Université Pierre et Marie Curie

Mémoire en vue de l'obtention du :
M2 « Algèbre et Géométrie »
(Anc. DEA « Méthodes Algébriques »)

Sur quelques lemmes de Thue-Siegel

Par Manuel Pégourié-Gonnard
Sous la direction de Patrice Philippon

Soutenu le 23 juin 2005
Jury de soutenance composé de :
Sinnou David et Patrice Philippon

Introduction

Une technique courante en approximation diophantienne et en théorie des nombres transcendants nécessite de pouvoir disposer de polynômes non nuls, vérifiant par exemple certaines conditions d'annulation, et ayant de « petits » coefficients. Pour les obtenir, on regarde leurs coefficients comme les inconnues d'un système linéaire homogène exprimant les conditions requises. L'idée, utilisée d'abord par Thue [13] en 1909, puis clairement formalisée par Siegel [12] en 1929, est alors la suivante : un système à « petits » coefficients (entiers) doit posséder de « petites » solutions (entières) non nulles. Nous appellerons donc lemme de Thue-Siegel (LTS) un énoncé majorant la taille minimale d'une solution non nulle en fonction de celle des coefficients des équations, ou de manière équivalente, la hauteur d'un vecteur non nul d'un sous-espace (l'espace des solutions) en fonction de la hauteur de cet espace.

La preuve originale du résultat de Thue et Siegel repose sur le principe de Dirichlet, et permet de contrôler la taille d'une solution. On peut obtenir le même résultat en appliquant le premier théorème de Minkowski sur les corps convexes. Le réseau est l'ensemble des points entiers du sous-espace, et le volume minimal du corps convexe convenablement choisi mesure la taille du petit vecteur. En fait, il est facile de démontrer un LTS « classique », i.e. ne contrôlant qu'une solution, sans faire appel à ces techniques de géométrie des nombres. Leur avantage est de pouvoir s'étendre : ainsi, il suffit d'invoquer le second théorème de Minkowski dans le raisonnement précédent pour obtenir un LTS « fort », i.e. contrôlant la taille de toute une base de l'espace des solutions.

D'autre part, certaines applications nécessitent de disposer d'analogues du LTS sur un corps de nombres. Une première idée est de voir un corps de nombres comme un espace vectoriel de dimension finie sur $\mathbf{Q}$, et d'appliquer alors le LTS sur $\mathbf{Q}$. Ceci impose des conditions contraignantes sur le nombre d'équations, et fournit de plus des majorations qui ne sont pas toujours assez fines. Des améliorations sont obtenues dans des cas particuliers, mais les techniques employées ne sont pas généralisables. En 1983, Bombieri et Vaaler ([2] Thm 9) démontrent une version générale de LTS sur les corps de nombres, grâce à une version adélique du second théorème de Minkowski. Le contexte adélique permet notamment de faire apparaître l'ensemble des solutions sur K comme un sous-groupe discret de $K_{\mathbf{A}}^n$, tout en interprétant en termes de hauteur le volume du « corps convexe adélique ».

Le calcul du volume fait apparaître le discriminant du corps K, avec un certain exposant, qui subsiste dans la constante finale. Ceci est déplaisant car, si l'on considère notre problème dans une extension finie E de K la constante semble croître de manière incontrôlée. On peut alors se demander si cette dépendance en $\mathrm{Disc}(K)$ de la constante est bien nécessaire. En 1995, Roy et Thunder [10] montrent que c'est en effet le cas en minorant la hauteur des vecteurs non nuls de certains espaces par une constante faisant apparaître

une certaine puissance de $\mathrm{Disc}(K)$. Ils notent toutefois sur un exemple que le résultat ne s'applique qu'aux vecteurs à coordonnées dans K, laissant donc ouverte la possibilité de trouver des vecteurs plus petits définis sur $\overline{K}$.

Parallèlement, on a aussi cherché à établir des LTS sur d'autres types de corps. Ainsi, en 1941, Mahler prouve un équivalent du LTS sur un corps de fonctions rationnelles. Par une méthode différente et plus rapide, Thunder démontre à nouveau ce résultat en 1995, en l'étendant du même coup aux corps de fonctions algébriques ([14] Thm 1). Il utilise pour cela une extension du théorème de Riemann-Roch, qui joue dans sa démonstration un rôle analogue au premier théorème de Minkowski, avant d'étendre sa majoration à toute une base par une méthode astucieuse. Analogue du $\mathrm{Disc}(K)$ des corps de nombres, le genre g du corps de base apparaît dans la constante finale. De même, il est également indispensable pour les solutions définies sur K, mais ne l'est potentiellement plus si l'on supprime cette condition.

En 1996, Roy et Thunder démontrent en effet une version « absolue » du LTS (i.e. solutions définies sur $\overline{K}$) dans laquelle n'apparaît plus de facteur $C(K)$ dépendant du corps de définition du sous-espace. Leur résultat ([9] Thm 2.2) est valable pour les corps de nombres ou les corps de fonctions. Pour cela, il utilisent les résultats « relatifs » (i.e. solutions su K) précédents, et en dimension 2, font progressivement disparaître $C(K)$ quand croit le degré des solutions, en utilisant adroitement la décomposition en facteurs linéaires des polynômes homogènes dans $\overline{K}[X, Y]$. Ils étendent ensuite ce résultat à la dimension quelconque. Un des outils essentiels pour ceci est la notion de hauteur tordue, une extension de la notion usuelle de hauteur, déjà utilisée par Thunder pour les corps de fonctions.

Nous nous proposons ici, après avoir dans une première partie défini la notion de hauteur tordue et établi sa propriété fondamentale, de démontrer quelques lemmes de Thue-Siegel à l'aide de cet outil. Nous nous intéresserons donc en premier lieu aux résultats relatifs. Nous démontrerons un analogue du résultat de Bombieri et Vaaler sur les corps de nombres, ainsi que le résultat de Thunder sur les corps de fonctions, avant d'énoncer des résultats d'optimalité de ces lemmes. Dans une dernière partie, nous démontrerons enfin le lemme « absolu » de Roy et Thunder, avant d'évoquer la question de son optimalité.

Les théorèmes sont numérotés à part des lemmes, propositions et corollaires, qui pour leur part sont numérotés linéairement.

1 Préliminaires

1.1 Notations et définitions

Afin de présenter de manière unifiée les résultats sur les corps de nombres ou de fonctions, la lettre k désignera $\mathbf{Q}$ ou $k_0(T)$, où k_0 est un corps quelconque. Nous prendrons pour K une extension finie de degré $d = [K : k]$, toutes les extensions considérées sont supposées incluses dans une clôture algébrique $\bar{k}$ fixée. Nous nous intéresserons aux bases d'un sous-espace V de $\bar{k}^n$, de dimension m, défini sur K.

Nous noterons $M(K)$ l'ensemble des places de K, supposées triviales sur k_0 si $k = k_0(T)$. Pour chaque place $v \in M(K)$, on note K_v le complété de K en v. S'il est non-archimédien, on note $\mathcal{O}_v$ son anneau des entiers, M_v l'idéal maximal de celui-ci, et π_v un générateur de ce dernier. Par ailleurs, k peut être naturellement muni (nous y reviendrons) d'un ensemble de valeurs absolues standard $\{\, |\cdot|_w \,, w \in M(k)\,\}$. Si $v \in M(K)$ divise w, on supposera $|\cdot|_v$ normalisée de façon à prolonger $|\cdot|_w$. Par abus, on notera encore v la place de $M(k)$ que divise $v \in M(K)$. Ainsi, on définit par $d_v = [K_v : k_v]$ le degré local en v. Au vu des normalisations choisies, la formule du produit s'écrit :

$$\forall x \in K, \prod_{v \in M(K)} |x|_v^{d_v} = 1\,.$$

Pour chaque corps local $(L, |\cdot|)$, on définit des normes sur L^n par :
$||x||_\infty = \max_i(|x^i|)$ où $x = (x^1, \ldots, x^n)$,

$$||x|| = \begin{cases} \max_i(|x^i|) & \text{si } |\cdot| \text{ est ultramétrique,} \\ \big(\sum_i |x^i|^2\big)^{1/2} & \text{sinon.} \end{cases}$$

Sauf précision contraire, on utilisera $||\cdot||$ et, si L est archimédien, on le supposera muni de la forme bilinéaire symétrique usuelle dont découle $||\cdot||$.

On définit alors la hauteur multiplicative de Weil sur K^n par :

$$H(x) = \Big[\prod_{v \in M(K)} ||x||_v^{d_v} \Big]^{1/d}.$$

Cette hauteur est projective par la formule du produit, et, grâce à la normalisation $(\cdot)^{1/d}$, absolue dans le sens où, si $x \in \bar{k}^n$, elle ne dépend pas du choix de $K \supset k(x)$ utilisé pour son calcul. On utilise de même la norme $||\cdot||_\infty$ pour définir une hauteur notée H^∞. Elle est reliée à H par :

$$H^\infty(x) \leqslant H(x) \leqslant \sqrt{n} H^\infty(x) \quad \forall x \in \bar{k}^n.$$

Dans le cadre des corps de fonctions, nous utiliserons également une hauteur additive h, qui est un log de H ($= H^\infty$ ici), et sera normalisée comme suit. On prend en chaque place la valuation normalisée ord_v d'image

$\mathbf{Z}$, puis pour chaque $x \in K^n$ on définit le diviseur $\mathrm{div}(x) = \sum_v \mathrm{ord}_v(x) \cdot v$, où $\mathrm{ord}_v(x) = \min_i(\mathrm{ord}_v(x^i))$. Notant alors $\mathfrak{m}(K,k) = [K:k]/[K_0:k_0]$ le degré effectif, où K_0, corps des constantes de K, est la clôture algébrique de k_0 dans K, on définit par $h(x) = -\deg(\mathrm{div}(x))/\mathfrak{m}(K,k)$ la hauteur logarithmique absolue de x. Cette lecture de la définition nous permettra d'utiliser plus facilement des arguments de type Riemann-Roch.

