\|\cdot\|_2$ désigne la norme 2 associée au produit scalaire canonique sur \mathbb{K}^N (dans cette dernière formule, on a implicitement choisi une base \mathfrak{B} orthonormée pour identifier E à \mathbb{K}^n et $\bigwedge^m(E)$ à \mathbb{R}^N , le résultat ne dépendant pas du choix de \mathfrak{B}).

Preuve Par (2.27) et par définition du déterminant généralisé, on a :

$$\begin{aligned} D^2(X_1, \dots, X_m) &= \det(MM^*) = \sum_{i=1}^N |M|_{\underline{1}, i}^{(m)} |M^*|_{i, \underline{1}}^{(m)} \quad \text{par (2.7)} \\ &= \sum_{i=1}^N |M|_{\underline{1}, i}^{(m)} \overline{|M|_{\underline{1}, i}^{(m)}}, \end{aligned}$$

d'où le résultat. □

Proposition 2.3.5. *Soit X, Y et X_2, \dots, X_m des vecteurs de E . Alors*

$$D(X + Y, X_2, \dots, X_m) \leq D(X, X_2, \dots, X_m) + D(Y, X_2, \dots, X_m),$$

avec égalité si et seulement si (X, Y, X_2, \dots, X_m) est liée.

Preuve La première assertion provient de l'inégalité triangulaire appliquée à

$$(X + Y) \wedge X_2 \wedge \dots \wedge X_m = X \wedge X_2 \wedge \dots \wedge X_m + Y \wedge X_2 \wedge \dots \wedge X_m,$$

en utilisant (2.30).

Si (X, X_2, \dots, X_m) est liée, alors $X \wedge X_2 \wedge \dots \wedge X_m = 0$ et l'égalité est triviale. De même si (Y, X_2, \dots, X_m) est liée.

Dans le cas où (X, X_2, \dots, X_m) et (Y, X_2, \dots, X_m) sont libres, alors il y a égalité si et seulement si il existe un $\alpha \neq 0$ tel que $X \wedge X_2 \wedge \dots \wedge X_m = \alpha Y \wedge X_2 \wedge \dots \wedge X_m$, donc par la proposition 2.2.3 (i), $\text{Vect}(X, X_2, \dots, X_m) = \text{Vect}(Y, X_2, \dots, X_m)$, donc (X, Y, X_2, \dots, X_m) est liée. □

Proposition 2.3.6 (Inégalité de Hadamard généralisée).

On a la formule

$$D(X_1, \dots, X_m, Y_1, \dots, Y_k) \leq D(X_1, \dots, X_m) D(Y_1, \dots, Y_k),$$

avec égalité si et seulement si l'un des deux facteurs du membre de droite est nul (i.e. (X_1, \dots, X_m) ou (Y_1, \dots, Y_k) est liée) ou si les sous-espaces $\text{Vect}(X_1, \dots, X_m)$ et $\text{Vect}(Y_1, \dots, Y_k)$ sont orthogonaux.

Preuve Si (X_1, \dots, X_m) ou (Y_1, \dots, Y_k) est liée, l'égalité est immédiate ; on supposera ces deux familles libres. On note π la projection orthogonale sur l'orthogonal de l'espace vectoriel engendré par les X_i . Alors d'après le point (iv) de la proposition 2.3.2, on a

$$\begin{aligned} D(X_1, \dots, X_m, Y_1, \dots, Y_k) &= D(X_1, \dots, X_m, \pi(Y_1), \dots, \pi(Y_k)) \\ &= D(X_1, \dots, X_m) D(\pi(Y_1), \dots, \pi(Y_k)), \end{aligned}$$

la dernière égalité étant obtenue à partir de (2.28) (la matrice des produits scalaires de la famille considérée est alors une matrice diagonale par blocs).

Remarquons maintenant que $\pi^{(p)}$ est également une projection orthogonale (cela apparaît de manière immédiate si on écrit π dans une base orthonormée adaptée). En particulier,

$$\begin{aligned}
D(\pi(Y_1), \dots, \pi(Y_k)) &= \|\pi(Y_1) \wedge \dots \wedge \pi(Y_k)\|_2 \\
&= \|\pi^{(p)}(Y_1 \wedge \dots \wedge Y_k)\|_2 \\
&\leq \|Y_1 \wedge \dots \wedge Y_k\|_2,
\end{aligned}$$

avec égalité si et seulement si $Y_1 \wedge \dots \wedge Y_k \in \text{Im } \pi^{(p)} = (\text{Im } \pi)^{(p)}$ (par (2.25)). Finalement, par (2.26), on a égalité si et seulement si $Y_i \in \text{Im } \pi$ pour tout i (en particulier π est alors de rang $\geq p$).

□

2.4 Géométrie des nombres - Corps convexe composé

Rappelons qu'un corps convexe symétrique de \mathbb{R}^n est un ensemble K compact, convexe, symétrique par rapport à l'origine (i.e. $x \in K \Rightarrow -x \in K$) et d'intérieur non vide. La jauge j_K d'un corps convexe est la fonction $\mathbb{R}^n \rightarrow \mathbb{R}^+$ définie par $j_K(x) = \inf\{\lambda > 0 \mid \text{tel que } x \in \lambda K\}$.

Définition 2.4.1. On appelle réseau de \mathbb{R}^n tout sous-groupe discret Λ de \mathbb{R}^n . La dimension du réseau est alors la dimension m de l'espace vectoriel qu'il engendre. Son covolume - appelé également déterminant - que l'on note $\text{covol}(\Lambda)$ ou $d(\Lambda)$ est le nombre positif

$$d(\Lambda) := D(X_1, \dots, X_m) = \|X_1 \wedge \dots \wedge X_m\|_2, \quad (2.31)$$

où (X_1, \dots, X_m) est n'importe quelle base de Λ . Par convention, si $\Lambda = \{0\}$, on pose $d(\Lambda) = 1$.

Remarques :

1) Le déterminant est bien défini puisque si (X_1, \dots, X_m) et (Y_1, \dots, Y_m) sont deux bases de Λ , que M_X (resp. M_Y) désigne la matrice dont les colonnes sont les X_i (resp. les Y_i), alors il existe $A \in \text{GL}_m(\mathbb{Z})$ (donc de déterminant ± 1) telle que $M_X = M_Y A$. En particulier,

$$X_1 \wedge \dots \wedge X_m = M_X^{(p)} = (\det A) M_Y^{(p)} = \pm Y_1 \wedge \dots \wedge Y_m.$$

2) Si $m = n$, le covolume est alors le volume pour la mesure de Lebesgue de n'importe quel parallélogramme fondamental.

3) Habituellement un réseau est plutôt défini comme étant un sous-groupe discret de rang maximal (i.e. de dimension égale à la dimension de l'espace).

Définition 2.4.2. Etant donné un réseau Λ de dimension m de \mathbb{R}^n et un corps convexe symétrique K , on note $\lambda_i(K, \Lambda)$ le i -ème minimum de Λ pour K . Par définition, c'est

$$\inf\{r > 0 \mid (rK) \cap \Lambda \text{ contient } i \text{ vecteurs linéairement indépendants}\}.$$

Les réels $\lambda_1(K, \Lambda) \leq \dots \leq \lambda_m(K, \Lambda)$ sont appelés les minima successifs de Λ pour K .

Rappelons enfin qu'on identifie l'algèbre extérieure p -ème $\bigwedge^p \mathbb{R}^n$ à \mathbb{R}^N via (2.24) (où on a posé $N := \binom{n}{p}$).

Mahler introduit la notion de corps convexe composé dans [11]; son article fut repris plus tard par Lekkerkerker et c'est cette référence que nous utiliserons. La majorité de ce paragraphe reprend donc le chapitre 2 section 15 de [10]. On fixe $1 \leq p \leq n$ deux

entiers et on pose $N := \binom{n}{p}$. Soit K_1, \dots, K_p des corps convexes symétriques de \mathbb{R}^n . De même qu'à partir des coordonnées de Grassmann on peut associer à un sous-espace de \mathbb{R}^n un sous-espace de \mathbb{R}^N , de même on peut associer à un corps convexe symétrique et un réseau de \mathbb{R}^n un corps convexe symétrique et un réseau de \mathbb{R}^N respectivement. Les théorèmes 2.4.6 et 2.4.9 sont les deux résultats principaux de ce paragraphe et permettent de donner des expressions (à une constante multiplicative près) du volume du corps convexe composé en fonction de celui du corps convexe initial, et des minima successifs du corps convexe composé pour le réseau composé en fonctions de ceux du corps convexe et du réseau initiaux. Ce qui est remarquable c'est que les constantes multiplicatives mises en jeu ne dépendent ni du corps convexe, ni du réseau considéré, ce qui permet d'importantes applications en géométrie des nombres, notamment dans [19] où Schmidt et Summerer étudient les minima successifs d'une famille de corps convexes symétriques composés dépendant d'un paramètre q .

Définition 2.4.3. On pose $\Sigma := \{x_1 \wedge \dots \wedge x_p \mid x_1 \in K_1, \dots, x_p \in K_p\} \subset \mathbb{R}^N$. On note $\mathcal{K} = \langle K_1, \dots, K_p \rangle$ - et on appelle corps convexe composé des K_i - l'enveloppe convexe de Σ .

Proposition 2.4.4. \mathcal{K} est un corps convexe symétrique de \mathbb{R}^N .

Preuve Par définition \mathcal{K} est convexe. De plus, Σ est l'image du compact $K_1 \times \dots \times K_p$ de \mathbb{R}^{np} par l'application continue (car polynomiale en ses coordonnées) $(x_1, \dots, x_p) \mapsto x_1 \wedge \dots \wedge x_p$, donc compact. Il en va de même de \mathcal{K} . Reste à voir que \mathcal{K} est symétrique et d'intérieur non vide.

K_1 est symétrique donc, si $x_1 \in K_1$, on a $-x_1 \in K_1$. On en déduit que pour tout $(x_1, \dots, x_p) \in K_1 \times \dots \times K_p$ on a $-x_1 \wedge \dots \wedge x_p \in \Sigma$. L'ensemble Σ est donc symétrique, et par conséquent \mathcal{K} l'est aussi.

On note $(e_j)_{1 \leq j \leq n}$ et $(E_i)_{1 \leq i \leq N}$ les bases canoniques respectives de \mathbb{R}^n et \mathbb{R}^N . Avec les notations de la section précédente, si $i = (i_1 < \dots < i_p)$ alors $e_{i_1} \wedge \dots \wedge e_{i_p} = E_i$. Or il existe $\delta > 0$ tel que $\delta e_i \in K_j$ pour tout i et pour tout j . On en déduit que Σ contient $\delta^p E_i$ pour tout i , et donc \mathcal{K} est d'intérieur non vide. □

Définition 2.4.5. Dans le cas où $K_1 = \dots = K_p := K$, on écrit $\mathcal{K} = K^{(p)}$, appelé composé p -ème de K .

Théorème 2.4.6. Soit K un corps convexe symétrique de \mathbb{R}^n . On note \mathcal{K} son composé p -ème ($0 < p < n$). On pose $P := \binom{n-1}{p-1}$. Alors il existe deux constantes $\alpha_1, \alpha_2 > 0$ ne dépendant que de n et p telles que

$$\alpha_1 \leq \text{Vol}(\mathcal{K}) \text{Vol}(K)^{-P} \leq \alpha_2,$$

où $\text{Vol}(\mathcal{K})$ et $\text{Vol}(K)$ désignent les volumes respectifs de \mathcal{K} et K .

Preuve On note S_n la sphère unité de \mathbb{R}^n . On va se ramener au cas où K est un ellipsoïde et montrer que dans ce cas la quantité $\text{Vol}(\mathcal{K}) \text{Vol}(K)^{-P}$ ne dépend que de n et p . Par le théorème de John (théorème 8 du chapitre I de [10]), il existe un ellipsoïde symétrique E tel que

$$E \subset K \subset n^{1/2} E.$$

Par définition d'un ellipsoïde il existe donc $A = (a_{ij})_{i,j}$ inversible telle que $E = AS_n$. On pose

$$\mathcal{K}_{n,p} := S_n^{(p)},$$

et

$$\mathcal{E} := E^{(p)} = (AS_n)^{(p)}.$$

D'après les résultats sur les matrices composées, il est immédiat que

$$\mathcal{E} = A^{(p)}\mathcal{K}_{n,p}.$$

Par ailleurs, d'après (2.16) on a $\det A^{(p)} = (\det A)^P$. On en déduit que

$$\text{Vol}(\mathcal{E})\text{Vol}(E)^{-P} = \text{Vol}(\mathcal{K}_{n,p})\text{Vol}(S_n)^{-P} =: \alpha,$$

qui ne dépend que de n et p .

Enfin

$$E \subset K \subset n^{1/2}E \Rightarrow \mathcal{E} \subset \mathcal{K} \subset n^{p/2}\mathcal{E},$$

et en posant $\alpha_1 := n^{-nP/2}\alpha$ et $\alpha_2 = n^{N/2}\alpha$ on a bien le résultat annoncé. \square

Lekkerkerker ne traite dans son livre que le cas des réseaux classiques (i.e. de rang maximal). La définition et la propriété qui suivent généralisent au cas d'un réseau de dimension quelconque certains résultats de [10].

Définition 2.4.7. Soit Λ un réseau de \mathbb{R}^n de dimension $k \geq p$. On définit $\Lambda^{(p)}$ comme étant le sous-groupe de \mathbb{R}^N engendré par les $\lambda_1 \wedge \cdots \wedge \lambda_p$, $\lambda_i \in \Lambda$ ($i = 1, \dots, p$).

Proposition 2.4.8. $\Lambda^{(p)}$ est un réseau de \mathbb{R}^N de dimension $K := \binom{k}{p}$, et si $(u_i)_{1 \leq i \leq k}$ est une base de Λ , alors $(u_{i_1} \wedge \cdots \wedge u_{i_p})_{i_1 < \cdots < i_p \leq k}$ est une base de $\Lambda^{(p)}$.

De plus, si $\Lambda = A(e_1\mathbb{Z} \oplus \cdots \oplus e_k\mathbb{Z})$, alors il existe i_1, \dots, i_K tels que $\Lambda^{(p)} = A^{(p)}(E_{i_1}\mathbb{Z} \oplus \cdots \oplus E_{i_K}\mathbb{Z})$. En particulier,

$$\text{covol}(\Lambda^{(p)}) = \text{covol}(\Lambda)^P.$$

(Rappelons que $P := \binom{n-1}{p-1}$).

Preuve On pose $\Lambda_0 := e_1\mathbb{Z} \oplus \cdots \oplus e_k\mathbb{Z}$.

Si $\Lambda = A\Lambda_0$, par définition de $\Lambda^{(p)}$ et (2.14), $\Lambda^{(p)}$ est l'image de $\Lambda_0^{(p)}$ par $A^{(p)}$. Or $\Lambda_0^{(p)}$ est un réseau de base $(e_{i_1} \wedge \cdots \wedge e_{i_p})_{i_1 < \cdots < i_p \leq k}$ donc $\Lambda^{(p)}$ est bien un réseau de \mathbb{R}^N et la propriété sur les bases (et la dimension) en découle immédiatement. L'assertion sur le covolume provient directement de (2.16), de la définition du covolume (2.31), et en remarquant qu'un réseau dont une base est constituée de vecteurs de la base canonique est de déterminant 1 (on a donc $d(\Lambda_0) = d(\Lambda_0^{(p)}) = 1$). \square

Théorème 2.4.9. Soit \mathcal{K} le composé p -ème d'un corps convexe symétrique K et soit Λ un réseau de \mathbb{R}^n de dimension n . On note $\lambda_i := \lambda_i(K, \Lambda)$ et $\mu_j := \lambda_j(\mathcal{K}, \Lambda^{(p)})$ les minima successifs de K et \mathcal{K} pour Λ et $\Lambda^{(p)}$ respectivement. Pour $\underline{i} = (i_1 < \cdots < i_p)$ on note $m_i := \lambda_{i_1} \cdots \lambda_{i_p}$. On note enfin $M_1 \leq \cdots \leq M_N$ les m_i ordonnés dans l'ordre croissant.

Alors il existe $\alpha_3 > 0$ ne dépendant que de n et p telle que

$$\alpha_3 M_j \leq \mu_j \leq M_j \quad (j = 1, \dots, N).$$

Preuve La preuve est issue de [10].

Pour commencer on se ramène au cas où $\Lambda = \mathbb{Z}^n$ (quitte à changer K en $A^{-1}K$).

Si $\Lambda = AZ^n$, puisque λK contient i vecteurs indépendants de Λ si et seulement si $\lambda A^{-1}K$ contient i vecteurs indépendants de \mathbb{Z}^n , on a $\lambda_i(K, \Lambda) = \lambda_i(A^{-1}K, \mathbb{Z}^n)$. De plus

$$\lambda_j(\mathcal{K}, \Lambda^{(p)}) = \lambda_j((A^{(p)})^{-1}\mathcal{K}, \mathbb{Z}^N) = \lambda_j((A^{-1}K)^{(p)}, \mathbb{Z}^N).$$

On suppose désormais que $\Lambda = \mathbb{Z}^n$.

Montrons l'inégalité de droite. On note f la jauge de K et u_1, \dots, u_n n points linéairement indépendants de \mathbb{Z}^n tels que $f(u_i) = \lambda_i$.

Si $\underline{i} = (i_1 < \dots < i_p)$, on pose $U_i = u_{i_1} \wedge \dots \wedge u_{i_p} \in \mathbb{Z}^N$ ($i = 1, \dots, N$).

Les U_i sont linéairement indépendants et $U_i \in m_i \mathcal{K}$ (puisque $U_i/m_i = (u_{i_1}/\lambda_{i_1}) \wedge \dots \wedge (u_{i_p}/\lambda_{i_p}) \in \mathcal{K}$).

On note φ la jauge de \mathcal{K} et on réarrange les U_i en une suite V_1, \dots, V_N telle que $\varphi(V_1) \leq \dots \leq \varphi(V_N)$. Avec ce qui précède on a $\varphi(U_i) \leq m_i$ et pour un r donné et pour tous i_1, \dots, i_r deux à deux distincts, $\varphi(V_r)$ est inférieur ou égal au maximum des r valeurs $\varphi(U_{i_1}), \dots, \varphi(U_{i_r})$. En choisissant j_1, \dots, j_r tel que $m_{j_1} = M_1, \dots, m_{j_r} = M_r$, on a

$$\varphi(V_1) \leq \dots \leq \varphi(V_r) \leq \max(\varphi(U_{j_1}), \dots, \varphi(U_{j_r})) \leq \max(m_{j_1}, \dots, m_{j_r}) = M_r.$$

Puisque les V_i sont linéairement indépendants cela prouve bien que

$$\mu_r \leq M_r.$$

Montrons l'inégalité de gauche. Par le théorème de Minkowski (cf [10] chapitre 2), on a

$$\frac{2^n}{n!} \leq \lambda_1 \dots \lambda_n \text{Vol}(K) \leq 2^n \quad \text{et} \quad \frac{2^N}{N!} \leq \mu_1 \dots \mu_N \text{Vol}(\mathcal{K}) \leq 2^N.$$

Donc

$$\mu_1 \dots \mu_N \text{Vol}(\mathcal{K}) \times (\lambda_1 \dots \lambda_n \text{Vol}(K))^{-P} \geq (N!)^{-1} 2^N 2^{-nP} = (N! 2^{(p-1)N})^{-1}.$$

Par le théorème 2.4.6 et en utilisant la relation $(\lambda_1 \dots \lambda_n)^P = M_1 \dots M_N$, on en déduit que

$$\prod_{j=1}^N \frac{\mu_j}{M_j} \geq (N! 2^{(p-1)N})^{-1} \alpha_1 =: \alpha_3.$$

Or on a prouvé que $\mu_j/M_j \leq 1$ pour tout j ; on a donc pour tout j ,

$$\frac{\mu_j}{M_j} \geq \alpha_3,$$

ce qui achève la démonstration du théorème. □

3 Angles d'inclinaison entre deux sous-espaces

La majeure partie de cette partie est issue de la partie II de [17] et reprend sa structure.

Il sera commode à partir de maintenant de noter la dimension d'un espace vectoriel en exposant : ainsi, si E, A, \dots sont des espaces vectoriels de dimensions n, d, \dots , on les

notera E^n, A^d, \dots (respectivement).

Ici, E^n (resp. U^n) désigne un espace euclidien (resp. hermitien) dont on note $\langle \cdot, \cdot \rangle$ le produit scalaire (qu'on supposera linéaire en sa première variable et anti-linéaire en sa seconde variable dans le cas hermitien) et $\|\cdot\|$ la norme associée. Tous les résultats qui suivent sont énoncés dans le cadre hermitien mais restent vrais dans le cadre euclidien en remplaçant simplement U^n par E^n et "transformation unitaire" par "transformation orthogonale". Les angles d'inclinaison sont une famille finie de réels associée à un couple de sous-espaces (A^d, B^e) et qui permet de définir mathématiquement la proximité entre ces sous-espaces.

3.1 Produits scalaires successifs $\lambda_1, \dots, \lambda_f$

Définition 3.1.1. Pour X et Y non nuls de U^n , on note

$$\lambda(X, Y) := \frac{|\langle X, Y \rangle|}{\|X\| \|Y\|}.$$

Théorème 3.1.2. Soient A^d et B^e deux sous-espaces vectoriels non nuls de U^n . On pose $f := \min(d, e) > 0$. Il existe des bases orthonormales (X_1, \dots, X_d) et (Y_1, \dots, Y_e) de A^d et B^e respectivement, et des réels $1 \geq \lambda_1 \geq \dots \geq \lambda_f \geq 0$ tels que pour tous $1 \leq i \leq d$ et $1 \leq j \leq e$

$$\langle X_i, Y_j \rangle = \lambda_i \delta_{ij}.$$

Les nombres $\lambda_1, \dots, \lambda_f$ ne dépendent pas du choix des deux bases considérées et sont invariants par transformations unitaires appliquées simultanément à A^d et B^e .

Définition 3.1.3. Les nombres $\lambda_1 \dots \lambda_f$ sont appelés les produits scalaires successifs de A^d et B^e . Si S est un sous-espace non nul et $X \neq 0$, on note $\lambda(X, S) = \lambda_1(\text{Vect}(X), S)$.

Preuve On renvoie à l'article de Schmidt [17] pour l'existence des deux bases et des λ_i . La proposition 3.1.6 ci-dessous permet de démontrer l'indépendance des λ_i du choix de ces bases. □

Définition 3.1.4. Si $\underline{Z} = (Z_1, \dots, Z_k)$ et $\underline{T} = (T_1, \dots, T_l)$ sont deux familles de U^n , on note

$$M(\underline{Z}, \underline{T}) := (\langle Z_i, T_j \rangle)_{\substack{1 \leq i \leq k \\ 1 \leq j \leq l}}.$$

Dans le cas particulier $\underline{Z} = \underline{T}$ on note plutôt

$$M(\underline{Z}) := M(\underline{Z}, \underline{Z})$$

(c'est la matrice considérée au début du paragraphe 2.3).

Lemme 3.1.5. Soit \mathfrak{B} une base orthonormée fixée de U^n . Si $\underline{Z} = (Z_1, \dots, Z_k)$ et $\underline{T} = (T_1, \dots, T_l)$ sont deux familles de U^n et qu'on note $M_{\underline{Z}}$ (resp. $M_{\underline{T}}$) la matrice dont la i -ème ligne est le vecteur Z_i (resp. T_i) écrit dans la base \mathfrak{B} , alors on a

$$M(\underline{Z}, \underline{T}) = M_{\underline{Z}} M_{\underline{T}}^*. \quad (3.1)$$

(Rappel : M^* désigne la trans-conjuguée de la matrice M .)

Preuve Calcul direct. □

Remarque : cette formule généralise la formule (2.27).

Proposition 3.1.6. Soient A^d et B^e deux sous-espaces vectoriels non nuls de U^n . On pose $f := \min(d, e) > 0$. Soient $\underline{X} = (X_1, \dots, X_d)$ et $\underline{Y} = (Y_1, \dots, Y_e)$ deux bases quelconques de A^d et B^e respectivement. On définit

$$P(\lambda) := D(X_1, \dots, X_d)^{-2} D(Y_1, \dots, Y_e)^{-2} \begin{vmatrix} \lambda M(\underline{X}) & M(\underline{X}, \underline{Y}) \\ M(\underline{Y}, \underline{X}) & \lambda M(\underline{Y}) \end{vmatrix}.$$

Alors $P(\lambda)$ ne dépend pas du choix des bases de A^d et B^e , et si on le note $P(A, B; \lambda)$, alors pour toute transformation unitaire τ de U^n on a

$$P(\tau A, \tau B; \lambda) = P(A, B; \lambda).$$

En particulier, si $f := \min(d, e)$, $h := \max(d, e)$ et $\lambda_1, \dots, \lambda_f$ sont définis comme dans le théorème 3.1.2, on a

$$P(\lambda) = (\lambda^2 - \lambda_1^2) \dots (\lambda^2 - \lambda_f^2) \lambda^{h-f}.$$

Remarques

- Ceci prouve l'assertion sur l'indépendance des λ_i du choix des bases dans le théorème 3.1.2.
- Le déterminant du membre de droite dans la définition de $P(\lambda)$ est le déterminant d'une matrice par blocs de taille $(d+e) \times (d+e)$.

Preuve Montrons que $P(\lambda)$ ne dépend que de A^d et B^e (les détails qui suivent ne figurent pas dans [17]). Soit \underline{X}' et \underline{Y}' deux autres bases de A^d et B^e respectivement. Alors il existe P et Q inversibles (de taille d et e respectivement) telles que

$$M_{\underline{X}'} = PM_{\underline{X}} \quad \text{et} \quad M_{\underline{Y}'} = QM_{\underline{Y}}.$$

Par (3.1) on en déduit immédiatement que

$$M(\underline{X}') = PM(\underline{X})P^*, \quad M(\underline{Y}') = QM(\underline{Y})Q^*,$$

et

$$\begin{pmatrix} \lambda M(\underline{X}') & M(\underline{X}', \underline{Y}') \\ M(\underline{Y}', \underline{X}') & \lambda M(\underline{Y}') \end{pmatrix} = \begin{pmatrix} P & 0 \\ 0 & Q \end{pmatrix} \begin{pmatrix} \lambda M(\underline{X}) & M(\underline{X}, \underline{Y}) \\ M(\underline{Y}, \underline{X}) & \lambda M(\underline{Y}) \end{pmatrix} \begin{pmatrix} P^* & 0 \\ 0 & Q^* \end{pmatrix},$$

donc

$$\begin{vmatrix} \lambda M(\underline{X}') & M(\underline{X}', \underline{Y}') \\ M(\underline{Y}', \underline{X}') & \lambda M(\underline{Y}') \end{vmatrix} = |\det P|^2 |\det Q|^2 \begin{vmatrix} \lambda M(\underline{X}) & M(\underline{X}, \underline{Y}) \\ M(\underline{Y}, \underline{X}) & \lambda M(\underline{Y}) \end{vmatrix}.$$

On conclut en remarquant que

$$D(X'_1, \dots, X'_d)^2 = \det M(\underline{X}') = \det(PM(\underline{X})P^*) = |\det P|^2 D(X_1, \dots, X_d),$$

et

$$D(Y'_1, \dots, Y'_e)^2 = |\det Q|^2 D(Y_1, \dots, Y_e).$$

Si τ est une transformation unitaire, il existe une matrice unitaire P telle que

$$M(\tau \underline{Z}, \tau \underline{T}) = (M_{\underline{Z}} P)(M_{\underline{T}} P)^* = M_{\underline{Z}} (P P^*) M_{\underline{T}}^* = M(\underline{Z}, \underline{T}),$$

ce qui prouve directement l'égalité

$$P(\tau A, \tau B; \lambda) = P(A, B; \lambda).$$

Pour finir, on choisit (X_1, \dots, X_d) et (Y_1, \dots, Y_e) comme dans le théorème 3.1.2. On a alors

$$M(\underline{X}) = I_d, \quad D(X_1, \dots, X_d) = 1$$

et

$$M(\underline{Y}) = I_e, \quad D(Y_1, \dots, Y_e) = 1$$

où I_d et I_e désignent les matrices identités de \mathcal{M}_d et \mathcal{M}_e respectivement. Il faut alors calculer le déterminant

$$P(\lambda) = \begin{vmatrix} \lambda M(\underline{X}) & M(\underline{X}, \underline{Y}) \\ M(\underline{Y}, \underline{X}) & \lambda M(\underline{Y}) \end{vmatrix}$$

en fonction de λ . Cette matrice est assez simple. Pour calculer son déterminant, il suffit de multiplier la i -ème colonne par λ et de lui retrancher la $(d+i)$ -ème colonne multipliée par λ_i ($i = 1, \dots, f$). Ainsi, le déterminant $\lambda^{-f} P(\lambda)$ est le déterminant d'une matrice par blocs égal à

$$(\lambda^2 - \lambda_1^2) \dots (\lambda^2 - \lambda_f^2) \lambda^{d-f} \lambda^e,$$

et on conclut avec l'égalité $d + e = h + f$. □

On peut caractériser les λ_i d'une autre manière.

Lemme 3.1.7. *Soient A^d et B^e deux sous-espaces non nuls de U^n . On pose $f = f(A, B) := \min(d, e)$ et on fixe $1 \leq i \leq f$. Alors $\lambda_i := \lambda_i(A^d, B^e)$ est le plus grand λ tel qu'il existe $A^i \subset A^d$ tel que pour tout $X \in A^i$ non nul il existe $Y \in B^e$ non nul vérifiant $\lambda(X, Y) \geq \lambda$. En d'autres termes :*

$$\lambda_i(A^d, B^e) = \sup_{A^i \subset A^d} \inf_{\substack{X \in A^i \\ X \neq 0}} \sup_{\substack{Y \in B^e \\ Y \neq 0}} \lambda(X, Y). \quad (3.2)$$

Preuve Cette preuve est laissée au lecteur par Schmidt, nous proposons ici une démonstration personnelle.

On définit λ'_i par l'égalité (3.2).

On prend (X_1, \dots, X_d) et (Y_1, \dots, Y_e) donnés par le théorème 3.1.2 et on pose $A^i := \text{Vect}(X_1, \dots, X_i)$. Soit $X = \sum_{k \leq i} \alpha_k X_k \in A^i$ de norme 1 (i.e. $\sum_k |\alpha_k|^2 = 1$). On pose $Y := \sum_{k \leq i} \alpha_k Y_k \in B^e$. Alors Y est de norme 1 et

$$\lambda(X, Y) = \sum_{k=1}^i |\alpha_k|^2 \lambda_k \geq \lambda_i \sum_{k=1}^i |\alpha_k|^2 = \lambda_i,$$

ce qui assure

$$\inf_{\substack{X \in A^i \\ X \neq 0}} \sup_{\substack{Y \in B^e \\ Y \neq 0}} \lambda(X, Y) \geq \lambda_i,$$

donc $\lambda'_i \geq \lambda_i$.

Réciproquement, (X_1, \dots, X_d) et (Y_1, \dots, Y_e) étant toujours donnés par le théorème 3.1.2, si $A^i \subset A^d$ alors il existe $X \in A^i \cap \text{Vect}(X_i, \dots, X_d)$ non nul de norme 1 qu'on écrit $X = \sum_k \alpha_k X_k$ (avec $\alpha_k = 0$ si $k < i$). Soit $Y = \sum_l \beta_l Y_l \in B^e$ de norme 1. Alors

$$\lambda(X, Y) = |\langle X, Y \rangle| = \left| \sum_{k=i}^f \lambda_k \alpha_k \beta_k \right| \leq \lambda_i \sum_k |\alpha_k| |\beta_k| \leq \lambda_i \|X\| \|Y\|,$$

(la dernière inégalité étant obtenue par Cauchy-Schwarz), ce qui prouve bien $\lambda'_i \leq \lambda_i$ (puisqu'on a X et Y de norme 1). □

Corollaire 3.1.8. Soient $A' \subset A$, $B' \subset B$ des sous-espaces. Alors $f' := f(A', B') \leq f := f(A, B)$ et

$$\lambda_i(A', B') \leq \lambda_i(A, B) \quad (i = 1, \dots, f').$$

Preuve Immédiate avec (3.2). □

3.2 Quantités ν_1, \dots, ν_t

Il est naturel de se poser la question de savoir quelles sont les valeurs possibles du f -uplet $(\lambda_1, \dots, \lambda_f)$ associé à un couple (A^d, B^e) . Remarquons que si $\dim(A \cap B) = c$ alors $\lambda_1(A, B) = \dots = \lambda_c(A, B) = 1$. Ceci impose des conditions sur les produits scalaires successifs : si $d + e > n$, alors $A^d \cap B^e$ est de dimension au moins $d + e - n > 0$ et $\lambda_1 = \dots = \lambda_{d+e-n} = 1$; décrire les λ_i ($i = 1, \dots, f$) revient à décrire les λ_i ($i > d + e - n$). C'est ce qui justifie l'introduction des ν_i . Le théorème 3.2.3 offre une description complète des ν_i . Le deuxième résultat principal de ce paragraphe est le théorème 3.2.4 qui assure que $\nu_i(A, B) = \nu_i(A^\perp, B^\perp)$. Tout ce qui suit est démontré par Schmidt dans [17], on ne fera que rappeler les résultats principaux.

Définition 3.2.1. Soient A^d et B^e deux sous-espaces. On pose $f := \min(d, e)$, $g := d + e - n$, et

$$t = t(A^d, B^e) := \min(d, e, n - d, n - e) = \min(d, e, e - g, d - g). \quad (3.3)$$

On définit alors ν_1, \dots, ν_t comme suit.

Si $d + e \leq n$ (i.e. $g \leq 0$) alors $t = f$ et on pose

$$\nu_i(A^d, B^e) := \lambda_i(A^d, B^e) \quad (i = 1, \dots, t).$$

Si $d + e > n$ (i.e. $g > 0$) alors $t = f - g$ et on pose

$$\nu_i(A^d, B^e) := \lambda_{i+g}(A^d, B^e) \quad (i = 1, \dots, t).$$

On a $1 \geq \nu_1 \geq \dots \geq \nu_t \geq 0$.

Définition 3.2.2. On dit que deux paires de sous-espaces (A^d, B^e) et (A'^d, B'^e) sont similaires s'il existe une transformation unitaire τ telle que $\tau(A^d) = A'^d$ et $\tau(B^e) = B'^e$. C'est une relation d'équivalence.

Le théorème 3.1.2 assure que les ν_i sont un invariant pour cette relation d'équivalence.

Théorème 3.2.3. Soient $0 < d < n$ et $0 < e < n$ et $1 \geq \nu_1 \geq \dots \geq \nu_t \geq 0$ donnés (où t est défini par (3.3)). Alors à similarité près il y a un unique couple (A^d, B^e) de sous-espaces tel que $\nu_i(A^d, B^e) = \nu_i$ ($i = 1, \dots, t$).

Preuve Nous ne donnons ici que le schéma de preuve. Schmidt procède en deux étapes : d'abord il prouve l'existence en construisant explicitement un couple d'espaces qui vérifie l'assertion du théorème (en distinguant les cas $g \leq 0$ et $g > 0$), puis pour l'unicité (à similarité près), il construit une transformation unitaire en utilisant les bases particulières fournies par le théorème 3.1.2. □

Théorème 3.2.4 (Principe de dualité).

Soient A et B deux sous-espaces de U^n d'orthogonaux respectifs A^\perp et B^\perp . Alors $t(A^\perp, B^\perp) = t(A, B) =: t$ et

$$\nu_i(A^\perp, B^\perp) = \nu_i(A, B) \quad (i = 1 \dots, t).$$

Preuve Là encore, seul un schéma de preuve est donné, le lecteur trouvera les détails dans [17].

Par symétrie, on peut supposer que $d + e = \dim A + \dim B \leq n$ et sans perte de généralité on peut également supposer $d \leq e$ (et donc $t = f = d$). Schmidt pose alors $V := A \cap B$ et écrit $A = A_1 \oplus^\perp V$ et $B = B_1 \oplus^\perp V$. Si $r = \dim V$ et $\lambda_1, \dots, \lambda_{d-r}$ sont les produits scalaires successifs de A_1 et B_1 alors $1, \dots, 1$ (r fois), $\lambda_1, \dots, \lambda_{d-r}$ sont les produits scalaires successifs de A et B . A l'aide des bases (X_1, \dots, X_{d-r}) et (Y_1, \dots, Y_{e-r}) fournies par le théorème 3.1.2 appliqué à A_1 et B_1 , Schmidt construit deux autres bases qui permettent de calculer $\nu_i(A^\perp, B^\perp)$. □

3.3 Angles d'inclinaison

Dans ce paragraphe sont introduites les dernières définitions dont nous avons besoin pour définir les angles d'inclinaison φ_i qui permettent de définir la proximité de deux sous-espaces A^d et B^e : intuitivement deux sous-espaces sont d'autant plus proches que leurs angles d'inclinaison sont petits. Ils se définissent complètement à partir des produits successifs λ_i . Les théorèmes d'approximation diophantienne de la suite n'utilisent toutefois pas directement les φ_i mais plutôt leur sinus, que l'on notera ω_i . Le lemme 3.3.6 permet de contrôler les angles d'inclinaison de l'image de deux sous-espaces par un même automorphisme en fonction des angles d'inclinaison des deux sous-espaces initiaux.