Remarquons que dans le cas des corps de nombres, on a une normalisation naturelle pour les valeurs absolues : aux places archimédiennes, $|n| = n$ pour $n \in \mathbf{N}$; aux places ultramétriques $|p|_p = p^{-1}$. Pour les corps de fonctions, il n'existe pas de telle normalisation canonique pour les valeurs absolues, par contre les valuations sont naturellement normalisées par $\mathrm{ord}_v(K_v^*) = \mathbf{Z}$. Pour passer aux valeurs absolues, on doit choisir une base d'exponentiation. On supposera ce choix fait une fois pour toutes ; les logarithmes et exponentielles seront donc supposés relatifs à cette base et on notera $h = \log(H)$, $|x|_v = \exp(\mathrm{ord}_v(x))$, etc. sans plus de précautions.

Si $\bar{k}^m \xrightarrow{\sim} V \subset \bar{k}^n$ est une paramétrisation de V, elle fournit sur $\bar{k}^m$ une hauteur induite par celle de $\bar{k}^n$. On va généraliser la notion de hauteur en s'inspirant de ceci. Plongeons K diagonalement dans son anneau des adèles $K_{\mathbf{A}}$. Remarquons que la définition de la hauteur s'étend immédiatement au cas où $x \in K_{\mathbf{A}}^n$. Pour chaque $A = (A_v) \in GL(K_{\mathbf{A}}^n)$, on introduit une hauteur tordue $H_A = H \circ A$ sur $\bar{k}^n$. Explicitement, si $x \in K^n$, on a :

$$H_A(x) = \Big[\prod_{v \in M(K)} ||A_v(x)||_v^{d_v} \Big]^{1/d}.$$

On peut, comme on l'a remarqué, définir ainsi la hauteur d'une droite vectorielle. Pour les sous-espaces de dimension supérieure, on utilise les coordonnées grassmanniennes. Plus précisément, pour tout anneau R, la base canonique de R^n induit une base sur $\bigwedge^r R$ que l'on peut par exemple ordonner lexicographiquement. Ceci fournit des isomorphismes $\bigwedge^r K_v^n \approx K_v^N$ ($N = \binom{n}{r}$), qui fournissent à leur tour une norme sur chaque complété de $\bigwedge^r K$. C'est précisément ce qu'il nous fallait pour définir une hauteur H sur $\bigwedge^r \bar{k}$, qu'on peut ensuite tordre en $H_{\bigwedge^r A}$ par $\bigwedge^r A \in GL(\bigwedge^r K_{\mathbf{A}})$. On pose alors $H_A(V) = H_{\bigwedge^m A}(\bigwedge^m V)$ si $m = \dim V > 0$; en particulier $H_A(\bar{k}^n) = H(\det A)$. Conventionnellement, $H_A(\{0\}) = 1$. Notant $A^* = ({}^tA)^{-1}$, on a ainsi l'importante formule de dualité suivante ([15] §2) :

$$H_{A^*}(V^\perp) = H(\det A)^{-1} H_A(V) \tag{1.1.1}$$

De même que sur $\bigwedge^r K_v$, on définit des normes sur chaque $\mathrm{S}^r K_v$, puis la hauteur tordue $H_{\mathrm{S}^r A}$ sur $\mathrm{S}^r \bar{k}^n$. On définit la norme de $P \in K_v[X_1, \ldots, X_n]$ via le vecteur de coordonnées les coefficient de P. Sur chaque partie homogène de $K_v[X_1, \ldots, X_n]$, l'isomorphisme $K_v[X_1, \ldots, X_n]_r \approx \mathrm{S}^r K_v$ est ainsi une isométrie. On définit enfin une action a de $GL(K_{\mathbf{A}}^n)$ sur $K_{\mathbf{A}}[X_1, \ldots, X_n]$ par

$a(A)\cdot P = P\circ {}^tA$. On récupère ainsi une hauteur H_A sur $\bar{k}[X_1,\ldots,X_n]$ tordue par $a(A)\in GL(K_{\mathbf{A}}[X_1,\ldots,X_n])$. Si l'isomorphisme $K[X_1,\ldots,X_n]_r \approx \mathrm{S}^rK$ associe P à α, on a ainsi :

$$H_A(P) = H_{\mathrm{S}^{r(A)}}(\alpha) \tag{1.1.2}$$

1.2 Propriétés des hauteurs tordues

On a introduit les hauteurs tordues comme un moyen de définir sur $\bar{k}^m \xrightarrow{\sim} V\subset \bar{k}^n$ une hauteur qui reflète la position de V dans $\bar{k}^n$. Le théorème suivant confirme et précise cette idée.

Théorème 1 *Soit $\phi:\bar{k}^m\to\bar{k}^n$ linéaire injective, définie sur K, d'image V. Pour chaque $A\in GL(K_{\mathbf{A}}^n)$, il existe un $B\in GL(K_{\mathbf{A}}^m)$ tel que :*
(i) $H_A(\phi(x)) = H_B(x)\quad \forall x\in\bar{k}^m$,
(ii) $H_A(V) = H_B(\bar{k}^m)$.
Le même B satisfait également :
(iii) $H_A(\phi(W)) = H_B(W)\quad \forall W\subset\bar{k}^m$,
(iv) $H_{\bigwedge^r A}((\bigwedge^r\phi)S) = H_{\bigwedge^r B}(S)\quad \forall S\subset\bigwedge^r\bar{k}^m$,
(v) $H_{\mathrm{S}^rA}((\mathrm{S}^r\phi)T) = H_{\mathrm{S}^rB}(T)\quad \forall T\subset \mathrm{S}^r\bar{k}^m$,
(vi) $H_{{}^tA}(P\circ {}^t\phi) = H_B(P)\quad \forall P\in\bar{k}[X_1,\ldots,X_m]_r$.

Avant de le démontrer, déduisons-en un principe pratique qui illustre son utilisation :

Corollaire 1 *Supposons (i) que pour tout $n\geqslant 1$, il existe une constante $C(K,n)$ telle que pour tout $A\in GL(K_{\mathbf{A}}^n)$, on puisse trouver une base (y_i) de $\bar{k}^n$ satisfaisant $H_A(y_1)\cdots H_A(y_n)\leqslant C(K,n)\,H_A(\bar{k}^n)$.*
Alors (ii) tout sous-espace $V\subset\bar{k}^n$, de dimension m, défini sur K, admet une base (x_j) satisfaisant $H(x_1)\cdots H(x_m)\leqslant C(K,m)\,H(V)$.
Si de plus en (i) les y_i peuvent être pris de degré $\leqslant r$ sur K, alors en (ii) il en est de même des x_j.

Démo : On peut choisir $\phi:\bar{k}^m\hookrightarrow\bar{k}^n$ définie sur K paramétrant V. On applique alors le théorème 1 à $A=\mathbf{Id}_{K_{\mathbf{A}}^n}$ pour obtenir $B\in GL(K_{\mathbf{A}}^m)$ telle que $H_B(\bar{k}^m)=H(V)$ et $H_B(y)=H(x)$ si $x=\phi(y)$. L'hypothèse (i) appliquée à B fournit alors des y_j dont les images par ϕ sont les x_i recherchés. L'assertion sur les degrés découle du fait que ϕ est définie sur K.

Pour démontrer le théorème 1, on fait d'abord un peu d'algèbre linéaire sur chaque facteur local :

Lemme 2 *Soit L un corps local, $\phi:L^m\to L^n$ linéaire injective, alors il existe $\lambda\in GL(L^m)$ telle que $\Phi=\phi\circ\lambda$ soit une isométrie.*

Démo : Supposons L non archimédien. Alors Φ est une isométrie si et seulement si $\Phi^{-1}(\mathcal{O}_L^n) = \mathcal{O}_L^m$. Or $\phi^{-1}(\mathcal{O}_L^n)$ est de type fini sans torsion sur $\mathcal{O}_L$ principal, donc libre, et de rang m car ϕ est injective. On en choisit donc une base, et on définit λ comme envoyant la base canonique de L^m sur cette dernière. Alors $\phi \circ \lambda$ est bien une isométrie.
Si maintenant L est archimédien, L^m est $\mathbf{R}^m$ ou $\mathbf{C}^m$ muni de sa structure euclidienne (resp. hermitienne) usuelle. On choisit une base orthonormale de $\phi(L^m)$ qu'on relève en une base de L^m par injectivité de ϕ. On choisit λ envoyant la base canonique sur celle-ci.

Lemme 3 *Si $\phi : L^m \to L^n$ est une isométrie, les applications suivantes sont également des isométries :*
(i) $\phi_F : F^m \to F^n$ pour toute extension finie F/L
(ii) $\bigwedge^r \phi$ et $\mathrm{S}^r\phi$.

Démo : Dans le cas non-archimédien, on commence par étendre en une base de $\mathcal{O}_L^n$ l'image par ϕ de la base canonique de L^m, on note ψ envoyant la base canonique de L^n dessus, de sorte que $\phi = \psi \circ \iota$, où $\iota : L^m \to L^n$ est la première injection naturelle $x \mapsto (x, 0)$. Clairement, ι_F, $\bigwedge^r \iota$ et $\mathrm{S}^r\iota$ sont des isométries. De plus, ψ et ψ^{-1} envoient les points entiers sur des points entiers, il en est donc de même de ψ_F, $\bigwedge^r \psi$, $\mathrm{S}^r\psi$ ainsi que de leurs inverses, ce qui montre bien que ce sont des isométries.
Pour le cas archimédien, on commence par décomposer de même, en utilisant cette fois une base orthonormale. Ici, ι_F, $\bigwedge^r \iota$ et $\mathrm{S}^r\iota$ sont toujours des isométries. Pour ψ_F, on voit que $\mathbf{R} \subset L \subset F \subset \mathbf{C}$, et le seul cas non trivial s'obtient en remarquant que $||x + iy||^2 = ||x||^2 + ||y||^2$ pour x et y dans $\mathbf{R}^n$. On peut ensuite supposer $L = \mathbf{C}$, et que ψ conserve le produit hermitien. On vérifie ensuite (tester sur la base canonique et étendre par multilinéarité) que les produits hermitiens induits par $\bigwedge^r \mathbf{C}^n \approx \mathbf{C}^N$ et $\mathrm{S}^r\mathbf{C}^n \approx \mathbf{C}^M$ sont donnés par :

$$(x_1 \wedge \cdots \wedge x_r | y_1 \wedge \cdots \wedge y_r) = \det[(x_i|y_j)]$$

$$(x_1 \cdots x_r | y_1 \cdots y_r) = \mathrm{perm}[(x_i|y_j)]$$

donc conservés par $\bigwedge^r \psi$ et $\mathrm{S}^r\psi$, qui sont ainsi des isométries.

Il faut ensuite recoller ces informations locales pour obtenir le théorème 1. On remarque d'abord que $A = (A_v) \in \prod GL(K_v^n)$ est dans $GL(K_{\mathbf{A}}^n)$ si et seulement si les A_v sont des isométries sauf pour un nombre fini de v.

Démo (du théorème) : Comme précédemment, on décompose $\phi = \psi \circ \iota$. On note $\phi_{\mathbf{A}}$, $\psi_{\mathbf{A}}$, $\iota_{\mathbf{A}}$ leur images par le changement de base $K_{\mathbf{A}} \leftarrow K$. Clairement, les $\iota_v (= (\iota_{\mathbf{A}})_v)$ sont isométriques, et $\psi_{\mathbf{A}} \in GL(K_{\mathbf{A}}^n)$. Ainsi, $\phi_{\mathbf{A}}$ est une isométrie sur presque chaque facteur local.
On construit alors B comme suit : si $A_v \circ \phi_v$ est une isométrie, on

prend $B_v = \mathbf{Id}_{K_v}$, sinon le lemme 2 permet de choisir B_v de sorte que $A_v \circ \phi_v \circ B_v^{-1}$ soit une isométrie. Par les remarques précédentes, B est bien dans $GL(K_{\mathbf{A}}^n)$. Comme c'est une isométrie sur chaque facteur local, on a, pour tout $y \in K^n$, $H(A \circ \phi \circ B^{-1}(y)) = H(y)$, donc $H_A(\phi(x)) = H_B(x)$ pour tout $x(= B^{-1}(y))$ de K^n, et même de $\bar{k}^n$ par le lemme 3 (i).

Maintenant, le lemme 3 (ii) assure que $\bigwedge^s(A \circ \phi \circ B^{-1}) = \bigwedge^s A \circ \bigwedge^s \phi \circ \bigwedge^s B^{-1}$ est une isométrie sur chaque facteur local. On en déduit $H_{\bigwedge^s A}((\bigwedge^s \phi)(z)) = H_{\bigwedge^s B}(z)$ pour $z \in \bigwedge^s \bar{k}^n$. Il suffit alors de prendre pour s la dimension de W et pour z un vecteur directeur de la droite $\bigwedge^s W$. Il vient : $H_A(\phi(W)) = H_{\bigwedge^s A}(\bigwedge^s(\phi(W))) = H_{\bigwedge^s A}((\bigwedge^s \phi)(\bigwedge^s W)) = H_{\bigwedge^s B}(\bigwedge^s W) = H_B(W)$.

On a ainsi montré (iii), dont (ii) est un cas particulier. On montre (iv) (resp. (v)) par des considérations similaires sur $\bigwedge^s(\bigwedge^r(A \circ \phi \circ B^{-1}))$ (resp. $\bigwedge^s(\mathrm{S}^r(A \circ \phi \circ B^{-1}))$), où $s = \dim S$ (resp. $\dim T$). Enfin, (vi) est une simple réécriture de (v) utilisant (1.1.2).

On rappelle par ailleurs l'inégalité suivante, connue pour les hauteurs classiques (p. ex. [11] lemma 8A, p.28), et qui se démontre de même pour les hauteurs tordues :

Proposition 4 *Pour toute base $x_1, \ldots, x_m$ de $V \subset \bar{k}^n$, on a :*

$$H_A(x_1) \cdots H_A(x_m) \geqslant H_A(V)\,.$$

On va en déduire une minoration sur la hauteur tordue de tout vecteur, non triviale *a priori*. En effet, la hauteur classique satisfait $H(x) \geqslant 1$ pour tout x, ce qu'on vérifie facilement par la formule du produit. Or, les coordonnées de $A(x)$ étant adéliques ne vérifient plus cette formule. On a bien sur $H_A(x) > 0$ pour x non nul, mais il peut être intéressant de savoir que $H_A(x)$ n'est pas arbitrairement proche de 0. C'est l'objet du corollaire suivant :

Lemme 5 *Pour tout $A \in GL(K_{\mathbf{A}}^n)$, il existe $\nu(A) > 0$ tel que $H_A(x) > 0$ pour tout $x \in \bar{k}^n$*

Démo : On considère la base canonique $e_1, \ldots, e_n$ de $\bar{k}^n$. On pose $M(A) = \max_i(H_A(e_i))$. On vérifie que $\nu(A) = H(\det A)/M^{n-1}$ convient. En effet, considérons $x \in \bar{k}^n$. On peut supposer, quitte à renuméroter, que $e_1, \ldots, e_{n-1}, x$ est une base de $\bar{k}^n$. Appliquant la proposition 4, on a $H_A(x)M^{n-1} \geqslant H_A(x)\prod_{i=1}^{n-1} H_A(e_i) \geqslant H_A(\bar{k}^n) = H(\det A)$ et la conclusion souhaitée.

2 Lemmes relatifs

2.1 Corps de nombres

L'objet du problème est donc ici de trouver une « petite » base de V, à coordonnées dans K extension finie de degré d de $\mathbf{Q}$. L'idée principale, due à Bombieri et Vaaler, est d'utiliser une version adélique du second théorème de Minkowski sur les corps convexes. Usuellement, un corps convexe est un voisinage compact de l'origine, convexe et symétrique par rapport à l'origine. Dans le contexte adélique, nous appellerons corps convexe adélique tout ensemble

$$\mathcal{S} = \prod_{v|\infty} \mathrm{S}_v \times \prod_{v\nmid\infty} \mathrm{T}_v \subset K_{\mathbf{A}}$$

tel que

- S_v est un corps convexe usuel,
- $\mathrm{T}_v = \mathcal{O}_v^n$ sauf pour un nombre fini de places où T_v reste toutefois un sous-$\mathcal{O}_v$-module de type fini de K_v^n.

Munissons $K_{\mathbf{A}}$ d'une multiplication externe par les réels $(\lambda, a) \mapsto \lambda \cdot a$ (non compatible avec l'addition) qui agit de la manière évidente aux places archimédiennes et est l'identité partout ailleurs. Ainsi on constate qu'un corps convexe adélique est bien un voisinage compact de l'origine, qui vérifie également : $x \in \mathcal{S} \implies -x \in \mathcal{S}$, et $\forall x, y \in \mathcal{S},\ \forall t \in [0;1],\ t\cdot x + (1-t)\cdot y \in \mathcal{S}$. On définit alors les minimums successifs par :

$$\lambda_i = \inf\{\lambda > 0 \mid \exists\, x_1, \ldots, x_i \in \lambda \cdot \mathcal{S} \cap K,\ x_1 \wedge \cdots \wedge x_i \neq 0\}.$$

La prochaine étape est maintenant de pouvoir mesurer le volume d'un corps convexe adélique. On choisit sur $\mathbf{R}$ la mesure de Lebesgue, sur $\mathbf{C} \approx \mathbf{R}^2$ le double de la mesure de Lebesgue, et sur chaque K_v (v finie) la mesure de Haar normalisée par $\mathrm{vol}(\mathcal{O}_v) = |\mathcal{D}_v|_v^{1/2}$, où $\mathcal{D}_v$ est la différente locale en v. On en déduit sur $K_{\mathbf{A}}$ une mesure de Haar telle que $\mathrm{vol}(K_{\mathbf{A}}/K) = 1$ ([16] pp. 90-91). On prend sur $K_{\mathbf{A}}^n$ la mesure produit. On peut alors énoncer l'analogue de la partie difficile, seule utile ici, du second théorème de Minkowski :

Théorème 2 *Les minimums λ_i d'un corps convexe adélique $\mathcal{S}$ vérifient :*

$$(\lambda_1 \cdots \lambda_n)^d \,\mathrm{vol}(\mathcal{S}) \leqslant 2^{dn}.$$

Le théorème de Minkowski adélique a été obtenu en 1969 par McFeat dans sa thèse de doctorat, et publié en 1971 ([8] Thm 5). Il a été obtenu indépendamment de McFeat par Bombieri et Vaaler en 1983 ([2] Thm 3), dans le but justement de démontrer un lemme de Thue-Siegel. Signalons que McFeat, ainsi que Bombieri et Vaaler démontrent également l'analogue de la « partie facile » (minoration) du théorème de Minkowski. Comme signalé dans [3], leurs « parties faciles » diffèrent : Bombieri et Vaaler obtiennent

une meilleure minoration mais sous des hypothèses plus fortes. Pour une démonstration, que nous admettrons ici, on peut également consulter la revue de l'article de Bombieri et Vaaler par Gramain [5].

On déduit alors assez facilement du théorème 2 la version suivante de LTS fort relatif sur un corps de nombres :

Théorème 3 *Tout sous-espace V de K^n admet une base $x_1, \ldots, x_m$ avec :*

$$H(x_1) \cdots H(x_m) \leqslant m^{m/2}\, C(K)^m\, H(V)$$

où $C(K) = |\mathrm{Disc}(K)|^{1/2d}$ et $d = [K : \mathbf{Q}]$.

On va en fait montrer un résultat ([9] lemma 5.2) qui, conformément au corollaire 1, est plus fort :

Lemme 6 *Pour tout $A \in GL(K_{\mathbf{A}}^n)$, il existe une base $x_1, \ldots, x_n$ de K^n telle que : $H_A(x_1) \cdots H_A(x_n) \leqslant n^{n/2}\, C(K)^n\, H_A(K^n)$.*

Démo : Il s'agit simplement d'appliquer le théorème 2 au corps convexe adélique $\mathcal{S} = A^{-1}(\mathcal{C})$, où $\mathcal{C} = \{x \in K_{\mathbf{A}} \mid ||x||_v \leqslant 1 \quad \forall v \in M(K)\}$ est la « boule unité adélique ». En effet on a $x \in \lambda \cdot \mathcal{S} \cap K^n \implies H_A(x) \leqslant \lambda$. Pour minorer $\mathrm{vol}(\mathcal{S})$, on utilise les faits suivants :

- $\mathrm{vol}(\mathcal{S}) = H(\det A)^{-d}\, \mathrm{vol}(\mathcal{C})$ ([16] chap 4 prop 3)
- $\prod_{v \nmid \infty} |\mathcal{D}_v|_v = |\mathrm{Disc}(K)|^{1/d}$
- le volume $V(n)$ de la boule unité euclidienne de $\mathbf{R}^n$ est minoré par $(4/n)^{n/2}$

Notant s le nombre de places complexes de K, il vient alors :

$$\begin{aligned} \mathrm{vol}(\mathcal{S}) &\geqslant V(n)^{d-2s}\, V(2n)^s\, 2^{ns}\, |\mathrm{Disc}(K)|^{-n/2}\, H(\det A)^{-d} \\ &\geqslant \left(\frac{4}{n}\right)^{dn/2} C(K)^{-dn}\, H(\det A)^{-d} \end{aligned}$$

d'où $\lambda_1 \cdots \lambda_n \leqslant n^{n/2}\, C(K)^n\, H(\det A)$ et le lemme.