Définition 3.3.1. Soient X et Y deux vecteurs non nuls de U^n . On pose

$$\omega(X, Y) := (1 - \lambda^2(X, Y))^{1/2}.$$

Proposition 3.3.2. Soient X, Y, Z non nuls. Alors

$$\omega(X, Z) \leq \omega(X, Y) + \omega(Y, Z).$$

Preuve Schmidt laisse cette preuve au lecteur dans son article, nous proposons ici une démonstration.

Soient X, Y, Z non nuls. Soit π la projection orthogonale sur $\text{Vect}(X, Z)$. Alors on a

$$\lambda(X, Y) = \frac{|\langle X, Y \rangle|}{\|X\| \|Y\|} = \frac{|\langle X, \pi(Y) \rangle|}{\|X\| \|Y\|} \leq \frac{|\langle X, \pi(Y) \rangle|}{\|X\| \|\pi(Y)\|} = \lambda(X, \pi(Y)),$$

donc $\omega(X, Y) \geq \omega(X, \pi(Y))$. De même $\omega(Z, Y) \geq \omega(Z, \pi(Y))$. Il suffit donc de montrer le résultat pour $\pi(Y)$ et on est ramené à un problème du plan (le résultat étant évident si (X, Z) est liée). On suppose X, Y, Z de norme 1, on pose $X_0 = X$ qu'on complète en (X_0, X_1) une base orthonormée de $\text{Vect}(X, Z)$. On écrit alors $Y = y_0 X_0 + y_1 X_1$ et $Z = z_0 X_0 + z_1 X_1$. Puisque $|y_0|^2 + |y_1|^2$ et $|z_0|^2 + |z_1|^2$ sont égaux à 1, il existe θ_y et θ_z dans $[0, \pi/2]$ tels que

$$(|y_0|, |y_1|) = (\cos \theta_y, \sin \theta_y) \quad \text{et} \quad (|z_0|, |z_1|) = (\cos \theta_z, \sin \theta_z).$$

On a alors :

$$\omega(X, Y) = \sin \theta_y, \quad \omega(X, Z) = \sin \theta_z,$$

et

$$\lambda(Y, Z) = |y_0 \bar{z}_0 + y_1 \bar{z}_1| \leq |y_0 \bar{z}_0| + |y_1 \bar{z}_1| = \cos(\theta_y) \cos(\theta_z) + \sin(\theta_y) \sin(\theta_z) = \cos(\theta_z - \theta_y),$$

donc

$$|\sin(\theta_z - \theta_y)| \leq \omega(Y, Z).$$

Finalement, on a

$$\begin{aligned} \omega(X, Z) &= \sin(\theta_z) = \sin(\theta_y + \theta_z - \theta_y) \\ &= \cos(\theta_z - \theta_y) \sin(\theta_y) + \cos(\theta_y) \sin(\theta_z - \theta_y) \\ &\leq \sin(\theta_y) + |\sin(\theta_z - \theta_y)| \\ &\leq \omega(X, Y) + \omega(Y, Z). \end{aligned}$$

□

Définition 3.3.3. Soient A^d et B^e deux sous-espaces non nuls de U^n . On pose $f := \min(d, e)$. Pour $1 \leq i \leq f$ on pose

$$\omega_i(A^d, B^e) := (1 - \lambda_i^2(A^d, B^e))^{1/2}.$$

Si X est non nul, on pose

$$\omega(X, B) := \omega_1(\text{Vect}(X), B).$$

On définit $\varphi_i(A^d, B^e) \in [0; \pi/2]$ par

$$\cos(\varphi_i) = \lambda_i \quad \text{et} \quad \sin(\varphi_i) = \omega_i.$$

Les φ_i sont appelés les *angles d'inclinaison* de A^d, B^e .

On définit également

$$\psi_i(A^d, B^e) := (1 - \nu_i^2)^{1/2} \quad (i = 1 \dots t),$$

où t est défini par (3.3). Si $d + e \leq n$, on a $\psi_i = \omega_i$ pour tout i , sinon $\psi_i = \omega_{i+g}$ pour tout i (rappelons que $g = d + e - n$). Enfin, on introduit

$$\mu(A^d, B^e) := \psi_1 \dots \psi_t.$$

Proposition 3.3.4. Soient A^d et B^e deux sous-espaces avec $d + e \leq n$ (on a alors $t = f$) et $P(\lambda)$ le polynôme de la proposition 3.1.6. Alors

$$\begin{aligned} \mu &= ((1 - \lambda_1^2) \dots (1 - \lambda_f^2))^{1/2} = P(1)^{1/2} \\ &= D(X_1, \dots, X_d, Y_1, \dots, Y_e) D(X_1, \dots, X_d)^{-1} D(Y_1, \dots, Y_e)^{-1} \\ &= \|X_1 \wedge \dots \wedge X_d \wedge Y_1 \wedge \dots \wedge Y_e\|_2 \|X_1 \wedge \dots \wedge X_d\|_2^{-1} \|Y_1 \wedge \dots \wedge Y_e\|_2^{-1}, \end{aligned}$$

où (X_1, \dots, X_d) (resp. (Y_1, \dots, Y_e)) est n'importe quelle base de A^d (resp. B^e) (et en utilisant l'identification (2.24) pour la dernière égalité).

Preuve Immédiate en revenant aux définitions.

□

Proposition 3.3.5. *Soit $X \neq 0$ et B^e un sous-espace de dimension $0 < e < n$. Alors*

$$\omega(X, B) + \omega(X, B^\perp) = 1.$$

De manière équivalente

$$\lambda(X, B) + \lambda(X, B^\perp) = 1.$$

Preuve Schmidt laisse la démonstration au lecteur dans [17]. L'équivalence entre les deux formules est facile. Montrons la deuxième : on suppose X de norme 1 et on le décompose en $X = Y + Z$ avec $Y \in B$ et $Z \in B^\perp$. On a : $1 = \|Y\|^2 + \|Z\|^2$, $\lambda(X, B) = \|Y\|^2$ et $\lambda(X, B^\perp) = \|Z\|^2$, ce qui permet de conclure. \square

Lemme 3.3.6. *Soient $0 < d < n$, $0 < e < n$, $f := \min(d, e)$ et σ un automorphisme de U^n . Alors il existe une constante $c = c(\sigma) > 0$ telle que pour tous sous-espaces A^d , B^e on ait*

$$\omega_i(\sigma(A^d), \sigma(B^e)) \leq c(\sigma)\omega_i(A^d, B^e) \quad (i = 1, \dots, f).$$

Preuve Cf la démonstration du Lemme 13 section 8 de [17]. \square

La définition des angles d'inclinaison (ou, de manière équivalente, des ω_i) n'est pas très pratique pour effectuer des calculs explicites lorsqu'on ne dispose pas des bases spéciales fournies par le théorème 3.1.2. Le théorème suivant contourne ce problème et permet de majorer les angles d'inclinaison de deux sous-espaces "proches" (i.e. avec de petits angles d'inclinaison) à partir de familles explicites de vecteurs satisfaisant certaines conditions plus faibles que celles du théorème 3.1.2.

Théorème 3.3.7. *Soit A^d , B^e deux sous-espaces de U^n , et $1 \leq i \leq f = \min(d, e)$. Soient $0 < \delta, \omega < 1$ donnés. On suppose qu'il existe i vecteurs linéairement indépendants $X_1, \dots, X_i \in A^d$ tels qu'en posant $A^j := \text{Vect}(X_1, \dots, X_j)$ ($j = 1, \dots, i$) on ait :*

$$\lambda(X_j, A^{j-1}) \leq 1 - \delta < 1 \quad (j = 2, \dots, i),$$

et qu'il existe i vecteurs $Y_1, \dots, Y_i \in B^e$ non nuls tels que

$$\omega(X_j, Y_j) \leq \omega \quad (j = 1, \dots, i).$$

Alors

$$\omega_j(A^j, B^e) \leq \left(\frac{2}{\delta}\right)^{(j-1)/2} \times \omega \quad (j = 1, \dots, i),$$

en particulier on retiendra

$$\omega_i(A^d, B^e) \leq \left(\frac{2}{\delta}\right)^{i/2} \times \omega.$$

Remarques :

- Pour que ce théorème apporte une véritable information, il faut que $(2/\delta)^{i/2}\omega < 1$, en particulier il est intéressant pour δ proche de 1 et ω proche de 0.
- A la limite $\delta = 1$, la première condition sur les X_j est une condition d'orthogonalité. On impose une minoration des angles entre les X_j (les X_j doivent être "assez éloignés")

les uns des autres).

- A la limite $\omega = 0$, la seconde condition sur les Y_j est une condition de colinéarité (Y_j et X_j colinéaires). On impose cette fois-ci que Y_j et X_j soient au contraire "proches".

Preuve Le lecteur est renvoyé à [17] (Part II, section 9, théorème 7) pour une preuve de ce théorème. Schmidt procède par récurrence sur j . L'initialisation se fait pour $j = 1$ (évident) et $j = 2$ (la plus grosse partie de la preuve). □

Corollaire 3.3.8. Soient A^d, B^e , et $1 < i \leq f := \min(d, e)$. Soient $X_1, \dots, X_i \in A^d$ et $Y_1, \dots, Y_i \in B^e$ des vecteurs non nuls. On suppose que

$$\begin{aligned} \lambda(X_j, X_k) &\leq 4^{-i} & (1 \leq j < k \leq i), \\ \omega(X_j, Y_j) &\leq \omega & (j = 1 \dots i). \end{aligned}$$

Alors

$$\omega_i(A^d, B^e) \leq 2^i \omega.$$

Preuve Cf [17] (Part II, section 9, Corollaire du théorème 7). □

4 Hauteur d'un sous-espace

Dans cette partie E^n désigne un espace vectoriel euclidien dont on notera $\langle \cdot, \cdot \rangle$ le produit scalaire et $\|\cdot\|$ la norme associée.

Dans le paragraphe 4.1 on donne une première définition de la hauteur d'un sous-espace de K^n et les premières propriétés que l'on peut en tirer. Intuitivement, la hauteur d'un sous-espace représente sa complexité. Le paragraphe 4.2 propose une définition alternative et plus géométrique de la hauteur qui s'exprime en termes de propriétés sur des réseaux. C'est plutôt cette définition ci qu'on utilisera en géométrie paramétrique des nombres.

4.1 Définition de la hauteur et premières propriétés

Ce paragraphe reprend la section 1 de l'article de Schmidt [17].

Cadre :

Soit K un corps de nombre de degré $[K : \mathbb{Q}] = p$. On note \mathcal{O}_K l'anneau des entiers de K (Schmidt le note I dans [17]). Tout idéal fractionnaire de K (i.e. un sous- \mathcal{O}_K -module $I \neq 0$ tel qu'il existe $a \in \mathcal{O}_K$ non nul vérifiant $aI \subset \mathcal{O}_K$) est un \mathcal{O}_K -module de type fini. On note $\sigma_1, \dots, \sigma_p$ les différents \mathbb{Q} -morphisms de corps distincts de K dans \mathbb{C} . Pour $\xi \in K$, on note $\xi^{(i)} = \sigma_i(\xi)$ et $\nu_i(\xi) = |\xi^{(i)}|$ ($i = 1, \dots, p$). De même, si $X = (\xi_1, \dots, \xi_n) \in K^n$, on note aussi $X^{(i)} = (\xi_1^{(i)}, \dots, \xi_n^{(i)})$ et $\nu_i(X) = (\nu_i(\xi_1), \dots, \nu_i(\xi_n))$, $i = 1, \dots, p$ (remarque : dans le cas $i = 1$ il y a a priori une ambiguïté avec la notation de la composé i -ème d'une matrice (cf définition 2.1.10) ; le contexte ne laissera pas d'ambiguïté, et dans le cas où K est un sous-corps de \mathbb{C} , en choisissant $\sigma_1 = \text{id}_{\mathbb{C}}$ les deux définitions coïncident. Pour $i \geq 2$ la composée i -ème de X n'est pas définie).

On notera également φ la forme bilinéaire symétrique non dégénérée sur K^n définie par

$$\varphi(X, Y) = \sum_{i=1}^n x_i y_i.$$

et $S^{\perp, \varphi}$ l'orthogonal d'un sous-espace S^d de K^n pour φ ($S^{\perp, \varphi}$ est un sous-espace vectoriel de codimension d) (remarque : il n'y aura pas d'ambiguïté sur le corps de nombres K dans le contexte où on utilisera φ).

Rappel sur la norme d'un élément ou d'un idéal :

Si $x \in K$, la norme de x est $N(x) := \prod_{j=1}^p \sigma_j(x) \in \mathbb{Q}$. La norme d'un idéal fractionnaire I de K est l'idéal fractionnaire de \mathbb{Q} engendré par les $N(x)$, $x \in I$. Comme il est principal, on identifie $N(I)$ avec son unique générateur positif $\in \mathbb{Q}$. Rappelons qu'on a une structure de groupe sur les idéaux fractionnaires et que si I_1 et I_2 sont deux idéaux fractionnaires, alors $N(I_1 I_2) = N(I_1) N(I_2)$. Si \mathfrak{p} est un idéal premier non nul de \mathcal{O}_K au-dessus d'un idéal $\mathfrak{q} \subset \mathbb{Q}$ de degré résiduel en \mathfrak{p} égal à $f_{\mathfrak{p}}$, alors $N(\mathfrak{p}) = \mathfrak{q}^{f_{\mathfrak{p}}}$. Notons enfin que si I est un idéal de \mathcal{O}_K , alors $N(I) = |\mathcal{O}_K/I|$.

Définition 4.1.1. On appelle fonction distance une fonction F vérifiant pour tous $x \in E^n$ et $t \in \mathbb{R}$

- (i) $F(x) \geq 0$ avec égalité si et seulement si $x = 0$
- (ii) $F(tx) = |t|F(x)$
- (iii) F est continue.

Typiquement, si N est une distance sur E^n , alors $x \mapsto N(x)$ est une fonction distance.

Proposition 4.1.2. *Etant données deux fonctions distance F_1, F_2 , il existe une constante $c = c(F_1, F_2) > 0$ telle que $F_1(x) \leq cF_2(x)$ pour tout x .*

Preuve Sur la sphère unité S de E^n , la fonction F_1/F_2 est bien définie et continue. Par compacité il existe donc c tel que pour tout $x \neq 0$ $F_1(x/||x||) \leq cF_2(x/||x||)$, et on conclut avec les points (i) et (ii) de la définition d'une fonction distance. □

Définition 4.1.3 (Hauteur d'une droite vectorielle de K^n).

Soit F une fonction distance sur \mathbb{R}^n .

Soit L un sous-espace vectoriel de dimension 1 de K^n . On choisit un vecteur directeur $X = (\xi_1, \dots, \xi_n) \neq 0$ de L . On écrit $\mathfrak{a} := \mathfrak{a}(X)$ l'idéal engendré par ξ_1, \dots, ξ_n . On définit alors la F -hauteur de L par

$$H_F(L) := N(\mathfrak{a})^{-1} \prod_{j=1}^p F(\nu_j(\xi_1), \dots, \nu_j(\xi_n)).$$

La quantité précédente est indépendante du choix de X eu égard aux propriétés de la norme d'un idéal et au point (ii) de la définition d'une fonction distance.

Proposition 4.1.4. *Etant données deux fonctions distance F_1, F_2 et L une droite vectorielle de K^n , alors*

$$H_{F_1}(L) \leq c(F_1, F_2)^p H_{F_2}(L).$$

Donc, à un facteur borné près, toutes les hauteurs d'un sous-espace de dimension 1 sont les mêmes.

La définition de la hauteur d'un sous-espace de dimension quelconque requiert les outils introduits dans la partie 2.2, en particulier la définition 2.2.4 (rappel : la droite S^* est l'espace vectoriel engendré par les coordonnées de Grassmann de S) et la définition de la fonction τ du théorème 2.2.6.

Définition 4.1.5 (Hauteur d'un sous-espace de dimension quelconque).

Soit d un entier, $0 \leq d \leq n$. Soit F une fonction distance τ -invariante sur \mathbb{R}^N avec $N := \binom{n}{d}$ (i.e. $F(\tau(x)) = F(x)$ pour tout x). Soit S^d un sous-espace de K^n de dimension d . Si $d = 0$ ou $d = n$ on pose

$$H_F(S) = 1.$$

Si $0 < d < n$, on définit la F -hauteur de S par

$$H_F(S) := H_F(S^*).$$

Proposition 4.1.6. *Etant données deux fonctions distance F_1, F_2 et S^d , alors*

$$H_{F_1}(S) \leq c(F_1, F_2)^p H_{F_2}(S).$$

Donc, à un facteur borné près, toutes les hauteurs d'un sous-espace donné de K^n sont les mêmes.

Théorème 4.1.7. *Soit F une fonction distance τ -invariante et S un sous-espace vectoriel de K^n . Alors*

$$H_F(S) = H_F(S^{\perp, \varphi}).$$

Preuve Si $d = 0$ ou n le résultat est trivial. Dans les autres cas c'est une application directe du théorème 2.2.6. □

Proposition 4.1.8. *Soit u un automorphisme de K^n . Alors il existe une constante strictement positive $c_F^{(d)}(u)$ telle que pour tout sous-espace S^d de K^n de dimension d on ait :*

$$H_F(u(S^d)) \leq c_F^{(d)}(u) H_F(S^d).$$

Preuve Soient u un automorphisme de K^n et S^d un sous-espace vectoriel de K^n . Par la proposition 2.2.3, on a $u(S)^* = u^{(d)}(S^*)$. On est ramené au cas $d = 1$ dans K^N (où on a posé $N := \binom{n}{d}$).

Schmidt laisse au lecteur la démonstration du cas $d = 1$, nous en proposons une ici.

Soit $X = (\xi_1, \dots, \xi_n)$ un vecteur directeur d'un sous-espace L^1 de dimension 1. On pose $X' := u(X) = (\xi'_1, \dots, \xi'_n)$ et on note \mathfrak{a} (resp. \mathfrak{a}') l'idéal fractionnaire de K engendré par les coordonnées de X (resp. de X').

Si on note \mathfrak{b} l'idéal fractionnaire engendré par les coefficients de la matrice de u dans la base canonique, alors on a $\mathfrak{a}' \subset \mathfrak{b}\mathfrak{a}$, ce qui assure

$$N(\mathfrak{a}) \leq N(\mathfrak{a}')/N(\mathfrak{b}).$$

Finalement

$$H_F(u(L^1)) \leq N(\mathfrak{b}^{-1}) \left(\prod_{j=1}^p \frac{F(\nu_j(\xi'_1), \dots, \nu_j(\xi'_n))}{F(\nu_j(\xi_1), \dots, \nu_j(\xi_n))} \right) H_F(L^1).$$

Fixons $1 \leq j \leq p$. Si on note $M := \max_{\|Y\|=1} F(Y) > 0$ et $m := \min_{\|Y\|=1} F(Y) > 0$ (qui ne dépendent que de F), on a

$$F(\nu_j(u(X)))F(\nu_j(X))^{-1} \leq Mm^{-1} \frac{\|\nu_j(u(X))\|}{\|\nu_j(X)\|}.$$

Remarquons maintenant que si on note encore $\|\cdot\|$ la norme hermitienne sur \mathbb{C}^n , alors pour tout $Y \in K^n$ on a $\|\nu_j(Y)\| = \|Y^{(j)}\|$. Ainsi, si on note $\tilde{u}^{(j)}$ l'automorphisme de \mathbb{C}^n correspondant à u transporté par σ_j (i.e. si $(u_{kl})_{k,l}$ est la matrice de u dans la base canonique de K^n , alors la matrice de $\tilde{u}^{(j)}$ dans la base canonique de \mathbb{C}^n est $(\sigma_j(u_{kl}))_{k,l}$), alors

$$\frac{\|\nu_j(u(X))\|}{\|\nu_j(X)\|} = \frac{\|\tilde{u}^{(j)}(X^{(j)})\|}{\|X^{(j)}\|} \leq \|\tilde{u}^{(j)}\|,$$

où $\|\cdot\|$ désigne la norme d'opérateur associée à la norme hermitienne sur \mathbb{C}^n (qui ne dépend plus de X). Finalement, on a :

$$H_F(u(L^1)) \leq \left(N(\mathfrak{b}^{-1}) \left(\frac{M}{m} \right)^p \prod_{j=1}^p \|\tilde{u}^{(j)}\| \right) H_F(L^1),$$

ce qui achève la démonstration. □

Définition 4.1.9. Soit S^d un sous-espace de K^n de dimension d . On pose $N := \binom{n}{d}$ et on note F_E la fonction distance associée à la norme euclidienne sur \mathbb{R}^N . On définit alors

$$H(S) := H_{F_E}(S).$$

Remarque A partir de maintenant, on ne considérera donc que les hauteurs issues des normes euclidiennes.

4.2 Seconde définition de la hauteur

Ce paragraphe reprend les sections 3 et 4 de l'article de Schmidt [17]. On donne une définition plus géométrique de la hauteur d'un sous-espace - le théorème 4.2.3 assure que les deux définitions coïncident - en la reliant au déterminant d'un certain réseau de \mathbb{R}^{np} attaché au sous-espace S^d . La preuve du théorème 4.2.6, l'un des résultats principaux de cette sous-partie, utilise cette nouvelle définition et des outils de géométrie des nombres.

Dans ce paragraphe K désigne encore un corps de nombres de degré p et on conserve les notations de la partie précédente (section 4.1) pour $\sigma_i, \xi^{(i)}$ et $X^{(i)}, i = 1, \dots, p$. On écrit $p = r_1 + 2r_2$ où r_1 est le nombre de plongements réels et r_2 le nombre de plongements complexes (non réels) à conjugaison près. On note δ le discriminant du corps K . Si on note Tr la forme trace de l'extension K/\mathbb{Q} , rappelons que $\delta = \det(\text{Tr}(v_i v_j))_{i,j}$, où $(v_i)_i$ est n'importe quelle \mathbb{Z} -base de \mathcal{O}_K . On a aussi $\delta = \det(\sigma_i(v_j))_{i,j}^2$. On écrit alors $\Delta := 2^{-r_2} |\delta|^{1/2}$. On note encore $F = F_E$ les fonctions distance associées aux normes euclidiennes sur les espaces $\mathbb{R}^{\binom{n}{d}}$. On reprendra un certain nombre des résultats établis dans les paragraphes 2.2 et 2.3.

Soient S^d un sous-espace de K^n , X_1, \dots, X_d des points linéairement indépendants de S^d . On pose $N := \binom{n}{d}$. Notons $Y := (\eta_1, \dots, \eta_N)$ les coordonnées de Grassmann de

S^d obtenues à partir des X_i (cf le paragraphe 2.2). Par définition des coordonnées de Grassmann et par (2.30) on a

$$F(\nu_j(\eta_1), \dots, \nu_j(\eta_N)) = D(X_1^{(j)}, \dots, X_d^{(j)}).$$

En notant \mathfrak{a} l'idéal engendré par les η_1, \dots, η_N , alors on a

$$H(S^d) = N(\mathfrak{a})^{-1} \prod_{j=1}^p D(X_1^{(j)}, \dots, X_d^{(j)}). \quad (4.1)$$

Plongement de K^n dans \mathbb{R}^{np}

Pour tout $\xi \in K$ on suppose que $\xi^{(i)}$ est réel pour $1 \leq i \leq r_1$ et que $\xi^{(r_1+r_2+j)}$ est le complexe conjugué de $\xi^{(r_1+j)}$ pour tout $1 \leq j \leq r_2$. On définit alors

$$\xi^{[i]} = \begin{cases} \xi^{(i)} & \text{si } 1 \leq i \leq r_1 \\ \operatorname{Re} \xi^{(i)} & \text{si } r_1 < i \leq r_1 + r_2 \\ \operatorname{Im} \xi^{(i)} & \text{si } r_1 + r_2 < i \leq p, \end{cases}$$

où Re et Im désignent respectivement les parties réelle et imaginaire. Etant donné $X = (\xi_1, \dots, \xi_n) \in K^n$, on pose

$$X^{[i]} = (\xi_1^{[i]}, \dots, \xi_n^{[i]}) \quad (i = 1, \dots, p).$$

On définit alors

$$\rho(X) = (\xi_1^{[1]}, \dots, \xi_1^{[p]}, \dots, \xi_n^{[1]}, \dots, \xi_n^{[p]}),$$

qu'on interprète comme un point de l'espace euclidien \mathbb{R}^{np} . Remarquons que ρ est un plongement \mathbb{Q} -linéaire. Si Σ est un ensemble de points de K^n , on écrit $\mathcal{O}_K(\Sigma) = \Sigma \cap \mathcal{O}_K^n$, l'ensemble des points de Σ à coordonnées dans \mathcal{O}_K .

Rappelons-nous des définitions introduites dans le paragraphe 2.4, notamment de la définition d'un réseau (qui n'est pas pour nous nécessairement de rang maximal). Soit S^d un sous-espace de K^n de dimension d . On a $\dim_{\mathbb{Q}}(S^d) = dp$, donc $\mathcal{O}_K(S)$ (sous- \mathbb{Z} -module de \mathcal{O}_K^n) et par conséquent $\rho(\mathcal{O}_K(S))$ contiennent exactement dp vecteurs linéairement indépendants. En écrivant alors

$$\Lambda(S) := \rho(\mathcal{O}_K(S)),$$

c'est un réseau de \mathbb{R}^{np} de dimension dp . Rappelons que $d(\Lambda(S))$ désigne son covolume.

Définition 4.2.1. Soit S^d un sous-espace vectoriel de K^n . Si $d = 0$ on pose $H'(S) = 1$, sinon on définit

$$H'(S) = \Delta^{-d} d(\Lambda(S)).$$

Proposition 4.2.2. On a la formule

$$H'(K^n) = 1,$$

ou, de manière équivalente

$$d(\Lambda(K^n)) = \Delta^n.$$

Preuve On suit ici les indications données par Schmidt dans [17] en donnant quelques détails explicites supplémentaires.

Supposons pour commencer que $n = 1$. Dans ce cas

$$\Lambda(S) = \{(\xi^{[1]}, \dots, \xi^{[p]}) \mid \xi \in \mathcal{O}_K\} \subset \mathbb{R}^p.$$

Soit (v_1, \dots, v_p) une base intégrale de K sur \mathbb{Q} (i.e. une \mathbb{Z} -base de \mathcal{O}_K). Alors $(\rho(v_1), \dots, \rho(v_p))$ est une \mathbb{Z} -base de $\Lambda(S)$. On a

$$\delta = \begin{vmatrix} v_1^{(1)} & \dots & v_p^{(1)} \\ \vdots & & \vdots \\ v_1^{(p)} & \dots & v_p^{(p)} \end{vmatrix}^2.$$

Si on note L_1, \dots, L_p les lignes du déterminant de droite et qu'on fait les manipulations (dans l'ordre) :

$$\begin{aligned} L_{r_1+j} &\leftarrow \frac{1}{2}(L_{r_1+j} + L_{r_1+r_2+j}) & j = 1, \dots, r_2 \\ L_{r_1+r_2+j} &\leftarrow -i(L_{r_1+r_2+j} - L_{r_1+j}) & j = 1, \dots, r_2, \end{aligned}$$

on trouve :

$$2^{-2r_2} |\delta| = \begin{vmatrix} v_1^{[1]} & \dots & v_p^{[1]} \\ \vdots & & \vdots \\ v_1^{[p]} & \dots & v_p^{[p]} \end{vmatrix}^2 = D(\rho(v_1), \dots, \rho(v_p))^2 = d(\Lambda(S))^2.$$

Donc $d(\Lambda(S)) = \Delta$. Pour $n \geq 1$, les np vecteurs $\rho(0, \dots, v_j, \dots, 0)$ ($j = 1, \dots, p$) forment une \mathbb{Z} -base de $\Lambda(S) = \rho(\mathcal{O}_K^n)$, et leur matrice est une matrice diagonale par blocs dont chaque bloc est de la forme

$$\begin{pmatrix} v_1^{[1]} & \dots & v_p^{[1]} \\ \vdots & & \vdots \\ v_1^{[p]} & \dots & v_p^{[p]} \end{pmatrix}.$$

On en déduit facilement que

$$d(\Lambda(K^n)) = \Delta^n.$$

□

Théorème 4.2.3. *Soit S un sous-espace vectoriel de K^n . Alors*

$$H(S) = H'(S).$$

Cela nous donne une nouvelle définition plus géométrique de la hauteur.

Preuve La preuve de Schmidt repose sur deux lemmes. Nous proposons pour le lemme 4.2.4 une preuve différente de celle de Schmidt et ne démontrerons pas le lemme 4.2.5 en entier. Le lecteur est renvoyé à [17] (théorème 2) pour les détails (Schmidt utilise d'ailleurs pour achever la démonstration du lemme 4.2.5 un troisième lemme technique que nous passons sous silence).

Soit S^d un sous-espace de K^n . Si $d = 0$ ou $d = n$ on a $H(S) = H'(S) = 1$; dans le premier cas c'est par définition, dans le second cas c'est un corollaire direct de la proposition 4.2.2. On suppose maintenant $0 < d < n$.

Soit $X_1, \dots, X_d \in \mathcal{O}_K(S)$ linéairement indépendants sur K . On désigne par $\mathcal{O}_K(X_1, \dots, X_d)$ le \mathcal{O}_K -module engendré par X_1, \dots, X_d et on note $\Lambda(X_1, \dots, X_d)$ son image par ρ : c'est un réseau de \mathbb{R}^{np} - sous-réseau de $\Lambda(S)$ - de dimension dp . Enfin, on note (η_1, \dots, η_N) les coordonnées de Grassmann de (X_1, \dots, X_d) (où on a posé $N = \binom{n}{d}$).

Lemme 4.2.4. *On a l'égalité*

$$d(\Lambda(X_1, \dots, X_d)) = \Delta^d \prod_{j=1}^p D(X_1^{(j)}, \dots, X_d^{(j)}).$$

Preuve On commence par remarquer que si $(\beta_1, \dots, \beta_p)$ est une base intégrale de K sur \mathbb{Q} alors les $Z_{ji} := \rho(\beta_j X_i)$ ($1 \leq j \leq p$, $1 \leq i \leq d$) forment une base du réseau $\Lambda(X_1, \dots, X_d)$. On a alors

$$d(\Lambda(X_1, \dots, X_d)) = D(Z_{11}, \dots, Z_{p1}, \dots, Z_{1d}, \dots, Z_{pd}).$$

Si on note $X_k := (\xi_{k1}, \dots, \xi_{kn})$ ($1 \leq k \leq d$) et

$$Z'_{ji} = ((\beta_j \xi_{i1})^{(1)}, \dots, (\beta_j \xi_{i1})^{(p)}, \dots, (\beta_j \xi_{in})^{(1)}, \dots, (\beta_j \xi_{in})^{(p)}),$$

on a

$$d(\Lambda(X_1, \dots, X_d)) = 2^{-r_2 d} D(Z'_{11}, \dots, Z'_{p1}, \dots, Z'_{1d}, \dots, Z'_{pd}).$$

Schmidt effectue ensuite des manipulations fines de déterminants et utilise quelques arguments algébriques (il considère les $\beta_i^{(j)}$ comme des variables) pour arriver au résultat (cf [17]). Nous suivrons un raisonnement différent.

Permuter les coordonnées des Z'_{ji} ne change pas le covolume du réseau ainsi obtenu. Nous pouvons donc remplacer les Z'_{ji} par les Y_{ji} où on a posé $Y_{ji} = ((\beta_j X_i)^{(1)}, \dots, (\beta_j X_i)^{(p)})$.

On note M_Y la matrice dont les lignes sont les vecteurs $Y_{11}, \dots, Y_{p1}, \dots, Y_{1p}, \dots, Y_{pd}$. D'après le paragraphe 2.3 on sait que $D(Y_{11}, \dots, Y_{pd})^2 = \det M_Y M_Y^*$. On remarque maintenant que si on pose $D = (\beta_i^{(j)})_{ij} \in \mathcal{M}_p(\mathbb{C})$, qu'on définit D' comme étant la matrice carrée de taille dp diagonale par blocs égaux à D , et enfin qu'on appelle M' la matrice de taille $dp \times np$ dont la $((k-1)d+i)$ -ème ligne ($1 \leq k \leq d$, $1 \leq i \leq p$) est

$$(\delta_{1i} \xi_{k1}^{(i)}, \dots, \delta_{1i} \xi_{kn}^{(i)}, \dots, \delta_{ni} \xi_{k1}^{(i)}, \dots, \delta_{ni} \xi_{kn}^{(i)}) = (0, \dots, 0, \underbrace{\xi_{k1}^{(i)}, \dots, \xi_{kn}^{(i)}}_{i\text{-ème bloc de taille } n}, 0, \dots, 0),$$

alors $M_Y = D' M'$ et par conséquent $\det(M_Y M_Y^*) = |\delta|^d \det(M' (M')^*)$. Pour calculer le déterminant du membre de droite, on permute les lignes de la matrice M' pour obtenir une matrice B diagonale par blocs (plus précisément, on réordonne les lignes de manière à ce que la $((k-1)d+i)$ -ème ligne devienne la $((i-1)p+k)$ -ème ligne. Le lecteur est invité à représenter lui-même les matrices mises en jeu). Le i -ème bloc diagonal de taille $d \times n$ ($1 \leq i \leq p$) est alors la matrice dont la k -ème ligne est $X_k^{(i)}$. On en déduit immédiatement que

$$\det(M' (M')^*) = \det(BB^*) = \left(\prod_{j=1}^d D(X_1^{(j)}, \dots, X_d^{(j)}) \right)^2,$$

et finalement

$$d(\Lambda(X_1, \dots, X_d))^2 = 2^{-2r_2 d} |\delta|^d \left(\prod_{j=1}^d D(X_1^{(j)}, \dots, X_d^{(j)}) \right)^2.$$

□

Lemme 4.2.5. *L'indice de $\Lambda(X_1, \dots, X_d)$ dans $\Lambda(S)$ est égal à $N(\mathfrak{a})$, où \mathfrak{a} est l'idéal engendré par η_1, \dots, η_N .*

Preuve Schmidt se ramène à un problème de congruence avec des idéaux de \mathcal{O}_K . Nous ne présenterons ici qu'une esquisse du schéma de preuve, le lecteur est renvoyé à [17] (lemme 5) pour la fin de la preuve.

L'indice de $\Lambda(X_1, \dots, X_d)$ dans $\Lambda(S)$ est égal à l'indice de $\mathcal{O}_K(X_1, \dots, X_d)$ dans $\mathcal{O}_K(S)$. Soit $Y = \sum_{i=1}^d \alpha_i X_i \in \mathcal{O}_K(S)$. Comme chaque mineur de taille d de la matrice dont la k -ème colonne ($k \neq i$) est X_k et la i -ème colonne est Y est dans \mathcal{O}_K , on en déduit que $\alpha_i \eta_j \in \mathcal{O}_K$ pour tous $1 \leq i \leq d$ et $1 \leq j \leq N$. Ainsi, si $\rho \neq 0$ est dans \mathfrak{a} et qu'on écrit $\beta_i := \rho \alpha_i$, on a $\beta_i \in \mathcal{O}_K$, $1 \leq i \leq d$. On est ramené à calculer l'indice de $\rho \mathcal{O}_K(X_1, \dots, X_d)$ dans $\rho \mathcal{O}_K(S)$. Si on note Φ l'isomorphisme de S^d vers K^d qui à un vecteur X associe ses coordonnées dans la base (X_1, \dots, X_d) , le travail précédent nous assure que l'image par Φ de $\rho \mathcal{O}_K(X_1, \dots, X_d)$ est égale à $(\rho \mathcal{O}_K)^d$ et l'image de $\rho \mathcal{O}_K(S)$ est incluse dans $\rho \mathcal{O}_K^d$. En transportant le problème par Φ , on est ramené à trouver le nombre de classes $(\beta_1, \dots, \beta_d) + (\rho \mathcal{O}_K)^d$ avec $(\beta_1, \dots, \beta_d)$ dans $\Phi(\rho \mathcal{O}_K(S))$, en d'autres termes on cherche le nombre de d -uplets $(\beta_1, \dots, \beta_d)$ à coordonnées dans $\mathcal{O}_K \pmod{(\rho)}$ (où (ρ) est l'idéal engendré par ρ) tels que le vecteur $\beta_1 X_1 + \dots + \beta_d X_d$ soit ρ fois un élément de $\mathcal{O}_K(S)$. Il s'agit donc de calculer le nombre de solutions de l'équation

$$\sum_{i=1}^d \beta_i X_i \equiv 0 \pmod{\rho},$$

et de montrer qu'il est égal à $N(\mathfrak{a})$. En décomposant les idéaux \mathfrak{a} et (ρ) en produit d'idéaux premiers

$$\mathfrak{a} = \mathfrak{p}_s^{e_1} \dots \mathfrak{p}_s^{e_s}, \quad (\rho) = \mathfrak{p}_s^{f_1} \dots \mathfrak{p}_s^{f_s} \quad (e_i \leq f_i),$$

par multiplicativité de la norme et par le lemme chinois, il suffit de montrer que le nombre de solutions entières $(\beta_1, \dots, \beta_d)$ distinctes modulo $\mathfrak{p}_i^{f_i}$ de

$$\beta_1 \xi_{1k} + \dots + \beta_d \xi_{dk} \equiv 0 \pmod{\mathfrak{p}_i^{f_i}} \quad (k = 1, \dots, n),$$

est égal à $N(\mathfrak{p}_i)^{f_i}$.

Schmidt démontre un résultat un peu plus général sur le nombre de solutions de problèmes de congruences similaires (le Lemme 6 section 3 de [17]) pour conclure. \square

Le lemme 4.2.5 implique que $d(\Lambda(S)) = N(\mathfrak{a})^{-1} d(\Lambda(X_1, \dots, X_d))$ et on conclut avec le lemme 4.2.4 et l'égalité (4.1). \square

Grâce à cette nouvelle définition, nous sommes en mesure de démontrer le résultat suivant.

Théorème 4.2.6. *Soit S^d un sous-espace de K^n .*

(a) *Soit $0 \leq d < e \leq n$. Alors il existe un sous-espace S^e tel que*

$$S^d \subset S^e \quad \text{et} \quad H(S^e) \leq c_1 H(S^d)^{(n-e)/(n-d)}.$$

(b) *Soit $0 \leq f < d \leq n$. Alors il existe un sous-espace S^f tel que*

$$S^f \subset S^d \quad \text{et} \quad H(S^f) \leq c_1 H(S^d)^{f/d}.$$

Ici c_1 est une constante > 0 qui ne dépend que de n et K .

Preuve La démonstration proposée ici est celle de Schmidt [17] (théorème 2). L'ordre des arguments a parfois été un peu changé.

Commençons avec le cas (a). Par récurrence on voit qu'il suffit de traiter le cas $e = d+1$. Dans cette preuve c_2, \dots seront toutes des constantes > 0 qui ne dépendent que de n et K .

Le réseau $\Lambda(S^d)$ est un réseau de \mathbb{R}^{np} de dimension dp et covolume $H(S^d)\Delta^d$ (par le théorème 4.2.3). Soit $F := \text{Vect}(\Lambda(S))$ et $E^{p(n-d)} := F^\perp$. Soit Λ' l'image de $\Lambda(K^n)$ par la projection orthogonale sur $E^{p(n-d)}$. Alors Λ' est un réseau de covolume

$$d(\Lambda') = d(\Lambda(K^n))/d(\Lambda(S^d)) = \Delta^{n-d}H(S^d)^\perp.$$

En effet, on a

$$d(\Lambda(K^n)) = d(\Lambda')d(\Lambda(K^n) \cap F),$$

et comme $\Lambda(S^d)$ est saturé dans $\Lambda(K^n)$, on a $\Lambda(K^n) \cap F = \Lambda(S^d)$.

On note $V(m)$ le volume de la boule unité de \mathbb{R}^m . La boule de $E^{p(n-d)}$ de centre 0 et de rayon ρ a pour volume $V(p(n-d))\rho^{p(n-d)}$. Par le premier théorème de Minkowski il existe un point G' non nul de Λ' de norme

$$\|G'\| \leq c_2 H(S^d)^{-\frac{1}{p(n-d)}}. \quad (4.2)$$

C'est la projection orthogonale d'un $G = \rho(X) \in \Lambda(K^n)$, où G (resp. X) n'est pas dans F (resp. S^d). On définit alors S^{d+1} par

$$S^{d+1} = S^d \oplus \text{Vect}(X).$$

Montrons que S^{d+1} convient. Pour montrer cela, par le théorème 4.2.3, il suffit de montrer que

$$d(\Lambda(S^{d+1})) \leq c_3 H(S^d)^{(n-d-1)/(n-d)}.$$

Si on appelle Λ^* le sous-réseau de $\Lambda(S^{d+1})$ engendré par $\Lambda(S^d)$ et $\rho(\beta_1 X), \dots, \rho(\beta_p X)$ (où on a pris $(\beta_1, \dots, \beta_p)$ une base intégrale de K/\mathbb{Q}), alors on peut se contenter de montrer

$$d(\Lambda^*) \leq c_3 H(S^d)^{(n-d-1)/(n-d)}.$$

On pose maintenant $G_j := \rho(\beta_j X)$ et on désigne par G'_j la projection de G_j sur $E^{p(n-d)}$ ($j = 1, \dots, p$). Soit H_1, \dots, H_{pd} une base de $\Lambda(S^d)$. Alors $H_1, \dots, H_{pd}, G_1, \dots, G_p$ est une base de Λ^* et on a

$$\begin{aligned} d(\Lambda^*) &= D(H_1, \dots, H_{pd}, G_1, \dots, G_p) = D(H_1, \dots, H_{pd}, G'_1, \dots, G'_p) \\ &\leq D(H_1, \dots, H_{pd}) \prod_{j=1}^p \|G'_j\| \quad (\text{par la proposition 2.3.6}) \\ &= \Delta^d H(S^d) \prod_{j=1}^p \|G'_j\|. \end{aligned}$$

Il s'agit donc de majorer judicieusement la norme des G'_j .

Par définition de G , $(G - G') \in F$ et comme le \mathbb{Q} -espace vectoriel $\rho(S^d)$ est dense dans F , par l'inégalité (4.2) il existe $X^* \in S^d$ tel que

$$\|\rho(X - X^*)\| = \|G - \rho(X^*)\| \leq 2c_2 H(S^d)^{-\frac{1}{p(n-d)}}.$$

Or, il est clair qu'il existe c_4 telle que $\|\rho(\beta_j Y)\| \leq c_4 \|\rho(Y)\|$ pour tout Y , donc on peut tirer de l'inégalité précédente les inégalités

$$\|\rho(\beta_j(X - X^*))\| = \|G_j - \rho(\beta_j X^*)\| \leq c_5 H(S^d)^{-\frac{1}{p(n-d)}} \quad (j = 1, \dots, p).$$

Mais maintenant, comme $\beta_j X^* \in S^d$, on a $\rho(\beta_j X^*) \in F$ et par conséquent

$$\|G'_j\| \leq c_5 H(S^d)^{-\frac{1}{p(n-d)}} \quad (j = 1, \dots, p).$$

Finalement on en déduit que

$$d(\Lambda^*) \leq \Delta^d H(S^d) \prod_{j=1}^p \|G'_j\| \leq c_6 H(S^d) H(S^d)^{-\frac{1}{n-d}} = c_6 H(S^d)^{(n-d-1)/(n-d)},$$

ce qui achève la démonstration du point (a).

Le cas (b) est prouvé par dualité : $T := (S^d)^{\perp, \varphi}$ est de dimension $n - d$ et par (a) est contenu dans un sous-espace T' de dimension $n - f$ avec $H(T') \leq c_1 H(T)^{f/d} = c_1 H(S^d)^{f/d}$. Alors $S^f := (T')^{\perp, \varphi}$ a les propriétés requises. \square

4.3 Nombre de sous-espaces de hauteur $\leq H$

Cette partie reprend la partie 5 de [17]. On conserve les notations du paragraphe précédent pour K . Le résultat principal de ce paragraphe est le théorème 4.3.3 qui donne le comportement asymptotique du nombre de sous-espaces de dimension d et de hauteur inférieure ou égale à H en fonction de H .

Lemme 4.3.1. *Soit $0 \leq d \leq n$. Alors $H(S^d) \geq 1$. De plus, il y a exactement $\binom{n}{d}$ sous-espaces $S^d \subset K^n$ de hauteur $H(S^d) = 1$: ce sont les sous-espaces S^d engendrés par d vecteurs distincts de la base canonique de K^n .*

Preuve La preuve est laissée au lecteur par Schmidt, nous en proposons une ici.

Notant X_1, \dots, X_d une base de S^d , la formule (4.1) utilisée avec la proposition 2.3.4 donne :

$$N(\mathbf{a})^2 H(S^d)^2 = \prod_{j=1}^p \left(\sum_{i=1}^N |\eta_i^{(j)}|^2 \right) \geq \sum_{i=1}^N \left| \prod_{j=1}^p \eta_i^{(j)} \right|^2 = \sum_{i=1}^N N(\eta_i)^2,$$

où on a noté $N = \binom{n}{d}$. Comme $N(\mathbf{a}) \leq N(\eta_i)$ pour tout i tel que $\eta_i \neq 0$, on en déduit que

$$N(S^d) \geq 1,$$

avec égalité si et seulement si $N(\eta_i) = 0$ pour tout i sauf un unique i_0 pour lequel $N(\eta_{i_0}) = 1$. L'équivalence (2.26) permet de conclure que S^d est égal à $\text{Vect}(e_{i_1}, \dots, e_{i_d})$ où (i_1, \dots, i_d) est le i_0 -ème d -uplet quand on les a ordonnés par l'ordre lexicographique. \square

Définition 4.3.2. On note $N(n, d, K, H)$ ou plus simplement $N(d, H)$ le nombre de sous-espaces de dimension d de K^n de hauteur inférieure ou égale à H .

Théorème 4.3.3. *Soit $H \geq 1$ et $1 \leq d < n$. Alors il existe des constantes c_6 et c_7 qui ne dépendent que de n et K telles que*

$$c_6 H^n \leq N(d, H) \leq c_7 H^n. \quad (4.3)$$

Preuve Le lecteur est renvoyé à [17] (théorème 3) pour la démonstration de ce théorème. La preuve se fait par récurrence sur la dimension d . L'initialisation découle d'un article de Schanuel [16] (qui fournit en réalité un résultat bien plus précis). Notons toutefois que le Lemme 10 de [17] avec $d = 1$ assure déjà que $N(1, H) \leq c_7 H^n$. \square

5 Approximations diophantiennes

Cette partie reprend la partie III de [17]. Les résultats principaux d'approximations diophantiennes sont résumés dans le paragraphe 5.3; les paragraphes 5.1 et 5.2 sont une préparation technique nécessaire pour les preuves desdits résultats. Enfin, le paragraphe 5.4 montre dans quelle mesure les inégalités obtenues sont optimales. On distinguera deux cas (a) et (b).

Cas (a) : G^n désigne un espace euclidien E^n , $q = 1$ et K est un corps de nombres réel (on voit donc K comme un sous-corps de \mathbb{R}).

Cas (b) : $q = 2$, G^n désigne un espace hermitien U^n et K est un corps de nombres complexe non réel (on voit \mathbb{K} comme un sous-corps de \mathbb{C}).

Dans les deux cas, C_1, C_2, \dots désignent des constantes > 0 qui ne dépendent que de K , n (mais qui sont indépendantes des sous-espaces A^d, \dots considérés). On utilisera les notations et notions introduites dans la partie 3.

5.1 Travail préliminaire

Le seul résultat de ce paragraphe - qui reprend la section 10 de [17] - est le théorème 5.1.1 dont nous nous efforçons de fournir une preuve dans le cas (b).

Théorème 5.1.1. *Soit $0 < d < n$, $c := n - d$ et $u := \min(c, d)$. Soit A^d un sous-espace de G^n et $H \geq 1$. Alors il existe des sous-espaces $B^1 \subset \dots \subset B^u$ de G^n , chacun défini sur K , tels que pour $i = 1, \dots, u$*

$$(1) \quad H(B^i) \leq H^i$$

$$(2) \quad H(B^1)\psi_i^q(A^d, B^i) \leq C_1 H^{-d/c} \leq C_1 H(B^i)^{-d/(ic)}$$

Preuve Dans [17] (théorème 8), Schmidt ne démontre le théorème que dans le cas (a). Ici nous allons au contraire traiter en détails le cas (b) dont il ne fournit qu'un schéma de preuve.

On note $\sigma_1, \dots, \sigma_p$ les plongements de K dans \mathbb{C} et $\xi^{(i)} = \sigma_i(\xi)$ pour $\xi \in K$. On écrit $p = r_1 + 2r_2$ et on suppose que $\xi^{(2j)}$ est le complexe conjugué de $\xi^{(2j-1)}$ pour $1 \leq j \leq r_2$ et que $\xi^{(2r_2+j)}$ est réel pour $1 \leq j \leq r_1$, pour tout $\xi \in K$. Puisque l'on est dans le cas (b), on peut supposer que σ_1 est l'identité Id_K (σ_2 est donc son complexe conjugué). Puisque $\mathbb{K} \subset \mathbb{C}$, on peut prolonger σ_1 (resp. σ_2) en l'identité de \mathbb{C} (resp. la conjugaison complexe) de sorte que $\xi^{(1)}$ (resp. $\xi^{(2)}$) ait un sens pour tout $\xi \in \mathbb{C}$.

On note $\xi \mapsto \xi^{[i]}$ l'application $K \rightarrow \mathbb{R}$ définie par

$$\xi^{[i]} = \begin{cases} \text{Re } \xi^{(i)} & \text{si } 1 \leq i \leq 2r_2 \text{ et } i \text{ impair,} \\ \text{Im } \xi^{(i)} & \text{si } 1 \leq i \leq 2r_2 \text{ et } i \text{ pair} \\ \xi^{(i)} & \text{si } 2r_2 + 1 \leq i \leq p \end{cases}$$

Comme on a prolongé σ_1 (resp. σ_2) à \mathbb{C} , $\xi^{[1]}$ (resp. $\xi^{[2]}$) a un sens pour tout $\xi \in \mathbb{C}$. Etant donné un vecteur $X = (\xi_1, \dots, \xi_n) \in K^n$, on pose $X^{(i)} := (\xi_1^{(i)}, \dots, \xi_n^{(i)}) \in \mathbb{C}^n$

$X^{[i]} := (\xi_1^{[i]}, \dots, \xi_n^{[i]})$. Remarquons que $X = X^{[1]} - iX^{[2]}$ et que cette égalité fait sens pour tout $X \in \mathbb{C}^n$. L'anneau \mathcal{O}_K désigne toujours l'anneau des entiers de K . Si $X \in K^n$, on note également

$$\mathfrak{X} = \rho(X) = (X^{[1]}, \dots, X^{[p]}) \in E^{np}. \quad (5.1)$$

Notez que la fonction ρ définie ici est égale à la fonction ρ du paragraphe 4.2 à une permutation des coordonnées près (de même pour les fonctions $\xi \mapsto \xi^{[i]}$), ce qui ne change pas les propriétés et les résultats obtenus dans ce précédent paragraphe. On note encore $\Lambda = \Lambda(K^n)$ l'ensemble des $\rho(X)$ tels que $X \in \mathcal{O}_K^n$. C'est un réseau de dimension np et de déterminant Δ^n (cf paragraphe 4.2).

Soient B^1, \dots, B^u tels que pour tout i $B^i = \text{Vect}_{\mathbb{C}}(X_1, \dots, X_i)$ avec $X_i = (\xi_{i1}, \dots, \xi_{in})$ et $\xi_{ij} \in K$. L'idéal fractionnaire \mathfrak{a}_i engendré par les coordonnées de $X_1 \wedge \dots \wedge X_i$ (cf paragraphe 2.2) est un idéal de \mathcal{O}_K donc de norme ≥ 1 . On en déduit par (4.1) que $H(B^i) \leq \prod_{j=1}^i D(X_1^{(j)}, \dots, X_i^{(j)})$. Ainsi, pour montrer le point (1) du théorème, il suffit de montrer

$$D(X_1^{(j)}, \dots, X_i^{(j)}) \leq H^{i/p} \quad (1 \leq i \leq u, 1 \leq j \leq p). \quad (5.2)$$

Un vecteur $X \in U^n$ s'écrit de manière unique $X = X^A + X^*$ avec $X^A \in A$ et $X^* \in A^\perp$ (rappelons que A^\perp est l'orthogonal de A dans U^n pour le produit scalaire hermitien; c'est un sous-espace de dimension c). Nous allons construire par récurrence des vecteurs $\mathfrak{X}_1 = \rho(X_1), \dots, \mathfrak{X}_u = \rho(X_u) \in \Lambda \setminus \{0\}$ vérifiant :

$$\begin{aligned} (i) \quad & \|X_j^*\| \leq C_2 H^{-n/(2c)+1/p} \quad (j = 1, \dots, u) \\ (ii) \quad & |\langle X_i^A, X_j^A \rangle| \leq (8^u u!)^{-1} \|X_i^A\|^2 \quad (1 \leq i < j \leq u) \\ (iii) \quad & D(X_1^A, \dots, X_j^A) = \|X_1^A \wedge \dots \wedge X_j^A\| \leq H^{j/p} 4^{-j} \quad (j = 1, \dots, u) \\ (iv) \quad & \|X_j^{[k]}\| \leq \frac{1}{2} H^{1/p} \quad (j = 1, \dots, u; k = 3, \dots, p) \end{aligned}$$

De plus pour $j = 1, \dots, u$ on va choisir \mathfrak{X}_j tel que $\|X_1^A \wedge \dots \wedge X_j^A\|$ soit minimale. (Rappel : on identifie l'algèbre extérieure $\Lambda^{(j)}(U^n)$ à $U^{\binom{n}{j}}$ comme dans le paragraphe 2.2). C'est pour cette construction qu'il y a le plus de différences entre le cas (a) et le cas (b). Dans le cas (b), Schmidt fournit uniquement les équations (i) – (iv); les calculs sous-jacents sont implicites et prennent la plus grosse partie de la preuve.

Pour $j = 1$ l'inégalité (ii) n'a pas besoin d'être satisfaite et l'inégalité (iii) devient $\|X_1^A\| \leq H^{1/p} 4^{-1}$. Pour $\mathfrak{X} = (X_{11}, \dots, X_{1n}, \dots, X_{p1}, \dots, X_{pn}) \in E^{pn}$ on pose $X := (X_{11} - iX_{21}, \dots, X_{1n} - iX_{2n}) \in U^n$. Les inégalités

$$\begin{aligned} \|X^*\| &\leq C_2 H^{-n/(2c)+1/p} \\ \|X^A\| &\leq H^{1/p} 4^{-1} \\ \|(X_{k1}, \dots, X_{kn})\| &\leq \frac{1}{2} H^{1/p} \quad (k = 3, \dots, p) \end{aligned}$$

définissent un corps convexe symétrique de E^{pn} . Afin de calculer le volume de ce convexe, notons Φ l'isomorphisme \mathbb{R} -linéaire entre E^{2n} et U^n défini par $\Phi(Y, Z) = Y - iZ$ (remarque : $\Phi^{-1}(X) = (X^{[1]}, X^{[2]})$). Ainsi, notant $\tilde{X} := (X_{11}, \dots, X_{1n}, X_{21}, \dots, X_{2n})$, on a $X = \Phi(\tilde{X})$. On pose également $\tilde{X}^A := \Phi^{-1}(X^A)$

et $\tilde{X}^* := \Phi^{-1}(X^*)$. Comme pour n'importe quel sous-espace vectoriel S de U^n on a $\Phi^{-1}(S^\perp) = \Phi^{-1}(S)^\perp$ et que $\|(Y, Z)\| = \|\Phi(Y, Z)\|$ pour tous $Y, Z \in E^n$, les égalités précédentes se réécrivent

$$\begin{aligned} \|\tilde{X}^*\| &\leq C_2 H^{-n/(2c)+1/p} \\ \|\tilde{X}^A\| &\leq H^{1/p} 4^{-1} \\ \|(X_{k1}, \dots, X_{kn})\| &\leq \frac{1}{2} H^{1/p} \quad (k = 3, \dots, p) \end{aligned}$$

avec $\tilde{X}^A \in \Phi^{-1}(A)$ (de dimension réelle $2d$) et $\tilde{X}^* \in \Phi^{-1}(A)^\perp$ (de dimension réelle $2c$). Le volume de ce convexe est donc égal à

$$\begin{aligned} (C_2^{2c} H^{-n+2c/p} V(2c)) \times (4^{-2d} H^{2d/p} V(2d)) \times (2^{-n} H^{n/p} V(n))^{p-2} \\ > C_2^{2c} 4^{-2d} 2^{-n(p-2)} V(2c) V(2d) V(n)^{p-2} > 2^{np} \Delta^n \end{aligned}$$

pour C_2 assez grand en fonction de n et K (où on a noté $V(m)$ le volume de la boule unité de E^m). Par le premier théorème de Minkowski il existe alors un $\mathfrak{X}_1 \in \Lambda$ non nul avec les propriétés requises. On le choisit tel que X_1^A soit de norme minimale.

Supposons maintenant construits $\mathfrak{X}_1, \dots, \mathfrak{X}_{j-1}$ pour un $j \geq 2$. Les inégalités (i) à (iv) en j définissent un convexe symétrique pour \mathfrak{X}_j . Afin d'appliquer le premier théorème de Minkowski, il suffit de montrer que son volume est $> 2^{np} \Delta^n$.

En décomposant $X_j^A = Y_j^A + Y_j^{A,*}$ avec $Y_j^A \in \text{Vect}_{\mathbb{C}}(X_1^A, \dots, X_{j-1}^A)$ et $Y_j^{A,*} \in \text{Vect}_{\mathbb{C}}(X_1^A, \dots, X_{j-1}^A)^\perp \cap A$, les inégalités (i) – (iv) se réécrivent :

$$\begin{aligned} (i') \quad \|X_j^*\| &\leq C_2 H^{-n/(2c)+1/p} \\ (ii') \quad |\langle X_i^A, Y_j^A \rangle| &\leq (8^u u!)^{-1} \|X_i^A\|^2 \quad (1 \leq i \leq j-1) \\ (iii') \quad \|X_1^A \wedge \dots \wedge X_{j-1}^A\| \times \|Y_j^{A,*}\| &\leq H^{j/p} 4^{-j} \\ (iv') \quad \|X_j^{[k]}\| &\leq \frac{1}{2} H^{1/p} \quad (k = 3, \dots, p) \end{aligned}$$

Cet ensemble est le produit de quatre ensembles convexes symétriques de sous-espaces deux à deux orthogonaux de E^{pn} dans lesquels varient respectivement $((X_j^*)^{[1]}, (X_j^*)^{[2]}, 0, \dots, 0)$, $((Y_j^A)^{[1]}, (Y_j^A)^{[2]}, 0, \dots, 0)$, $((Y_j^{A,*})^{[1]}, (Y_j^{A,*})^{[2]}, 0, \dots, 0)$ et les $(0, \dots, 0, \underbrace{X_j^{[k]}}_{\substack{k\text{-ème bloc de} \\ n \text{ coordonnées}}}, 0, \dots, 0)$.

Si X_1^A, \dots, X_{j-1}^A sont linéairement dépendants, (ii') et (iii') définissent un ensemble convexe de dimension (réelle) $2d$ et de volume infini. On peut donc les supposer linéairement indépendants.

L'inégalité (i') définit un ensemble convexe de $\Phi^{-1}(A)^\perp$ (de dimension (réelle) $2c$) de volume $C_2^{2c} H^{-n+2c/p} V(2c)$.

L'inégalité (ii') définit un ensemble convexe de $\Phi^{-1}(\text{Vect}_{\mathbb{C}}(X_1^A, \dots, X_{j-1}^A))$ (de dimension (réelle) $2(j-1)$). Afin de calculer son volume, introduisons Ψ , l'isomorphisme entre E^{2n} et U^n défini par $\Psi(Y, Z) = Z + iY$ (remarque : on a $\Psi = i\Phi$ et $\Psi^{-1}(X) = (-X^{[2]}, X^{[1]})$). Si on note $\tilde{Y}_j^A := \Phi^{-1}(Y_j^A)$, $\tilde{X}_i^A := \Phi^{-1}(X_i^A)$ et $\hat{X}_i^A := \Psi^{-1}(X_i^A)$ ($1 \leq i < j$), l'inégalité précédente se réécrit :

$$\sqrt{\left| \langle \tilde{X}_i^A, \tilde{Y}_j^A \rangle \right|^2 + \left| \langle \hat{X}_i^A, \tilde{Y}_j^A \rangle \right|^2} \leq (8^u u!)^{-1} \|X_i^A\|^2 \quad (1 \leq i \leq j-1).$$

L'ensemble défini par ces inégalités contient le sous-ensemble convexe de $\Phi^{-1}(\text{Vect}_{\mathbb{C}}(X_1^A, \dots, X_{j-1}^A))$ défini par

$$\begin{aligned} \left| \left\langle \tilde{X}_i^A, \tilde{Y}_j^A \right\rangle \right| &\leq (2^{1/2} 8^u u!)^{-1} \|X_i^A\|^2 \quad (1 \leq i \leq j-1) \\ \left| \left\langle \hat{X}_i^A, \tilde{Y}_j^A \right\rangle \right| &\leq (2^{1/2} 8^u u!)^{-1} \|X_i^A\|^2 \quad (1 \leq i \leq j-1), \end{aligned}$$

dont le volume est, puisque $\|\tilde{X}_i^A\| = \|\hat{X}_i^A\| = \|X_i^A\|$,

$$\begin{aligned} 2^{2(j-1)} (2^{1/2} 8^u u!)^{-2(j-1)} \prod_{i=1}^{j-1} \left(\|\tilde{X}_i^A\|^2 \|\hat{X}_i^A\|^2 \right) D(\tilde{X}_1^A, \dots, \tilde{X}_{j-1}^A, \hat{X}_1^A, \dots, \hat{X}_{j-1}^A)^{-1} \\ > (8^u u!)^{-2(j-1)} D(\tilde{X}_1^A, \dots, \tilde{X}_{j-1}^A, \hat{X}_1^A, \dots, \hat{X}_{j-1}^A) \end{aligned}$$

par l'inégalité de Hadamard. Remarquons maintenant que

$$D(\tilde{X}_1^A, \dots, \tilde{X}_{j-1}^A, \hat{X}_1^A, \dots, \hat{X}_{j-1}^A) = D(X_1^A, \dots, X_{j-1}^A)^2.$$

Cela provient de deux faits : d'abord si $Y \in U^n$ et si on pose $\tilde{Y} = \Phi^{-1}(Y)$, $\hat{Y} = \Psi^{-1}(Y)$, alors \tilde{Y} et \hat{Y} sont orthogonaux et $D(\tilde{Y}, \hat{Y}) = D(\tilde{Y})D(\hat{Y}) = D(Y)^2$. Ensuite, si $(Y, Z) \in U^{2n}$ est une paire de vecteurs orthogonaux entre eux, alors il en va de même des paires (\tilde{Y}, \tilde{Z}) et (\hat{Y}, \hat{Z}) . Cela assure que si S est un sous-espace vectoriel de U^n , si \tilde{p} et \hat{p} désignent les projections orthogonales sur S , $\Phi^{-1}(S)$ et $\Psi^{-1}(S)$ respectivement, alors on a $\tilde{p} \circ \Phi^{-1} = \Phi^{-1} \circ p$ et $\hat{p} \circ \Psi^{-1} = \Psi^{-1} \circ p$. On conclut alors en orthogonalisant les familles $(X_1^A, \dots, X_{j-1}^A)$ et $(\tilde{X}_1^A, \dots, \tilde{X}_{j-1}^A, \hat{X}_1^A, \dots, \hat{X}_{j-1}^A)$. Ainsi, le volume du convexe défini par (ii') est plus grand que

$$(8^u u!)^{-2(j-1)} D(X_1^A, \dots, X_{j-1}^A)^2.$$

L'inégalité (iii') définit un ensemble convexe de $\Phi^{-1}(\text{Vect}_{\mathbb{C}}(X_1^A, \dots, X_{j-1}^A))^{\perp} \cap \Phi^{-1}(A)$ (de dimension $2(d-j+1)$) de volume

$$\begin{aligned} (H^{j/p} 4^{-j} D(X_1^A, \dots, X_{j-1}^A)^{-1})^{2(d-j+1)} V(2(d-j+1)) \\ \geq (H^{j/p} 4^{-j} D(X_1^A, \dots, X_{j-1}^A)^{-1})^2 (H^{j/p} 4^{-j} (4^{-(j-1)} H^{(j-1)/p})^{-1})^{2(d-j)} V(2(d-j+1)) \\ \geq H^{2j/p} 4^{-2j} D(X_1^A, \dots, X_{j-1}^A)^{-2} (4^{-1} H^{1/p})^{2(d-j)} V(2(d-j+1)) \\ \geq H^{2d/p} D(X_1^A, \dots, X_{j-1}^A)^{-2} 4^{-2d} V(2(d-j+1)), \end{aligned}$$

où on a utilisé (iii) avec $j-1$ pour passer de la première ligne à la deuxième.

Enfin l'inégalité (iv') définit un ensemble convexe de $E^{n(p-2)}$ de volume

$$(2^{-n} H^{n/p} V(n))^{p-2} > 2^{-np} H^{n-2n/p} V(n)^{p-2}.$$

Ces quatre inégalités montrent que le volume du convexe défini par les quatre équations ensemble est strictement supérieur à $C_2^{2c} \times \lambda$ où λ est une constante > 0 qui ne dépend que de n et p . Donc pour C_2 assez grand, on a bien un volume $> 2^{np} \Delta^n$ et le premier théorème de Minkowski donne \mathfrak{X}_j non nul vérifiant (i) – (iv). On le choisit de sorte que $D(X_1^A, \dots, X_j^A)$ soit minimal, ce qui achève notre construction des $\mathfrak{X}_1, \dots, \mathfrak{X}_u$.

Pour chaque i il existe $X_i \in \mathcal{O}_K^n$ tel que $\mathfrak{X}_i = \rho(X_i)$ ($1 \leq i \leq u$). On pose

$$B^i := \text{Vect}_{\mathbb{C}}(X_1, \dots, X_i).$$

Vérifions que les B^i conviennent. La suite de la preuve est presque identique dans les deux cas (a) et (b); les arguments sont les mêmes, seules les inégalités mises en jeu

sont légèrement différentes.

Montrons pour commencer que B^i est bien de dimension i . Pour cela, il suffit de montrer que les X_k^A sont linéairement indépendants. Afin de montrer la non nullité des X_j^A , montrons que

$$\|X_1^A\| \geq 2^n C_2 H^{-n/(2c)+1/p}, \quad (5.3)$$

pour H assez grand. Si ce n'était pas le cas, en utilisant (i) il existerait des H arbitrairement grands pour lesquels

$$\|X_1\| \leq 2^{n+1} C_2 H^{-n/(2c)+1/p} \quad \text{et donc } |\xi_{1k}|, |\overline{\xi_{1k}}| \leq 2^{n+1} C_2 H^{-n/(2c)+1/p} \quad (1 \leq k \leq n)$$

(en écrivant $X = (\xi_{11}, \dots, \xi_{1n})$).

En utilisant (iv) on a aussi :

$$|\xi_{1k}^{(j)}| \leq H^{1/p} \quad (1 \leq k \leq n; 3 \leq j \leq p).$$

On en déduit qu'il y aurait des H arbitrairement grands pour lesquels

$$N(\xi_{1k}) \leq 2^{2(n+1)} C_2^2 H^{1-n/c} < 1,$$

si H est assez grand, ce qui impliquerait $\xi_{1k} = 0$ ($k = 1, \dots, n$), et donc $X_1 = 0$ et par suite $\mathfrak{X}_1 = 0$ (pour des H arbitrairement grands), ce qui n'est pas. Donc (5.3) est vraie.

Soit $1 \leq k < j \leq u$. Alors \mathfrak{X}_j vérifie les trois inégalités (i), (ii) et (iv) satisfaites par \mathfrak{X}_k ((ii) est plus restrictive pour \mathfrak{X}_j que pour \mathfrak{X}_k). Donc, par minimalité de $D(X_1^A, \dots, X_k^A)$ on a

$$D(X_1^A, \dots, X_{k-1}^A, X_k^A) \leq D(X_1^A, \dots, X_{k-1}^A, X_j^A) \quad (1 \leq k < j \leq u). \quad (5.4)$$

En particulier, pour $k = 1$ on obtient $\|X_1^A\| \leq \|X_j^A\|$. Par (5.3) cela assure la non nullité des $\|X_j^A\|$.

Montrons maintenant par récurrence sur j les inégalités

$$\|X_j^A\| \geq \frac{1}{2} \|X_{j-1}^A\| \quad \text{et } \lambda(X_k^A, X_j^A) \leq (4^u u!)^{-1} \quad (1 \leq k < j \leq u). \quad (5.5)$$

On définit Y_j tel que $X_j^A = \|X_j^A\| Y_j$ ($j = 1, \dots, u$). Par (5.4) appliqué avec $k = j-1$, on trouve après simplification par $\|X_1^A\| \dots \|X_{j-2}^A\|$ que

$$\|X_{j-1}^A\| D(Y_1^A, \dots, Y_{j-2}^A, Y_{j-1}^A) \leq \|X_j^A\| D(Y_1^A, \dots, Y_{j-2}^A, Y_j^A) \leq \|X_j^A\|.$$

Par hypothèse de récurrence et en développant le déterminant (2.28) on a

$$D^2(Y_1^A, \dots, Y_{j-1}^A) \geq 1 - \sum_{\substack{\sigma \in \mathfrak{S}_{j-1} \\ \sigma \neq \text{Id}}} (4^u u!)^{-1} \geq 1 - (j-1)! (4^u u!)^{-1} \geq \frac{1}{4}, \quad (5.6)$$

et donc $\|X_j^A\| \geq \|X_{j-1}^A\|/2$. Par (ii),

$$\lambda(X_k^A, X_j^A) \leq (8^u u!)^{-1} \|X_k^A\| / \|X_j^A\| \leq (8^u u!)^{-1} 2^{j-k} \leq (4^u u!)^{-1} \quad (1 \leq k < j),$$

ce qui achève la démonstration de (5.5). Un petit argument (cf [17], preuve du corollaire du Théorème 7 p. 448-449) permet de montrer que X_1^A, \dots, X_u^A sont linéairement

indépendants. Il en va donc de même de X_1, \dots, X_u et ceci assure que B^i est bien de dimension i .

Montrons maintenant (5.2) (ce qui prouvera l'assertion (1) du théorème). Par (iv), il suffit de le vérifier pour $j = 1$ (et ce sera également vrai pour $j = 2$). Par (i), (5.3) et (5.5),

$$\|X_i^*\| \leq \|X_i^A\| \quad (i = 1, \dots, u),$$

en particulier

$$\|X_i\| \leq 2\|X_i^A\| \quad (i = 1, \dots, u). \quad (5.7)$$

L'inégalité (5.5) et le calcul (5.6) assurent aussi $\|X_1^A\| \dots \|X_i^A\| \leq 2D(X_1^A, \dots, X_i^A)$. Donc par (iii),

$$\begin{aligned} D(X_1, \dots, X_i) &\leq \|X_1\| \dots \|X_i\| \leq 2^i \|X_1^A\| \dots \|X_i^A\| \\ &\leq 2^{i+1} D(X_1^A, \dots, X_i^A) \\ &\leq H^{i/p} 2^{-i+1} \leq H^{i/p}. \end{aligned}$$

D'où l'assertion (1) du théorème.

Reste à prouver l'assertion (2) du théorème (nous détaillons ici un peu plus les arguments de Schmidt). Pour cela, remarquons déjà que $\|X_1\| \leq 2\|X_1^A\| \leq 2^i \|X_i^A\|$ par (5.7) et (5.5), et que $\|X_i\| \omega(A, X_i) = \|X_i^*\| \leq C_2 H^{-n/(2c)+1/p}$ par (i) (remarque : en revenant aux définitions des fonctions ω et λ définies dans la partie 3, on obtient $\lambda(X_i, A^d) = \|X_i^A\|/\|X_i\|$ et $\omega(X_i, A^d) = \|X_i^*\|/\|X_i\|$). En combinant ces deux inégalités et en majorant l'exposant $2i$ qui apparaît par $2n$, on obtient finalement $\|X_1\|^2 \omega^2(A, X_i) \leq 2^{2n} C_2^2 H^{-n/c+2/p}$ ($i = 1, \dots, u$).

Maintenant, si on revient à la définition de la hauteur pour une droite (cf Définition 4.1.3) et puisque X_1 est à coordonnées dans l'anneau des entiers, on a

$$H(B^1) \leq \prod_{k=1}^p \|X_1^{(k)}\| = \|X_1\|^2 \prod_{k=3}^p \|X_1^{(k)}\|.$$

Or par (iv) on a pour $k = 3, \dots, p$ l'inégalité $\|X_1^{(k)}\| \leq H^{1/p}$. En combinant les estimations précédentes on trouve finalement, puisque $n - c = d$, $H(B^1) \omega^2(A^d, X_i) \leq 2^{2n} C_2^2 H^{-d/c}$ ($i = 1, \dots, u$). Pour achever la démonstration de la deuxième assertion du théorème on va utiliser le corollaire du théorème 3.3.7. Les inégalités que l'on vient de montrer nous assurent l'existence de $Y_j \in A^d$ non nul tel que $\omega(X_j, Y_j)^2 \leq 2^{2n} C_2^2 H^{-d/c} H(B^1)^{-1}$ pour $j = 1, \dots, u$. L'inégalité (5.5) implique quant à elle que $\lambda(X_j, X_k) \leq 4^{-u}$ (avec $1 \leq j < k \leq u$). Toutes les conditions du corollaire mentionné sont vérifiées avec $\omega = 2^n C_2 H^{-d/(2c)} H(B^1)^{-1/2}$ et l'on conclut facilement en remarquant qu'ici $\psi_i(A^d, B^i) = \omega_i(A^d, B^i)$ pour $i = 1, \dots, u$ (puisque $d + i \leq d + u \leq n$, cf §3.2). □

5.2 Théorèmes du Going-up et du Going-down

Ce paragraphe reprend la section 11 de [17]. Il contient entre autres les théorèmes du going-up et du going-down. Le deuxième est beaucoup plus technique que le premier, nous le prouvons dans le cas (b).

Théorème 5.2.1 (Going-up).

Soient $d + e < n$, A^d , B^e des sous-espaces de G^n et $H \geq 1$ et $c > 0$ fixés. On suppose B^e défini sur K et de hauteur $H(B^e) \leq H$. On suppose de plus que pour $i = 1, \dots, t = \min(d, e)$ existent des constantes $x_i, y_i \geq 0$ telles que

$$H(B^e)^{x_i} \psi_i(A^d, B^e) \leq cH^{-y_i}.$$

Alors il existe des constantes $C_3 = C_3(n, e, K)$ et $C_4 = C_4(n, e, K, x_i, y_i)$ telles qu'en posant $H' := C_3 H^{(n-e-1)/(n-e)}$ il existe un sous-espace $B^{e+1} \supset B^e$ défini sur K , de hauteur $H(B^{e+1}) \leq H'$ et satisfaisant

$$H(B^{e+1})^{x_i(n-e)/(n-e-1)} \psi_i(A^d, B^{e+1}) \leq cC_4 H'^{-y_i(n-e)/(n-e-1)}. \quad (5.8)$$

Preuve La preuve de ce théorème est celle de Schmidt [17] (théorème 9).