Comme on l'a signalé, Bombieri et Vaaler ont les premiers montré un LTS général sur les corps de nombres. Plus précisément, sous les hypothèses du théorème 3, ils obtiennent $H^\infty(x_1) \ldots H^\infty(x_m) \leqslant (2/\pi)^{sn/d}\, C(K)^m\, H^\infty(V)$, soit en comparant H à H^∞ et en majorant brutalement $(2/\pi)^{sn/d}$:

$$H(x_1) \cdots H(x_m) \leqslant n^{m/2}\, C(K)^m\, H(V)$$

comparable à notre résultat.

On peut déjà remarquer le $n^{m/2}$ qui intervient au lieu de notre $m^{m/2}$. Cela vient du fait que Bombieri et Vaaler travaillent à l'intérieur de K^n et non intrinsèquement dans $K^m \approx V$. De plus, leur méthode ne semble pas

s'adapter aux hauteurs tordues, et ne fournit en particulier aucun équivalent de notre lemme 6. Or, on le verra, c'est bien de ce résultat qu'on a besoin pour établir la proposition 19, qui est le LTS classique absolu montré par Roy et Thunder, et dont ils déduisent leur LTS fort.

Signalons qu'en revanche le concept de hauteur tordue ne semble pas faire un mariage heureux avec H^∞. En effet, on ne dispose *a priori* pas d'un équivalent du théorème 1 dont la conclusion serait :
(i) $H_A^\infty(\phi(x)) = H_B^\infty(x)$
(ii) $H_A(V) = H_B(K^m)$,
le problème résidant dans le conflit aux place infinies entre les normes $||\cdot||_\infty$, utilisée pour les points, et $||\cdot||$ pour les sous-espaces. Il semble d'ailleurs que les hauteurs H^∞ soient généralement moins aisées à manipuler, le résultat de Bombieri et Vaaler reposant notamment sur la délicate inégalité de sciage du cube de Vaaler.

2.2 Corps de fonctions

On va utiliser une version étendue du théorème de Riemann-Roch. Commençons par rappeler puis étendre certaines notions relatives aux diviseurs. Le groupe $\mathrm{Div}(K)$ des diviseurs de K, que nous noterons additivement, est ordonné naturellement par l'ordre produit de $\mathbf{Z}^{(M(K))}$. Ainsi les diviseurs positifs sont les diviseurs effectifs. On a par ailleurs un morphisme naturel $K^* \to \mathrm{Div}(K)$ donné par $\mathrm{div}(x) = \sum \mathrm{ord}_v(x) \cdot v$, qui se prolonge à $K_\mathbf{A}\backslash\{0\}$ par la même formule.

On introduit alors $\Lambda(D) = \{x \in K_\mathbf{A} \mid x = 0 \text{ ou } \mathrm{div}(x) + D \geqslant 0\}$, puis $L(D) = \Lambda(D) \cap K$. On remarque que : $D \leqslant D' \Leftrightarrow \Lambda(D) \subset \Lambda(D')$. Par ailleurs, K, $K_\mathbf{A}$, $\Lambda(D)$ et $L(D)$ sont des K_0-espaces vectoriels. Sauf précision contraire les dimensions sont désormais considérées sur K_0. On notera ainsi $\ell(D) = \dim L(D)$. Ce nombre est fini. Par ailleurs, si $U \subset V$ sont deux espaces, on notera $(V : U)$ la dimension du quotient. On définit ainsi $c(D) = (K_\mathbf{A} : \Lambda(D) + K)$. Si κ désigne un diviseur de la classe canonique, on a $c(D) = \ell(\kappa - D)$ ([1], p. 264).

Nous allons maintenant généraliser ces notions avec $A \in GL(K_\mathbf{A}^n)$ en lieu et place de $D \in \mathrm{Div}(K)$. Cette idée est justifiée par le morphisme surjectif $GL(K_\mathbf{A}) \twoheadrightarrow \mathrm{Div}(K)$. Ainsi, si $D \in \mathrm{Div}(K)$ provient de $A \in GL(K_\mathbf{A})$, on a $\Lambda(D) = \{x \in K_\mathbf{A} \mid x = 0 \text{ ou } \mathrm{div}(A(x)) \geqslant 0\}$. On se rappelle alors la définition du « cube unité adélique » (cf. démo du lemme 6) : $\mathcal{C}$ $(= \mathcal{C}^n) = \prod \mathcal{O}_v = \{x \in K_\mathbf{A}^n \mid \mathrm{div}(x) \geqslant 0\}$. On est donc naturellement conduit à poser $\Lambda(A) = A^{-1}(\mathcal{C})$, puis $L(A) = \Lambda(A) \cap K^n$, $\ell(A) = \dim L(A)$ et $c(A) = (K_\mathbf{A}^n : \Lambda(A) + K^n)$. Ces deux nombres sont également finis ; la finitude de $\ell(A)$ est l'objet du lemme suivant, tandis que celle de $c(A)$ viendra naturellement au cours de la démonstration du théorème 4.

Lemme 7 *Pour tout $A \in GL(K_\mathbf{A})^n$, $\ell(A)$ est fini.*

Démo : Nous allons déduire la finitude de $\ell(A)$ de celle, supposée bien connue, de $\ell(D)$ pour $D \in \mathrm{Div}(K)$. On pose $C = A^{-1}$ et $C^i = \pi_i \circ C$, où π_i est la i-ème projection naturelle $K_{\mathbf{A}}^n \to K_{\mathbf{A}}$. On remarque que $C_v^i(\mathcal{O}_v^n)$ est un $\mathcal{O}_v$-module libre de rang 1, égal à $\mathcal{O}_v$ en presque toute place v. Ainsi on peut trouver $c_i \in K_{\mathbf{A}}^*$ tel que $C^i(\prod \mathcal{O}_v^n) = c_i \prod \mathcal{O}_v$, donc $\Lambda(A) = C(\mathcal{C}) \subset \oplus_i \Lambda(c_i^{-1})$ et $\ell(A) \leqslant \sum_i \ell(\mathrm{div}(c_i^{-1}))$.

On est maintenant en mesure d'énoncer le théorème suivant, dont le cas $n = 1$ n'est autre, en vertu des remarques précédentes, que le théorème classique de Riemann-Roch :

Théorème 4 *Si K est de genre g, pour tout $A \in GL(K_{\mathbf{A}}^n)$, on a :*

$$\ell(A) - c(A) = \deg(\mathrm{div}(\det A)) + n(1-g)\,.$$

Pour la démonstration, on suivra [14], c'est-à-dire qu'on reprendra la structure de la première preuve du théorème classique de Riemann-Roch exposée par Artin ([1] Chap 14 §1-2). Une preuve d'apparence très différente est donnée, dans le cas k_0 fini, par Weil ([16] Chap 6 Thm 1), qui donne par ailleurs des indications pour le cas général. Thunder ([14] thm 3), ainsi que Weil (pour le cas k_0 fini) donnent un énoncé apparemment plus général, avec des sous-espaces de K^n. Par des considérations analogues au théorème 1, il se ramène à l'énoncé précédent, seul utile ici. La structure de la preuve sera la suivante : on montre que la quantité $\ell(A) - c(A) - \deg(\mathrm{div}(\det A))$ ne dépend pas de A, puis on la calcule pour un A agréable. Pour cela, on aura besoin de vérifier que cette quantité est bien définie, c'est-à-dire que $c(A)$ est effectivement fini.

Démo : On abrège $\deg(\mathrm{div}(\det(A)))$ en $\mathrm{ddd}(A)$; on va calculer $\mathrm{ddd}(A) - \mathrm{ddd}(B)$. Pour commencer, on prolonge sur $GL(K_{\mathbf{A}}^n)$ l'ordre naturel des diviseurs, en convenant que $A \geqslant B \Leftrightarrow \Lambda(A) \supset \Lambda(B)$. Le pré-ordre ainsi défini n'est bien sûr pas total. Cependant, on peut toujours trouver C tel que $\Lambda(C) \supset \Lambda(A) + \Lambda(B)$ (prendre p. ex. sur chaque facteur local une homothétie de rapport assez petit). Un tel C est supérieur, et en particulier comparable, à A et B. Ainsi, quitte à intercaler C on peut supposer $A \geqslant B$ dans le calcul de $\mathrm{ddd}(A) - \mathrm{ddd}(B)$.
On va d'abord montrer que $\mathrm{ddd}(A) - \mathrm{ddd}(B) = (\Lambda(A) : \Lambda(B))$. On part de $(\Lambda(A) : \Lambda(B)) = \sum_v (A_v^{-1}(\mathcal{O}_v^n) : B_v^{-1}(\mathcal{O}_v^n))$, et on regarde sur chaque facteur local. On considère $C\ (= C_v) = A_v B_v^{-1}$. Ainsi, $C(\mathcal{O}_v^n)$ est un sous-$\mathcal{O}_v$-module de $\mathcal{O}_v^n$, de rang n. Il s'écrit donc $\oplus_{i=1}^n C_i x_i \mathcal{O}_v$, avec $C_i \in \mathcal{O}_v$, dans une base (x_i) convenable de $\mathcal{O}_v^n$. Maintenant, si $D : x_i \mapsto x_i/C_i$, on a $DC(\mathcal{O}_v^n) = \mathcal{O}_v^n$, donc DC est de déterminant inversible dans $\mathcal{O}_v$ et $\det(C) = C_1 \cdots C_n$ à un inversible près. Posant