On définit C_3 comme étant la constante c_1 du théorème 4.2.6. D'après la première assertion de ce théorème il existe $B^{e+1} \supset B^e$ défini sur K , de hauteur $H(B^{e+1}) \leq C_3 H(B^e)^{(n-e-1)/(n-e)} \leq H'$. Par le corollaire 3.1.8 on a aussi les inégalités $\psi_i(A^d, B^{e+1}) \leq \psi_i(A^d, B^e)$ ($i = 1, \dots, t$). En combinant ces inégalités on a le résultat escompté pour un C_4 convenable. \square

Théorème 5.2.2 (Going-down).

Soient A^d , B^e des sous-espaces de G^n et $H \geq 1$. On suppose B^e défini sur K et de hauteur $H(B^e) \leq H$. Soient $1 \leq h \leq f' = \min(d, e-1)$, $c \geq 1$ et $y_1 \geq \dots \geq y_h \geq (qh)^{-1}$ (rappel : q a été défini au début de la partie 5). On suppose

$$H(B^e) \omega_i^q(A^d, B^e) \leq c^q H^{-(qy_i-1)} \quad (i = 1, \dots, h). \quad (5.9)$$

On pose $y := y_1 + \dots + y_h$ et on suppose que

$$y'_i := y_i e(qy + e - 1)^{-1} \geq q^{-1} \quad (i = 1, \dots, h). \quad (5.10)$$

Alors il existe un sous-espace $B^{e-1} \subset B^e$ défini sur K , de hauteur

$$H(B^{e-1}) \leq C_5 H(B^e) H^{(qy-1)/e} \leq C_5 H^{(e+qy-1)/e} =: H'$$

et satisfaisant à la relation

$$H(B^{e-1}) \omega_i^q(A^d, B^{e-1}) \leq C_6 c^q H^{-(qy'_i-1)(qy+e-1)/e} = C_7 c^q H'^{-(qy'_i-1)} \quad (i = 1, \dots, h),$$

et donc (puisque $-(qy'_i-1) \leq 0$)

$$\omega_i(A^d, B^{e-1}) \leq C_8 c H(B^{e-1})^{-y'_i} \quad (i = 1, \dots, h).$$

Si à la place de (5.9) on a

$$\omega_i(A^d, B^e) = 0 \quad (i = 1, \dots, h), \quad (5.11)$$

on pose

$$y'_0 := e(qh)^{-1}.$$

Alors pour tout $H' \geq C_9 H$, il existe un sous-espace $B^{e-1} \subset B^e$ défini sur K , de hauteur $H(B^{e-1}) \leq H'$ et satisfaisant

$$H(B^{e-1})\omega_i^q(A^d, B^{e-1}) \leq C_{10} H^{qy'_0} H'^{-(qy'_0-1)} \quad (i = 1, \dots, h),$$

et donc (puisque $-(qy'_0 - 1) \leq 0$)

$$\omega_i(A^d, B^{e-1}) \leq C_{11} H^{y'_0} H(B^{e-1})^{-qy'_0} \quad (i = 1, \dots, h).$$

Dans ce théorème les constantes C_5, \dots, C_{11} dépendent uniquement de n, K, y_1, \dots, y_h mais pas de A^d, B^e et H .

Preuve Nous ne traitons ici que le cas (b) (K est donc un sous-corps de \mathbb{C} non réel), le lecteur est invité à consulter [17] (théorème 10) pour le cas (a). La différence majeure entre les deux cas est le fait que dans le cas (a) la forme bilinéaire φ du paragraphe 2.2 coïncide avec le produit scalaire, ce qui n'est plus vrai dans notre cas. Il est à noter qu'une partie de la preuve de Schmidt est fautive à cause d'une confusion entre la forme bilinéaire et le produit scalaire hermitien (Schmidt a une même notation pour ces deux formes et jongle une bonne partie de la preuve avec les deux).

Ici, Δ est défini comme au paragraphe 4.2 (avec K). Dans cette preuve, on travaillera avec le corps \overline{K} défini comme le conjugué complexe de K (attention, ici cette notion ne renvoie pas à la clôture algébrique de K).

Supposons pour commencer que (5.9) est vérifié. On pose $m = n - e$ et $B^{e,\perp} := (B^e)^\perp = \text{Vect}(Z_1, \dots, Z_m)$ avec $Z_i \in \overline{K}^n$ (attention, comme on a pris l'orthogonal hermitien ce sous-espace est désormais défini sur le conjugué complexe \overline{K} de K , a priori distinct de K). Nous allons construire un vecteur $W \in \overline{K}^n \setminus (B^e)^\perp$ et nous vérifierons que

$$B^{e-1} := \text{Vect}(W, Z_1, \dots, Z_m)^\perp, \quad (5.12)$$

qui est défini sur K , possède toutes les propriétés requises.

Notons en premier lieu que pour tout sous-espace S défini sur K , on a $H(S) = H(S^\perp)$. Ceci vient directement du fait que $H(S) = H(\overline{S})$, du théorème 4.1.7 et du fait que $\overline{S}^\perp = S^{\perp, \varphi}$.

On pose $\lambda_i := \lambda_i(A^d, B^e)$ (pour $i = 1, \dots, f := \min(d, e)$) et on choisit (X_1, \dots, X_d) , (Y_1, \dots, Y_e) bases orthonormées respectives de A^d et B^e vérifiant $\langle X_i, Y_j \rangle = \delta_{ij} \lambda_i$, l'existence de ces deux bases étant fournie par le théorème 3.1.2. Ainsi pour $i = 1, \dots, f$ on a $\omega(X_i, Y_j) = \omega_i(A^d, B^e)$. Notons que $\langle Y_i, Z_j \rangle = 0$ ($i = 1, \dots, e$ et $j = 1, \dots, m$).

On va utiliser les résultats du paragraphe 4.2 avec le corps de nombres \overline{K} et le plongement ρ défini comme dans la preuve du théorème 5.1.1. Pour cette preuve uniquement, les p plongements $\sigma_1, \dots, \sigma_p$ seront ceux de \overline{K} dans \mathbb{C} (et l'on utilisera les notations $Y^{(i)}$, $Y^{[i]}$, ... rattachés à ces plongements). Le réseau $\Lambda(B^{e,\perp}) \subset E^{pm}$ (construit à partir de \overline{K} et ρ) est un réseau de dimension pm et de déterminant $\Delta^m H(B^{e,\perp}) = \Delta^m H(B^e)$ (puisque K et \overline{K} ont même discriminant en valeur absolue). On choisit $\mathfrak{J}_1, \dots, \mathfrak{J}_{pm}$ une base de ce réseau et on note Π l'ensemble des vecteurs $\mathfrak{J} = \sum c_i \mathfrak{J}_i$ avec $c_i \in [-1/2, 1/2]$. Cet ensemble a pour volume $\Delta^m H(B^e)$ (lorsqu'on le voit dans l'espace vectoriel de dimension pm qu'il engendre) et ne contient aucun vecteur du réseau $\Lambda := \Lambda(\overline{K}^n)$ à l'exception du vecteur nul. On note enfin $S^* := \text{Vect}(\Lambda(B^{e,\perp})) \subset E^{pn}$.

Construction de W

Comme précisé plus haut, on suppose que ρ est défini comme dans la preuve du théorème 5.1.1 avec \overline{K} à la place de K . Rappelons en particulier que σ_1 est l'identité de \overline{K} et σ_2 son conjugué complexe. Nous les prolongeons en l'identité et la conjugaison complexe de \mathbb{C} de sorte que $Y^{[1]}$ et $Y^{[2]}$ ont un sens pour n'importe quel vecteur $Y \in G^n$. Si $Y \in G^n$ et $Z \in \overline{K}^n$, on a

$$\langle Y, Z \rangle = \left\langle (Y^{[1]}, Y^{[2]}, 0, \dots, 0), \rho(Z) \right\rangle + i \left\langle (-Y^{[2]}, Y^{[1]}, 0, \dots, 0), \rho(Z) \right\rangle.$$

Comme $(Y_j)_j$ est une famille orthonormale et que chacun de ses vecteurs est orthogonal à $B^{e,\perp}$, on en déduit que les vecteurs

$$\mathfrak{Y}_j^1 := (Y_j^{[1]}, Y_j^{[2]}, 0, \dots, 0) \quad (j = 1, \dots, h),$$

et

$$\mathfrak{Y}_j^2 := (-Y_j^{[2]}, Y_j^{[1]}, 0, \dots, 0) \quad (j = 1, \dots, h),$$

forment une famille orthonormée de E^{pn} orthogonale à S^* . On pose alors $T_j := \text{Vect}(\mathfrak{Y}_1^1, \mathfrak{Y}_1^2, \dots, \mathfrak{Y}_j^1, \mathfrak{Y}_j^2)$ ($j = 1, \dots, h$), sous-espace de dimension $2j$.

Un vecteur $\mathfrak{X} \in E^{pn}$ se décompose de manière unique

$$\mathfrak{X} = \mathfrak{X}^* + \mathfrak{X}_T + \mathfrak{X}_0,$$

avec $\mathfrak{X}^* \in S^*$, $\mathfrak{X}_T \in T_h$ et \mathfrak{X}_0 orthogonal à S^* et à T_h . L'ensemble des \mathfrak{X} vérifiant

$$(i) \quad \mathfrak{X}^* \in \Pi$$

$$(ii) \quad |\langle \mathfrak{X}_T, \mathfrak{Y}_j^i \rangle| \leq H^{-(y_j - (2y-1)/(ep))} (H/H(B^e))^{1/(2h)} \quad (j = 1, \dots, h; i = 1, 2)$$

$$(iii) \quad \|\mathfrak{X}_0\| \leq C_{12} H^{(2y-1)/(ep)}$$

est un corps convexe symétrique qui est le produit de trois corps convexes symétriques appartenant à des sous-espaces orthogonaux entre eux; son volume est donc le produit des volumes de ces trois convexes et est égal à :

$$\begin{aligned} \Delta^m H(B^e) \times \left(\prod_{j=1}^h (2H^{-(y_j - (2y-1)/(ep))} (H/H(B^e))^{1/(2h)})^2 \right) \times \\ \times (C_{12} H^{(2y-1)/(ep)})^{pe-2h} V(pe-2h) \end{aligned}$$

(rappel : on a noté $V(l)$ le volume de la boule unité de E^l , c'est-à-dire à $\Delta^m 4^h V(pe-2h) C_{12}^{ep-2h} > 2^{pn} \Delta^n$ si C_{12} est assez grand. Par le premier théorème de Minkowski il existe ainsi un $\mathfrak{X} \in \Lambda$ non nul dans cet ensemble. Soit $W \in \mathcal{O}_{\overline{K}}^n$ tel que $\mathfrak{X} = \rho(W)$).

Nous allons maintenant établir un certain nombre de propriétés de W dans le but de montrer que B^{e-1} défini par (5.12) possède toutes les propriétés requises. Notamment, il nous faut contrôler $|\langle W, Y_j \rangle|$ et $\|V_j\|$ (où V_j est la projection orthogonale de $W^{(j)}$ sur $\text{Vect}(Z_1^{(j)}, \dots, Z_m^{(j)})^\perp$) car ces quantités interviendront directement dans le calcul de la hauteur $H(B^{e-1})$ de B^{e-1} .

Quel que soit $1 \leq j \leq h$ on a

$$\begin{aligned} |\langle Y_j, W \rangle| &= |\langle \mathfrak{Y}_j^1, \rho(W) \rangle + i \langle \mathfrak{Y}_j^2, \rho(W) \rangle| = |\langle \mathfrak{Y}_j^1, \mathfrak{X} \rangle + i \langle \mathfrak{Y}_j^2, \mathfrak{X} \rangle| \\ &\leq |\langle \mathfrak{Y}_j^1, \mathfrak{X}_T \rangle| + |\langle \mathfrak{Y}_j^2, \mathfrak{X}_T \rangle| \\ &\leq 2H^{-(y_j - (2y-1)/(ep))} (H/H(B^e))^{1/(2h)}, \end{aligned} \tag{5.13}$$

par (ii).

Notons aussi que (ii) et (iii) impliquent

$$\|\mathfrak{X} - \mathfrak{X}^*\| \leq C_{14} H^{(2y-1)/(ep)},$$

puisque par hypothèse $y_i \geq 1/(2h)$ (on peut majorer brutalement $H^{-(y_j - (2y-1)/(ep))} \times (H/H(B^e))^{1/(2h)}$ par $H^{(2y-1)/(ep)}$).

Pour $1 \leq j \leq p$, on écrit $W^{(j)} = U_j + V_j$ avec $U_j \in \text{Vect}(Z_1^{(j)}, \dots, Z_m^{(j)})$ et V_j orthogonal à $\text{Vect}(Z_1^{(j)}, \dots, Z_m^{(j)})$ (remarque : c'est entre autres ici qu'il y a des confusions entre bilinéaire et hermitien dans [17] et que les arguments de Schmidt ne fonctionnent pas exactement de la façon dont ils sont écrits).

Si maintenant σ_j est réel, on pose $\mathfrak{W}_j := (0, \dots, \underbrace{V_j}_{j\text{-ème bloc}}, \dots, 0)$. Alors \mathfrak{W}_j est orthogonal à S^* , donc à \mathfrak{X}^* , et à $\mathfrak{X} - \mathfrak{W}_j$ (par définition de V_j et U_j). On en déduit que $\|\mathfrak{W}_j\|^2 = |\langle \mathfrak{W}_j, \mathfrak{X} \rangle| = |\langle \mathfrak{W}_j, \mathfrak{X} - \mathfrak{X}^* \rangle| \leq \|\mathfrak{W}_j\| \times C_{14} H^{(2y-1)/(ep)}$. D'où

$$\|V_j\| = \|\mathfrak{W}_j\| \leq C_{14} H^{(2y-1)/(ep)}.$$

Si σ_j et σ_{j+1} sont complexes conjugués, alors

$$\mathfrak{W}_j^1 := (0, \dots, 0, \underbrace{V_j^{[1]}, V_j^{[2]}}_{\text{blocs } j \text{ et } j+1}, 0, \dots, 0) \text{ et } \mathfrak{W}_j^2 := (0, \dots, 0, \underbrace{-V_j^{[2]}, V_j^{[1]}}_{\text{blocs } j \text{ et } j+1}, 0, \dots, 0),$$

sont orthogonaux à S^* , et donc en particulier à \mathfrak{X}^* . On a alors

$$\begin{aligned} \left| \langle W^{[j]}, V_j^{[1]} \rangle + \langle W^{[j+1]}, V_j^{[2]} \rangle \right| &= |\langle \mathfrak{X}, \mathfrak{W}_j^1 \rangle| = |\langle \mathfrak{X} - \mathfrak{X}^*, \mathfrak{W}_j^1 \rangle| \leq C_{14} H^{(2y-1)/(ep)} \|\mathfrak{W}_j^1\| \\ \left| \langle W^{[j]}, V_j^{[2]} \rangle - \langle W^{[j+1]}, V_j^{[1]} \rangle \right| &= |\langle \mathfrak{X}, \mathfrak{W}_j^2 \rangle| = |\langle \mathfrak{X} - \mathfrak{X}^*, \mathfrak{W}_j^2 \rangle| \leq C_{14} H^{(2y-1)/(ep)} \|\mathfrak{W}_j^2\|, \end{aligned}$$

et comme

$$\left| \langle W^{(j)}, V_j \rangle \right| = \left| \langle W^{[j]}, V_j^{[1]} \rangle + \langle W^{[j+1]}, V_j^{[2]} \rangle + i(\langle W^{[j]}, V_j^{[2]} \rangle - \langle W^{[j+1]}, V_j^{[1]} \rangle) \right|,$$

on en déduit

$$\|V_j\|^2 = |\langle W^{(j)}, V_j \rangle| \leq 2C_{14} H^{(2y-1)/(ep)} \|V_j\|,$$

et par suite

$$\|V_j\| \leq 2C_{14} H^{(2y-1)/(ep)}. \quad (5.14)$$

Cette égalité est vraie pour $j = 1, \dots, p$.

Comme annoncé on définit B^{e-1} par (5.12). Montrons maintenant que $H(B^{e-1}) \leq H'$. Soit \mathfrak{a} l'idéal engendré par les coordonnées de Grassmann de la famille (Z_1, \dots, Z_m) (i.e. par les coordonnées du vecteur $Z_1 \wedge \dots \wedge Z_m$, cf §2.2), et \mathfrak{b} l'idéal engendré par les coordonnées de Grassmann de la famille (W, Z_1, \dots, Z_m) . Comme les coordonnées de W sont dans $\mathcal{O}_{\overline{K}}$ on a $\mathfrak{b} \subset \mathfrak{a}$ (en effet, si on revient à la définition des coordonnées de Grassmann, \mathfrak{b} est engendré par les mineurs d'ordre $m+1$ de la matrice dont les lignes sont W, Z_1, \dots, Z_m ; le résultat est alors évident en développant un tel mineur par rapport à la première ligne), et donc $N(\mathfrak{b}) \geq N(\mathfrak{a})$. De plus, $H(B^e) = H(B^{e,\perp}) = N(\mathfrak{a})^{-1} \prod_{i=1}^p D(Z_1^{(i)}, \dots, Z_m^{(i)})$. On a alors

$$\begin{aligned} H(B^{e-1}) &= H((B^{e-1})^\perp) = N(\mathfrak{b})^{-1} \prod_{i=1}^p D(W^{(i)}, Z_1^{(i)}, \dots, Z_m^{(i)}) \\ &= N(\mathfrak{b})^{-1} \prod_{i=1}^p \|V_i\| D(Z_1^{(i)}, \dots, Z_m^{(i)}) \leq H(B^e) \prod_{i=1}^p \|V_i\| \\ &\leq C_5 H(B^e) H^{(2y-1)/e} \leq C_5 H^{(e+2y-1)/e} = H', \end{aligned} \quad (5.15)$$

pour un C_5 convenable.

Démontrons maintenant l'inégalité du théorème sur $H(B^{e-1})\omega_i^2(A^d, B^{e-1})$ ($i = 1, \dots, h$). Puisque $\langle Y_i, Z_j \rangle = 0$ ($j = 1, \dots, m$) et $\|Y_i\| = 1$, on a

$$\begin{aligned} & \omega^2((B^{e-1})^\perp, Y_i) \\ &= \mu^2((B^{e-1})^\perp, Y_i) \\ &= D^2(Y_i, W, Z_1, \dots, Z_m) \|Y_i\|^{-2} D(W, Z_1, \dots, Z_m)^{-2} \\ &= \left(\|Y_i\|^2 D^2(W, Z_1, \dots, Z_m) - |\langle Y_i, W \rangle|^2 D^2(Z_1, \dots, Z_m) \right) D(W, Z_1, \dots, Z_m)^{-2} \\ &= 1 - |\langle Y_i, W \rangle|^2 D^2(Z_1, \dots, Z_m) D(W, Z_1, \dots, Z_m)^{-2}, \end{aligned}$$

la première égalité venant de la définition de ω et de μ dans ce cas particulier (cf définition 3.3.3), la deuxième de la proposition suivant la définition 3.3.3, la troisième s'obtenant enfin en développant une première fois le déterminant $D^2(Y_i, W, Z_1, \dots, Z_m)$ selon la première colonne, puis en redéveloppant le second déterminant non nul obtenu selon sa première ligne.

La proposition 3.3.5 permet alors de conclure que

$$\omega(B^{e-1}, Y_i) = |\langle Y_i, W \rangle| D(Z_1, \dots, Z_m) / D(W, Z_1, \dots, Z_m),$$

et par suite

$$D^2(W, Z_1, \dots, Z_m) \omega^2(B^{e-1}, Y_i) = |\langle Y_i, W \rangle|^2 D^2(Z_1, \dots, Z_m).$$

On obtient alors

$$\begin{aligned} & H(B^{e-1}) \omega^2(B^{e-1}, Y_i) \\ &= N(\mathfrak{b})^{-1} \left(\prod_{j=3}^p D(W^{(j)}, Z_1^{(j)}, \dots, Z_m^{(j)}) \right) D^2(W, Z_1, \dots, Z_m) \omega^2(B^{e-1}, Y_i) \\ &= N(\mathfrak{b})^{-1} \left(\prod_{j=3}^p D(W^{(j)}, Z_1^{(j)}, \dots, Z_m^{(j)}) \right) D^2(Z_1, \dots, Z_m) |\langle Y_i, W \rangle|^2 \\ &\leq C_{15} N(\mathfrak{a})^{-1} \left(\prod_{j=1}^p D(Z_1^{(j)}, \dots, Z_m^{(j)}) \right) \left(\prod_{j=3}^p \|V_j\| \right) H^{-2(y_i - (2y-1)/(ep))} (H/H(B^e))^{1/h} \\ &\leq C_{16} H(B^e) H^{(p-2)(2y-1)/(ep) - 2[y_i - (2y-1)/(ep)]} H/H(B^e) \\ &\leq C_{16} H^{(e+2y-1)/e - 2y_i}. \end{aligned}$$

la troisième ligne s'obtenant à partir de (5.13) et la quatrième à partir de (5.14). Il existe donc un vecteur $R_i \in B^{e-1}$ tel que

$$H(B^{e-1}) \omega^2(R_i, Y_i) \leq C_{16} H^{-(2y'_i - 1)(2y + e - 1)/e},$$

(Rappelons qu'on avait posé $y'_i := (y_i e)/(qy + e - 1)$).

Par hypothèse, on a aussi

$$H(B^e) \omega^2(X_i, Y_i) \leq c^2 H^{-2y_i + 1},$$

donc, avec (5.15) :

$$H(B^{e-1}) \omega^2(X_i, Y_i) \leq c^2 C_5 H^{-2y_i + 1 + (2y-1)/e} = c^2 C_5 H^{-(2y'_i - 1)(2y + e - 1)/e}.$$

L'inégalité de la proposition 3.3.2 nous assure alors

$$H(B^{e-1}) \omega^q(R_i, X_i) \leq C_{17} H^{-qy_i + 1 + (qy-1)/e} = C_{17} H^{-(qy'_i - 1)(qy + e - 1)/e} \quad (i = 1, \dots, h),$$

et le théorème 3.3.7 (avec $\delta = 1$ par exemple, puisque la famille des (X_j) est orthonormée) fournit l'inégalité

$$H(B^{e-1})\omega_i^2(A^d, B^{e-1}) \leq c^2 C_6 H^{-(2y'_i-1)(2y+e-1)/e} \quad (q = 1, \dots, h),$$

pour C_6 assez grand. Ceci achève la démonstration de la première partie du théorème. Schmidt ne donne qu'une esquisse de démonstration dans le cas où (5.11) est vérifiée, nous détaillons un peu plus ici. On définit $H_1 \geq 1$ par $H' = C_9 H H_1^{2h/e}$ (C_9 sera précisé plus tard). Dans la construction de W on remplace les équations du convexe par

$$\begin{aligned} (i') \quad \mathfrak{X}^* &\in \Pi \\ (ii') \quad |\langle \mathfrak{X}_T, \mathfrak{Y}_j^i \rangle| &\leq H_1^{-(1-2h/(ep))} \quad (j = 1, \dots, h; i = 1, 2) \\ (iii') \quad \|\mathfrak{X}_0\| &\leq C_{18} H_1^{2h/(ep)}. \end{aligned}$$

Ces équations définissent un corps convexe symétrique de volume

$$\begin{aligned} \Delta^m H(B^e) \times (2H_1^{-(1-2h/(ep))})^{2h} \times (C_{18} H_1^{2h/(ep)})^{ep-2h} V(ep-2h) \\ = \Delta^m H(B^e) 4^h C_{18}^{ep-2h} V(ep-2h) \\ > 2^{pn} \Delta^n, \end{aligned}$$

pour C_{18} assez grand. Par le premier théorème de Minkowski il existe ainsi un $\mathfrak{X} \in \Lambda$ non nul dans cet ensemble. Soit $W \in \mathcal{O}_{\mathbb{K}}^n$ tel que $\mathfrak{X} = \rho(W)$.

A la place de (5.13) on obtient l'inégalité

$$|\langle W, Y_j \rangle| \leq 2H_1^{-(1-2h/(ep))}.$$

On a aussi l'inégalité

$$\|\mathfrak{X} - \mathfrak{X}^*\| \leq C_{19} H_1^{2h/(ep)},$$

qui permet de remplacer (5.14) par

$$\|V_j\| \leq 2C_{19} H_1^{2h/(ep)}.$$

Le calcul de $H(B^{e-1})$ conduit alors à

$$H(B^{e-1}) \leq H(B^e) \prod_{j=1}^p \|V_j\| \leq C H(B^e) H_1^{2h/e}.$$

En choisissant maintenant $C_9 := C$ on obtient bien $H(B^{e-1}) \leq H'$. Pour la fin de la preuve, en reprenant les calculs précédents avec les nouvelles estimations, on trouve

$$\begin{aligned} H(B^{e-1})\omega^2(B^{e-1}, Y_i) &\leq C_{20} H(B^e) \times H_1^{(p-2)2h/(ep)} \times H_1^{-2(1-2h/(ep))} \\ &\leq C_{21} H H_1^{2h/e-2} = C_{21} H H_1^{2h/e(1-e/h)} \\ &= C_{10} H^{e/h} H'^{1-e/h} = C_{10} H^{2y'_0} H'^{-(2y'_0-1)}, \end{aligned}$$

ce qui achève la démonstration du théorème. □

Nous terminons ce paragraphe par un troisième théorème. On se rappellera la définition de μ (cf définition 3.3.3).

Théorème 5.2.3. *Soient $0 < d, e$ tels que $d+e < n$, $H \geq 1$ et A^d, B^e des sous-espaces de G^n avec B^e défini sur K et de hauteur $H(B^e) \leq H$. On suppose que*

$$H(B^e) \leq \mu^q(A^d, B^e) \leq cH^{-x},$$

(avec $x \geq 0$ et $c > 0$ fixés).

Alors il existe un sous-espace $B^{e+1} \supset B^e$ défini sur K de hauteur

$$H(B^{e+1}) \leq H^{(e+1)/e} =: H',$$

et satisfaisant

$$\begin{aligned} H(B^{e+1})\mu^q(A^d, B^{e+1}) &\leq cC_{22}H^{-x-(d+e)/(e(n-d-e))} \\ &= cC_{22}H'^{-\left(\frac{ex+(d+e)/(n-d-e)}{e+1}\right)}. \end{aligned}$$

Preuve Cf [17] (théorème 11). Schmidt se restreint au cas (a). □

5.3 Approximations diophantiennes

En combinant les théorèmes du paragraphe précédent nous sommes en mesure d'établir un certain nombre de résultats d'approximation diophantienne. Ce paragraphe reprend la section 12 de [17]. On conservera les notations du paragraphe 5.2 pour G^n , K , q ...

Théorème 5.3.1. *Soient $d, e > 0$ tels que $d+e \leq n$ et $1 \leq j \leq t := \min(d, e)$. Soit A^d un sous-espace de G^n et $H \geq 1$. Alors il existe un sous-espace B^e défini sur K , de hauteur $H(B^e) \leq H$ et vérifiant :*

$$\psi_j^q(A^d, B^e) \leq C_{24}H^{-d(n-j)/(j(n-d)(n-e))}.$$

Si $j = 1$, on peut améliorer l'inégalité précédente en

$$H(B^e)^{(n-1)/(n-e)}\psi_1^q(A^d, B^e) \leq C_{24}H^{-d(n-1)/((n-d)(n-e))}.$$

Corollaire 5.3.2. *En plus des hypothèses du théorème, on suppose qu'il n'existe aucun B^e défini sur K tel que $\dim A^d \cap B^e \geq j$ (en d'autres termes, on suppose qu'il n'existe aucun B^e défini sur K tel que $\psi_j(A^d, B^e) = 0$). Alors il existe une infinité de sous-espaces B^e définis sur K tels que*

$$\psi_j^q(A^d, B^e) \leq C_{24}H(B^e)^{-d(n-j)/(j(n-d)(n-e))}.$$

Si $j = 1$, on peut améliorer l'inégalité précédente en

$$\psi_1^q(A^d, B^e) \leq C_{24}H(B^e)^{-n(n-1)/((n-d)(n-e))}.$$

Preuve Cf théorème 12 de [17]. Schmidt utilise une récurrence sur e à j fixé en initialisant à $e = j$ grâce au théorème 5.1.1. La suite de la preuve vient essentiellement des théorèmes 5.1.1 et 5.8 (going-up). □

Théorème 5.3.3. Soient $d, e > 0$ tels que $d + e \leq n$ et $1 \leq j \leq t := \min(d, e)$ vérifiant

$$j + n - t \geq j(j + n - d - e).$$

Soit A^d un sous-espace de G^n et $H \geq 1$. Alors il existe un sous-espace B^e défini sur K , de hauteur $H(B^e) \leq H$ et vérifiant :

$$H(B^e)\psi_j^q(A^d, B^e) \leq C_{26}H^{1-(j+n-t)/(j(j+n-d-e))}.$$

Corollaire 5.3.4. En plus des hypothèses du théorème, on suppose qu'il n'existe aucun B^e défini sur K tel que $\dim A^d \cap B^e \geq j$ (en d'autres termes, on suppose qu'il n'existe aucun B^e défini sur K tel que $\psi_j(A^d, B^e) = 0$). Alors il existe une infinité de sous-espaces B^e définis sur K tels que

$$\psi_j^q(A^d, B^e) \leq C_{26}H(B^e)^{-(j+n-t)/(j(j+n-d-e))}.$$

Preuve Cf théorème 13 de [17].

Schmidt suppose d'abord $e \leq d$ (de sorte que $t = e$). Par dualité, il suffit alors de montrer que si $d, e > 0$ sont tels que $d + e \geq n$, $d \leq e$, et vérifient $j + e \geq j(j + d + e - n)$, et que A^d est un sous-espace de G^n , alors il existe un sous-espace B^e défini sur K , de hauteur $H(B^e) \leq H$ vérifiant

$$H(B^e)\psi_j^q(A^d, B^e) \leq C_{26}H^{1-(j+e)/(j(j+d+e-n))}.$$

Dans ce cas la preuve se fait par récurrence sur j . Pour le cas $j = 1$ on utilise la deuxième partie du théorème 5.2.2 (going-down). Pour la suite de la récurrence Schmidt fait encore usage du théorème du going-down.

Dans le cas $e \geq d$ (on a alors $t = d$ et $d \leq n - d$), Schmidt utilise une récurrence descendante sur e de $n - d$ à d . Le cas $e = n - d$ vient du cas précédent. Le passage de $e + 1$ à e vient essentiellement du théorème du going-down. □

On rappelle que μ a été défini dans la définition 3.3.3.

Théorème 5.3.5. Soit $0 < d < n$; on pose $u = \min(d, n - d)$. Soit A^d un sous-espace de G^n et $H \geq 1$. Alors il existe des sous-espaces $B^1 \subset \dots \subset B^u$ définis sur K de hauteur $H(B^i) \leq H^i$ ($i = 1, \dots, u$) et vérifiant

$$H(B^i)\mu^q(A^d, B^i) \leq C_{27}H^{-[d/(n-d)+(d+1)/(n-d-1)+\dots+(d+i-1)/(n-d-i+1)]},$$

d'où

$$\mu^q(A^d, B^i) \leq C_{27}H(B^i)^{-(n/i)[1/(n-d)+1/(n-d-1)+\dots+1/(n-d-i+1)]} \quad (i = 1, \dots, u).$$

Preuve Cf théorème 14 de [17]. Il faut essentiellement utiliser les théorèmes 5.1.1 et 5.2.3. □

Corollaire 5.3.6. Soient $d, e > 0$ tels que $d + e \leq n$; on pose $t = \min(d, e)$. Soient A^d un sous-espace de G^n et $H \geq 1$. Alors il existe un sous-espace B^e défini sur K de hauteur $H(B^e) \leq H$ et vérifiant

$$H(B^e)^{(n-t)/(n-e)}\mu^q(A^d, B^e) \leq C_{28}H^{-(n-t)/(t(n-e))[d/(n-d)+\dots+(d+t-1)/(n-d-t+1)]},$$

d'où

$$\mu^q(A^d, B^e) \leq C_{28}H(B^e)^{-n(n-t)/(t(n-e))[1/(n-d)+\dots+1/(n-d-t+1)]}.$$

Preuve Cf [17] (corollaire du théorème 14). \square

5.4 De la qualité de ces approximations

Ce paragraphe reprend la section 13 de [17].

Théorème 5.4.1. *Soient $d, e > 0$ tels que $d + e \leq n$; on pose $t = \min(d, e)$. Soit L un corps de nombres contenant K tel que $[L : K] = n$. Dans le cas (a) on suppose que L et toutes ses images par un K -plongement dans \mathbb{C} sont réels, et dans le cas (b) on suppose simplement $L \subset \mathbb{C}$. Alors il existe un sous-espace A^d de G^n défini sur la clôture normale de L/\mathbb{Q} (dans \mathbb{C}) et une constante C_{29} tels que pour tout B^e défini sur K*

$$\psi_t(A^d, B^e) \geq C_{29} H(B^e)^{-n/(qt(n+t-d-e))}.$$

Corollaire 5.4.2. *Soient $d, e > 0$ tels que $d + e \leq n$; on pose $t = \min(d, e)$ et on suppose*

$$n \geq t(t + n - d - e). \quad (5.16)$$

Soit A^d un sous-espace de G^n tel que $\dim A^d \cap B^e < t$ (ou $\dim A^d \cap B^e < t - 1$ en cas d'égalité dans (5.16)) pour tout B^e défini sur K . Alors il existe une infinité de sous-espaces B^e définis sur K vérifiant

$$\psi_t(A^d, B^e) \leq C_{26} H(B^e)^{-n/(qt(n+t-d-e))}.$$

L'exposant $-n/(qt(n+t-d-e))$ est le meilleur possible. (cf l'inégalité du corollaire du théorème 5.3.3 avec $j = t$). C'est un problème ouvert de déterminer le meilleur exposant possible dans ce corollaire lorsque la condition (3.3.3) n'est pas vérifiée.

Preuve Cf théorème 15 de [17]. \square

Théorème 5.4.3. *Etant donné un réel α , on note $\{\alpha\}$ l'entier vérifiant $\alpha \leq \{\alpha\} < \alpha + 1$.*

Soient $d, e > 0$ tels que $d + e \leq n$. On pose

$$m := \{(e(n - e) + 1)/(n + 1 - d - e)\}.$$

Soit L un corps de nombres contenant K de degré $[L : K] =: l \geq m$, tel que L soit réel dans le cas (a) et complexe dans le cas (b) (cf début de la partie 5). Alors il existe un sous-espace A^d de G^n défini sur L tels que pour tout B^e défini sur K on ait

$$\mu^q(A^d, B^e) \geq C_{40} H(B^e)^{-l},$$

où $C_{40} = C_{40}(n, d, e, K, L, A^d)$.

Corollaire 5.4.4. *Soient n, d, e, m vérifiant les hypothèses du théorème. Alors il existe un sous-espace A^d de G^n tel que pour tout B^e sur K on ait $\mu^q(A^d, B^e) \geq C_{41} H(B^e)^{-m}$.*

Corollaire 5.4.5. *Soient n, d, e, m vérifiant les hypothèses du théorème et $1 \leq i \leq t := \min(d, e)$. Alors il existe un sous-espace A^d de G^n tel que pour tout B^e défini sur K on ait*

$$\psi_i^q(A^d, B^e) \geq C_{42} H(B^e)^{-m/i}.$$

Preuve Cf théorème 16 de [17] (preuve assez longue et technique). \square

6 Approximation diophantienne et géométrie paramétrique des nombres

L'idée de la géométrie paramétrique des nombres développée par Schmidt et Summerer dans [18] et [19] est d'étudier la suite des minima successifs d'une famille de corps convexes paramétrée par un réel $Q > 1$ pour un réseau fixé construit à partir d'un vecteur $u = (1, \xi_1, \dots, \xi_{n-1}) \in \mathbb{R}^n$ dont les coordonnées sont linéairement indépendantes sur \mathbb{Q} . Le log en base Q de ces minima est intimement lié à des exposants classiques d'approximation diophantienne (approximations rationnelles simultanées notamment). Les deux paragraphes suivants explicitent les correspondances entre ces deux sortes d'exposants.

6.1 Exposants d'approximation classiques

Soit $n \geq 2$. On fixe des réels ξ_1, \dots, ξ_{n-1} tels que $1, \xi_1, \dots, \xi_{n-1}$ soient linéairement indépendants sur \mathbb{Q} . Dans le cadre d'approximations rationnelles on peut considérer pour $Q > 1$ deux problèmes duaux :

Problème E (approximation rationnelle simultanée)

On cherche les solutions $(x, y_1, \dots, y_{n-1}) \in \mathbb{Z}^n$ non nulles du système

$$\begin{aligned} |x| &\leq Q \\ |\xi_j x - y_j| &\leq Q^{-\eta} \quad j = 1, \dots, n-1. \end{aligned}$$

Problème E^*

On cherche les solutions $(x, y_1, \dots, y_{n-1}) \in \mathbb{Z}^n$ non nulles du système

$$\begin{aligned} \left| x - \sum_{j=1}^{n-1} \xi_j y_j \right| &\leq Q^{-\eta} \\ |y_j| &\leq Q \quad j = 1, \dots, n-1. \end{aligned}$$

Pour $\eta = 1/(n-1)$, le théorème de Dirichlet sur les approximations simultanées (qui peut être aisément prouvé à l'aide du premier théorème de Minkowski) nous assure que le problème E possède toujours une solution non nulle, tandis que pour le problème E^* avec $\eta = n-1$, c'est un théorème de Dirichlet sur les formes linéaires (dont la démonstration se résume aussi essentiellement au premier théorème de Minkowski) qui fournit l'existence d'une solution non triviale.

On peut alors définir des exposants d'approximation classiques étudiés notamment par Khinchin [7], [8], Jarník [6],... et plus récemment par Roy [13], Bugeaud et Laurent [2], [3] :

Définition 6.1.1. On définit ω (resp. $\widehat{\omega}$) comme étant la borne supérieure de l'ensemble des nombres η tels que le problème E possède une solution non triviale $(x, y_1, \dots, y_{n-1}) \in \mathbb{Z}^n$ pour des Q arbitrairement grands (resp. pour tout Q assez grand).

Définition 6.1.2. On définit ω^* (resp. $\widehat{\omega}^*$) comme étant la borne supérieure de l'ensemble des nombres η tels que le problème E^* possède une solution non triviale $(x, y_1, \dots, y_{n-1}) \in \mathbb{Z}^n$ pour des Q arbitrairement grands (resp. pour tout Q assez grand).

On peut alors établir un certain nombre de résultats et d'inégalités qui décrivent le comportement de ces exposants.

6.2 Approche du problème d'approximation simultanée par la géométrie paramétrique des nombres

On conserve les notations du paragraphe précédent pour n et ξ_1, \dots, ξ_n . Schmidt et Summerer ont développé de nouveaux outils pour aborder le problème précédent [18].

Problème E

Pour $Q > 1$ on définit l'application linéaire $T_Q : \mathbb{R}^n \rightarrow \mathbb{R}^n$ par $T_Q(p_1, \dots, p_n) = (Qp_1, Q^{-1/(n-1)}p_2, \dots, Q^{-1/(n-1)}p_n)$. On note \mathcal{C} la boule unité pour la norme $\|(p_1, \dots, p_n)\|_\infty = \max_{1 \leq i \leq n} |p_i|$ et $\Lambda = \Lambda(\xi)$ le réseau des points de la forme $(x, \xi_1 x - y_1, \dots, \xi_{n-1} x - y_{n-1})$ avec $(x, y_1, \dots, y_{n-1}) \in \mathbb{Z}$. On considère alors les minima successifs $\lambda_1(Q) \leq \dots \leq \lambda_n(Q)$ de Λ pour $\mathcal{C}(Q) := T_Q(\mathcal{C})$.

Le théorème de Dirichlet sur les approximations simultanées implique $\lambda_1(Q) \leq 1$.

Problème E^*

Pour $Q > 1$ on définit l'application linéaire $T_Q^* : \mathbb{R}^n \rightarrow \mathbb{R}^n$ par $T_Q^*(p_1, \dots, p_n) = (Q^{-1}p_1, Q^{1/(n-1)}p_2, \dots, Q^{1/(n-1)}p_n)$ (remarque : $T_Q^* = T_Q^{-1}$ avec les notations précédentes). Rappelons que \mathcal{C}^* , le polaire de \mathcal{C} , est la boule unité pour la norme $\|(p_1, \dots, p_n)\|_1 = \sum_{i=1}^n |p_i|$. Un petit calcul montre que le réseau dual du réseau Λ est le réseau $\Lambda^* = \Lambda^*(\xi)$ des points de la forme $(x - \sum_{i=1}^{n-1} \xi_i y_i, y_1, \dots, y_{n-1})$ avec $(x, y_1, \dots, y_{n-1}) \in \mathbb{Z}$. On considère alors les minima successifs $\lambda_1^*(Q) \leq \dots \leq \lambda_n^*(Q)$ de Λ^* pour $\mathcal{C}^*(Q) := T_Q^*(\mathcal{C})$. On s'intéresse également aux minima successifs $\nu_1(Q) \leq \dots \leq \nu_n(Q)$ du réseau Λ^* pour $\mathcal{C}(Q)$.

Le théorème de Dirichlet sur les formes linéaires implique $\nu_1(Q) \leq 1$ et l'inclusion $\mathcal{C}^* \subset \mathcal{C} \subset n\mathcal{C}^*$ montre que $\nu_i(Q) \leq \lambda_i^*(Q) \leq n\nu_i(Q)$ ($i = 1, \dots, n$). L'inégalité de Mahler (combinée avec l'inégalité précédente) montre aussi qu'il existe des constantes $c_0, c_1 > 0$ indépendantes de Q telles que

$$c_0 \leq \lambda_i(Q)\nu_{n+1-i}(Q) \leq c_1 \quad i = 1, \dots, n.$$

Notons que pour tout i , $\lambda_i(Q)$, $\lambda_i^*(Q)$ et $\nu_i(Q)$ sont des fonctions continues de Q . Schmidt et Summerer montrent également que l'indépendance sur \mathbb{Q} de $1, \xi_1, \dots, \xi_{n-1}$ assure que pour $1 \leq s < n$ fixé, il existe Q, Q' et Q'' arbitrairement grands tels que $\lambda_s(Q) = \lambda_{s+1}(Q)$, $\lambda_s^*(Q') = \lambda_{s+1}^*(Q')$ et $\nu_s(Q'') = \nu_{s+1}(Q'')$.

A l'aide du second théorème de Minkowski, Schmidt et Summerer montrent une autre propriété intéressante qui motive la définition 6.2.1 : il existe des constantes $c_2, c_3 > 0$ indépendantes de Q telles que pour tout Q

$$\begin{aligned} c_2 Q^{-1} &\leq \lambda_1(Q) \leq \lambda_n(Q) \leq c_3 Q^{n-1} \\ c_2 Q^{-1/(n-1)} &\leq \nu_1(Q) \leq \nu_n(Q) \leq c_3 Q. \end{aligned}$$

Définition 6.2.1. Pour $1 \leq i \leq n$ et $Q > 1$ on pose $\psi_i(Q) := \log(\lambda_i(Q))/\log Q$, de sorte que

$$\lambda_i(Q) = Q^{\psi_i(Q)}.$$

On pose également

$$\bar{\psi}_i := \limsup_{Q \rightarrow \infty} \psi_i(Q) \quad \text{et} \quad \underline{\psi}_i := \liminf_{Q \rightarrow \infty} \psi_i(Q).$$

Le travail de Schmidt et Summerer dans [18] et [19] et l'enjeu de la géométrie paramétrique des nombres consistent à étudier finement le comportement de fonctions similaires aux fonctions ψ_i . Le théorème suivant fait le lien entre ces exposants et ceux définis dans le paragraphe précédent.

Théorème 6.2.2. *On a les formules*

$$(\omega + 1)(1 + \underline{\psi}_1) = (\widehat{\omega} + 1)(1 + \overline{\psi}_1) = \frac{n}{n-1},$$

et

$$(\omega^* + 1)\left(\frac{1}{n-1} - \overline{\psi}_n\right) = (\widehat{\omega}^* + 1)\left(\frac{1}{n-1} - \underline{\psi}_n\right) = \frac{n}{n-1}.$$

Ainsi, ω (resp. $\widehat{\omega}, \omega^*, \widehat{\omega}^*$) détermine entièrement $\overline{\psi}_1$ (resp. $\underline{\psi}_1, \overline{\psi}_n, \underline{\psi}_n$) et réciproquement. Ces formules permettent en outre de traduire certaines des inégalités que l'on obtiendra (celles faisant intervenir ces exposants) en des inégalités faisant intervenir les exposants classiques. En fait, on peut aussi traduire en termes d'exposants "classiques" toute équation faisant intervenir les $\underline{\psi}_j, \overline{\psi}_j$ grâce à la propriété 8.2.2.

Ainsi, avec ce nouvel outil d'étude, Schmidt et Summerer redémontrent beaucoup de résultats classiques : le principe de transfert de Khinchin [7], [8], les identités de Jarník [6], le raffinement de Laurent des relations de Khinchin et d'autres résultats récents de Bugeaud, Laurent et Moshchevitin [4], [12], [9], ainsi que de nouvelles inégalités.

7 Une conjecture résolue par Roy

Dans [19] Schmidt et Summerer montrent que le n -uplet des minima successifs d'une famille de corps convexes symétriques par rapport à un certain réseau $\Lambda(\mathbf{u})$ (avec $\mathbf{u} \in \mathbb{R}^n$) peut être approximé par une fonction d'une certaine classe. Les auteurs conjecturent que réciproquement, à toute fonction de cette classe correspondrait un vecteur \mathbf{u} de sorte qu'elle approcherait les minima successifs associés au réseau $\Lambda(\mathbf{u})$ et à la famille de corps convexes paramétrés. Roy démontre dans un article récent [14] cette conjecture en simplifiant la classe de fonctions considérée. Le but de cette partie est d'introduire cette nouvelle classe de fonctions utilisée par Roy et de présenter la structure de sa preuve permettant de résoudre le problème posé par Schmidt et Summerer.

7.1 Définitions et formulation du problème

Soit $n \geq 2$ un entier. La lettre q désignera toujours un réel ≥ 0 . Pour tous vecteurs $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ on notera $\mathbf{x} \cdot \mathbf{y} = \langle \mathbf{x}, \mathbf{y} \rangle$ le produit scalaire de \mathbf{x} avec \mathbf{y} . On fixe \mathbf{u} un vecteur unitaire de \mathbb{R}^n . On pose alors

$$\mathcal{C}_{\mathbf{u}}(e^q) := \{\mathbf{x} \in \mathbb{R}^n ; \|\mathbf{x}\| \leq 1, |\mathbf{x} \cdot \mathbf{u}| \leq e^{-q}\}.$$

Pour $j = 1, \dots, n$, on note $L_{\mathbf{u},j}(q)$ le plus petit des nombres réels $L \geq 0$ tels que $e^L \mathcal{C}_{\mathbf{u}}(e^q)$ contienne au moins j vecteurs linéairement indépendants de \mathbb{Z}^n . Plus classiquement, $L_{\mathbf{u},j}(q) = \log \lambda_j(\mathcal{C}_{\mathbf{u}}(e^q), \mathbb{Z}^n)$ (cf partie 2.4). On regroupe ces minima successifs en une unique fonction $\mathbf{L}_{\mathbf{u}} : [0, +\infty[\rightarrow \mathbb{R}^n$ en posant

$$\mathbf{L}_{\mathbf{u}}(q) = (L_{\mathbf{u},1}(q), \dots, L_{\mathbf{u},n}(q)).$$

Il y a quelques différences mineures entre cette fonction et celle qu'étudient Schmidt et Summerer dans leurs articles, puisque ces derniers considèrent plutôt la famille de corps convexes définis dans la partie précédente et le réseau $\Lambda(\xi)$ plutôt que \mathbb{Z}^n .

Cependant on peut établir des propriétés similaires dans les deux cas.

On pose

$$\Delta_n := \{(x_1, \dots, x_n) \in \mathbb{R}^n ; x_1 \leq \dots \leq x_n\}$$

et

$$\Phi_n : \mathbb{R}^n \rightarrow \Delta_n$$

l'application continue qui réarrange les coordonnées d'un vecteur dans l'ordre croissant.

On définit le *graphe combiné* d'un ensemble de fonctions à valeurs réelles définies sur un intervalle I comme étant l'union de leur graphe dans $I \times \mathbb{R}$. Pour une fonction $\mathbf{P} : [c, +\infty[\rightarrow \Delta_n$ et un sous-intervalle $I \subset [c, +\infty[$, on définit également le *graphe combiné de \mathbf{P} sur I* comme étant le graphe combiné de ses composantes P_1, \dots, P_n restreintes à I .

Remarquons que si \mathbf{P} est continue et que l'ensemble des réels $q \geq c$ tels que $P_1(q), \dots, P_n(q)$ ne soient pas tous distincts est discret, alors l'application \mathbf{P} est uniquement déterminée par son graphe combiné sur $[c, +\infty[$.

On peut alors introduire l'objet combinatoire de base de Roy.

Définition 7.1.1. Soit $\delta \in]0, +\infty[$ et $s \in \mathbb{N}^* \cup \{\infty\}$. On appelle *canevas de maille δ et cardinal s* dans \mathbb{R}^n la donnée d'une suite $(\mathbf{a}^{(i)})_{0 \leq i < s}$ de vecteurs de Δ_n et de deux suites de s entiers $(k_i)_{0 \leq i < s}$ et $(l_i)_{0 \leq i < s}$ vérifiant pour tout $0 \leq i < s$ les conditions suivantes :

(C1) les coordonnées $(a_1^{(i)}, \dots, a_n^{(i)})$ de $\mathbf{a}^{(i)}$ forment une suite strictement croissante de multiples entiers de $\delta > 0$ (i.e. de la forme $n\delta$, $n \in \mathbb{N}^*$).

(C2) on a $1 \leq k_0 \leq l_0 = n$ et $1 \leq k_i < l_i \leq n$ si $i \geq 1$.

(C3) si $i + 1 < s$ alors $k_i \leq l_{i+1}$, $a_{l_{i+1}}^{(i)} + \delta \leq a_{l_{i+1}}^{(i+1)}$ et

$$(a_1^{(i)}, \dots, \widehat{a_{k_i}^{(i)}}, \dots, a_n^{(i)}) = (a_1^{(i+1)}, \dots, \widehat{a_{l_{i+1}}^{(i+1)}}, \dots, a_n^{(i+1)}),$$

où le chapeau sur une coordonnée indique qu'elle a été omise.

Dans le cas où $s < \infty$, on pose $l_s := n$ et on définit

$$\mathbf{a}^{(s)} = (a_1^{(s)}, \dots, a_{n-1}^{(s)}, \infty) \in \mathbb{R}^{n-1} \times \{\infty\},$$

de telle sorte que la condition (C3) soit vérifiée pour $i = s - 1$.

Ainsi dans une telle suite $(\mathbf{a}^{(i)})_{0 \leq i < s}$, chaque vecteur $\mathbf{a}^{(i+1)}$ (avec $i \leq s - 1$) est obtenu à partir du vecteur précédent $\mathbf{a}^{(i)}$ en remplaçant l'une de ses coordonnées par un multiple de δ strictement plus grand, différent de toutes les autres coordonnées de $\mathbf{a}^{(i)}$ et en réordonnant dans l'ordre croissant les coordonnées du nouveau n -uplet ainsi obtenu. En particulier, la suite de ces vecteurs $\mathbf{a}^{(i)}$ détermine entièrement la suite des entiers $(k_i)_{0 \leq i < s}$ et $(l_i)_{0 \leq i < s}$.

La condition (C2) sera illustrée ci-dessous juste avant l'énoncé du théorème 7.1.3.

Définition 7.1.2 (n -système rigide de maille δ). A chaque canevas de maille $\delta > 0$ défini comme dans la définition 7.1.1, on associe une fonction $\mathbf{P} : [q_0, +\infty[\rightarrow \Delta_n$ définie par

$$\mathbf{P}(q) = \Phi_n(a_1^{(i)}, \dots, \widehat{a_{k_i}^{(i)}}, \dots, a_n^{(i)}, a_{k_i}^{(i)} + q - q_i) \quad (0 \leq i < s, q_i \leq q < q_{i+1}),$$

où on a posé $q_i := a_1^{(i)} + \dots + a_n^{(i)}$ ($0 \leq i < s$) et $q_s = \infty$ si $s < \infty$. On dit que cette fonction est un n -système rigide de maille δ et que $(q_i)_{0 \leq i < s}$ est la suite de ses *nombre de transition* (en anglais : *switch numbers*).

Puisque $a_{k_i}^{(i)} + q_{i+1} - q_i = a_{i+1}^{(i+1)}$ pour $i + 1 < s$, une telle fonction \mathbf{P} est continue. Son graphe combiné sur l'intervalle $[q_i, q_{i+1}[$ (avec $0 \leq i < s$) consiste en $n - 1$ segments horizontaux (semi-ouverts) et un segment de pente 1 (semi-ouvert). Leurs extrémités gauches sont les points $(q_i, a_j^{(i)})$ ($1 \leq j \leq n$) et leurs extrémités droites sont les points $(q_{i+1}, a_j^{(i+1)})$ ($1 \leq j \leq n$). La figure suivante issue de [14] montre le graphe combiné d'un 5-système rigide de maille 1 attaché au canevas $\{(1, 2, 4, 5, 8), (1, 2, 4, 7, 8), (1, 4, 5, 7, 8)\}$ de cardinal $s = 3$ avec $k_2 = 1$.

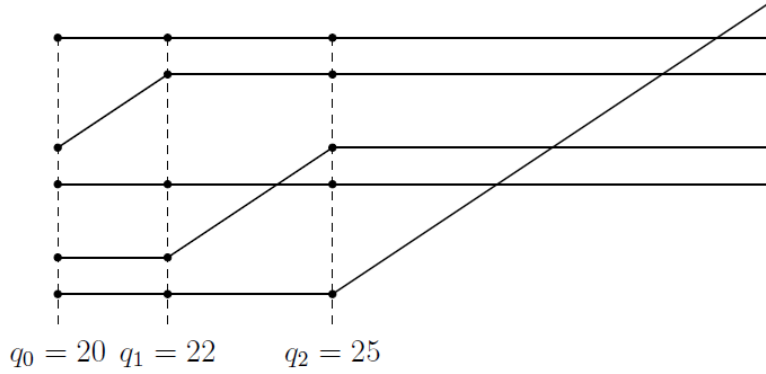


FIGURE 1 – Graphe combiné d'un 5-système rigide

La condition (C2) imposée sur le canevas traduit le fait que pour tout indice i ($1 \leq i < s$) la droite qui contient le segment de pente 1 sur $[q_i, q_{i+1}[$ se situe à droite de la droite contenant le segment de pente 1 de l'intervalle précédent $[q_{i-1}, q_i[$.

Le théorème principal de [14] (qui prouve la conjecture de Schmidt et Summerer) est le suivant.

Théorème 7.1.3 (Roy, 2014).

Soit $n \geq 2$ un entier et soit $\delta \in]0, +\infty[$. Pour tout vecteur unitaire $\mathbf{u} \in \mathbb{R}^n$ il existe un n -système rigide $\mathbf{P} : [q_0, \infty[\rightarrow \Delta_n$ de maille δ tel que $\mathbf{L}_{\mathbf{u}} - \mathbf{P}$ soit bornée sur $[q_0, \infty[$. Réciproquement, pour tout n -système rigide $\mathbf{P} : [q_0, \infty[\rightarrow \Delta_n$ de maille δ , il existe un vecteur unitaire $\mathbf{u} \in \mathbb{R}^n$ tel que $\mathbf{L}_{\mathbf{u}} - \mathbf{P}$ soit bornée sur $[q_0, \infty[$.

7.2 Liens avec les exposants "classiques" d'approximation

Nous allons de nouveau établir les liens entre les exposants précédemment définis et les exposants plus "classiques" qu'on peut associer à un vecteur \mathbf{u} de \mathbb{R}^n et qui mesurent la qualité d'approximation de \mathbf{u} par des sous-espaces définis sur \mathbb{Q} de dimension d (voir [4], [9], [17]). Comme Roy dans [14] nous nous restreignons ici à deux cas en particulier. Le corollaire 7.2.4 donne une première application intéressante du théorème 7.1.3.

Dans le cas $d = n - 1$:

Définition 7.2.1. On note $\tau(\mathbf{u})$ (resp. $\hat{\tau}(\mathbf{u})$) la borne supérieure des réels $\tau > 0$ tels que le système d'inégalités

$$\|\mathbf{x}\| \leq X \quad \text{et} \quad |\mathbf{x} \cdot \mathbf{u}| \leq X^{-\tau}$$

possède des solutions $\mathbf{x} \in \mathbb{Z}^n$ non nulles pour des valeurs arbitrairement grandes de X (resp. pour tout X assez grand).

Dans le cas $d = 1$:

Définition 7.2.2. On note $\lambda(\mathbf{u})$ (resp. $\widehat{\lambda}(\mathbf{u})$) la borne supérieure des réels $\lambda > 0$ tels que le système d'inégalités

$$\|\mathbf{x}\| \leq X \quad \text{et} \quad \|\mathbf{x} \wedge \mathbf{u}\| \leq X^{-\lambda}$$

possède des solutions $\mathbf{x} \in \mathbb{Z}^n$ non nulles pour des valeurs arbitrairement grandes de X (resp. pour tout X assez grand), où on a identifié $\bigwedge^2 \mathbb{R}^n$ à l'espace euclidien \mathbb{R}^N (avec $N = \binom{n}{2}$) comme dans la partie 2.2.

La proposition qui suit explicite la correspondance entre ces deux définitions et les définitions 6.1.1 et 6.1.2. Nous en proposons une démonstration personnelle.

Proposition 7.2.3. Soit $\mathbf{u} := (1, \xi_1, \dots, \xi_{n-1}) \in \mathbb{R}^n$ dont on suppose les coordonnées linéairement indépendantes sur \mathbb{Q} . Alors

$$(\omega, \widehat{\omega}, \omega^*, \widehat{\omega}^*) = (\lambda(\mathbf{u}), \widehat{\lambda}(\mathbf{u}), \tau(\mathbf{u}), \widehat{\tau}(\mathbf{u})),$$

où $\omega, \widehat{\omega}, \omega^*, \widehat{\omega}^*$ sont les constantes des définitions 6.1.1 et 6.1.2.

Preuve Revenons aux définitions 6.1.1 et 6.1.2.

Remarquons que l'ensemble des $\eta > 0$ tels que le problème E (resp. E^*) possède une solution non triviale dans \mathbb{Z}^n pour des Q arbitrairement grands est un intervalle. Il en va de même lorsque l'on impose que cela soit vrai pour tout Q assez grand.

Montrons pour commencer que $\omega^* = \tau(\mathbf{u})$.

Fixons maintenant $\eta < \omega^*$ et $\varepsilon > 0$ tel que $\eta + \varepsilon < \omega^*$. Alors il existe des Q arbitrairement grands tels qu'il existe $(x, y_1, \dots, y_{n-1}) \in \mathbb{Z}^n$ non nul vérifiant

$$\left| x - \sum_{j=1}^{n-1} \xi_j y_j \right| \leq Q^{-\eta-\varepsilon} \quad \text{et} \quad |y_j| \leq Q \quad (j = 1, \dots, n-1).$$

On pose $\mathbf{x} := (-x, y_1, \dots, y_{n-1})$. On a alors $\|\mathbf{x}\| \leq c_1 Q$ où $c_1 = c_1(n, \xi)$ ne dépend que de $n, \xi_1, \dots, \xi_{n-1}$ mais pas de Q , et

$$|\mathbf{x} \cdot \mathbf{u}| = \left| x - \sum_{j=1}^{n-1} \xi_j y_j \right| \leq Q^{-\eta-\varepsilon} \leq (c_1 Q)^{-\eta}$$

pour Q assez grand. On en déduit que $\eta \leq \tau(\mathbf{u})$, et par suite que $\omega^* \leq \tau(\mathbf{u})$.

Réciproquement, si $\tau < \tau(\mathbf{u})$, alors une solution $\mathbf{x} = (-x, y_1, \dots, y_{n-1})$ du système

$$\|\mathbf{x}\| \leq X \quad \text{et} \quad |\mathbf{x} \cdot \mathbf{u}| \leq X^{-\tau}$$

fournit directement une solution (x, y_1, \dots, y_{n-1}) du problème E^* avec $Q = X$ et $\eta = \tau$. Donc $\tau \leq \omega^*$ et par suite $\tau(\mathbf{u}) = \omega^*$.

Un raisonnement analogue permet de montrer que $\widehat{\tau}(\mathbf{u}) = \widehat{\omega}^*$.

Montrons maintenant que $\omega = \lambda(\mathbf{u})$ et $\widehat{\omega} = \widehat{\lambda}(\mathbf{u})$.

L'idée est la même que précédemment, nous allons un peu plus vite pour traiter les deux cas ensemble. Soit $\mathbf{x} = (x, y_1, \dots, y_{n-1}) \in \mathbb{Z}^n$. Commençons par remarquer que les coordonnées du vecteurs $\mathbf{x} \wedge \mathbf{u}$ sont les mineurs d'ordre 2 de la matrice dont la première ligne est \mathbf{x} et la deuxième est \mathbf{u} ; ce sont donc les réels de la forme $x\xi_i - y_i$ ($i = 1, \dots, n-1$) et $y_i \xi_j - y_j \xi_i$ ($1 \leq i < j \leq n-1$).

Si $\mathbf{x} = (x, y_1, \dots, y_{n-1})$ vérifie les inégalités

$$\|\mathbf{x}\| \leq X \quad \text{et} \quad \|\mathbf{x} \wedge \mathbf{u}\| \leq X^{-\lambda}$$

pour $X \geq 1$ et $\lambda > 0$, alors en particulier \mathbf{x} est solution du problème E avec $Q = X$ et $\eta = \lambda$ puisque $|x| \leq \|\mathbf{x}\| \leq X$ et $|\xi_j x - y_j| \leq \|\mathbf{x} \wedge \mathbf{u}\| \leq X^{-\lambda}$ ($j = 1, \dots, n-1$). Cela implique $\omega \geq \lambda(\mathbf{u})$ et $\widehat{\omega} \geq \widehat{\lambda}(\mathbf{u})$.

Supposons maintenant que \mathbf{x} est une solution non nulle du problème E avec $Q \geq 1$ et $\eta + \varepsilon$ (avec $\eta, \varepsilon > 0$; en particulier $x \neq 0$). Alors

$$\|\mathbf{x}\| \leq c_2 Q$$

où $c_2 = c_2(n, \xi)$ est une constante qui ne dépend que de $n, \xi_1, \dots, \xi_{n-1}$ mais pas de Q . Par ailleurs puisque pour $1 \leq i < j \leq n-1$ on a

$$\begin{aligned} |y_i \xi_j - y_j \xi_i| &= \left| \frac{y_i}{x} (x \xi_j - y_j) - \frac{y_j}{x} (x \xi_i - y_i) \right| \\ &\leq c_3 Q^{-\eta-\varepsilon} \end{aligned}$$

où $c_3 = c_3(n, \xi)$ ne dépend que de $n, \xi_1, \dots, \xi_{n-1}$ et pas de Q , et que pour $1 \leq i \leq n-1$ on a aussi

$$|\xi_i x - y_i| \leq Q^{-\eta-\varepsilon},$$

on en déduit finalement

$$\|\mathbf{x} \wedge \mathbf{u}\| \leq c_4 Q^{-\eta-\varepsilon}$$

avec $c_4 = c_4(n, \xi)$ constante qui ne dépend que de $n, \xi_1, \dots, \xi_{n-1}$ et pas de Q . En particulier, si Q est assez grand on a

$$\|\mathbf{x} \wedge \mathbf{u}\| \leq (c_2 Q)^{-\eta}.$$

En résumé, cela implique que \mathbf{x} vérifie les inégalités

$$\|\mathbf{x}\| \leq X \quad \text{et} \quad \|\mathbf{x} \wedge \mathbf{u}\| \leq X^{-\eta}$$

où on a posé $X := c_2 Q$. Cela implique que $\omega \leq \lambda(\mathbf{u})$ et $\widehat{\omega} \leq \widehat{\lambda}(\mathbf{u})$, d'où l'égalité. \square

On peut alors établir un premier corollaire du théorème 7.1.3.

Corollaire 7.2.4. *Soit $n \geq 2$ un entier et $\delta > 0$. L'application $\theta : \mathbb{R}^4 \rightarrow \mathbb{R}^4$ définie par*

$$\theta(\tau, \widehat{\tau}, \widehat{\lambda}, \lambda) \rightarrow \left(\frac{1}{\tau+1}, \frac{1}{\widehat{\tau}+1}, \frac{\widehat{\lambda}}{\widehat{\lambda}+1}, \frac{\lambda}{\lambda+1} \right)$$

établit une bijection entre l'ensemble des quadruplets $(\tau(\mathbf{u}), \widehat{\tau}(\mathbf{u}), \widehat{\lambda}(\mathbf{u}), \lambda(\mathbf{u}))$ où \mathbf{u} parcourt l'ensemble des vecteurs unitaires de \mathbb{R}^n dont les coordonnées sont linéairement indépendantes sur \mathbb{Q} , et l'ensemble des quadruplets de la forme

$$\left(\liminf_{q \rightarrow \infty} \frac{P_1(q)}{q}, \limsup_{q \rightarrow \infty} \frac{P_1(q)}{q}, \liminf_{q \rightarrow \infty} \frac{P_n(q)}{q}, \limsup_{q \rightarrow \infty} \frac{P_n(q)}{q} \right)$$

où $\mathbf{P} = (P_1, \dots, P_n)$ parcourt l'ensemble des n -systèmes rigides de maille δ pour lesquels P_1 n'est pas borné.

Preuve Dans [19] Schmidt et Summerer montrent que si \mathbf{u} est un vecteur non nul dont les coordonnées sont Q -linéairement indépendantes, alors on a

$$\begin{aligned} &\left(\frac{1}{\tau(\mathbf{u})+1}, \frac{1}{\widehat{\tau}(\mathbf{u})+1}, \frac{\widehat{\lambda}(\mathbf{u})}{\widehat{\lambda}(\mathbf{u})+1}, \frac{\lambda(\mathbf{u})}{\lambda(\mathbf{u})+1} \right) = \\ &\left(\liminf_{q \rightarrow \infty} \frac{L_{\mathbf{u},1}(q)}{q}, \limsup_{q \rightarrow \infty} \frac{L_{\mathbf{u},1}(q)}{q}, \liminf_{q \rightarrow \infty} \frac{L_{\mathbf{u},n}(q)}{q}, \limsup_{q \rightarrow \infty} \frac{L_{\mathbf{u},n}(q)}{q} \right). \end{aligned}$$

La conclusion est immédiate avec le théorème de Roy puisque si $\mathbf{P} = (P_1, \dots, P_n)$ est un n -système rigide de maille δ tel que la différence $\mathbf{L}_{\mathbf{u}} - \mathbf{P}$ soit bornée, alors

$$\liminf_{q \rightarrow \infty} \frac{P_1(q)}{q} = \liminf_{q \rightarrow \infty} \frac{L_{\mathbf{u},1}(q)}{q},$$

et on a des inégalités similaires pour les autres composantes. Cf également la remarque ci-dessous. □

Remarque : Si \mathbf{u} est un vecteur unitaire de \mathbb{R}^n et $\mathbf{P} = (P_1, \dots, P_n)$ un n -système rigide de maille δ tel que $\mathbf{L}_{\mathbf{u}} - \mathbf{P}$ soit bornée, alors les coordonnées de \mathbf{u} sont linéairement indépendantes si et seulement si P_1 n'est pas bornée.

7.3 La théorie de Schmidt et Summerer

Nous reprenons ici la partie 2 de l'article de Roy [14] qui reprend elle-même les résultats de Schmidt et Summerer [19] dans un langage un peu adapté. Nous ne donnerons pas la preuve des résultats énoncés. L'une des définitions les plus importantes est celle de (n, γ) -système (cf définition 7.3.7) et le théorème principal de cette partie est le théorème 7.3.8 qui affirme que toute fonction $\mathbf{L}_{\mathbf{u}}$ peut être approchée à une différence bornée près par un (n, γ) -système. Il est à noter aussi l'importance du lemme 7.3.9 qui décrit les n -systèmes rigides de maille δ comme un sous-ensemble particulier des $(n, 0)$ -systèmes. Dans les deux premiers paragraphes 7.3.1 et 7.3.2 on introduit les définitions de base et les premières propriétés sur les familles de corps convexes paramétrés dans un cadre assez général. Le paragraphe 7.3.3 spécifie ces résultats pour la famille principale qui nous intéressera (famille de corps convexes de \mathbb{R}^n). Dans le paragraphe 7.3.4 nous étudions des familles de corps convexes pseudo-composés (dans $\bigwedge^k \mathbb{R}^n$) dont les propriétés seront importantes pour la construction d'un (n, γ) -système convenable vérifiant le théorème 7.3.8.

Le lecteur est invité à lire la partie 2.4 où il trouvera des rappels de géométrie des nombres (définition d'un corps convexe, des minima successifs etc). Dans cette partie, V désignera un espace vectoriel euclidien de dimension finie $N \geq 1$. Un réseau Λ de V désignera un sous-groupe discret de V de rang N (attention, dans la définition proposée dans la partie 2.4, un réseau n'était pas forcément de rang maximal, ce qu'on suppose dorénavant). Supposons que \mathcal{C} soit un corps convexe symétrique et Λ un réseau de V . Pour simplifier on notera $\lambda_1(\mathcal{C}) \leq \dots \leq \lambda_n(\mathcal{C})$ les minima successifs de Λ pour \mathcal{C} (Λ n'apparaît pas dans cette notation mais il n'y aura pas d'ambiguïté sur le réseau dans le contexte où on utilisera cette notation). Pour chaque $\mathbf{x} \in V$, on définit $\lambda_{\mathbf{x}}(\mathcal{C}) = j_{\mathcal{C}}(\mathbf{x})$ où $j_{\mathcal{C}}$ est la jauge de \mathcal{C} . Autrement dit, c'est le plus petit réel $\lambda \geq 0$ tel que $\mathbf{x} \in \lambda\mathcal{C}$. On le notera aussi parfois $\lambda(\mathbf{x}, \mathcal{C})$. On sait qu'il existe $(\mathbf{x}_1, \dots, \mathbf{x}_n)$ des vecteurs linéairement indépendants de Λ tels que $\lambda(\mathbf{x}_j, \mathcal{C}) = \lambda_j(\mathcal{C})$.

7.3.1 Une famille générale de corps convexes

Soit V défini comme au-dessus et Λ un réseau de V qu'on suppose de covolume égal à 1. On suppose que $V = U \oplus W$ avec $K := \dim(W) > 0$, et on note $N := \dim(V)$. On considère alors la famille de corps convexes

$$\mathcal{C}(Q) := \{\mathbf{x} \in V ; \|\mathbf{x}\| \leq 1 \text{ et } \|\text{proj}_W(\mathbf{x})\| \leq Q^{-1}\} \quad (Q \geq 1),$$

où proj_W est la projection orthogonale sur W . Pour chaque $j = 1, \dots, N$ on définit alors une fonction $L_j : [0, \infty[\rightarrow \mathbb{R}$ par

$$L_j(q) := \log \lambda_j(\mathcal{C}(e^q)) \quad (q \geq 0).$$

On a clairement $L_1(q) \leq \dots \leq L_N(q)$ pour tout $q \geq 0$; on regroupe alors ces fonctions en une unique application $\mathbf{L} : [0, \infty[\rightarrow \Delta_N$ en posant

$$\mathbf{L}(q) := (L_1(q), \dots, L_N(q)) \quad (q \geq 0).$$

Le second théorème de Minkowski permet de démontrer le résultat suivant :

Lemme 7.3.1. *Pour tout $q \geq 0$ on a*

$$|L_1(q) + \dots + L_N(q) - Kq| \leq N \log(N).$$

Preuve Cf [14], Lemme 2.1. □

Nous utiliserons ces résultats généraux dans l'espace euclidien \mathbb{R}^n mais également dans des espaces de la forme $\bigwedge^{(p)} \mathbb{R}^n$ (avec des corps convexes composés).

7.3.2 Trajectoires de points et graphes combinés

On garde les notations du paragraphe précédent. Pour tout $\mathbf{x} \in V$ et pour tout $Q \geq 1$, on a

$$\lambda_{\mathbf{x}}(\mathcal{C}(Q)) = \lambda(\mathbf{x}, \mathcal{C}(Q)) = \max\{\|\mathbf{x}\|, Q\|\text{proj}_W(\mathbf{x})\|\}.$$

Si $\mathbf{x} \neq 0$ ce nombre est strictement supérieur à 0 et on a alors (en posant $\log(0) := -\infty$)

$$L_{\mathbf{x}}(q) := \log \lambda_{\mathbf{x}}(\mathcal{C}(e^q)) = \max\{\log \|\mathbf{x}\|, q + \log \|\text{proj}_W(\mathbf{x})\|\} \quad (q \geq 0). \quad (7.1)$$

En particulier, c'est une fonction continue, affine par morceaux. Si $\text{proj}_W(\mathbf{x}) = 0$, $L_{\mathbf{x}}$ est constante égale à $\log \|\mathbf{x}\|$, sinon elle a pente 0 puis 1. Notons aussi que si \mathbf{x} et \mathbf{y} sont deux vecteurs non nuls de V linéairement dépendants, alors $L_{\mathbf{x}} - L_{\mathbf{y}}$ est constante. En particulier, elles ont la même dérivée aux points $q > 0$ en lesquels elles sont dérivables.

Définition 7.3.2. Si \mathbf{x} est un vecteur non nul de V , le graphe de la fonction $L_{\mathbf{x}}$ est appelé la *trajectoire* de \mathbf{x} .

Explicitement, la trajectoire de \mathbf{x} est l'ensemble $\{(q, L_{\mathbf{x}}(q)) ; q \geq 0\}$. Remarquons qu'on a l'inclusion

$$\{(q, L_j(q)) ; q \geq 0, 1 \leq j \leq N\} \subset \{(q, L_{\mathbf{x}}(q)) ; q \geq 0, \mathbf{x} \in \Lambda \setminus \{0\}\},$$

ce qui signifie en d'autres termes que le graphe combiné de L_1, \dots, L_n est contenu dans le graphe combiné des fonctions $L_{\mathbf{x}}$ avec $\mathbf{x} \in \Lambda \setminus \{0\}$, i.e. l'union des trajectoires de ces vecteurs. On peut déjà remarquer que

$$L_1(q) = \inf\{L_{\mathbf{x}}(q) ; \mathbf{x} \in \Lambda \setminus \{0\}\} \quad (q \geq 0).$$

Le lemme suivant dû à Schmidt et Summerer donne une première propriété fondamentale.