$c_{v,i} = \mathrm{ord}_v(C_i)$, on a ainsi $\sum_i c_{v,i} = \mathrm{ord}_v(\det(A_v B_v^{-1}))$. Il vient alors :

$$\begin{aligned}(\Lambda(A) : \Lambda(B)) &= \sum_{v \in M(K)} (\mathcal{O}_v^n : C_v(\mathcal{O}_v)^n) \\ &= \sum_v \sum_{i=1}^n (\mathcal{O}_v : \pi_v^{c_{v,i}} \mathcal{O}_v) \\ &= \sum_v f_v \sum_{i=1}^n c_{v,i} \\ &= \mathrm{ddd}(A) - \mathrm{ddd}(B).\end{aligned}$$

où on a noté $f_v = (\mathcal{O}_v : \pi_v \mathcal{O}_v)$ le degré local effectif.
On utilise maintenant $\Lambda(A) \supset \Lambda(B)$ et les isomorphismes classiques $(U+V)/V \approx V/(U \cap V)$ et $(U/W)/(V/W) \approx U/V$ pour calculer :

$$\begin{aligned}\left(\frac{K_{\mathbf{A}}^n}{\Lambda(B)+K^n}\right) / \left(\frac{K_{\mathbf{A}}^n}{\Lambda(A)+K^n}\right) &\approx \frac{\Lambda(A)+K^n}{\Lambda(B)+K^n} \\ &\approx \left(\frac{\Lambda(A)}{\Lambda(B)}\right) / \left(\frac{\Lambda(B)+\Lambda(A)\cap K^n}{\Lambda(B)}\right) \\ &\approx \left(\frac{\Lambda(A)}{\Lambda(B)}\right) / \left(\frac{L(A)}{L(B)}\right),\end{aligned}$$

soit en prenant les dimensions :

$$c(B) = c(A) + \mathrm{ddd}(A) - \mathrm{ddd}(B) + \ell(B) - \ell(A) \qquad (2.2.1)$$

Il est alors temps de s'intéresser à la finitude de $c(A)$. Pour cela, on utilise l'existence ([1] Chap 13 Lemma 2) d'un $x_0 \in K_{\mathbf{A}}^*$ tel que $\Lambda(x_0) + K = K_{\mathbf{A}}$. On note $I = \mathbf{Id}_{K_{\mathbf{A}}^n}$, et on choisit C tel que $C \geqslant A$ et $C \geqslant x_0 I$; en particulier $\Lambda(C) + K^n = K_{\mathbf{A}}^n$, soit $c(C) = 0$. Comme $C \geqslant A$, on peut substituer $(C; A)$ à $(A; B)$ dans (2.2.1) et $c(A) = \mathrm{ddd}(C) - \mathrm{ddd}(A) + \ell(A) - \ell(C)$ est fini.
On a donc prouvé $\ell(A) - c(A) - \mathrm{ddd}(A) = \ell(B) - c(B) - \mathrm{ddd}(B)$ dès que $A \geqslant B$, et en fait pour A et B quelconques. Ainsi on calcule :

$$\begin{aligned}\ell(A) - c(A) - \mathrm{ddd}(A) &= \ell(I) - c(I) - \mathrm{ddd}(I) \\ &= \dim(K_0^n) - (K_{\mathbf{A}}^n : \mathcal{C}^n + K^n) \\ &= n\big(\dim(K_0) - (K_{\mathbf{A}} : \mathcal{C}^1 + K)\big) \\ &= n(1 - c(0)) = n(1-g).\end{aligned}$$

Le théorème 4 admet le corollaire suivant, immédiat mais fondamental car il permet de lui faire jouer un rôle analogue, sur les corps de fonctions rationnelles, au premier théorème de Minkowski :

Corollaire 8 *Si $K = k_0(T)$ est rationnel et A satisfait $h_A(K^n) < n$, alors on peut trouver x non nul dans K^n avec $h_A(x) \leqslant 0$.*

Démo : Les deux points sont : $x \in L(A) \implies h_A(x) \leqslant 0$ et, dans le cas rationnel, $-\deg(\mathrm{div}(\det A)) = h_A(K^n)$. Or on a $\ell(A) = n - h_A(K^n) + c(A) > 0$ d'où la non-nullité de $L(A)$.

On introduit maintenant les minimums successifs relatifs d'une hauteur tordue additive h_A :

$$\rho_i(A) = \inf\{\rho \in \mathbf{R} \mid \exists\, x_1, \ldots, x_i \in K^n,\ x_1 \wedge \cdots \wedge x_i \neq 0 \text{ et } h_A(x_j) \leqslant \rho\}.$$

On fait de suite quelques remarques. On a bien sûr $\rho_1(A) \leqslant \cdots \leqslant \rho_n(A) < +\infty$. De plus, le corollaire 5 montre, en passant aux logarithmes, que $\rho_1(A) > -\infty$. Enfin les ρ_i sont atteints, ils sont même entiers si $\mathfrak{m}(K, k) = 1$, ce qui est par exemple le cas si K est un corps de fonctions rationnelles.

On se place pour l'instant dans le cas rationnel, où le corollaire 8 assure un certain contrôle du premier minimum de certaines hauteurs tordues h_A. On veut préciser ceci, et étendre notre contrôle à tous les minimums successifs. On va pour cela utiliser une application B qui va « coder » l'information adéquate sur les minimums de A pour que l'application à B du corollaire 8 permette de conclure sur A. La proposition suivante est un LTS fort sur les corps de fonctions rationnelles :

Proposition 9 *Soit $K = k_0(T)$ et $A \in GL(K^n_{\mathbf{A}})$, alors*

$$\rho_1(A) + \cdots + \rho_n(A) \leqslant h_A(K^n).$$

En fait, on a déjà l'inégalité inverse par la proposition 4 donc, dans le cas rationnel, une égalité, qui ne subsistera pas dans cas général. La démonstration, dont la stratégie sera reprise pour la proposition 22, utilise les deux lemmes suivants :

Lemme 10 *Dans K^n muni d'une hauteur tordue h_A, il existe une suite de sous-espaces $\{0\} = V_0 \subset V_1 \subset \cdots \subset V_n = K^n$ avec $\dim(V_i) = i$ satisfaisant :*

$$x \notin V_{i-1} \implies h_A(x) \geqslant \rho_i(A)$$

Démo : On va procéder par récurrence sur i, et exploiter le fait que les $\rho_i(A)$ sont atteints. On a déjà $V_0 = \{0\}$. On suppose ensuite V_i défini par récurrence ($1 \leqslant i < n$), et on choisit $y_1, \ldots, y_{i+1}$ linéairement indépendants de hauteurs bornées par $\rho_{i+1}(A)$. L'un (au moins) des y_j n'est pas dans V_i, on l'appelle x_{i+1} et on pose $V_{i+1} = V_i \oplus Kx_{i+1}$. On a ainsi obtenu une base (x_i) de K^n satisfaisant $h_A(x_i) \leqslant \rho_i(A)$ et $V_i = \langle x_1, \ldots, x_i \rangle$.
Considérons alors $x \notin V_{i-1}$. Quitte à remplacer i par le plus petit

j satisfaisant $\rho_j(A) = \rho_i(A)$, on peut supposer $\rho_{i-1}(A) < \rho_i(A)$. Si $i = 1$, alors certainement $h_A(x) \geqslant \rho_1(A)$. Si $i > 1$ et que l'on suppose de plus $h_A(x) < \rho_i(A)$, on obtient alors un système $x_1, \ldots, x_{i-1}, x$ de i vecteurs linéairement indépendants de hauteurs strictement inférieures à $\rho_i(A)$, ce qui est absurde. Ainsi $h_A(x) \geqslant \rho_i(A)$

Lemme 11 *Soient L un corps local non archimédien de valuation* $\mathrm{ord}(\cdot)$*, et des sous-espaces*

$$\{0\} \subset W_1 \subset \cdots \subset W_n = L^n \text{ avec } \dim W_i = i\,.$$

On se donne $a_1, \ldots a_n \in L$ tels que $\mathrm{ord}(a_1) \leqslant \cdots \leqslant \mathrm{ord}(a_n)$*. Alors on peut trouver $\phi \in GL(L^n)$ tel que :*
(i) $\mathrm{ord}(\det \phi) = \mathrm{ord}(a_1) + \cdots + \mathrm{ord}(a_n)$,
(ii) pour tout $x \in W_i$ on a : $\mathrm{ord}(\phi(x)) \leqslant \mathrm{ord}(a_i) + \mathrm{ord}(x)$.

Démo : On va se ramener au cas $W_i = \langle e_1, \ldots, e_i \rangle$, où (e_j) est la base canonique de K^n. On peut trouver des x_j tels que $x_1, \ldots, x_i$ forment une base sur $\mathcal{O}$ du $\mathcal{O}$-module libre $W_i \cap \mathcal{O}^n$. On définit alors $\psi \in GL(\mathcal{O}^n)$ comme appliquant x_j sur e_j. Ainsi ψ est une isométrie et $\mathrm{ord}(\det \psi) = 0$. On peut donc composer à droite par ψ sans modifier les conditions (i) et (ii), et ainsi supposer $W_i = \langle e_1, \ldots, e_i \rangle$. On pose alors $\phi : e_i \mapsto a_i e_i$. On a bien $\det \phi = a_1 \cdots a_n$. De plus si $x \in W_i$ on écrit $x = (x^1, \ldots, x^i, 0, \ldots, 0)$, $\mathrm{ord}(x) = \mathrm{ord}(x^{j_0})$, et on vérifie immédiatement :

$$\begin{aligned} \mathrm{ord}(\phi(x)) &= \min_j(\mathrm{ord}(a_1 x^1), \ldots, \mathrm{ord}(a_i x^i) \\ &\leqslant \mathrm{ord}(a_{j_0} x^{j_0}) \\ &\leqslant \mathrm{ord}(a_i) + \mathrm{ord}(x). \end{aligned}$$

Démo (de la proposition) : On commence par choisir arbitrairement une place $v \in M(K)$, sur laquelle on va travailler avec le lemme 11 de façon à coder dans une application B les minimums successifs de A. Comme K est supposé rationnel, les $\rho_i(A)$ sont dans $\mathbf{Z}$ et on peut choisir des a_i dans K_v tels que $\mathrm{ord}_v(a_i) = \rho_i(A) - 1$. On pose aussi $W_i = A_v(V_i) \subset K_v^n$ et on applique le lemme 11 qui fournit $\phi_v \in GL(K_v^n)$. On pose $\phi_w = \mathbf{Id}_{K_w}$ si $w \neq v$, puis $B = \phi \circ A$. Vérifions que B a les propriétés utiles.

Le (i) du lemme 11 donne $\mathrm{ord}_v(\det \phi) = \sum \rho_i(A) - n$. Comme de plus $\mathrm{ord}_w(\det \phi) = 0$ si $w \neq v$, on a $h_A(K^n) - \sum \rho_i(A) = h_B(K^n) - n$. On veut donc montrer que $h_B(K^n) \geqslant n$, ce qui sera donné par le corollaire 8 si l'on montre que $h_B(x) \geqslant 0$ pour tout x non nul. On peut supposer $x \in V_i \backslash V_{i-1}$, ce qui donne $A_v(x) \in W_i$. Le (ii) du lemme 11 s'écrit alors $\mathrm{ord}_v(B_v(x)) \leqslant \mathrm{ord}_v(A_v(x)) + \rho_i(A) - 1$; aux autres places on a $\mathrm{ord}_w(B_w(x)) = \mathrm{ord}_w(A_w(x))$. Sommant ceci, on trouve $h_B(x) \geqslant h_A(x) + 1 - \rho_i(A)$. Or on a supposé $x \notin V_{i-1}$, donc $h_A(x) \geqslant \rho_i$ par le lemme 10. Ainsi on a bien $h_B(x) > 0$ et la conclusion annoncée.

La proposition suivante étend ce résultat aux corps de fonctions algébriques. Pour cela, on va en fait voir K comme un espace vectoriel de dimension finie d sur $k = k_0(T)$. On ne va pas brutalement regarder le système de $n-m$ équations à coefficients dans K définissant V comme un système de $d(n-m)$ équations à coefficients dans k, car, comme on l'a mentionné, cela ne donne pas de bons résultats. On va plutôt utiliser à nouveau le théorème 4.

Proposition 12 *Pour tout $A \in GL(K_{\mathbf{A}}^n)$, il existe une base $x_1, \ldots, x_n$ de K^n telle que :*

$$h_A(x_1) + \cdots + h_A(x_n) \leqslant n \left(\frac{g - 1 + \mathfrak{m}(K,k)}{\mathfrak{m}(K,k)} \right) + h_A(K^n) \, .$$

Démo : On commence donc par fixer un isomorphisme $k^d \xrightarrow{\sim} K$, qui donne lieu à $\phi : k_{\mathbf{A}}^{nd} \xrightarrow{\sim} K_{\mathbf{A}}^n$, isomorphisme de $k_{\mathbf{A}}$-modules topologiques ([1] Chap 13 §2). On va construire $\tilde{A} \in GL(k_{\mathbf{A}}^{nd})$ tel que $\Lambda(\tilde{A}) = \phi^{-1}(\Lambda(A))$. Chaque facteur local de $\phi^{-1}(\Lambda(A))$ s'écrit $\phi_v^{-1}(\prod_{w|v} A_w^{-1}(\mathcal{O}_w^n))$, donc est un sous-$\mathcal{O}_v$-module de k_v^{nd}. Il est libre (de type fini sans torsion sur $\mathcal{O}_v$ principal), de rang nd (ϕ_v est un isomorphisme), et on en choisit une base. On prend $\tilde{A}_v$ l'appliquant sur la base canonique de k_v^{nd}, c'est une isométrie pour presque tout v, on obtient ainsi $\tilde{A} \in GL(k_{\mathbf{A}}^{nd})$, qui satisfait $\Lambda(\tilde{A}) = \phi^{-1}(\Lambda(A))$.
Comme $[K_0 : k_0] = \mathfrak{m}(K,k)/d$, on vérifie :

$$\begin{aligned}
\dim_{K_0} \left(\frac{K_{\mathbf{A}}^n}{\Lambda(A) + K^n} \right) &= \frac{\mathfrak{m}(K,k)}{d} \dim_{k_0} \left(\frac{k_{\mathbf{A}}^{nd}}{\Lambda(\tilde{A}) + k^{nd}} \right) \\
\dim_{K_0} \left(\Lambda(A) \cap K^n \right) &= \frac{\mathfrak{m}(K,k)}{d} \dim_{k_0} \left(\Lambda(\tilde{A}) \cap k^{nd} \right) \\
\mathfrak{m}(K,k) h_A(K^n) + n(g-1) &= \frac{\mathfrak{m}(K,k)}{d} \left(h_{\tilde{A}}(k^{nd}) - nd \right)
\end{aligned}$$

où la troisième égalité vient en soustrayant membre à membre les deux premières puis en appliquant le théorème 4 dans chaque membre.
On utilise alors la proposition 9 pour obtenir une base (y_j) de k^{nd} satisfaisant $\sum h_{\tilde{A}}(y_j) \leqslant h_{\tilde{A}}(k^{nd})$. On choisit y_{j_1} de hauteur minimale, on pose $x_1 = \phi(y_{j_1})$. Au plus d des y_j ont leur image par ϕ colinéaire à x_1, on choisit y_{j_2} de hauteur minimale parmi les autres, on pose $x_2 = \phi(y_{j_2})$. En on construit ainsi de proche en proche une base (x_i) de K^n qui satisfait :

$$\sum_{i=1}^{n} h_{\tilde{A}}(\phi^{-1}(x_i)) \leqslant \frac{h_{\tilde{A}}(k^{nd})}{d} = h_A(K^n) + n \left(\frac{g - 1 + \mathfrak{m}(K,k)}{\mathfrak{m}(K,k)} \right) .$$

La conclusion souhaitée est alors conséquence du lemme suivant :

Lemme 13 *Avec les notations précédentes, pour tout $x \in K^n$ non nul, on a $h_A(x) \leqslant h_{\tilde{A}}(\phi^{-1}(x))$.*

Démo : On pose $y = \phi^{-1}(x)$ et $\tilde{D} = -\mathrm{div}(\tilde{A}(y)) \in \mathrm{Div}(k)$, de sorte que $h_{\tilde{A}}(y) = \deg(\tilde{D})$. On peut supposer que $\tilde{D}$ provient de $a \in k_{\mathbf{A}}^*$. On a alors $\mathrm{div}(\tilde{A}(ay)) = 0$ donc $ay \in \Lambda(\tilde{A})$ et $ax = \phi(ay) \in \Lambda(A)$. Or ceci signifie $\mathrm{div}(A(ax)) \geqslant 0$ ou encore $\mathrm{div}(A(x)) \geqslant -D$, où $D = \mathrm{div}(a) \in \mathrm{Div}(K)$. Mais D et $\tilde{D}$ sont reliés par $\deg(D) = \mathfrak{m}(K,k)\deg(\tilde{D})$ ([1] Chap 15 Thm 9), d'où $h_A(x) \leqslant \deg(D)/\mathfrak{m}(K,k) = h_{\tilde{A}}(y)$.

On pose $C(K) = \exp\left(\frac{g-1+\mathfrak{m}(K,k)}{\mathfrak{m}(K,k)}\right)$ si K est un corps de fonctions. On relit alors la proposition précédente multiplicativement avant de lui appliquer le corollaire 1. On a ainsi montré la version suivante de LTS fort relatif sur les corps de fonctions :

Théorème 5 *Tout sous-espace V de K^n admet une base $x_1, \ldots, x_m$ avec :*

$$H(x_1) \cdots H(x_m) \leqslant C(K)^m H(V).$$

2.3 Résultats d'optimalité

On énonce ici sans démonstration deux résultats, dus à Roy et Thunder [10] (resp. Thunder [14]), qui montrent une certaine optimalité des LTS relatifs présentés ci-dessus. On peut en fait voir les constantes obtenues comme composées de deux morceaux : une partie en $C(K)$ dépendant du corps de base au travers de son discriminant ou de son genre, et une partie ne dépendant que de la dimension m. Le facteur en m est d'ailleurs éventuellement transparent (égal à 1) pour les corps de fonctions, ou pour les corps de nombres si on utilise H^∞ (résultat de Bombieri et Vaaler). Les résultats suivants ne donnent aucune indication sur le facteur en m. Leur but est de montrer que la dépendance gênante en $C(K)$ est bien indispensable.

Proposition 14 *Pour tous entiers m, n, d vérifiant $d > 1$ et $2 \leqslant m \leqslant n-1$, pour tout corps de nombres K de degré d, il existe un sous-espace $V \subset K^n$ de dimension m et une constante $C_1 = C_1(n,d)$ telle que toute base $x_1, \ldots, x_m$ de V satisfait :*

$$H(x_1) \cdots H(x_m) \geqslant C_1\, C(K)^{\frac{\lfloor (m-1)/2 \rfloor}{d-1}}\, H(V).$$

Proposition 15 *Pour tous entiers m, n, d vérifiant $d > 1$ et $2 < m \leqslant n-1$, pour tout K extension séparable de $k_0(T)$ de degré d, il existe un sous-espace $V \subset K^n$ de dimension m et une constante $C_2 = C_2(n,d)$ telle que toute base $x_1, \ldots, x_m$ de V satisfait :*

$$H(x_1) \cdots H(x_m) \geqslant C_2\, C(K)^{\frac{\lfloor (m-1)/2 \rfloor}{d-1}}\, H(V).$$

3 Lemmes absolus

3.1 Un LTS absolu

Nous suivons ici Roy et Thunder [9] pour établir un LTS fort sur $\bar{k}$, où k désigne $\mathbf{Q}$ ou $k_0(T)$. Afin de présenter simultanément les deux cas, on introduit δ valant 1 si $k = \mathbf{Q}$, et 0 sinon. On rappelle :

$$C(K) = \begin{cases} \left(\sqrt{|\mathrm{Disc}(K)|}\right)^{1/d} & \text{si } k = \mathbf{Q}, \\ \exp\left(\dfrac{g(K) - 1 + \mathfrak{m}(K,k)}{\mathfrak{m}(K,k)}\right) & \text{si } k = k_0(T). \end{cases}$$

La proposition suivante résume ainsi le lemme 6 et la proposition 12 :

Proposition 16 *Pour tout $A \in GL(K_{\mathbf{A}}^n)$, il existe une base $x_1, \cdots, x_n$ de K^n telle que : $H_A(x_1) \cdots H_A(x_n) \leqslant n^{\delta n/2}\, C(K)^n\, H_A(K^n)$.*

Nous allons montrer que l'on peut faire disparaître le $C(K)$ déplaisant en prenant les x_i dans $\bar{k}$. En contrepartie, on va récolter un $2^{\delta n(n-1)/2}$ moins bon que le $n^{\delta n/2}$ précédent, nous y reviendrons. On établit d'abord un résultat plus précis en dimension deux, qui montre comment $C(K)$ disparaît asymptotiquement quand le degré sur K des x_i tend vers l'infini.