Lemme 7.3.3. *Les fonctions L_1, \dots, L_N sont continues, affines par morceaux avec pente 0 et 1. En tout point $q > 0$ en lequel L_1 change de pente 1 à 0 on a $L_1(q) = L_2(q)$.*

Preuve Cf [14]. □

7.3.3 La famille principale de corps convexes

Soit $n \geq 2$ un entier et u un vecteur unitaire de \mathbb{R}^n . On applique les résultats précédents avec $\mathbb{R}^n = U \overset{\perp}{\oplus} W$ où $W := \text{Vect}(u)$ et $U := W^\perp = \{\mathbf{x} \in \mathbb{R}^n ; \mathbf{x} \cdot \mathbf{u} = 0\}$, en utilisant le réseau $\Lambda := \mathbb{Z}^n$. Dans ce cas, puisque pour tout $\mathbf{x} \in \mathbb{R}^n$ on a $\|\text{proj}_W(\mathbf{x})\| = |\mathbf{x} \cdot \mathbf{u}|$, cela donne la famille de corps convexes

$$\mathcal{C}_{\mathbf{u}}(Q) := \{\mathbf{x} \in \mathbb{R}^n ; \|\mathbf{x}\| \leq 1 \text{ et } |\mathbf{x} \cdot \mathbf{u}| \leq Q^{-1}\} \quad (Q \geq 1),$$

et son application associée $\mathbf{L}_{\mathbf{u}} = (L_{\mathbf{u},1}, \dots, L_{\mathbf{u},n}) : [0, \infty[\rightarrow \Delta_n$ où $L_{\mathbf{u},j} := \log \lambda_j(\mathcal{C}_{\mathbf{u}}(e^q))$. Par les lemmes précédents, les fonctions $L_{\mathbf{u},j}$ sont continues, affines par morceaux avec pentes 0 et 1 et satisfont à l'inégalité

$$|L_{\mathbf{u},1}(q) + \dots + L_{\mathbf{u},n}(q) - q| \leq n \log(n) \quad (q \geq 0). \quad (7.2)$$

En particulier ce sont des fonctions croissantes. Par (7.1), la trajectoire d'un vecteur $\mathbf{x} \in \mathbb{Z}^n$ non nul pour cette famille de corps convexes $\mathcal{C}_{\mathbf{u}}$ est le graphe de la fonction $L_{\mathbf{x}} : [0, \infty[\rightarrow \mathbb{R}$ donnée par

$$L_{\mathbf{x}}(q) = L(\mathbf{x}, q) := \max\{\log \|\mathbf{x}\|, q + \log |\mathbf{x} \cdot \mathbf{u}|\} \quad (q \geq 0). \quad (7.3)$$

De manière équivalente, on peut noter que $\lambda(\mathbf{x}, \mathcal{C}_{\mathbf{u}}(Q)) = \max\{\|\mathbf{x}\|, Q|\mathbf{x} \cdot \mathbf{u}|\} \quad (Q \geq 1)$.

7.3.4 Familles de corps convexes pseudo-composés

On utilise ici certaines notations et résultats de la partie 2.4.

Soit \mathbf{u} et U définis comme dans le paragraphe précédent. On fixe un entier $k \in \{1, \dots, n\}$ et on munit l'espace vectoriel $V := \bigwedge^k \mathbb{R}^n$ de l'unique structure euclidienne telle que pour toute base orthonormée $(\mathbf{e}_1, \dots, \mathbf{e}_n)$ de \mathbb{R}^n , les produits extérieurs $\mathbf{e}_{j_1} \wedge \dots \wedge \mathbf{e}_{j_k}$ avec $1 \leq j_1 < \dots < j_k \leq n$ forment une base orthonormale de $\bigwedge^k \mathbb{R}^n$. On définit aussi $\Lambda := (\mathbb{Z}^n)^{(k)}$ (avec les notations de la partie 2.4) le réseau de covolume 1 engendré par les produits extérieurs de la forme $\mathbf{x}_1 \wedge \dots \wedge \mathbf{x}_k$ avec $\mathbf{x}_1, \dots, \mathbf{x}_k \in \mathbb{Z}^n$. On a une décomposition en somme directe orthogonale :

$$\bigwedge^k \mathbb{R}^n = U_k \overset{\perp}{\oplus} W_k,$$

où $U_k = U^{(k)} = \bigwedge^k U$ et $W_k = \bigwedge^{k-1} U \wedge \text{Vect}(\mathbf{u})$. On pose alors

$$N := \dim V = \binom{n}{k} \quad \text{et} \quad K := \dim(W_k) = \binom{n-1}{k-1},$$

et pour $Q \geq 1$ on définit

$$\mathcal{C}_{\mathbf{u}}^{[k]}(Q) := \{\omega \in \bigwedge^k \mathbb{R}^n ; \|\omega\| \leq 1 \text{ et } \|\text{proj}_{W_k}(\omega)\| \leq Q^{-1}\}.$$

On note $\mathbf{L}_{\mathbf{u}}^{(k)} = (L_{\mathbf{u},1}^{(k)}, \dots, L_{\mathbf{u},N}^{(k)}) : [0, \infty[\rightarrow \Delta_N$ l'application associée donnée par

$$L_{\mathbf{u},j}^{(k)}(q) = \log \lambda_j(\mathcal{C}_{\mathbf{u}}^{[k]}(e^q)) \quad (q \geq 0, 1 \leq j \leq N).$$

Par les lemmes 7.3.1 et 7.3.3, ces fonctions sont continues, affines par morceaux de pentes 0 et 1, et satisfont l'inégalité :

$$|L_{\mathbf{u},1}^{(k)}(q) + \dots + L_{\mathbf{u},N}^{(k)}(q) - Kq| \leq N \log(N) \quad (q \geq 0). \quad (7.4)$$

De plus on a $L_{\mathbf{u},1}^{(k)}(q) = L_{\mathbf{u},2}^{(k)}(q)$ en tout point $q > 0$ tel que $L_{\mathbf{u},1}^{(k)}(q)$ passe de pente de 1 à 0.

Quand $k = 1$ on a $\bigwedge^1 \mathbb{R}^n = \mathbb{R}^n$ et $\mathbf{L}_{\mathbf{u}}^{(1)} = \mathbf{L}_{\mathbf{u}}$. Dans le cas général, on utilise $\mathcal{C}_{\mathbf{u}}^{[k]}(Q)$ comme approximation du corps convexe composé k -ème $\mathcal{C}_{\mathbf{u}}(Q)^{(k)}$ de $\mathcal{C}_{\mathbf{u}}(Q)$ (cf partie 2.4). Le lemme suivant montre $\mathcal{C}_{\mathbf{u}}(Q)^{(k)} \subset k\mathcal{C}_{\mathbf{u}}^{[k]}(Q)$.

Lemme 7.3.4. Soit $Q \geq 1$ et $\mathbf{x}_1, \dots, \mathbf{x}_k \in \mathcal{C}_{\mathbf{u}}(Q)$. Alors $\mathbf{x}_1 \wedge \dots \wedge \mathbf{x}_k \in k\mathcal{C}_{\mathbf{u}}^{[k]}(Q)$.

Preuve Cf [14], Lemme 2.3. □

D'après (7.1), la trajectoire d'un vecteur non nul $\omega \in \bigwedge^k \mathbb{Z}^n$ pour la famille de corps convexes $\mathcal{C}_{\mathbf{u}}^{[k]}(Q)$ est le graphe de la fonction $L_\omega : [0, \infty[\rightarrow \mathbb{R}$ donnée par

$$L_\omega(q) = L(\omega, q) := \max\{\log \|\omega\|, q + \log \|\text{proj}_{W_k}(\omega)\|\} \quad (q \geq 0). \quad (7.5)$$

Le lemme suivant est presque une conséquence directe - les constantes mises en jeu étant un peu différentes - du théorème 2.4.9.

Lemme 7.3.5. Soit $q \geq 0$. On note $(S_{\mathbf{u},1}^{(k)}(q), \dots, S_{\mathbf{u},N}^{(k)}(q))$ la suite de toutes les sommes de la forme $L_{\mathbf{u},j_1}(q) + \dots + L_{\mathbf{u},j_k}(q)$ avec $1 \leq j_1 < \dots < j_k \leq n$ rangées dans l'ordre croissant. Alors on a

$$-\log(n) \leq S_{\mathbf{u},j}^{(k)}(q) - L_{\mathbf{u},j}^{(k)}(q) \leq 2^n n \log(n) \quad (q \geq 0, 1 \leq j \leq N).$$

Preuve Cf [14], Lemme 2.6. □

On pourra retenir notamment les conséquences suivantes :

Lemme 7.3.6. On pose $c_1 := 2^n n \log(n)$. Alors pour tout $q \geq 0$ on a

- (i) $|L_{\mathbf{u},1}^{(k)}(q) - L_{\mathbf{u},1}(q) - \dots - L_{\mathbf{u},k}(q)| \leq c_1$,
- (ii) $|L_{\mathbf{u},2}^{(k)}(q) - L_{\mathbf{u},1}(q) - \dots - L_{\mathbf{u},k-1}(q) - L_{\mathbf{u},k+1}(q)| \leq c_1$ si $1 < k < n$,
- (iii) $|L_{\mathbf{u},j}^{(n-1)}(q) + L_{\mathbf{u},n+1-j}(q) - q| \leq c_1 + n \log(n)$ pour $(j = 1, \dots, n)$.

Preuve Cela découle directement du lemme précédent et de l'inégalité (7.2). □

7.3.5 Le théorème d'approximation de Schmidt et Summerer

La définition qui suit est issue de [19] et reformulée par Roy pour l'adapter à son contexte. C'est cette dernière que nous présentons.

Définition 7.3.7. Soit $\gamma, q_0 \geq 0$. Un (n, γ) -système sur l'intervalle $[q_0, \infty[$ est une fonction $\mathbf{P} = (P_1, \dots, P_n) : [q_0, \infty[\rightarrow \mathbb{R}^n$ vérifiant les cinq conditions suivantes :

- (S1) $-\gamma \leq P_j(q) \leq P_{j+1}(q) + \gamma$ ($1 \leq j < n, q_0 \leq q$).
- (S2) $P_j(q_1) \leq P_j(q_2) + \gamma$ ($1 \leq j \leq n, q_0 \leq q_1 \leq q_2$).
- (S3) La fonction $M_j := P_1 + \dots + P_j : [q_0, \infty[\rightarrow \mathbb{R}$ est continue, affine par morceaux de pentes 0 et 1 (pour $j = 1, \dots, n$).
- (S4) $M_n(q) = q$ ($q_0 \leq q$).
- (S5) Si pour un certain $j \in \{1, \dots, n-1\}$ la fonction M_j passe d'une pente 1 à une pente 0 au point $q > q_0$ alors $P_{j+1}(q) \leq P_j(q) + \gamma$.

Remarque : Les $(n, 0)$ -systèmes ont des propriétés beaucoup plus simples et agréables que les (n, γ) -systèmes généraux. Remarquons aussi que si \mathbf{P} est un (n, γ) -système sur

$[q_0, \infty[$ alors pour tout $\delta > 0$ la fonction $q \mapsto \delta^{-1}\mathbf{P}(\delta q)$ est un $(n, \gamma/\delta)$ -système sur $[q_0/\delta, \infty[$.

Le résultat suivant est issu de [19] ; nous donnons la version reformulée de Roy. Il montre l'importance de la notion de (n, γ) -système.

Théorème 7.3.8 (Schmidt-Summerer, 2013). *Soit $\gamma := 6n2^n \log(n)$. Alors pour tout \mathbf{u} vecteur unitaire de \mathbb{R}^n il existe un (n, γ) -système $\mathbf{P} : [0, \infty[\rightarrow \mathbb{R}^n$ tel que*

$$\sup_{q \geq 0} \|\mathbf{P}(q) - \mathbf{L}_{\mathbf{u}}(q)\|_{\infty} \leq \gamma.$$

Preuve Nous ne donnons ici que la construction du (n, γ) -système qui convient, le lecteur est renvoyé à [14] (théorème 2.9) pour les détails.

On pose $M_0 := 0$ et $M_k := L_{\mathbf{u},1}^{(k)}$ pour $k = 1, \dots, n$. Par le lemme 7.3.6 (i), on a

$$|M_k(q) - L_{\mathbf{u},1}(q) - \dots - L_{\mathbf{u},k}(q)| \leq c_1 \quad (1 \leq k \leq n, 0 \leq q).$$

On définit alors $P_k := M_k - M_{k-1}$ pour $k = 1, \dots, n$ et on obtient

$$|P_k(q) - L_{\mathbf{u},k}(q)| \leq 2c_1 \quad (1 \leq k \leq n, 0 \leq q).$$

Alors $\mathbf{P} := (P_1, \dots, P_n)$ est un (n, γ) -système avec toutes les propriétés requises. \square

Terminons cette partie par un résultat important qui permet de voir les n -systèmes rigides comme un sous-ensemble des $(n, 0)$ -systèmes.

Lemme 7.3.9. *Soit $\delta \in]0, \infty[$. Les n -systèmes rigides de maille δ sont exactement les $(n, 0)$ -systèmes $(P_1, \dots, P_n) : [q_0, \infty[\rightarrow \mathbb{R}^n$ qui vérifient que pour $q = q_0$ et pour tout $q > q_0$ pour lequel au moins l'une des fonctions $P_1 + \dots + P_j$ ($1 \leq j < n$) passe de pente 0 à pente 1, les nombres $P_1(q), \dots, P_n(q)$ sont des multiples deux à deux distincts de δ .*

En particulier, un n -système rigide $(P_1, \dots, P_n) : [q_0, \infty[\rightarrow \mathbb{R}^n$ vérifie toujours

$$P_1(q) + \dots + P_n(q) = q \quad (q \geq q_0).$$

7.4 Esquisse de la construction de Roy

Nous reprenons ici la structure de la preuve de Roy du théorème 7.1.3. Le lecteur est renvoyé à [14] pour les détails manquants des démonstrations des résultats intermédiaires. Dans cette partie n désigne un entier ≥ 2 . Les paragraphes 7.4.1, 7.4.2 et 7.4.3 introduisent tous les outils nécessaires à la démonstration de la seconde assertion du théorème de Roy. Elle repose sur la construction d'une famille de bases de \mathbb{Z}^n qui possèdent de fortes propriétés. Le paragraphe 7.4.4 fournit quant à lui les outils pour étendre le théorème 7.3.8 de Schmidt et Summerer et prouver la première assertion du théorème de Roy ; l'auteur introduit la notion de (n, γ) -systèmes réduits et montre que sous certaines conditions tout n -système rigide peut être approché - à une différence bornée près - par un tel (n, γ) -système réduit. Cette classe de fonctions intermédiaires permettra de passer des (n, γ) -systèmes aux n -systèmes rigides. Dans un dernier paragraphe 7.4.5 on combine les différents résultats intermédiaires des paragraphes précédents pour achever la démonstration.

Notations : Si $\mathbf{x}_1, \dots, \mathbf{x}_k \in V$ où V est un K -espace vectoriel, on note $\text{Vect}_K(\mathbf{x}_1, \dots, \mathbf{x}_k) = \text{Vect}(\mathbf{x}_1, \dots, \mathbf{x}_k)$ l'espace vectoriel engendré par $\mathbf{x}_1, \dots, \mathbf{x}_k$.

On note aussi $\text{Vect}_{\mathbb{Z}}(\mathbf{x}_1, \dots, \mathbf{x}_k)$ le \mathbb{Z} -module engendré par $\mathbf{x}_1, \dots, \mathbf{x}_k$. Roy utilise les notations $\langle \mathbf{x}_1, \dots, \mathbf{x}_k \rangle_K$ et $\langle \mathbf{x}_1, \dots, \mathbf{x}_k \rangle_{\mathbb{Z}}$, mais celles-ci auraient pu engendrer des confusions avec la notation du produit scalaire définie au début de la partie 3.

Rappelons que $\|\cdot\|$ désigne la norme euclidienne. On fera également usage de la norme $\|\cdot\|_{\infty}$ définie sur \mathbb{R}^n par $\|(x_1, \dots, x_n)\|_{\infty} = \max_{1 \leq j \leq n} |x_j|$.

7.4.1 Un premier résultat

Nous commençons par nous intéresser à la deuxième assertion du théorème de Roy. Afin de construire un vecteur unitaire \mathbf{u} tel que $\mathbf{L}_{\mathbf{u}} - \mathbf{P}$ soit bornée dans le cas où \mathbf{P} est un n -système rigide de maille suffisamment grande, Roy commence par supposer que \mathbf{u} existe, et il en déduit l'existence d'une suite de n -uplets de vecteurs de \mathbb{Z}^n possédant de fortes propriétés. C'est ce qu'exprime le théorème suivant.

Théorème 7.4.1. *Soit $\delta, \varepsilon > 0$ tels que $\delta \geq 6(n\varepsilon + c_1)$ où $c_1 = 2^n n \log(n)$ comme dans le lemme 7.3.6. Soit $\mathbf{P} = (P_1, \dots, P_n) : [q_0, \infty[\rightarrow \mathbb{R}^n$ un n -système rigide de maille δ . On suppose qu'il existe un vecteur unitaire $\mathbf{u} \in \mathbb{R}^n$ tel que $\|\mathbf{P}(q) - \mathbf{L}_{\mathbf{u}}(q)\|_{\infty} \leq \varepsilon$ pour tout $q \geq q_0$. On considère les suites $(q_i)_{0 \leq i < s}$, $(k_i)_{0 \leq i < s}$ et $(l_i)_{0 \leq i < s}$ attachées à \mathbf{P} comme dans la définition 7.1.2 et on pose $q_s = \infty$ si $s < \infty$. Alors, pour tout entier i tel que $0 \leq i < s$, il existe un n -uplet $(\mathbf{x}_1^{(i)}, \dots, \mathbf{x}_n^{(i)})$ de vecteurs linéairement indépendants de \mathbb{Z}^n vérifiant :*

- 1) $\|\mathbf{P}(q) - \Phi_n(L(\mathbf{x}_1^{(i)}, q), \dots, L(\mathbf{x}_n^{(i)}, q))\|_{\infty} \leq \varepsilon \quad (q_i \leq q < q_{i+1})$,
- 2) $(\mathbf{x}_1^{(i)}, \dots, \widehat{\mathbf{x}_{k_i}^{(i)}}, \dots, \mathbf{x}_n^{(i)}) = (\mathbf{x}_1^{(i+1)}, \dots, \widehat{\mathbf{x}_{l_{i+1}}^{(i+1)}}, \dots, \mathbf{x}_n^{(i+1)})$ si $i + 1 < s$,
- 3) $|\log \|\mathbf{x}_j^{(i)}\| - P_j(q_i)| \leq \varepsilon \quad (j = 1, \dots, n)$ si $i \geq 1$,
- 4) $\mathbf{x}_{l_{i+1}}^{(i+1)} \in \text{Vect}(\mathbf{x}_1^{(i)}, \dots, \mathbf{x}_{l_{i+1}}^{(i)})$ si $i + 1 < s$,
- 5) $0 \leq \log |\det(\mathbf{x}_1^{(i)}, \dots, \mathbf{x}_n^{(i)})| \leq n\varepsilon + \log(n)$,
- 6) $|\log \|\mathbf{x}_1^{(i)} \wedge \dots \wedge \widehat{\mathbf{x}_{k_i}^{(i)}} \wedge \dots \wedge \mathbf{x}_n^{(i)}\| - \sum_{j \neq k_i} \log \|\mathbf{x}_j^{(i)}\|| \leq n\varepsilon + 2c_1$ si $i \geq 1$.

La propriété 1) signifie que sur chaque intervalle $[q_i, q_{i+1}[$, le graphe combiné de \mathbf{P} est contenu dans un ε -voisinage de l'union des trajectoires de $\mathbf{x}_1^{(i)}, \dots, \mathbf{x}_n^{(i)}$. En utilisant en plus l'hypothèse de l'énoncé, cela implique que pour chaque q dans cet intervalle, à 2ε près $\mathbf{x}_1^{(i)}, \dots, \mathbf{x}_n^{(i)}$ réalisent les logarithmes des minima successifs de $\mathcal{C}_{\mathbf{u}}(e^q)$. A cause du graphe combiné particulier d'un n -système rigide, on déduit aussi de 1) qu'il y a exactement un vecteur $\mathbf{x}_j^{(i)}$ dont la trajectoire a pente 1 sur $[q_i + 2\varepsilon, \infty[$ alors que les autres ont pente 0 sur $[0, q_{i+1} - 2\varepsilon[$. Si $i + 1 < s$, 2) assure que ce vecteur particulier est en fait $\mathbf{x}_{k_i}^{(i)}$. Sur l'intervalle suivant $[q_{i+1}, q_{i+2}[$, sa trajectoire est remplacée par celle d'un nouveau vecteur $\mathbf{x}_{l_{i+1}}^{(i+1)}$ alors que les autres trajectoires sont conservées. La figure suivante issue de [14] illustre ceci sur un exemple avec $n = 5$. Les lignes continues représentent le graphe combiné de \mathbf{P} et celles en pointillés la trajectoire des points $\mathbf{x}_j^{(i)}$.

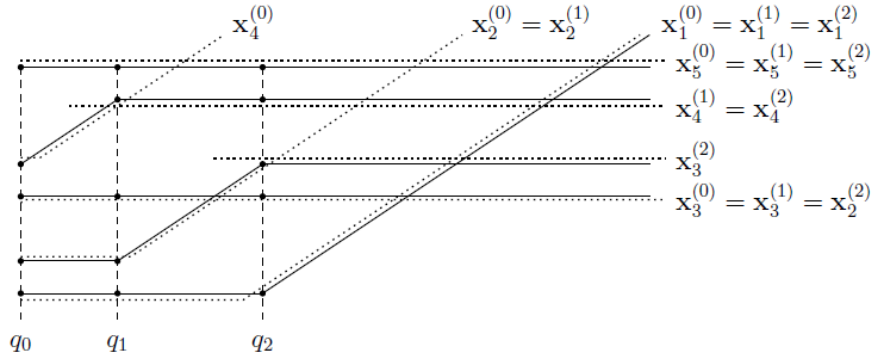


FIGURE 2 – Graphe combiné d'un 5-système rigide et les trajectoires des vecteurs entiers qui l'approximent à ε près.

Les propriétés 3) à 6) donnent des informations supplémentaires sur les $\mathbf{x}_j^{(i)}$. Par exemple, 5) assure que chaque n -uplet $(\mathbf{x}_1^{(i)}, \dots, \mathbf{x}_n^{(i)})$ engendre un sous-groupe de \mathbb{Z}^n dont l'indice est uniformément borné en i . La propriété 6) implique le résultat (beaucoup plus faible) que les angles entre deux vecteurs parmi $\mathbf{x}_1^{(i)}, \dots, \widehat{\mathbf{x}_{k_i}^{(i)}}, \dots, \mathbf{x}_n^{(i)}$ sont minorés (uniformément en i). Dans la construction de la preuve, on imposera ici que la famille $\mathbf{x}_1^{(i)}, \dots, \widehat{\mathbf{x}_{k_i}^{(i)}}, \dots, \mathbf{x}_n^{(i)}$ soit presque orthogonale (cf définition 7.4.12).

Préliminaires pour la construction des points $x_j^{(i)}$

Roy établit d'abord quelques lemmes intermédiaires que nous retranscrivons ici. Supposons que les hypothèses du théorème 7.4.1 soient satisfaites. Pour $\mathbf{x} \in \mathbb{R}^n$ non nul et $j \in \{1, \dots, n\}$ on pose

$$L(\mathbf{x}, \infty) := \lim_{q \rightarrow \infty} L(\mathbf{x}, q), \quad L_{\mathbf{u}, j}(\infty) := \lim_{q \rightarrow \infty} L_{\mathbf{u}, j}(q), \quad P_j(\infty) := \lim_{q \rightarrow \infty} P_j(q).$$

On définit aussi

$$V_j(q) := \text{Vect}(\{\mathbf{x} \in \mathbb{Z}^n \setminus \{0\} ; L(\mathbf{x}, q) \leq L_{\mathbf{u}, j}(q)\})_{\mathbb{R}} \quad (1 \leq j \leq n, 0 \leq q \leq \infty),$$

$$V_0(q) := 0 \quad (0 \leq q \leq \infty).$$

L'espace vectoriel $V_j(\infty)$ n'est intéressant que quand $L_{\mathbf{u}, j}(\infty) < \infty$ (ce qui équivaut, d'après les hypothèses de l'énoncé, à $P_j(\infty) < \infty$). Dans ce cas, $V_j(\infty)$ est orthogonal à \mathbf{u} et on a $L(\mathbf{x}, q) = \log \|\mathbf{x}\|$ ($q \geq 0$) pour tout $\mathbf{x} \in V_j(\infty)$ non nul. Sinon, on a $V_j(\infty) = \mathbb{R}^n$.

Lemme 7.4.2. *Soit $j \in \{1, \dots, n-1\}$. Alors sur tout sous-intervalle I de $[q_0, \infty[$ sur lequel on a $P_{j+1}(q) > P_j(q) + 2\varepsilon$, la famille d'espaces vectoriels $V_j(q)$ est constante de dimension j .*

Lemme 7.4.3. *Soit i un entier tel que $0 \leq i < s$. Alors il existe un point $\mathbf{x} \in \mathbb{Z}^n$ qui n'appartient pas à $V_{k_i-1}(q_i)$ et tel que*

$$|L(\mathbf{x}, t) - \underbrace{(P_{k_i}(q_i) + \max(0, t - q_i))}_{=P_{k_i}(t) \text{ au voisinage de } q_i}| \leq \varepsilon \quad (t \geq 0).$$

En particulier la trajectoire de \mathbf{x} a pente 0 sur $[0, q_i - 2\varepsilon]$ si $i \geq 1$ et pente 1 sur $[q_i + 2\varepsilon, \infty[$.

Lemme 7.4.4. *Soit j un entier tel que $1 \leq j < n$. Supposons que $P_j(\infty) < \infty$. Alors $V_j(\infty)$ est de dimension j et il existe un réel $q \geq 0$ tel que $V_j(t) = V_j(\infty)$ pour tout $t \geq q$. De plus, il existe un vecteur $\mathbf{x} \in \mathbb{Z}^n \cap V_j(\infty)$ qui n'appartient pas à $V_{j-1}(\infty)$ et tel que*

$$|L(\mathbf{x}, t) - P_j(\infty)| \leq \varepsilon \quad (t \geq 0).$$

Notons aussi que ces lemmes impliquent que $V_j(q_i)$ est toujours de dimension j pour tout $1 \leq i \leq s$ ($i \neq \infty$) et $0 \leq j \leq n$.

Preuve [des trois lemmes] Cf Roy [14] (lemmes 3.2, 3.3 et 3.4). □

Construction des points $x_j^{(i)}$

On suit la construction proposée par Roy.

On note \mathcal{S} le point $(q_0, P_{k_0}(q_0))$ dont on fait l'union avec l'ensemble de tous les segments horizontaux maximaux contenus dans le graphe combiné de \mathbf{P} qui ne sont pas des points. A chaque $S \in \mathcal{S}$ on associe un vecteur $\mathbf{x}_S \in \mathbb{Z}^n$ de la manière suivante :

Cas n°1 : si S est borné. Dans ce cas son extrémité de droite est de la forme $(q_i, P_{k_i}(q_i))$ avec $0 \leq i < s$. On choisit alors pour \mathbf{x}_S un vecteur qui vérifie la propriété du lemme 7.4.3.

Cas n°2 : si S n'est pas borné. Dans ce cas, S est contenu dans $[0, \infty[\times\{P_j(\infty)\}$ pour un certain j ($1 \leq j < n$) tel que $P_j(\infty) < \infty$. On choisit alors pour \mathbf{x}_S un vecteur qui vérifie la propriété du lemme 7.4.4.

Maintenant, pour tous i, j tels que $0 \leq i < s$ et $1 \leq j \leq n$, il existe un unique segment horizontal $S \in \mathcal{S}$ contenant le point $(q_i, P_j(q_i))$ et on pose $\mathbf{x}_j^{(i)} := \mathbf{x}_S$. Dans le cas dégénéré où $s < \infty$, on a $P_j(\infty) < \infty$ pour tout $j = 1, \dots, n-1$. Pour ces j on pose alors $\mathbf{x}_j^{(s)} := \mathbf{x}_S$ où $S \in \mathcal{S}$ est l'unique segment non borné contenu dans $[0, \infty[\times\{P_j(\infty)\}$.

Ces vecteurs $\mathbf{x}_j^{(i)}$ vérifient les propriétés du théorème 7.4.1.

Preuve Cf [14] (p.15-18). L'un des points les plus délicats de la preuve est l'indépendance linéaire de la famille $(\mathbf{x}_1^{(i)}, \dots, \mathbf{x}_n^{(i)})$. □

7.4.2 Retour sur la hauteur de sous-espaces

Nous rappelons et complétons quelques définitions et résultats déjà présentés dans la partie 4 en introduisant la notion de distance entre sous-espaces et en énonçant quelques premières propriétés. La notion principale de ce paragraphe est celle de famille *presque orthogonale* (définition 7.4.12). Le lecteur est renvoyé à [14] §4 pour les preuves détaillées des lemmes présentés.

Nous sommes ici dans le cas particulier où le corps de nombres K est égal au corps \mathbb{Q} des nombres rationnels. Soit $n \geq 2$ un entier. On dit qu'un sous-espace vectoriel

$V \subset \mathbb{R}^n$ est défini sur \mathbb{Q} s'il est engendré par des éléments de \mathbb{Q}^n . Si $V \neq 0$, cela revient à imposer que $V \cap \mathbb{Z}^n$ est un réseau de V . On note alors $H(V)$ sa hauteur, définie par

$$H(V) = \|\mathbf{x}_1 \wedge \cdots \wedge \mathbf{x}_m\|,$$

où $(\mathbf{x}_1, \dots, \mathbf{x}_m)$ est n'importe quelle base de $V \cap \mathbb{Z}^n$ et $\|\cdot\|$ est la norme euclidienne de $\bigwedge^m \mathbb{R}^n$ qu'on a muni de la structure euclidienne comme dans le paragraphe 7.3.4. C'est le covolume du réseau $V \cap \mathbb{Z}^n$. Notons qu'en particulier on a $H(\mathbb{R}^n) = 1$ et qu'on avait posé $H(0) = 1$.

Le lemme suivant permet d'exprimer de manière simple la hauteur d'un hyperplan de \mathbb{R}^n dans un cas particulier.

Lemme 7.4.5. *Soit $(\mathbf{x}_1, \dots, \mathbf{x}_n)$ une base de \mathbb{Z}^n et \mathbf{u} un vecteur unitaire de \mathbb{R}^n orthogonal à $V := \text{Vect}_{\mathbb{R}}(\mathbf{x}_1, \dots, \mathbf{x}_{n-1})$. Alors*

$$H(V) = |\mathbf{x}_n \cdot \mathbf{u}|^{-1}.$$

Preuve On a $1 = \|\mathbf{x}_1 \wedge \cdots \wedge \mathbf{x}_n\| = \|\mathbf{x}_1 \wedge \cdots \wedge \mathbf{x}_{n-1}\| |\mathbf{x}_n \cdot \mathbf{u}| = H(V) |\mathbf{x}_n \cdot \mathbf{u}|$. □

Définition 7.4.6. Si \mathbf{x} et \mathbf{y} sont deux vecteurs non nuls de \mathbb{R}^n on définit

$$\text{dist}(\mathbf{x}, \mathbf{y}) := \frac{\|\mathbf{x} \wedge \mathbf{y}\|}{\|\mathbf{x}\| \|\mathbf{y}\|}$$

la distance (projective) entre \mathbf{x} et \mathbf{y} .

Notons que $\text{dist}(\mathbf{x}, \mathbf{y}) = \omega(\mathbf{x}, \mathbf{y})$ avec les notations de la partie 3. Il représente le sinus de l'angle aigu entre les deux droites engendrées par \mathbf{x} et \mathbf{y} . C'est une fonction de $(\mathbb{R}^n \setminus \{0\})^2$ dans $[0, 1]$ continue et symétrique. On a par ailleurs vu qu'elle satisfaisait l'inégalité triangulaire.

Définition 7.4.7. Pour tout $\mathbf{x} \in \mathbb{R}^n$ non nul et tout V sous-espace non nul de \mathbb{R}^n , on définit la distance de \mathbf{x} à V par

$$\text{dist}(\mathbf{x}, V) := \inf\{\text{dist}(\mathbf{x}, \mathbf{y}) ; \mathbf{y} \in V \setminus \{0\}\} = \inf\{\text{dist}(\mathbf{x}, \mathbf{y}) ; \mathbf{y} \in S^n \cap V\},$$

où S^n désigne la sphère unité de \mathbb{R}^n .

On pose $\text{dist}(\mathbf{x}, 0) = 1$.

On a alors le résultat suivant qui permet d'exprimer de manière plus explicite la distance d'un vecteur à un sous-espace :

Lemme 7.4.8. *Soit $\mathbf{x} \in \mathbb{R}^n \setminus \{0\}$ et V un sous-espace quelconque de \mathbb{R}^n . Alors*

$$\text{dist}(\mathbf{x}, V) = \frac{\|\text{proj}_{V^\perp}(\mathbf{x})\|}{\|\mathbf{x}\|}.$$

De plus, pour tout sous-espace U de \mathbb{R}^n contenant V on a $\text{dist}(\mathbf{x}, V) \geq \text{dist}(\mathbf{x}, U)$.

Définition 7.4.9. Soient V_1 et V_2 deux sous-espaces non nuls de \mathbb{R}^n . On définit la distance de V_1 à V_2 par

$$\text{dist}(V_1, V_2) := \sup\{\text{dist}(\mathbf{x}, V_2) ; \mathbf{x} \in V_1 \setminus \{0\}\} = \sup\{\text{dist}(\mathbf{x}, V_2) ; \mathbf{x} \in S^n \cap V_1\}.$$

Comme $S^n \cap V_1$ est compact et que la fonction $\text{dist}(\cdot, V_2)$ est continue par le lemme précédent, la borne supérieure est atteinte par un certain $\mathbf{x} \in S^n \cap V_1$. Notons que cette distance n'est en général pas symétrique en V_1, V_2 (par exemple, si $V_1 \subset V_2$ avec $V_1 \neq V_2$ alors on a $\text{dist}(V_1, V_2) = 0$ alors que $\text{dist}(V_2, V_1) > 0$). On a toutefois les propriétés suivantes :

Lemme 7.4.10. *Soit $\mathbf{x} \in \mathbb{R}^n \setminus \{0\}$ et V_1, V_2 des sous-espaces non nuls de \mathbb{R}^n . Alors on a*

$$\text{dist}(\mathbf{x}, V_2) \leq \text{dist}(\mathbf{x}, V_1) + \text{dist}(V_1, V_2).$$

De plus, si V est un sous-espace non nul de \mathbb{R}^n alors on a aussi l'inégalité triangulaire

$$\text{dist}(V, V_2) \leq \text{dist}(V, V_1) + \text{dist}(V_1, V_2).$$

Le lemme suivant traite d'un cas particulier où la fonction distance entre deux sous-espaces est symétrique.

Lemme 7.4.11. *Soient V_1, V_2 deux sous-espaces non nuls de codimension 1 dans un sous-espace $U \subset \mathbb{R}^n$. Pour $i = 1, 2$, on choisit un vecteur unitaire \mathbf{u}_i dans $U \cap V_i^\perp$. Alors on a*

$$\text{dist}(V_1, V_2) = \text{dist}(\mathbf{u}_1, \mathbf{u}_2).$$

Définition 7.4.12. Soit $m \geq 1$ un entier. Une suite $(\mathbf{x}_1, \dots, \mathbf{x}_m)$ de vecteurs de \mathbb{R}^n est dite *presque orthogonale* si elle est libre et vérifie les inégalités

$$\text{dist}(\mathbf{x}_j, \text{Vect}_{\mathbb{R}}(\mathbf{x}_1, \dots, \mathbf{x}_{j-1})) \geq 1 - \frac{1}{2^{j-1}} \quad (2 \leq j \leq m).$$

Notons que $(\mathbf{x}_1, \dots, \mathbf{x}_m)$ est orthogonale si et seulement si pour tout j tel que $2 \leq j \leq m$ on a $\text{dist}(\mathbf{x}_j, \text{Vect}_{\mathbb{R}}(\mathbf{x}_1, \dots, \mathbf{x}_{j-1})) = 1$.

Remarquons que la suite (\mathbf{x}) avec $\mathbf{x} \in \mathbb{R}^n$ non nul est presque orthogonale. Par le lemme 7.4.8, on a aussi que toute sous-suite non vide d'une suite presque orthogonale est elle-même presque orthogonale.

Nous terminons ce paragraphe par un résultat lié à cette notion.