Théorème 6 *Soient $A \in GL(K_{\mathbf{A}}^2)$ et $r > 0$, alors il existe une base y_1, y_2 de $\bar{k}^2$ avec $[K(y_i) : K] \leqslant r$ et*

$$H_A(y_1)\, H_A(y_2) \leqslant 2^{\delta}\, (r+1)^{\delta/r}\, C(K)^{2/r}\, H_A(\bar{k}^2).$$

La démonstration utilise les deux lemmes suivants :

Lemme 17 *Pour tout $r > 0$, on a $H_{\mathrm{S}^r A}(\mathrm{S}^r \bar{k}^n) = H_A(\bar{k}^n)^s$ avec $s = \binom{r+n-1}{n}$.*

Démo : Comme $H_A(\bar{k}^n) = H(\det A)$, il suffit bien sûr de voir que pour tout anneau R, $\det \mathrm{S}^r\phi = (\det \phi)^s$, où ϕ est un endomorphisme de R^n. Il s'agit de prouver une identité entre polynômes à coefficients entiers, on peut donc supposer $R = \mathbf{C}$. Supposons que la matrice de ϕ dans la base canonique admette une décomposition LU, L triangulaire inférieure, U supérieure. On constate que si la matrice de ψ est triangulaire dans une base, celle de $\mathrm{S}^r\psi$ l'est aussi dans la base associée, de sorte qu'on vérifie facilement l'égalité annoncée dans ce cas, donc pour toute matrice admettant une décomposition LU. Or ces matrices forment un ensemble dense de $GL(\mathbf{C}^n)$, d'où la conclusion.

Lemme 18 *Soient $P_1, \ldots, P_r \in \bar{k}[X_1, X_2]_1$ de produit Q et $A \in GL(K_{\mathbf{A}}^2)$, alors $H_A(P_1) \cdots H_A(P_r) \leqslant \sqrt{2}^{\,\delta r}\, H_A(Q)$.*

Démo : On regarde sur chaque facteur local. Aux places ultramétriques, on a $||P_1 \cdots P_r|| = ||Q||$, en invoquant le lemme de Gauss si on veut, ou par récurrence en considérant le coefficient de plus grande valeur absolue de chaque polynôme.
Aux places archimédiennes, on suppose $P \in \mathbf{C}[X_1, X_2]_1$ et on introduit la mesure de Mahler $\mathcal{M}$. On utilise alors sa multiplicativité et les trois faits suivants valables pour P linéaire homogène et Q quelconque :
– $\mathcal{M}(Q) \leqslant ||Q||$ ([6] lemma B.7.3.1(iii))
– $||P||_\infty \leqslant \mathcal{M}(P)$ ([6] lemma B.7.3.2)
– $||P|| \leqslant \sqrt{2}\,||P||_\infty$
Il vient alors :

$$\begin{aligned} ||P_1 \cdots P_r|| &\geqslant \mathcal{M}(P_1 \cdots P_r) = \mathcal{M}(P_1) \cdots \mathcal{M}(P_r) \\ &\geqslant ||P_1||_\infty \cdots ||P_r||_\infty \\ &\geqslant \sqrt{2}^{-r}\, ||P_1|| \cdots ||P_r|| \end{aligned}$$

Démo (du théorème) : Tout repose sur l'identification $\mathrm{S}^r R \approx R[X,Y]_r \approx R^{r+1}$. On identifie d'abord $\mathrm{S}^r K^2$ et K^{r+1} et on utilise la proposition 16 pour avoir une base $\alpha_0, \ldots, \alpha_r$ de $\mathrm{S}^r K^2$ telle que

$$H_{\mathrm{S}^r A}(\alpha_0) \cdots H_{\mathrm{S}^r A}(\alpha_r) \leqslant (r+1)^{\delta(r+1)/2}\, C(K)^{r+1}\, H_{\mathrm{S}^r A}(\mathrm{S}^r K^2) \quad (3.1.1)$$

Le point est qu'aux α_i correspondent des polynômes homogènes à deux variables qu'on va pouvoir décomposer en facteurs linéaires dans $\bar{k}[X,Y]$: $\alpha_i = \prod x_{ij}$. On voit alors les x_{ij} comme des vecteurs de $\bar{k}^2$. Ils vérifient $[K(x_{ij}) : K] \leqslant r$ et on va en extraire la base désirée.
On choisit pour y_1 un des x_{ij} de hauteur minimale. Le sous-espace de $\bar{k}[X,Y]_r$ formé des multiples de y_1^s est isomorphe à $\bar{k}[X,Y]_{r-s}$ de dimension $r+1-s$. Ainsi, au moins s des α_i ne sont pas divisibles par y_i^s et ont donc au plus $r+1-s$ facteurs x_{ij} associés y_1. Il y a donc au plus $\frac{r(r+1)}{2}$ vecteurs x_{ij} colinéaires à y_1. Parmi les au moins $\frac{r(r+1)}{2}$ autres, on choisit y_2 de hauteur minimale.
Mettant bout à bout ces décomptes, le lemme 18 et la minoration (3.1.1) où l'on a substitué le résultat du lemme 17, il vient :

$$\begin{aligned} \Big[H_A(y_1)\, H_A(y_2)\Big]^{r(r+1)/2} &\leqslant \prod_{i=0}^{r} \prod_{j=1}^{r} H_A(x_{ij}) \leqslant \prod_{i=0}^{r} \sqrt{2}^{\delta r}\, H_{\mathrm{S}^r A}(\alpha_i) \\ &\leqslant \Big[2^\delta\, (r+1)^{\delta/r}\, C(K)^{2/r}\, H_A(K^2)\Big]^{r(r+1)/2} \end{aligned}$$

et le théorème en prenant les racines.

Pour poursuivre, on définit les minimums successifs absolus d'une hauteur tordue multiplicative H_A :

$$\mu_i(A) = \inf\{\mu > 0 \mid \exists\, x_1, \ldots, x_i \in \bar{k}^n,\ x_1 \wedge \cdots \wedge x_i \neq 0 \text{ et } H_A(x_j) \leqslant \mu\}.$$

On remarque que $0 < \mu_1(A) \leqslant \cdots \leqslant \mu_n(A) < +\infty$, la stricte positivité étant donnée par le corollaire 5. On va d'abord majorer $\mu_1(A)$ par récurrence sur $n \geqslant 2$.

Proposition 19 *Soit $A \in GL(K_{\mathbf{A}}^n)$ pour $n \geqslant 2$ alors on a :*
(i) $\mu_1(A^*)^{n-1} \leqslant 2^{\delta(n-1)(n-2)/2}\, H(\det A)^{-1}\, \mu_1(A)$,
(ii) $\mu_1(A) \leqslant 2^{\delta(n-1)/2}\, H_A(\bar{k}^n)^{1/n}$.

Démo : Dans le cas $n = 2$, (i) est donné par la formule de dualité (1.1.1), et (ii) par le théorème 6.

Si $n > 2$, on procède par récurrence. On fixe $\varepsilon > 0$ et on choisit $x \in \bar{k}^n$ tel que $H_A(x) \leqslant \mu_1(A) + \varepsilon$. On pose $V = \bar{k}x$ et le théorème 1 fournit $B \in GL(K_{\mathbf{A}}^{n-1})$ qui reflète la hauteur H_{A^*} sur $V^\perp$. En particulier $H(\det B) = H_{A^*}(V^\perp) = H(\det A)^{-1}\, H_A(V)$ et $\mu_1(A^*) \leqslant \mu_1(B)$. Utilisant le (ii) de l'hypothèse de récurrence, puis la formule de dualité (1.1.1) il vient :

$$\begin{aligned}
\mu_1(A^*)^{n-1} &\leqslant \mu_1(B)^{n-1} \\
&\leqslant 2^{\delta(n-1)(n-2)/2}\, H(\det B) \\
&\leqslant 2^{\delta(n-1)(n-2)/2}\, H(\det A)^{-1}\, H_A(V) \\
&\leqslant 2^{\delta(n-1)(n-2)/2}\, H(\det A)^{-1}\, (\mu_1(A) + \varepsilon)
\end{aligned}$$

qui prouve (i) car ε est arbitrairement petit. Le (ii) s'en déduit en combinant (i) avec... lui-même, appliqué à A^* ! Plus explicitement :

$$\begin{aligned}
\mu_1(A)^{(n-1)^2} &\leqslant 2^{\delta(n-1)^2(n-2)/2}\, H(\det A^*)^{-(n-1)}\, \mu_1(A^*)^{n-1} \\
&\leqslant 2^{\delta(n-1)^2(n-2)/2+\delta(n-1)(n-2)/2}\, H(\det A)^{n-2}\, \mu_1(A)
\end{aligned}$$

d'où :

$$\mu_1(A)^{n(n-2)} \leqslant \left[2^{\delta(n-1)/2}\, H_A(\bar{k}^n)^{1/n}\right]^{n(n-2)}$$

qui donne (ii) en simplifiant.

On remarque qu'on vient en fait comme annoncé de prouver un LTS classique. Nous allons en déduire un LTS fort en utilisant la même stratégie qu'à la proposition 9, appliquée néanmoins avec plus de prudence et d'ε. En effet, les μ_i ne sont plus atteints, plus à valeurs entières, et on doit enfin prendre garde aux extensions de corps. On commence par deux lemmes analogues aux lemmes 10 et 11 :

Lemme 20 *Dans $\bar{k}^n$ muni d'une hauteur tordue H_A, il existe une suite de sous-espaces $\{0\} = V_0 \subset V_1 \subset \cdots \subset V_n = K^n$ avec $\dim(V_i) = i$ satisfaisant :*

$$x \notin V_{i-1} \implies h_A(x) \geqslant \mu_i(A).$$

Démo : On commence par choisir $\varepsilon > 0$ de sorte que $\mu_{i-1}(A) + \varepsilon < \mu_i(A)$ chaque fois que $\mu_{i-1}(A) \neq \mu_i(A)$. De façon analogue au lemme 10 on construit une base $x_1, \ldots, x_n$ de $\bar{k}^n$ telle que $H_A(x_i) \leqslant \mu_i(A) + \varepsilon$. Considérons alors $x \notin V_{i-1}$ et posons à nouveau j minimal avec $\mu_j(A) = \mu_i(A)$. Alors $x_1, \ldots x_{j-1}, x$, forment un système linéairement indépendant et on a toujours $H_A(x_k) < \mu_j(A)$ pour $1 \leqslant k < j$ grâce au choix de ε, assurant ainsi $H_A(x) \geqslant \mu_i(A)$.