Lemme 7.4.13. *Soit $(\mathbf{x}_1, \dots, \mathbf{x}_m)$ un m -uplet primitif de vecteurs de \mathbb{Z}^n presque orthogonal et $U := \text{Vect}_{\mathbb{R}}(\mathbf{x}_1, \dots, \mathbf{x}_m)$. Alors on a*

$$e^{-2} \|\mathbf{x}_1\| \dots \|\mathbf{x}_m\| \leq H(U) \leq \|\mathbf{x}_1\| \dots \|\mathbf{x}_m\|.$$

7.4.3 Cas des systèmes rigides de grande maille

Cette partie reprend la partie 5 de [14]. Roy démontre la seconde assertion du théorème 7.1.3 dans le cas des n -systèmes rigides de grande maille (théorème 7.4.16). Comme le suggère le théorème 5.3.3, on commence par construire par récurrence une suite de bases de \mathbb{Z}^n qui vérifient de fortes propriétés. Nous ne présenterons pas les preuves en détails ici, juste le début des constructions de Roy.

On fixe $s \in \mathbb{N}^* \cup \{\infty\}$ et on pose

$$C := 2^{n+3} e^4.$$

On suppose que pour tout entier i tel que $0 \leq i < s$ il existe un vecteur $\mathbf{A}^{(i)} = (A_1^{(i)}, \dots, A_n^{(i)}) \in \mathbb{R}^n$ et des entiers k_i et l_i vérifiant les conditions

$$1 \leq k_0 \leq l_0 = n \quad \text{et} \quad 1 \leq k_i < l_i \leq n \quad \text{si} \quad i \geq 1 \quad (7.6)$$

$$A_1^{(i)} \geq C, \quad A_j^{(i)} \geq A_{j-1}^{(i)} C \quad \text{pour} \quad j = 2, \dots, n, \quad (7.7)$$

$$k_{i-1} \leq l_i \quad \text{et} \quad A_{l_i}^{(i)} \geq A_{l_i}^{(i-1)} C \quad \text{si} \quad i \geq 1, \quad (7.8)$$

$$(A_1^{(i)}, \dots, \widehat{A_{l_i}^{(i)}}, \dots, A_n^{(i)}) = (A_1^{(i-1)}, \dots, \widehat{A_{k_{i-1}}^{(i-1)}}, \dots, A_n^{(i-1)}) \quad \text{si} \quad i \geq 1. \quad (7.9)$$

Ces conditions sont évidemment à mettre en lien avec celles intervenant dans la définition d'un canevas (cf définition 7.1.1). On a alors le résultat suivant :

Proposition 7.4.14. *Pour tout entier i tel que $0 \leq i < s$, il existe une base de \mathbb{Z}^n $(\mathbf{x}_1^{(i)}, \dots, \mathbf{x}_n^{(i)})$ qui vérifie les propriétés suivantes :*

- 1) $(\mathbf{x}_1^{(0)}, \dots, \mathbf{x}_{n-1}^{(0)})$ est presque orthogonale,
- 2) $(\mathbf{x}_1^{(i)}, \dots, \widehat{\mathbf{x}_{k_i}^{(i)}}, \dots, \mathbf{x}_n^{(i)})$ est presque orthogonale,
- 3) $A_j^{(i)} \leq \|\mathbf{x}_j^{(i)}\| \leq 2A_j^{(i)}$ pour $j = 1, \dots, n$,
- 4) $\text{dist}\left(\mathbf{x}_{l_i}^{(i)}, \text{Vect}_{\mathbb{R}}(\mathbf{x}_1^{(i)}, \dots, \widehat{\mathbf{x}_{k_i}^{(i)}}, \dots, \mathbf{x}_{l_{i-1}}^{(i)})\right) \geq 1 - \frac{1}{2^{l_i-1}}$ si $k_i < l_i$,
- 5) $\mathbf{x}_{l_i}^{(i)} \in \mathbf{x}_{k_{i-1}}^{(i-1)} + \text{Vect}_{\mathbb{Z}}(\mathbf{x}_1^{(i-1)}, \dots, \widehat{\mathbf{x}_{k_{i-1}}^{(i-1)}}, \dots, \mathbf{x}_{l_i}^{(i-1)})$ si $i \geq 1$,
- 6) $(\mathbf{x}_1^{(i)}, \dots, \widehat{\mathbf{x}_{l_i}^{(i)}}, \dots, \mathbf{x}_n^{(i)}) = (\mathbf{x}_1^{(i-1)}, \dots, \widehat{\mathbf{x}_{k_{i-1}}^{(i-1)}}, \dots, \mathbf{x}_n^{(i-1)})$ si $i \geq 1$.

De plus, si \mathbf{u}_i est un vecteur unitaire orthogonal à $\text{Vect}_{\mathbb{R}}(\mathbf{x}_1^{(i)}, \dots, \widehat{\mathbf{x}_{k_i}^{(i)}}, \dots, \mathbf{x}_n^{(i)})$, alors pour tous i et j tels que $0 \leq i < j < s$ on a

$$\text{dist}(\mathbf{u}_i, \mathbf{u}_j) \leq \frac{2e^4}{\|\mathbf{x}_1^{(i+1)}\| \dots \|\mathbf{x}_n^{(i+1)}\|}.$$

Preuve Cf proposition 5.3 de [14]. Roy utilise plusieurs lemmes intermédiaires (lemmes 5.1 et 5.2 de [14]) pour la construction (par récurrence) des bases qui conviennent. La presque orthogonalité est un point particulièrement délicat à montrer. \square

La proposition suivante complète la proposition ci-dessus en construisant un vecteur unitaire \mathbf{u} et en estimant la jauge des vecteurs $x_1^{(i)}, \dots, x_n^{(i)}$ par rapport au corps convexe $\mathcal{C}_{\mathbf{u}}(Q)$ (pour Q bien choisi).

Proposition 7.4.15. *On conserve les notations de la propositions précédentes et on pose*

$$Q_i := A_1^{(i)} \dots A_n^{(i)} \quad (0 \leq i < s),$$

et $Q_s := \infty$ si $s \neq \infty$. Alors il existe un vecteur unitaire $\mathbf{u} \in \mathbb{R}^n$ tel que pour tout entier i ($0 \leq i < s$) et tout $Q \in [Q_i, Q_{i+1}[$ on ait

- 1) $A_j^{(i)} \leq \lambda(\mathbf{x}_j^{(i)}, \mathcal{C}_{\mathbf{u}}(Q)) \leq 8e^4 A_j^{(i)}$ pour tout $j \in \{1, \dots, \widehat{k_i}, \dots, n\}$,
- 2) $\frac{A_{k_i}^{(i)} Q}{2^n Q_i} \leq \lambda(\mathbf{x}_{k_i}^{(i)}, \mathcal{C}_{\mathbf{u}}(Q)) \leq \frac{8A_{k_i}^{(i)} Q}{2^n Q_i}$.

Preuve Cf proposition 5.4 de [14].

La construction de \mathbf{u} provient de la seconde assertion de la proposition 7.4.14. Quand

$s = \infty$, elle assure en effet que l'image dans $\mathbb{P}^{n-1}(\mathbb{R})$ de la suite $(\mathbf{u}_i)_{i \geq 1}$ converge vers la classe d'un vecteur unitaire $\mathbf{u} \in \mathbb{R}^n$ tel

$$\text{dist}(\mathbf{u}_i, \mathbf{u}) \leq \frac{2e^4}{\|\mathbf{x}_1^{(i+1)}\| \dots \|\mathbf{x}_n^{(i+1)}\|} \quad (0 \leq i < s).$$

Quand $s \neq \infty$ ces inégalités restent vraies en prenant $\mathbf{u} := \mathbf{u}_{s-1}$ et en remplaçant le membre de droite de l'inégalité par 0 lorsque $i = s - 1$.

Roy montre alors que \mathbf{u} convient. □

Avec les résultats précédents on peut alors établir une version quantitative de la seconde assertion du théorème 7.1.3 pour des n -systèmes rigides de maille assez grande.

Théorème 7.4.16. *Soit $\delta \geq 4 + (n + 3) \log 2$ et $\mathbf{P} : [q_0, \infty[\rightarrow \mathbb{R}^n$ un n -système rigide de maille δ . Alors il existe un vecteur unitaire $\mathbf{u} \in \mathbb{R}^n$ tel que*

$$\sup_{q \geq q_0} \|\mathbf{P}(q) - \mathbf{L}_{\mathbf{u}}(q)\|_{\infty} \leq n \log(8e^4 n).$$

Preuve Cf [14] (théorème 5.5) pour les détails.

On considère la suite de vecteurs $(\mathbf{a}^{(i)})_{0 \leq i < s}$ et les suites d'entiers $(k_i)_{0 \leq i < s}$ et $(l_i)_{0 \leq i < s}$ qui définissent le canevas attaché à \mathbf{P} comme dans les définitions 7.1.1 et 7.1.2. Pour chaque i ($0 \leq i < s$) on écrit $\mathbf{a}^{(i)} = (a_1^{(i)}, \dots, a_n^{(i)})$ et on pose

$$A_j^{(i)} := \exp(a_j^{(i)}) \quad (1 \leq j \leq n).$$

Alors les conditions (C1)–(C3) de la définition 7.1.1 assurent que les conditions (7.6)–(7.9) sont également vérifiées et donc qu'on peut appliquer les propositions 7.4.14 et 7.4.15. On considère les bases $(\mathbf{x}_1^{(i)}, \dots, \mathbf{x}_n^{(i)})$ de \mathbb{Z}^n données par la proposition 7.4.14 pour $0 \leq i < s$, et $\mathbf{u} \in \mathbb{R}^n$ le vecteur donné par la proposition 7.4.15. Roy montre alors que ce vecteur unitaire convient. □

7.4.4 Systèmes réduits et approximations par des n -systèmes rigides

Ce paragraphe reprend les sections 6 et 7 de [14]. Les résultats principaux sont les propositions 7.4.18 et 7.4.19 (la plus délicate à montrer étant la deuxième) : elles permettent de renforcer le théorème 7.3.8 de Schmidt et Summerer et de prouver la première assertion du théorème 7.1.3.

Définition 7.4.17. Soit $q_0 \geq 0$. Un (n, γ) -système *réduit* sur $[q_0, \infty[$ est un (n, γ) -système $\mathbf{P} = (P_1, \dots, P_n) : [q_0, \infty[\rightarrow \mathbb{R}^n$ tel que si $j = 1, \dots, n - 1$, $a \geq q_0$ et $b \geq a + n\gamma$ sont tels que $P_1 + \dots + P_j$ est constant sur $[a, b]$, alors chaque fonction P_1, \dots, P_j est constante sur $[a, b - n\gamma]$.

Exemples : tout $(n, 0)$ -système est déjà un $(n, 0)$ -système réduit. Si \mathbf{P} est un (n, γ) système réduit sur $[q_0, \infty[$, alors $q \rightarrow \delta^{-1}\mathbf{P}(\delta q)$ est un $(n, \gamma/\delta)$ -système réduit sur $[q_0/\delta, \infty[$.

Proposition 7.4.18. *Soit $\mathbf{P} = (P_1, \dots, P_n) : [q_0, \infty[\rightarrow \mathbb{R}^n$ un (n, γ) -système sur $[0, \infty[$. Alors il existe un $(n, 2n\gamma)$ -système réduit $\tilde{\mathbf{P}} = (P_1, \dots, P_n) : [q_0, \infty[\rightarrow \mathbb{R}^n$ tel que $\|\mathbf{P} - \tilde{\mathbf{P}}\|_{\infty} \leq n\gamma$.*

Preuve Cf proposition 6.2 de [14] pour les détails. Roy construit $\tilde{\mathbf{P}}$ à l'aide de plusieurs lemmes techniques (lemmes 6.3, 6.4 et 6.5 de [14]). \square

L'étape suivante après avoir montré que tout (n, γ) -système s'approche par un (n, γ) -système réduit est de montrer que tout (n, γ) -système réduit peut être approché en un certain sens (et à une différence bornée près) par un n -système rigide. C'est ce que traduit la proposition suivante.

Proposition 7.4.19. *Soient γ, δ des réels tels que $0 \leq \gamma < \delta/(2n)^2$ et $\mathbf{P} : [0, \infty[\rightarrow \mathbb{R}^n$ un (n, γ) -système réduit. On pose $q_0 := n(n+1)\delta/2$. Alors il existe un n -système rigide $\mathbf{R} : [q_0, \infty[\rightarrow \mathbb{R}^n$ de maille δ tel que*

$$\|\mathbf{P}(q) - \mathbf{R}(q)\|_\infty \leq 3n^2\delta \quad (q \geq q_0).$$

Preuve Cf [14] proposition 7.1 de [14].

Roy se ramène à prouver cette proposition pour le cas d'un n -système rigide intégral - i.e. de maille 1 - en effectuant un changement de variable. En effet, $\tilde{\mathbf{P}}(q) := \delta^{-1}\mathbf{P}(q\delta)$ définit un $(n, \gamma/\delta)$ -système réduit sur $[0, \infty[$, et si $\tilde{\mathbf{R}}$ est un n -système rigide de maille 1 sur $[\tilde{q}_0, \infty[$ avec $\tilde{q}_0 := n(n+1)/2$, alors $\mathbf{R}(q) := \delta\tilde{\mathbf{R}}(q/\delta)$ définit un n -système rigide de maille δ sur $[q_0, \infty[$. Les inégalités requises se transfèrent immédiatement.

La construction de \mathbf{R} est délicate ; Roy utilise huit lemmes pour parvenir à ses fins.

7.4.5 Preuve du théorème de Roy

Nous avons désormais tous les outils pour démontrer le théorème 7.1.3. Ce paragraphe reprend la partie 8 de [14]. Elle est constituée de deux théorèmes : le premier prouve la seconde assertion du théorème 7.1.3, le deuxième prouve la première assertion de ce même théorème. Rappelons que tout n -système rigide est un $(n, 0)$ -système.

Théorème 7.4.20. *Soit $q_0 \geq 0$ et soit $\mathbf{P} : [q_0, \infty[\rightarrow \mathbb{R}^n$ un $(n, 0)$ -système. Alors il existe un vecteur unitaire $\mathbf{u} \in \mathbb{R}^n$ tel que*

$$\|\mathbf{P}(q) - \mathbf{L}_{\mathbf{u}}(q)\|_\infty \leq 3n^2(n+9) \quad (q \geq q_0).$$

Preuve Nous suivons ici la démonstration proposée par Roy dans [14] (théorème 8.1). La première chose que fait Roy est de prolonger \mathbf{P} en un $(n, 0)$ -système défini sur $[0, \infty[$. Pour cela, il pose $t_0 := 0$ et $t_i := P_1(q_0) + \dots + P_i(q_0)$ pour $i = 1, \dots, n$. On a alors $t_n = q_0$ et il définit

$$\mathbf{P}(q) = \Phi_n(0, \dots, 0, P_1(q_0), \dots, P_{i-1}(q_0), q - t_{i-1}) \quad (t_{i-1} \leq q \leq t_i, 1 \leq i \leq n),$$

de telle sorte que pour $i = 1, \dots, n$, le graphe combiné de \mathbf{P} sur $[t_{i-1}, t_i]$ consiste en $n-1$ segments horizontaux (pas forcément distincts) avec ordonnées $0, \dots, 0, P_1(q_0), \dots, P_{i-1}(q_0)$, et un segment de pente 1 d'extrémités $(t_{i-1}, 0)$ et $(t_i, P_i(q_0))$. Ces formules prolongent bien \mathbf{P} en un $(n, 0)$ -système sur $[0, \infty[$.

On peut donc supposer $q_0 = 0$. On pose $\delta := n+7$ et $\tilde{q}_0 := n(n+1)\delta/2$. Puisque tous les $(n, 0)$ -systèmes sont réduits, la proposition 7.4.19 fournit l'existence d'un n -système rigide $\mathbf{R} : [\tilde{q}_0, \infty[\rightarrow \mathbb{R}^n$ de maille δ tel que

$$\|\mathbf{P}(q) - \mathbf{R}(q)\|_\infty \leq 3n^2\delta \quad (q \geq \tilde{q}_0).$$

Le théorème 7.4.16 utilisé avec \mathbf{R} donne l'existence de $\mathbf{u} \in \mathbb{R}^n$ tel que

$$\|\mathbf{R}(q) - \mathbf{L}_{\mathbf{u}}(q)\|_\infty \leq n \log(8e^4 n) \quad (q \geq \tilde{q}_0).$$

On a alors en combinant ces inégalités

$$\|\mathbf{P}(q) - \mathbf{L}_{\mathbf{u}}(q)\|_{\infty} \leq 3n^2\delta + n \log(8e^4n) \quad (q \geq \tilde{q}_0). \quad (7.10)$$

Il reste à étendre cette inégalité pour $q \in [0, \tilde{q}_0]$. Sur cet intervalle, les coordonnées de $\mathbf{P}(q)$ sont positives (au sens large) et majorées par $P_1(q) + \dots + P_n(q) = q \leq \tilde{q}_0$ (car \mathbf{P} est un $(n, 0)$ -système). Les coordonnées de $\mathbf{L}_{\mathbf{u}}(q)$ sont également positives (au sens large) et majorées par $L_{\mathbf{u},1}(q) + \dots + L_{\mathbf{u},n}(q) \leq \tilde{q}_0 + n \log(n)$ par l'inégalité (7.2). Cela permet d'étendre l'inégalité (7.10) à tout $[0, \infty[$. On conclut en notant que $3n^2\delta + n \log(8e^4n) \leq 3n^2(n+9)$. \square

\square

Pour démontrer la première assertion du théorème 7.1.3, on remarque que tout n -système rigide de maille $\delta > 0$ est aussi un n -système rigide de maille δ/N pour tout entier $N \geq 1$. Cette remarque et le théorème suivant fournissent une version quantitative de la première assertion du théorème de Roy.

Théorème 7.4.21. *Soit $\delta > 24n^4 2^n \log(n)$ et $\mathbf{u} \in \mathbb{R}^n$ un vecteur unitaire de \mathbb{R}^n . On pose $q_0 := n(n+1)\delta/2$. Alors il existe un n -système rigide $\mathbf{R} : [q_0, \infty[\rightarrow \mathbb{R}^n$ de maille δ tel que*

$$\|\mathbf{L}_{\mathbf{u}}(q) - \mathbf{R}(q)\|_{\infty} \leq 4n^2\delta \quad (q \geq q_0).$$

Preuve On suit la preuve de Roy [14] (théorème 8.2). On pose $\gamma := 6n2^n \log(n)$. Par le théorème 7.3.8 il existe un (n, γ) -système $\mathbf{P} : [0, \infty[\rightarrow \mathbb{R}^n$ tel que

$$\|\mathbf{L}_{\mathbf{u}}(q) - \mathbf{P}(q)\|_{\infty} \leq \gamma \quad (q \geq 0).$$

La proposition 7.4.18 assure l'existence d'un $(n, 2n\gamma)$ -système réduit $\tilde{\mathbf{P}} : [0, \infty[\rightarrow \mathbb{R}^n$ tel que

$$\|\mathbf{P}(q) - \tilde{\mathbf{P}}(q)\|_{\infty} \leq n\gamma \quad (q \geq 0).$$

Finalement puisque $\delta > 4n^3\gamma$, la proposition 7.4.19 fournit l'existence d'un n -système rigide $\mathbf{R} : [q_0, \infty[\rightarrow \mathbb{R}^n$ de maille δ satisfaisant l'inégalité

$$\|\tilde{\mathbf{P}}(q) - \mathbf{R}(q)\|_{\infty} \leq 3n^2\delta \quad (q \geq q_0).$$

Combinant ces trois inégalités on trouve, pour $q \geq q_0$:

$$\|\mathbf{L}_{\mathbf{u}}(q) - \mathbf{R}(q)\|_{\infty} \leq (n+1)\gamma + 3n^2\delta \leq 4n^2\delta.$$

\square

8 Une application du théorème de Roy

Cette partie reprend l'article de Roy [15]. Le résultat principal est le théorème 8.1.4; il donne une description complète de certains exposants d'approximation diophantienne classiques en utilisant les outils de la géométrie paramétrique des nombres et en particulier le théorème 7.1.3.

8.1 Définitions et formulation du problème

Nous rappelons ici un certain nombre de définitions déjà introduites (notamment dans les paragraphes 7.1 et 7.2) et introduisons les exposants classiques d'approximation qui généralisent ceux des définitions 7.2.1 et 7.2.2.

Si $n \geq 1$ est un entier, on note $\mathbf{x} \cdot \mathbf{y}$ le produit scalaire usuel entre deux vecteurs \mathbf{x} et \mathbf{y} de l'espace euclidien \mathbb{R}^n . On munit également son algèbre extérieure $\bigwedge \mathbb{R}^n = \bigoplus_{k=1}^n \bigwedge^k \mathbb{R}^n$ de la structure d'espace euclidien telle que pour toute base orthonormée $(\mathbf{e}_1, \dots, \mathbf{e}_n)$ de \mathbb{R}^n , l'ensemble des produits extérieurs $\mathbf{e}_{i_1} \wedge \dots \wedge \mathbf{e}_{i_k}$ avec $0 \leq k \leq n$ et $1 \leq i_1 < \dots < i_k \leq n$ forme une base orthonormée de $\bigwedge \mathbb{R}^n$. Pour $k = 0, \dots, n$ on note $\bigwedge^k \mathbb{Z}^n$ le réseau de $\bigwedge^k \mathbb{R}^n$ engendré par les produits extérieurs de la forme $\mathbf{x}_1 \wedge \dots \wedge \mathbf{x}_k$ avec $\mathbf{x}_1, \dots, \mathbf{x}_k \in \mathbb{Z}^n$. Pour $k = 0$ on a $\bigwedge^k \mathbb{R}^n = \mathbb{R}$ et on pose $\bigwedge^0 \mathbb{Z}^n = \mathbb{Z}$. Si S^k est un sous-espace de \mathbb{R}^n défini sur \mathbb{Q} de dimension k , rappelons qu'on peut définir sa *hauteur* par la formule

$$H(S) = \|\mathbf{x}_1 \wedge \dots \wedge \mathbf{x}_k\|,$$

où $\mathbf{x}_1, \dots, \mathbf{x}_k$ est une base de $S \cap \mathbb{Z}^n$.

La distance (projective) $\text{dist}(\mathbf{u}, S)$ d'un vecteur $\mathbf{u} \in \mathbb{R}^n$ à un sous-espace S définie dans le paragraphe 7.4.2 est égale à

$$\text{dist}(\mathbf{u}, S) = \frac{\|\text{proj}_{S^\perp}(\mathbf{u})\|}{\|\mathbf{u}\|} = \frac{\|\mathbf{u} \wedge \mathbf{x}_1 \wedge \dots \wedge \mathbf{x}_k\|}{\|\mathbf{u}\| \|\mathbf{x}_1 \wedge \dots \wedge \mathbf{x}_k\|},$$

où $(\mathbf{x}_1, \dots, \mathbf{x}_k)$ est n'importe quelle base de S . Géométriquement, cela représente le sinus de l'angle le plus petit entre la droite engendrée par \mathbf{u} et S . Notons que cette quantité est aussi égale à $\psi_1(A, S) = \omega_1(A, S)$ avec les notations du paragraphe 3.3 où on a posé $A := \mathbb{R}\mathbf{u}$.

Définition 8.1.1. Soit $n \geq 1$ et $\mathbf{u} \in \mathbb{R}^{n+1} \setminus \{0\}$. Pour $j = 0, \dots, n-1$ on note $\omega_j(\mathbf{u})$ (resp. $\widehat{\omega}_j(\mathbf{u})$) la borne supérieure de l'ensemble des réels ω tels que pour des Q arbitrairement grands (resp. pour tout Q assez grand) il existe un sous-espace $S \subset \mathbb{R}^{n+1}$ défini sur \mathbb{Q} de dimension $j+1$ vérifiant

$$H(S) \leq Q \quad \text{et} \quad H(S)\text{dist}(\mathbf{u}, S) \leq Q^{-\omega}.$$

Remarque : En particulier on a $\omega_j(\mathbf{u}) = \infty$ si \mathbf{u} appartient à un sous-espace S de \mathbb{R}^{n+1} défini sur \mathbb{Q} de dimension $j+1$. Sinon $\omega_j(\mathbf{u})$ est la borne supérieure de l'ensemble des réels ω pour lesquels il existe une infinité de sous-espaces $S \subset \mathbb{R}^{n+1}$ définis sur \mathbb{Q} de dimension $j+1$ tels que

$$\text{dist}(\mathbf{u}, S) \leq H(S)^{-\omega-1}.$$

On peut donner une définition alternative de ces exposants d'approximation qui permettra de faire le lien dans le paragraphe suivant avec la géométrie paramétrique des nombres.

Lemme 8.1.2 (Bugeaud-Laurent). Soit $j \in \{0, \dots, n-1\}$. Alors $\omega_j(\mathbf{u})$ (resp. $\widehat{\omega}_j(\mathbf{u})$) est la borne supérieure des réels ω tels que les inégalités

$$\|\mathbf{z}\| \leq Q \quad \text{et} \quad \|\mathbf{z} \wedge \mathbf{u}\| \leq Q^{-\omega} \tag{8.1}$$

possèdent une solution non nulle $\mathbf{z} \in \bigwedge^{j+1} \mathbb{Z}^{n+1}$ pour des valeurs de Q arbitrairement grandes (resp. pour tout Q assez grand).

Preuve Ce résultat est prouvé dans [4] §4 pour $\omega_j(\mathbf{u})$ dans le cas où les coordonnées de \mathbf{u} sont linéairement indépendantes sur \mathbb{Q} mais un petit argument permet de rendre valable la preuve pour \mathbf{u} non nul, quelconque ainsi que pour $\widehat{\omega}_j(\mathbf{u})$. \square

Le théorème 5.3.3 utilisé avec $d := 1$, $e := j + 1$, $A^d := \mathbb{R}\mathbf{u}$ assure que

$$\omega_j(\mathbf{u}) \geq \widehat{\omega}_j(\mathbf{u}) \geq \frac{j+1}{n-j} \quad (0 \leq j \leq n-1). \quad (8.2)$$

Le résultat qui suit donne d'autres relations sur ces exposants.

Théorème 8.1.3 (Schmidt, Laurent). *Soit $n \geq 1$ un entier. Pour tout vecteur $\mathbf{u} \in \mathbb{R}^{n+1}$ non nul on a $\omega_0(\mathbf{u}) \geq 1/n$ et*

$$\frac{j\omega_j(\mathbf{u})}{\omega_j(\mathbf{u}) + j + 1} \leq \omega_{j-1}(\mathbf{u}) \leq \frac{(n-j)\omega_j(\mathbf{u}) - 1}{n-j+1} \quad (1 \leq j \leq n-1), \quad (8.3)$$

en définissant par convention le quotient de gauche comme étant égal à j et celui de droite à ∞ si $\omega_j(\mathbf{u}) = \infty$.

Preuve Nous proposons ici une démonstration de ces inégalités à partir des théorèmes 5.8 du going-up (pour l'inégalité de droite) et 5.2.2 du going-down (pour l'inégalité de gauche).

Commençons par l'inégalité de droite.

Soit $1 \leq j \leq n-1$. Si $\omega_{j-1}(\mathbf{u}) = \infty$ alors \mathbf{u} appartient à un sous-espace de dimension j défini sur \mathbb{Q} , donc a fortiori à un espace de dimension $j+1$ défini sur \mathbb{Q} . Donc $\omega_j(\mathbf{u}) = \infty$ et il n'y a rien à montrer. On suppose désormais $\omega_{j-1}(\mathbf{u}) < \infty$.

Soit $\omega < \omega_{j-1}(\mathbf{u})$ et soit S^j un sous-espace de \mathbb{R}^{n+1} défini sur \mathbb{Q} tel que

$$H(S^j)\text{dist}(\mathbf{u}, S^j) \leq H(S^j)^{-\omega}. \quad (8.4)$$

Le théorème 5.8 (going-up) appliqué avec $H := H(S^j)$, $A^d := \mathbb{R}\mathbf{u}$, $B^e := S^j$, $c = 1/C_4$ (où C_4 est définie comme dans le théorème 5.8), $x_1 = 1$ et $y_1 = \omega$ donne l'existence d'un sous-espace S^{j+1} défini sur \mathbb{Q} contenant S^j et vérifiant :

$$H(S^{j+1})^{(n+1-j)/(n-j)}\text{dist}(\mathbf{u}, S^{j+1}) \leq H^{-\omega(n+1-j)/(n-j)} \quad \text{et} \quad H(S^{j+1}) \leq H',$$

où on a posé $H' := C_3 H^{(n-j)/(n+1-j)}$. On en déduit que

$$H(S^{j+1})\text{dist}(\mathbf{u}, S^{j+1}) \leq H(S^{j+1})^{-[\omega(n+1-j)+1]/(n-j)}. \quad (8.5)$$

Or, comme il y a une infinité de sous-espaces S^j vérifiant l'inégalité (8.4) et que $H(S^{j+1}) \geq H(S^j)$ (et qu'il n'y a qu'un nombre fini de sous-espaces de hauteur inférieure à un réel H) cela implique qu'il y a un nombre infini de S^{j+1} vérifiant (8.5), et par suite

$$[\omega(n+1-j)+1]/(n-j) \leq \omega_j(\mathbf{u}).$$

On conclut en faisant tendre ω vers $\omega_{j-1}(\mathbf{u})$ et avec la remarque suivant la définition des $\omega_j(\mathbf{u})$.

Montrons maintenant l'inégalité de gauche.

Soit $\omega < \omega_j(\mathbf{u})$. Soient $Q \geq 1$ et S^{j+1} tels que

$$H(S^{j+1}) \leq Q \quad \text{et} \quad H(S^{j+1})\text{dist}(\mathbf{u}, S^{j+1}) \leq Q^{-\omega}. \quad (8.6)$$

On applique le théorème 5.2.2 (going-down) avec $H := Q$, $A^d := \mathbb{R}\mathbf{u}$, $B^e := S^{j+1}$, $y_1 := \omega + 1$ et $c := 1/C_7$ (où C_7 est définie comme dans l'énoncé du théorème du going-down). On a alors $y'_1 = y_1(j+1)/(y_1+j)$ et toutes les hypothèses de l'énoncé sont satisfaites. En particulier, en posant $Q' := C_5 Q^{(j+y_1)/(j+1)}$ on a l'existence d'un sous-espace $S^j \subset S^{j+1}$ défini sur \mathbb{Q} et vérifiant

$$H(S^j) \leq Q' \quad \text{et} \quad H(S^j) \text{dist}(\mathbf{u}, S^j) \leq Q'^{-(y'_1-1)}.$$

Cette inégalité est vérifiée pour des Q' arbitrairement grands car il existe des Q arbitrairement grands tels qu'il existe S^{j+1} vérifiant (8.6). Donc

$$y'_1 - 1 \leq \omega_{j-1}(\mathbf{u}).$$

En faisant tendre ω vers $\omega_j(\mathbf{u})$ on a $y'_1 - 1$ qui tend vers $j\omega_j(\mathbf{u})/(\omega_j(\mathbf{u}) + j + 1)$, d'où le résultat. □

Notons que Roy fournit une autre preuve de ce théorème en utilisant les outils de la géométrie paramétrique des nombres.

Ces inégalités ont été observées par Laurent dans [9] qui introduisit alors les exposants $\omega_j(\mathbf{u})$. Dans le même article il fait la remarque que chaque inégalité (8.3) prise individuellement est la meilleure possible car en les combinant on trouve les inégalités de transfert de Khinchine qui sont connues pour être optimales. On peut aussi montrer que l'ensemble des valeurs prises par $\omega_j(\mathbf{u})$ est l'intervalle entier $[(j+1)/(n-j), \infty[$. Le théorème suivant est le résultat principal de l'article de Roy [15] et décrit complètement le comportement du n -uplet $(\omega_0(\mathbf{u}), \dots, \omega_{n-1}(\mathbf{u}))$.

Théorème 8.1.4 (Roy, 2014). *Soit $n \geq 1$ un entier. Soient $\omega_0, \dots, \omega_{n-1} \in [0, \infty[$ vérifiant les inégalités $\omega_0 \geq 1/n$ et*

$$\frac{j\omega_j}{\omega_j + j + 1} \leq \omega_{j-1} \leq \frac{(n-j)\omega_j - 1}{n-j+1} \quad (1 \leq j \leq n-1). \quad (8.7)$$

Alors il existe un vecteur $\mathbf{u} \in \mathbb{R}^{n+1}$ dont les coordonnées sont \mathbb{Q} -linéairement indépendantes et tel que

$$\omega_j(\mathbf{u}) = \omega_j \quad \text{et} \quad \widehat{\omega}_j(\mathbf{u}) = \frac{j+1}{n-j} \quad (0 \leq j \leq n-1).$$

Preuve Cf paragraphe 8.5. □

8.2 Liens avec la géométrie paramétrique des nombres

Soit $n \geq 1$ un entier. Nous rappelons les notations introduites dans le paragraphe 7.3.3. Notons que dans cette partie on a $n+1$ à la place de n .

Soit $\mathbf{u} \in \mathbb{R}^{n+1}$ non nul. Pour tout réel $Q \geq 1$ on considère le corps convexe

$$\mathcal{C}_{\mathbf{u}}(Q) = \{\mathbf{x} \in \mathbb{R}^{n+1} ; \|\mathbf{x}\| \leq 1, |\mathbf{x} \cdot \mathbf{u}| \leq Q^{-1}\}$$

et on note $\lambda_1(\mathcal{C}_{\mathbf{u}}(Q)) \leq \dots \leq \lambda_{n+1}(\mathcal{C}_{\mathbf{u}}(Q))$ ses $n+1$ minima successifs (pour le réseau \mathbb{Z}^{n+1}). On pose aussi

$$L_{\mathbf{u},j}(q) = \log \lambda_j(\mathcal{C}_{\mathbf{u}}(e^q)) \quad (q \geq 0, 1 \leq j \leq n+1),$$

qu'on réunit dans une unique application $\mathbf{L}_{\mathbf{u}} : [0, \infty[\rightarrow \mathbb{R}^{n+1}$ en posant

$$\mathbf{L}_{\mathbf{u}}(q) = (L_{\mathbf{u},1}(q), \dots, L_{\mathbf{u},n+1}(q)) \quad (q \geq 0).$$

Définition 8.2.1. Pour $j = 1, \dots, n+1$ on définit

$$\underline{\psi}_j(\mathbf{u}) = \liminf_{q \rightarrow \infty} \frac{L_{\mathbf{u},1}(q) + \dots + L_{\mathbf{u},j}(q)}{q} \quad \text{et} \quad \overline{\psi}_j(\mathbf{u}) = \limsup_{q \rightarrow \infty} \frac{L_{\mathbf{u},1}(q) + \dots + L_{\mathbf{u},j}(q)}{q}.$$

La proposition suivante permet de faire le lien entre ces quantités et les exposants classiques d'approximation du paragraphe précédent.

Proposition 8.2.2. Pour tout $j = 0, \dots, n-1$ on a

$$\omega_j(\mathbf{u}) = \frac{1}{\underline{\psi}_{n-j}(\mathbf{u})} - 1 \quad \text{et} \quad \widehat{\omega}_j(\mathbf{u}) = \frac{1}{\overline{\psi}_{n-j}(\mathbf{u})} - 1.$$

Preuve On reprend la preuve de [15] (proposition 3.1) en passant vite sur certains détails. Pour cette démonstration on utilise la notion de corps convexes pseudo-composés et le théorème 2.4.9. On prendra la définition des $\omega_j(\mathbf{u})$ donnée par le lemme 8.1.2.

Pour $j = 0, \dots, n-1$ et $Q \geq 1$ on définit

$$\mathcal{K}_{\mathbf{u}}^{(j+1)}(Q) = \{\mathbf{z} \in \wedge^{j+1} \mathbb{R}^{n+1}; \|\mathbf{z}\| \leq Q, \|\mathbf{z} \wedge \mathbf{u}\| \leq 1\}.$$

D'après le lemme 3 du §4 de [4], $\mathcal{K}_{\mathbf{u}}^{(j+1)}(Q)$ est comparable au $(j+1)$ -ème corps composé de $\mathcal{K}_{\mathbf{u}}^{(1)}(Q)$. Ce dernier est lui-même comparable au polaire de $\mathcal{C}_{\mathbf{u}}(Q)$ et a pour volume $\text{vol}(\mathcal{C}_{\mathbf{u}}(Q)) \asymp Q^{-1}$, par conséquent le premier minimum de $\mathcal{K}_{\mathbf{u}}^{(j+1)}(Q)$ pour le réseau $\wedge^{j+1} \mathbb{Z}^{n+1}$ vérifie

$$\begin{aligned} \lambda_1(\mathcal{K}_{\mathbf{u}}^{(j+1)}(Q)) &\asymp \lambda_1(\mathcal{K}_{\mathbf{u}}^{(1)}(Q)) \dots \lambda_{j+1}(\mathcal{K}_{\mathbf{u}}^{(1)}(Q)) \\ &\asymp \lambda_{n+1}(\mathcal{C}_{\mathbf{u}}(Q))^{-1} \dots \lambda_{n-j+1}(\mathcal{C}_{\mathbf{u}}(Q))^{-1} \\ &\asymp Q^{-1} \lambda_1(\mathcal{C}_{\mathbf{u}}(Q)) \dots \lambda_{n-j}(\mathcal{C}_{\mathbf{u}}(Q)) \quad (0 \leq j \leq n-1, Q \geq 1), \end{aligned}$$

où les constantes sous-jacentes ne dépendent que de n et \mathbf{u} mais pas de Q (cf [10]; la première équation est donnée par le théorème 2.4.9, la deuxième équation provient du théorème de Mahler, la troisième du second théorème de Minkowski). Maintenant, le corps convexe symétrique de $\wedge^{j+1} \mathbb{R}^{n+1}$ défini par (8.1) est $Q^{-\omega} \mathcal{K}_{\mathbf{u}}^{(j+1)}(Q^{\omega+1})$, donc son premier minimum pour le réseau $\wedge^{j+1} \mathbb{Z}^{n+1}$ vérifie

$$\lambda_1(Q^{-\omega} \mathcal{K}_{\mathbf{u}}^{(j+1)}(Q^{\omega+1})) \asymp Q^{-1} \lambda_1(\mathcal{C}_{\mathbf{u}}(Q^{\omega+1})) \dots \lambda_{n-j}(\mathcal{C}_{\mathbf{u}}(Q^{\omega+1})) \quad (8.8)$$

où les constantes sous-entendues ne dépendent ni de Q , ni de ω (avec $Q \geq 1$ et $\omega+1 \geq 0$). Par le lemme 8.1.2, $\omega_j(\mathbf{u})$ (resp. $\widehat{\omega}_j(\mathbf{u})$) est la borne supérieure de l'ensemble des réels ω pour lesquels ce premier minimum est inférieur ou égal à 1 pour des Q arbitrairement grands (resp. pour tout Q assez grand). Par ailleurs, puisque $\omega_j(\mathbf{u}) \geq \widehat{\omega}_j(\mathbf{u}) \geq 0$ par (8.2), on peut considérer uniquement les $\omega > -1$. Si on écrit $Q = e^{q/(\omega+1)}$ avec $q > 0$ et qu'on passe au logarithme en divisant par q , on obtient que $\omega_j(\mathbf{u})$ (resp. $\widehat{\omega}_j(\mathbf{u})$) est la borne supérieure de l'ensemble des réels ω pour lesquels

$$\frac{\log(\lambda_1(Q^{-\omega} \mathcal{K}_{\mathbf{u}}^{(j+1)}(Q^{\omega+1})))}{q} \leq 0,$$

pour des Q arbitrairement grands (resp. pour tout Q assez grand). Or, l'estimation (8.8) devient

$$\frac{\log(\lambda_1(Q^{-\omega} \mathcal{K}_{\mathbf{u}}^{(j+1)}(Q^{\omega+1})))}{q} = \frac{L_{\mathbf{u},1}(q) + \dots + L_{\mathbf{u},n_j}(q)}{q} - \frac{1}{\omega+1} + \mathcal{O}\left(\frac{1}{q}\right)$$

(où les constantes asymptotiques ne dépendent pas de q ni de ω), ce qui implique que $\omega_j(\mathbf{u})$ (resp. $\widehat{\omega}_j(\mathbf{u})$) est la borne supérieure de l'ensemble des réels ω pour lesquels

$$\frac{L_{\mathbf{u},1}(q) + \cdots + L_{\mathbf{u},n-j}(q)}{q} \leq \frac{1}{\omega + 1}$$

pour des Q arbitrairement grands (resp. pour tout Q assez grand). □

Grâce à cette proposition importante, nous allons pouvoir transposer le théorème 8.1.4 dans le langage de la géométrie paramétrique des nombres : les inégalités (8.7) se transformeront en les inégalités (8.11) ci-dessous, et l'existence de $\mathbf{u} \in \mathbb{R}^{n+1}$ avec les bonnes propriétés reviendra à l'existence d'un n -système généralisé (cf définition 8.3.6) avec les "bonnes" propriétés.

8.3 La notion de n -systèmes généralisés

Ce paragraphe reprend la partie 4 de [15]. On fixe $n \geq 2$ un entier. Rappelons que précédemment nous avons déjà défini la notion de (n, γ) -systèmes (définition 7.3.7) et de (n, γ) -systèmes réduits (définition 7.4.17) ainsi que celle de n -systèmes rigides (définition 7.1.2) qui apparaissaient comme des $(n, 0)$ -systèmes particuliers. Toutes ces classes de fonctions permettent d'approcher les fonctions $\mathbf{L}_{\mathbf{u}}$ avec $\mathbf{u} \in \mathbb{R}^n$ non nul. Nous allons maintenant définir en suivant [15] deux nouvelles classes un peu différentes : les n -systèmes (définition 8.3.1) et les n -systèmes généralisés (définition 8.3.6). Tout $(n, 0)$ -système est un n -système (en fait les deux définitions coïncident quand on regarde les n -systèmes sur $[q_0, \infty[$ avec $q_0 \geq 0$) et tout n -système est aussi un n -système généralisé. Rappelons que $\|\cdot\|$ désigne la norme euclidienne. On fera également usage de la norme $\|\cdot\|_{\infty}$ définie sur \mathbb{R}^n par $\|(x_1, \dots, x_n)\|_{\infty} = \max_{1 \leq j \leq n} |x_j|$.

Définition 8.3.1. Soit I un sous-intervalle de $[0, \infty[$ d'intérieur non vide. Un n -système sur I est une fonction continue affine par morceaux $\mathbf{P} = (P_1, \dots, P_n) : I \rightarrow \mathbb{R}^n$ vérifiant les conditions suivantes :

(S1) Pour tout $q \in I$ on a $0 \leq P_1(q) \leq \cdots \leq P_n(q)$ et $P_1(q) + \cdots + P_n(q) = q$.

(S2) Si H est un sous-intervalle ouvert non vide de I sur lequel \mathbf{P} est différentiable, alors il existe un entier r ($1 \leq r \leq n$) tel que P_r est de pente 1 sur H et P_j est constante sur H pour tout $j \neq r$.

(S3) Si q est un point intérieur à I en lequel \mathbf{P} n'est pas différentiable et si les entiers r et s qui vérifient $P_r'(q^-) = P_s'(q^+) = 1$ sont tels que $r < s$, alors on a $P_r(q) = P_{r+1}(q) = \cdots = P_s(q)$.

Ici, la condition " $\mathbf{P} : I \rightarrow \mathbb{R}^n$ affine par morceaux" signifie que l'ensemble D des points de I en lesquels \mathbf{P} n'est pas différentiable (en incluant les extrémités de I qui sont dans I) est un sous-ensemble discret de I et que la différentielle de \mathbf{P} est localement constante sur $I \setminus D$. Une telle fonction admet toujours une dérivée à droite $\mathbf{P}'(q^+)$ en tout point $q \in I$, $q \neq \inf I$, et une dérivée à gauche $\mathbf{P}'(q^-)$ en tout point $q \in I$, $q \neq \sup I$. La pente d'une composante P_r de \mathbf{P} sur un sous-intervalle ouvert H de $I \setminus D$ est la valeur (constante) de sa dérivée sur H ou, de manière équivalente, la pente de son graphe sur H .

La figure suivante issue de [15] montre le graphe combiné des fonctions P_r, \dots, P_s sur un voisinage d'un point q sous les hypothèses de la condition (S3). Dans ce cas,

les fonctions P_{r+1}, \dots, P_s coïncident à gauche de q tandis que P_r, \dots, P_{s-1} coïncident à droite de q .

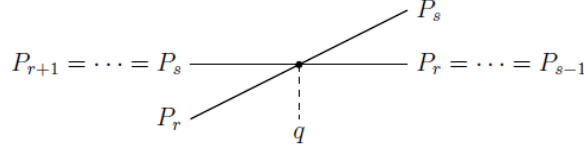


FIGURE 3 – Illustration de la condition (S3)

Proposition 8.3.2. *Supposons $I = [q_0, \infty[$ avec $q_0 \geq 0$. Alors les n -systèmes sur I sont exactement les $(n, 0)$ -systèmes sur I . En particulier, tout n -système rigide est un n -système.*

Preuve La démonstration est laissée au lecteur et ne présente pas de difficulté particulière. Il suffit de revenir à la définition 7.3.7 d'un $(n, 0)$ -système. □

Théorème 8.3.3. *Soit $\mathbf{u} \in \mathbb{R}^n$ non nul. Alors il existe $q_0 \geq 0$ et un n -système \mathbf{P} sur $[q_0, \infty[$ tel que $\mathbf{L}_{\mathbf{u}} - \mathbf{P}$ soit bornée sur $[q_0, \infty[$. Réciproquement, pour tout n -système \mathbf{P} sur un intervalle $[q_0, \infty[$ avec $q_0 \geq 0$, il existe un vecteur $\mathbf{u} \in \mathbb{R}^n$ non nul tel que $\mathbf{L}_{\mathbf{u}} - \mathbf{P}$ soit bornée sur $[q_0, \infty[$.*

Preuve La première assertion provient du théorème 7.1.3 (puisque tout n -système rigide est également un n -système). La seconde assertion vient du théorème 7.4.20 (puisque tout $(n, 0)$ -système est également un n -système). Notons que considérer un vecteur \mathbf{u} non nécessairement unitaire ne change rien. □

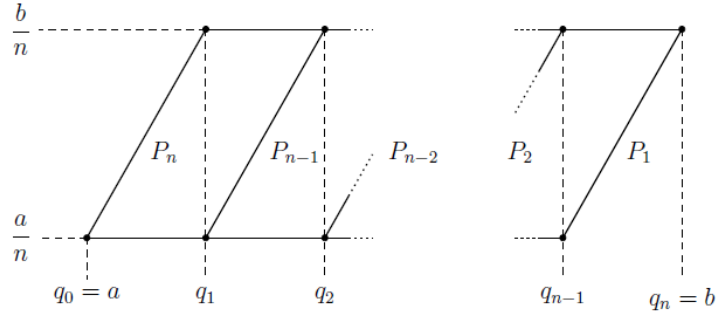
Le but de cette partie est d'étendre ce théorème d'approximation à une classe de fonctions plus grande : les n -systèmes généralisés. L'étude des sous-familles de $(\omega_0, \dots, \omega_{n-1}, \widehat{\omega}_0, \dots, \widehat{\omega}_{n-1})$ sur $\mathbb{R}^{n+1} \setminus \{0\}$ se ramènera à des problèmes sur des $(n+1)$ -systèmes généralisés.

Les deux lemmes ci-dessous donnent des exemples de n -systèmes. Le deuxième permet notamment le recollement de n -systèmes.

Lemme 8.3.4. *Soit $a, b \in \mathbb{R}$ des réels tels que $0 \leq a < b$. Alors il existe un n -système $\mathbf{P} : [a, b] \rightarrow \mathbb{R}^n$ tel que $\mathbf{P}(a) = (a/n, \dots, a/n)$ et $\mathbf{P}(b) = (b/n, \dots, b/n)$.*

Preuve Nous suivons la construction de [15]. On pose $q_i := (n-i)a/n + ib/n$ pour $i = 0, \dots, n$ et pour $j = 1, \dots, n$ on définit P_j comme l'unique fonction continue affine par morceaux sur $[a, b]$ qui est constante égale à a/n sur $[q_0, q_{n-j}]$, de pente 1 sur $[q_{n-j}, q_{n-j+1}]$, et constante égale à b/n sur $[q_{n-j+1}, q_n]$. Alors $\mathbf{P} = (P_1, \dots, P_n) : [a, b] \rightarrow \mathbb{R}^n$ est un n -système qui convient.

La figure suivante tirée de [15] illustre le graphe combiné de ce n -système.

FIGURE 4 – Graphe combiné d'un n -système pour le lemme 8.3.4

□

Lemme 8.3.5. Soit $a, b, c \in \mathbb{R}$ vérifiant $0 \leq a < b < c$. On suppose que $\mathbf{P}^{(1)} : [a, b] \rightarrow \mathbb{R}^n$ et $\mathbf{P}^{(2)} : [b, c] \rightarrow \mathbb{R}^n$ sont deux n -systèmes vérifiant $\mathbf{P}^{(1)}(b) = \mathbf{P}^{(2)}(b) = (b/n, \dots, b/n)$. Alors il existe un n -système \mathbf{P} sur $[a, c]$ qui prolonge $\mathbf{P}^{(1)}$ et $\mathbf{P}^{(2)}$.

Preuve Il existe une unique fonction $\mathbf{P} : [a, c] \rightarrow \mathbb{R}^n$ qui prolonge $\mathbf{P}^{(1)}$ et $\mathbf{P}^{(2)}$. Elle est continue mais pas nécessairement différentiable en b . Cependant, elle vérifie la condition (S3) en $q = b$ puisque toutes ses coordonnées sont égales en ce point. C'est donc un n -système.

□

Soit I un sous-intervalle de $[0, \infty[$ et D un sous-ensemble discret de I . En combinant les deux lemmes précédents on peut construire un n -système \mathbf{P} sur I tel que $\mathbf{P}(q) = (q/n, \dots, q/n)$ aux points $q \in D$. En choisissant D judicieusement, cela assure que $\sup_{q \in I} \|\mathbf{P}(q) - \tilde{\mathbf{P}}(q)\|_\infty$ est fini et arbitrairement petit, où on a posé $\tilde{\mathbf{P}}(q) = (q/n, \dots, q/n)$ pour tout $q \in I$. Un argument supplémentaire permettra de garder les propriétés d'approximation par des n -systèmes pour des fonctions plus générales ($\tilde{\mathbf{P}}$ en fera partie) : les n -systèmes généralisés. La proposition 8.3.7 et le corollaire 8.3.8 formalisent cette affirmation.

Définition 8.3.6. Soit I un sous-intervalle de $[0, \infty[$ d'intérieur non vide. Un n -système généralisé sur I est une fonction continue affine par morceaux $\mathbf{P} = (P_1, \dots, P_n) : I \rightarrow \mathbb{R}^n$ qui vérifie les propriétés suivantes :

(G₁) Pour tout $q \in I$ on a $0 \leq P_1(q) \leq \dots \leq P_n(q)$ et $P_1(q) + \dots + P_n(q) = q$.

(G₂) Si H est un sous-intervalle d'intérieur non vide de I sur lequel \mathbf{P} est différentiable, alors il existe des entiers \underline{r} et \bar{r} vérifiant $1 \leq \underline{r} \leq \bar{r} \leq n$ et tels que $P_{\underline{r}}, P_{\underline{r}+1}, \dots, P_{\bar{r}}$ coïncident sur tout H et sont de pente $1/(\bar{r} - \underline{r} + 1)$ tandis que les autres composantes P_j de \mathbf{P} sont constantes sur H .

(G₃) Si q est un point intérieur à I en lequel \mathbf{P} n'est pas différentiable, si $\underline{r}, \bar{r}, \underline{s}, \bar{s}$ sont les entiers pour lesquels

$$P'_j(q^-) = \frac{1}{\bar{r} - \underline{r} + 1} \quad (\underline{r} \leq j \leq \bar{r}) \quad \text{et} \quad P'_j(q^+) = \frac{1}{\bar{s} - \underline{s} + 1} \quad (\underline{s} \leq j \leq \bar{s}), \quad (8.9)$$

et si $\underline{r} < \bar{s}$, alors on a $P_{\underline{r}}(q) = P_{\underline{r}+1} = \dots = P_{\bar{s}}(q)$.

La figure qui suit (issue de [15]) est le graphe combiné des fonctions $P_{\underline{r}}, \dots, P_{\underline{s}}$ au voisinage d'un point q sous les hypothèses de la condition (G_3) quand $\bar{r} < \underline{s}$.

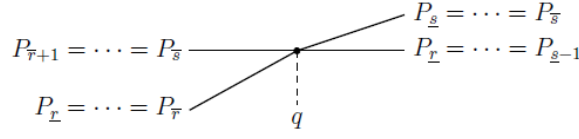


FIGURE 5 – Illustration de la condition (G_3) quand $\bar{r} < \underline{s}$

Tout n -système est aussi un n -système généralisé (avec $\underline{r} = \bar{r}$ dans la condition (G_2)).

Proposition 8.3.7. *Soient $\tilde{\mathbf{P}}$ un n -système généralisé sur un sous-intervalle d'intérieur non vide I de $[0, \infty[$ et D un sous-ensemble discret de I . Alors il existe un n -système \mathbf{P} sur I tel que $\mathbf{P}(t) = \tilde{\mathbf{P}}(t)$ pour tout $t \in D$.*

Preuve Nous ne faisons ici que la construction de Roy (cf [15], Proposition 4.6) sans vérifier qu'elle a les propriétés énoncées.

Soit D_0 l'ensemble des points de I où \mathbf{P} n'est pas différentiable auquel on a adjoint les frontières de I contenues dans I . Comme D_0 est un sous-ensemble discret de I , on peut supposer sans perte de généralité que D contient D_0 . On peut aussi supposer que I et D ont mêmes bornes supérieure et inférieure, quitte à ajouter des points à D .

Soit $]t_1, t_2[$ un intervalle maximal de $I \setminus D$ (nécessairement ouvert et borné d'après les hypothèses ci-dessus). Alors t_1 et t_2 appartiennent à D et comme $D \subset I$, $[t_1, t_2]$ est un sous-intervalle de I . Comme $\tilde{\mathbf{P}}$ est différentiable sur $]t_1, t_2[$, il existe des entiers \underline{r} et \bar{r} avec $1 \leq \underline{r} \leq \bar{r} \leq n$ tels que $\tilde{P}_{\underline{r}}, \dots, \tilde{P}_{\bar{r}}$ coïncident et sont de pente $1/(\bar{r} - \underline{r} + 1)$ sur $]t_1, t_2[$ tandis que les autres composantes \tilde{P}_j sont constantes égales à c_j sur ce même intervalle. On pose

$$m := \bar{r} - \underline{r} + 1 \quad \text{et} \quad c := (c_1 + \dots + c_{\underline{r}-1}) + (c_{\bar{r}+1} + \dots + c_n).$$

Par le lemme 8.3.4 il existe un m -système (A_1, \dots, A_m) sur $[t_1, t_2]$ vérifiant

$$(A_1(t_i), \dots, A_m(t_i)) = \left(\frac{t_i}{m}, \dots, \frac{t_i}{m} \right) \quad \text{pour } i = 1, 2.$$

Puisque $\tilde{P}_j(q) = (q-c)/m$ pour $j = \underline{r}, \dots, \bar{r}$ et $q \in [t_1, t_2]$, l'application $\mathbf{P} : [t_1, t_2] \rightarrow \mathbb{R}^n$ donnée par

$$\mathbf{P}(q) = \left(c_1, \dots, c_{\underline{r}-1}, A_1(q) - \frac{c}{m}, \dots, A_m(q) - \frac{c}{m}, c_{\bar{r}+1}, \dots, c_n \right) \quad (t_1 \leq q \leq t_2),$$

coïncide avec $\tilde{\mathbf{P}}$ aux points $q = t_1, t_2$. Comme les intervalles $[t_1, t_2]$ recouvrent I et n'ont pas de points intérieurs en commun, cela implique que $\mathbf{P} : I \rightarrow \mathbb{R}^n$ est une fonction affine par morceaux continue qui coïncide avec $\tilde{\mathbf{P}}$ sur D .

Roy montre alors que \mathbf{P} est un n -système (cf [15] pour la fin de la preuve). □

Corollaire 8.3.8. *Soit $\tilde{\mathbf{P}} : I \rightarrow \mathbb{R}^n$ un n -système généralisé et $\varepsilon > 0$. Alors il existe un n -système \mathbf{P} sur I tel que $\|\tilde{\mathbf{P}}(q) - \mathbf{P}(q)\|_\infty \leq \varepsilon$ pour tout $q \in I$.*

Preuve On reprend la preuve de Roy dans [15]. Soit D l'ensemble de tous les multiples $k\varepsilon/2$ ($k \in \mathbb{N}$) qui sont dans I . Par la proposition 8.3.7, il existe un n -système \mathbf{P} sur I tel que $\mathbf{P}(t) = \tilde{\mathbf{P}}(t)$ pour tout $t \in D$. Puisque les composantes de \mathbf{P} et $\tilde{\mathbf{P}}$ sont de pentes 0 et 1, on a pour tous $q \in I$ et $t \in D$

$$\|\tilde{\mathbf{P}}(q) - \mathbf{P}(q)\|_\infty \leq \|\tilde{\mathbf{P}}(q) - \tilde{\mathbf{P}}(t)\|_\infty + \|\mathbf{P}(q) - \mathbf{P}(t)\|_\infty \leq 2|q - t|.$$

On conclut en choisissant t tel que $|q - t| \leq \varepsilon/2$. □

Ainsi, le théorème 8.3.3 reste valable pour les n -systèmes généralisés.

8.4 Une famille de n -systèmes généralisés

Soit $n \geq 2$ un entier. On considère l'ensemble

$$\Delta^{(n)} = \{(a_1, \dots, a_n) \in \mathbb{R}^n ; 0 < a_1 < \dots < a_n \text{ et } a_1 + \dots + a_n = 1\}$$

et son adhérence

$$\overline{\Delta}^{(n)} = \{(a_1, \dots, a_n) \in \mathbb{R}^n ; 0 \leq a_1 \leq \dots \leq a_n \text{ et } a_1 + \dots + a_n = 1\}.$$

Si \mathbf{P} est un n -système généralisé et $q \in I \cap \mathbb{R}_+^*$, alors par la condition (G_1) de la définition 8.3.6 on a $q^{-1}\mathbf{P}(q) \in \overline{\Delta}^{(n)}$. Pour $j = 1, \dots, n$ on définit une fonction $\psi_j : \overline{\Delta}^{(n)} \rightarrow \mathbb{R}$ par

$$\psi_j(a_1, \dots, a_n) = a_1 + \dots + a_j.$$

Pour prouver le théorème 8.1.4 Roy utilise la construction qui suit.

Proposition 8.4.1. *Soit $\mathbf{a} = (a_1, \dots, a_n) \in \Delta^{(n)}$. On pose*

$$\begin{aligned} q_i &= a_1 + \dots + a_i + (n-i)a_i & (1 \leq i \leq n), \\ q_{n-1+i} &= (i-1)a_i + a_i + \dots + a_n & (1 \leq i \leq n). \end{aligned}$$

Alors il existe un unique n -système généralisé $\mathbf{P} = (P_1, \dots, P_n)$ sur $[q_1, q_{2n-1}] = [na_1, na_n]$ dont le graphe combiné est le suivant :

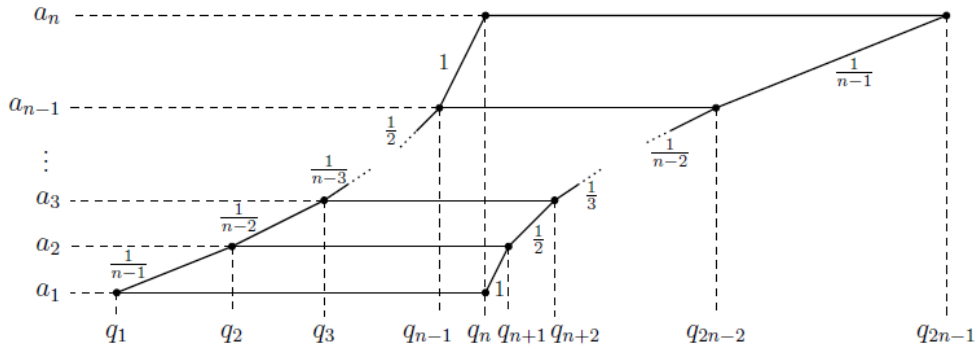


FIGURE 6 – Graphe combiné d'un n -système généralisé.

(Sur cette figure, chaque nombre $1/m$ désigne la pente d'un segment et m est le nombre de composantes de \mathbf{P} dont le graphe coïncide avec ce segment sur l'intervalle

$[q_i, q_{i+1}]$ correspondant.)

De plus, pour $j = 1, \dots, n-1$ on a

$$\inf \psi_j(q^{-1}\mathbf{P}(q)) = \psi_j(\mathbf{a}) \quad \text{et} \quad \sup \psi_j(q^{-1}\mathbf{P}(q)) = \frac{j}{n}$$

où les bornes inférieure et supérieure sont toutes deux prises sur l'ensemble des $q \in [na_1, na_n]$.

Preuve Cf [15] (proposition 5.1). □

Le prochain résultat permet de changer l'intervalle de définition d'un n -système généralisé tout en conservant certaines informations.

Lemme 8.4.2. *Si $\mathbf{P} : [a, b] \rightarrow \mathbb{R}^n$ est un n -système généralisé avec $b > a > 0$ alors, pour tout $d > c > 0$ vérifiant $d/c = b/a$, l'application $\tilde{\mathbf{P}} : [c, d] \rightarrow \mathbb{R}^n$ donnée par*

$$\tilde{\mathbf{P}}(q) = \frac{c}{a}\mathbf{P}\left(\frac{a}{c}q\right) \quad (q \in [c, d])$$

est aussi un n -système généralisé. De plus on a pour tout $j = 1, \dots, n$

$$\inf_{q \in [a, b]} \psi_j(q^{-1}\mathbf{P}(q)) = \inf_{q \in [c, d]} \psi_j(q^{-1}\tilde{\mathbf{P}}(q)) \quad \text{et} \quad \sup_{q \in [a, b]} \psi_j(q^{-1}\mathbf{P}(q)) = \sup_{q \in [c, d]} \psi_j(q^{-1}\tilde{\mathbf{P}}(q)).$$

En fait, le graphe combiné de $\tilde{\mathbf{P}}$ est l'image du graphe combiné de \mathbf{P} par un changement d'échelle de rapport c/a . Remarquons enfin que le lemme 8.3.5 reste vrai pour les n -systèmes généralisés, ce qui permet de montrer la proposition suivante.

Proposition 8.4.3. *Soit E un sous-ensemble non vide de $\overline{\Delta}^{(n)}$. Alors il existe un n -système généralisé \mathbf{P} sur $[1, \infty[$ qui vérifie pour $j = 1, \dots, n$*

$$\liminf_{q \rightarrow \infty} \psi_j(q^{-1}\mathbf{P}(q)) = \inf \psi_j(E) \quad \text{et} \quad \limsup_{q \rightarrow \infty} \psi_j(q^{-1}\mathbf{P}(q)) = \frac{j}{n}.$$

Preuve On reprend la construction de [15] (proposition 5.4).

Si $E = \{(1/n, \dots, 1/n)\}$, on prend simplement $\mathbf{P}(q) = (q/n, \dots, q/n)$ (avec $q \geq 1$). Sinon, on choisit une suite $(\mathbf{a}^{(i)})_{i \geq 1}$ de points de $\Delta^{(n)}$ dont l'ensemble des points d'accumulation est l'adhérence \overline{E} de E . Pour tout $i \geq 1$, on considère le n -système généralisé attaché au point $\mathbf{a}^{(i)}$ par la proposition 8.4.1 et, en partant de $q_1 = 1$, on utilise récursivement le lemme 8.4.2 pour le transformer en un n -système généralisé $\mathbf{P}^{(i)}$ défini sur un intervalle de la forme $[q_i, q_{i+1}]$. Par construction on a

$$\inf_{q \in [q_i, q_{i+1}]} \psi_j(q^{-1}\mathbf{P}^{(i)}(q)) = \psi_j(\mathbf{a}^{(i)}) \quad \text{et} \quad \sup_{q \in [q_i, q_{i+1}]} \psi_j(q^{-1}\mathbf{P}^{(i)}(q)) = \frac{j}{n}. \quad (8.10)$$

Notons aussi que

$$\mathbf{P}^{(i-1)}(q_i) = \mathbf{P}^{(i)}(q_i) = (q_i/n, \dots, q_i/n) \quad (i \geq 2)$$

et que $\limsup q_{i+1}/q_i > 1$ car E contient au moins un point (a_1, \dots, a_n) tel que $a_n > a_1$. Cela implique $\lim q_i = \infty$ et donc il existe un unique n -système généralisé \mathbf{P} sur $[1, \infty[$ dont la restriction à $[q_i, q_{i+1}]$ est $\mathbf{P}^{(i)}$ pour tout $i \geq 1$. Par (8.10), on obtient finalement

$$\liminf_{q \rightarrow \infty} \psi_j(q^{-1}\mathbf{P}(q)) = \liminf_{q \rightarrow \infty} \psi_j(\mathbf{a}^{(i)}) = \inf \psi_j(E) \quad \text{et} \quad \limsup_{q \rightarrow \infty} \psi_j(q^{-1}\mathbf{P}(q)) = \frac{j}{n}$$

pour $j = 1, \dots, n$. □

8.5 Preuve du théorème principal

Le résultat qui suit est dernier résultat intermédiaire dont nous avons besoin pour démontrer le théorème 8.1.4.

Proposition 8.5.1. *Soit $n \geq 1$ un entier. Supposons que $\underline{\psi}_1, \dots, \underline{\psi}_n \in [0, 1]$ vérifient les inégalités $\underline{\psi}_n \leq n/(n+1)$ et*

$$\frac{\underline{\psi}_j}{j} \leq \frac{\underline{\psi}_{j+1}}{j+1} \quad \text{et} \quad \frac{1 - \underline{\psi}_j}{n+1-j} \leq \frac{1 - \underline{\psi}_{j+1}}{n-j} \quad (1 \leq j \leq n-1). \quad (8.11)$$

Alors il existe un sous-ensemble fini non vide E de $\overline{\Delta}^{(n+1)}$ tel que

$$\underline{\psi}_j = \min \psi_j(E) \quad (1 \leq j \leq n).$$

Preuve On présente la construction de la proposition 6.1 de [15].

Si $n = 1$ on a $\underline{\psi}_1 \leq 1/2$ et on prend simplement $E = \{(\underline{\psi}_1, 1 - \underline{\psi}_1)\}$. Supposons maintenant que $n \geq 2$ et soit $k \in \{1, \dots, n-1\}$.

Affirmation : il existe un $\mathbf{a} \in \overline{\Delta}^{(n+1)}$ tel que $\psi_j(\mathbf{a}) \geq \underline{\psi}_j$ pour $j = 1, \dots, n$ avec égalité pour $j = k$ et $j = k+1$.

Pour montrer cela on pose $c = \underline{\psi}_k/k$ et $d = (1 - \underline{\psi}_{k+1})/(n-k)$, et on considère le point

$$\mathbf{a} = \underbrace{(c, \dots, c)}_{k \text{ fois}}, \underline{\psi}_{k+1} - \underline{\psi}_k, \underbrace{(d, \dots, d)}_{(n-k) \text{ fois}} \in \mathbb{R}^{n+1}.$$

Par hypothèse $c \geq 0$, et les inégalités (8.11) pour $j = k$ donnent $c \leq \frac{\underline{\psi}_{k+1} - \underline{\psi}_k}{n-k} \leq d$.

Comme la somme des coordonnées de \mathbf{a} vaut 1, cela assure que $\mathbf{a} \in \overline{\Delta}^{(n+1)}$. Notons aussi que les inégalités (8.11) montrent que les quotients $\underline{\psi}_j/j$ et $(1 - \underline{\psi}_j)/(n+1-j)$ sont des fonctions croissantes de $j \in \{1, \dots, n\}$. On en déduit que pour $j = 1, \dots, k$ on a $\underline{\psi}_j/j \leq c$ et finalement $\psi_j(\mathbf{a}) = jc \geq \underline{\psi}_j$ avec égalité si $j = k$. De même, pour $j = k+1, \dots, n$ on a $d \leq (1 - \underline{\psi}_j)/(n+1-j)$ et on obtient $\psi_j(\mathbf{a}) = 1 - (n+1-j)d \geq \underline{\psi}_j$, avec égalité si $j = k+1$. Cela prouve l'affirmation.

On conclut en faisant varier k , ce qui donne un ensemble fini de points avec les propriétés requises. □

Remarque : En général on ne peut pas espérer que E consiste en un unique élément. Roy fournit un contre-exemple assez simple dans son article. Il remarque en effet que tout point $\mathbf{a} = (a_1, \dots, a_{n+1}) \in \overline{\Delta}^{(n+1)}$ vérifie

$$\psi_{j-1}(\mathbf{a}) + \psi_{j+1}(\mathbf{a}) - 2\psi_j(\mathbf{a}) = a_{j+1} - a_j \geq 0 \quad (2 \leq j \leq n),$$

et par conséquent il n'y a par exemple aucun $\mathbf{a} \in \overline{\Delta}^4$ vérifiant $\psi_1(\mathbf{a}) = 0$, $\psi_2(\mathbf{a}) = 1/3$, $\psi_3(\mathbf{a}) = 1/2$ bien que $(\underline{\psi}_1, \underline{\psi}_2, \underline{\psi}_3) = (0, 1/3, 1/2)$ vérifie les inégalités (8.11) pour $n = 3$.

Preuve [du théorème 8.1.4]. Nous reprenons les arguments de [15]. Soit $\omega_0, \dots, \omega_{n-1} \in [0, \infty]$ vérifiant les conditions du théorème 8.1.4. Comme le suggère la proposition 8.2.2 on définit

$$\underline{\psi}_j := \frac{1}{\omega_{n-j} + 1} \in [0, 1] \quad (1 \leq j \leq n).$$

Ces nombres vérifient toutes les hypothèses de la proposition 8.5.1 ci-dessus. Soit E le sous-ensemble de $\overline{\Delta}^{(n+1)}$ donné par cette proposition et \mathbf{P} le $(n+1)$ -système généralisé sur $[1, \infty[$ construit dans la proposition 8.4.3 à partir de ce choix de E . Alors on a

$$\liminf_{q \rightarrow \infty} \psi_j(q^{-1}\mathbf{P}(q)) = \inf \psi_j(E) = \underline{\psi}_j \quad \text{et} \quad \limsup_{q \rightarrow \infty} \psi_j(q^{-1}\mathbf{P}(q)) = \frac{j}{n+1}.$$

Le corollaire 8.3.8 et le théorème d'approximation 8.3.3 fournissent l'existence d'un vecteur $\mathbf{u} \in \mathbb{R}^{n+1}$ non nul tel que la différence $\mathbf{P} - \mathbf{L}_{\mathbf{u}}$ soit bornée sur $[1, \infty[$. Cela implique, par définition des quantités $\underline{\psi}_j(\mathbf{u})$ et $\overline{\psi}_j(\mathbf{u})$,

$$\underline{\psi}_j(\mathbf{u}) = \underline{\psi}_j \quad \text{et} \quad \overline{\psi}_j(\mathbf{u}) = \frac{j}{n+1} \quad (1 \leq j \leq n).$$

Maintenant, en utilisant la proposition 8.2.2, cela implique que pour $j = 0, \dots, n-1$,

$$\omega_j(\mathbf{u}) = \frac{1}{\underline{\psi}_{n-j}(\mathbf{u})} - 1 = \omega_j \quad \text{et} \quad \widehat{\omega}_j(\mathbf{u}) = \frac{1}{\overline{\psi}_{n-j}(\mathbf{u})} - 1 = \frac{j+1}{n-j}.$$

Enfin, puisque $\widehat{\omega}_{n-1}(\mathbf{u}) = n < \infty$, les coordonnées de \mathbf{u} sont nécessairement linéairement indépendantes sur \mathbb{Q} (cf la remarque après la définition 8.1.1). □

Références

- [1] A. AITKEN : *Determinants and matrices*. London, 1951.
- [2] Y. BUGEAUD et M. LAURENT : On exponents of homogeneous and inhomogeneous Diophantine approximation. *Moscow Math. J*, 5(4):747–766, 2005.
- [3] Y. BUGEAUD et M. LAURENT : Exponents of Diophantine approximation. *Diophantine Geometry Proceedings, Scuola Normale Superiore Pisa, Ser. CRM*, 4:101–121, 2007.
- [4] Y. BUGEAUD et M. LAURENT : On transfer inequalities in Diophantine approximation, II. *Mathematische Zeitschrift*, 265(2):249–262, 2010.
- [5] W. HODGE et D. PEDOE : *Methods of Algebraic Geometry*. Camb. Univ. Press, 1947.
- [6] V. JARNÍK : Zum Khintchineschen "Übertragungssatz". 1938.
- [7] A. KHINTCHINE : Über eine Klasse linearer diophantischer Approximationen. *Rendiconti del Circolo Matematico di Palermo (1884-1940)*, 50(2):170–195, 1926.
- [8] A. KHINTCHINE : Zur metrischen Theorie der diophantischen Approximationen. *Mathematische Zeitschrift*, 24(1):706–714, 1926.
- [9] M. LAURENT : On transfer inequalities in Diophantine Approximation. *In Analytic Number Theory, Essays in honour of Klaus Roth*, p. 306–314. Cambridge U. Press, 2009.
- [10] C. LEKKERKERKER : *Geometry of Numbers*, vol. VIII de *Bibliotheca mathematica (Amsterdam. 1952)*. Wolter-Noordhoff, 1969.
- [11] K. MAHLER : On compound convex bodies I. *Proc. London Math. Soc. (3)*, 5:358–379, (1955).
- [12] N. MOSHCHEVITIN : Exponents for three-dimensional simultaneous Diophantine approximations. *Czechoslovak mathematical journal*, 62(1):127–137, 2012.
- [13] D. ROY : On Two Exponents of Approximation Related to a Real Number and Its Square. *Canad. J. Math*, 59(1):211–224, 2007.
- [14] D. ROY : On Schmidt and Summerer parametric geometry of numbers. arXiv :1406.3669 [math.NT] (Annals of Math., à paraître), 14 Jun 2014.

-
- [15] D. ROY : Spectrum of the exponents of best rational approximation. arXiv :1410.1007 [math.NT] (soumis), 4 Oct 2014.
 - [16] S. SCHANUEL : On heights in number fields. *Bulletin of the American Mathematical Society*, 70(2):262–263, 1964.
 - [17] W. M. SCHMIDT : On heights of algebraic subspaces and Diophantine approximations. *Ann. of Math.*, 85:430–472, 1967.
 - [18] W. M. SCHMIDT et L. SUMMERER : Parametric geometry of numbers and applications. *Acta Arith.*, 140:67–91, (2009).
 - [19] W. M. SCHMIDT et L. SUMMERER : Diophantine approximation and parametric geometry of numbers. *Monatsh. Math.*, 169:51–104, (2013).