Lemme 21 *On se donne $(L, |\cdot|)$ un corps local, des sous-espaces*

$$\{0\} \subset W_1 \subset \cdots \subset W_n = L^n \text{ avec } \dim W_i = i$$

et $a_1, \ldots a_n \in L$ tels que $|a_1| \geqslant \cdots \geqslant |a_n|$. Alors on peut trouver $\phi \in GL(L^n)$ tel que :
(i) $|\det \phi| = |a_1 \cdots a_n|$,
(ii) pour toute extension F/L finie et tout $x \in W_i \otimes_L F$ on a :
$||\phi(x)|| \geqslant |a_i|\,||x||$.

Démo : On se ramène comme précédemment au cas $W_i = \langle e_1, \ldots, e_i \rangle$ en composant par une isométrie ψ convenable, à condition de remarquer qu'alors ψ_F est aussi une isométrie, et ne modifie donc pas la condition (ii). On pose alors $\phi(e_i) = a_i e_i$ comme plus haut.
Le cas archimédien, d'ailleurs inutile ici, se traite de façon analogue.

Proposition 22 *Pour tout $A \in GL(K_{\mathbf{A}}^n)$ on a :*

$$\mu_1(A) \cdots \mu_n(A) \leqslant 2^{\delta n(n-1)/2}\, H_A(\bar{k}^n).$$

Démo : On considère une suite de sous-espaces V_i dont l'existence est assurée par le lemme 20. Soit E un corps de définition des V_i, où E/K est finie. On choisit $v \in M(E)$, on note encore v sa restriction à k, puis on prend $a \in E$ tel que $|a|_v > 1$. Quitte à étendre E et à extraire une racine d'ordre élevé de a, on peut même supposer que $1 < |a|_v^e < 1 + \varepsilon$ où $e = [E_v : k_v]/[E : k]$. Posons ensuite $p_i = \inf\{p \in \mathbf{Z} \mid \mu_i(A)|a^p|_v^e \geqslant 1\}$ et $a_i = a^{p_i}$ de sorte que $|a_1|_v \geqslant \cdots \geqslant |a_n|_v$ et :

$$1 \leqslant \mu_i(A)|a_i|_v^e \leqslant 1 + \varepsilon\,.$$

On définit enfin $W_i = A_v(V_i) \subset E_v^n$. On est alors dans les conditions idéales pour appliquer le lemme 21 qui fournit un $\phi_v \in GL(E_v^n)$. On complète par $\phi_w = \mathbf{Id}_{E_w}$ sur les autres facteurs pour avoir $\phi \in GL(E_{\mathbf{A}}^n)$ tel que notamment

$$H(\det \phi) = |a_1 \cdots a_n|_v^e \leqslant (1+\varepsilon)^n \Big(\prod_i \mu_i(A) \Big)^{-1}.$$

On pose $B = \phi \circ A$; on a $H_B(\bar{k}^n) = H(\det\phi) H_A(\bar{k}^n)$. On majore alors $\mu_i(B)$ par la proposition 19 (ii) :

$$\mu_1(B)^n \leqslant 2^{\delta n(n-1)/2}\, H_B(\bar{k}^n) \leqslant (1+\varepsilon)^n\, 2^{\delta n(n-1)/2}\, H_A(\bar{k}^n) / \prod_i \mu_i(A)$$

ce qui donne envie de vérifier si $\mu_1(B) \geqslant 1$.
Soit donc $x \in \bar{k}$ non nul, défini sur F extension finie de E. Alors

$$\begin{aligned} H_B(x) &= \prod_{\substack{w \in M(F) \\ w \mid v}} ||\phi_w(A_v x)||_w^{\frac{[F_w:E_v][E_v:k_v]}{[F:E][E:k]}} \prod_{\substack{w \in M(F) \\ w \nmid v}} ||A_w x||_w^{\frac{[F_w:k_w]}{[F:k]}} \\ &\geqslant |a_i|_v^e\, H_A(x) \quad \text{où } i \text{ est tel que } A_v x \in W_i \backslash W_{i-1} \\ &\geqslant \mu_i(A)^{-1}\, H_A(x) \\ &\geqslant 1 \end{aligned}$$

ce qui prouve bien $\mu_1(B) \geqslant 1$. La proposition s'en déduit en faisant tendre ε vers 0 dans

$$\mu_1(A) \cdots \mu_n(A) \leqslant 2^{\delta n(n-1)/2}\, H_A(\bar{k}^n)(1+\varepsilon)^n.$$

En vertu du corollaire 1, et vu la définition des minimums successifs, on a en fait montré le LTS fort absolu suivant :

Théorème 7 *Soit $V \subset \bar{k}^n$ un sous-espace de dimension m, alors pour tout $\varepsilon > 0$ il existe une base $x_1, \ldots, x_m$ de V telle que :*

$$H(x_1) \cdots H(x_m) \leqslant \big(2^{\delta m(m-1)/2} + \varepsilon\big) H(V)\,.$$

Il convient de noter que les x_i en question dépendent *a priori* du ε choisit.

3.2 Questions d'optimalité

Dans le cas des corps de fonctions, le résultat du théorème 7 se lit : $H(x_1) \cdots H(x_m) \leqslant (1+\varepsilon) H(V)$. Or la proposition 4 dit que les x_i vérifient également $H(x_1) \cdots H(x_m) \geqslant H(V)$. Le résultat est donc optimal. Rappelons que si k est un corps de fonctions rationnelles, il existe même une base (x_i) de V telle que $H(x_1) \cdots H(x_m) = H(V)$ par la proposition 9.

Concernant le cas de $\overline{\mathbf{Q}}$, on commence par montrer le résultat suivant :

Proposition 23 *Soient m et n deux entiers satisfaisant $2 \leqslant m \leqslant n-1$. Alors il existe un sous-espace $V \subset \overline{\mathbf{Q}}^n$ de dimension m tel que toute base (x_i) de V satisfait :*

$$H(x_1) \cdots H(x_m) \geqslant \frac{2^{m/2}}{\sqrt{m+1}}\, H(V)\,.$$

Démo : On définit V par les équations $x_1 + \cdots + x_{m+1} = 0$ et $x_{m+2} = \cdots = x_n = 0$. Ainsi, $H(V) = \sqrt{m+1}$ et tout vecteur non nul de V a au moins deux coordonnées non nulles, de sorte qu'on se ramène à montrer que $H(1,y) \geqslant \sqrt{2}$ pour tout $y \in \overline{\mathbf{Q}}$ non nul. On note $P_y = a_d X^d + \cdots + a_0 = a_d \prod_{i=1}^{d} (X - y_i) \in \mathbf{Z}[X]$ le polynôme minimal de y. Il vient alors :

$$\begin{aligned}
H(1,y) &= \left(\prod_{v|\infty} \max(1, |y|_v) \right)^{1/d} \left(\prod_{v|\infty} \sqrt{1 + |y|_v^2} \right)^{1/d} \\
&= a_d^{1/d} \left(\prod_{i=1}^{d} (1 + |y_i|^2)^{1/d} \right)^{1/2} \\
&= a_d^{1/d} \left(\exp \left(\frac{f\left(\log |y_1|^2\right) + \cdots + f\left(\log |y_d|^2\right)}{d} \right) \right)^{1/2} \\
&\geqslant a_d^{1/d} \left(\exp f\left(\log |y_1 \cdots y_d|^{2/d} \right) \right)^{1/2} \\
&\geqslant \left(|a_d|^{2/d} + |a_0|^{2/d} \right)^{1/2} \\
&\geqslant \sqrt{2}
\end{aligned}$$

où on a noté $f : x \mapsto \log(1 + e^x)$ et utilisé sa convexité.

Cette proposition montre donc que la constante optimale est minorée par $2^{m/2}/\sqrt{m+1}$ alors que le théorème 7 la majore par $2^{m(m+1)}/2$. Essentiellement, son comportement est donc « entre » $\sqrt{2}^m$ et $\sqrt{2}^{m^2}$, ce qui laisse de la place. En fait, une amélioration de la borne supérieure a été obtenue par S. David et P. Philippon en 1999 ([4] lemme 4.7 et remarque). Ils obtiennent un $\sqrt{m}^m$ en utilisant dans le cas linéaire un théorème général de Zhang sur les minimums successifs algébriques. Ils donnent par ailleurs en appendice une démonstration de la version « classique » (contrôle d'un seul vecteur) de leur LTS indépendante du résultat de Zhang.

On ne sait *a priori* pas si cette constante $\sqrt{m}^m$ est optimale, on peut en tout cas remarquer qu'elle coïncide, d'une part avec la partie en m de la constante obtenue sur un corps de nombres, d'autre part avec la constante du LTS initial sur $\mathbf{Q}$ ...

Références

[1] E. Artin, *Algebraic numbers and algebraic functions*, Gordon and Breach, 1967.

[2] E.Bombieri – J.D. Vaaler, *On Siegel's lemma*, Invent. Math. **73** (1983), 11-32.

[3] E.Bombieri – J.D. Vaaler, *Addendum to « On Siegel's lemma »*, Invent. Math. **75** (1984), 377.

[4] S. David – P. Philippon, *Minorations des hauteurs normalisées des sous-variétés des tores*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4) **XXVIII** (1999), 489-543.

[5] F. Gramain, *Sur le lemme de Siegel (D'après E. Bombieri et J.D. Vaaler)*, Publ. Math. Paris 6, Problèmes diophantiens 1983-1984.

[6] M. Hindry – J.H. Silverman, *Diophantine geometry : an introduction*, Springer GTM **201**, 2000.

[7] K. Mahler, *An analogue of Minkowski's geometry of numers in a field of series*, Ann. of Math. (2) **42** (1941), 488-522.

[8] R.B. McFeat, *Geometry of numbers in Adele spaces*, Diss. Math. **LXXXVIII** (1971), 54pp.

[9] D. Roy – J.L. Thunder, *An absolute Siegel's lemma*, J. Reine Angew. Math. **476** (1996), 1-26.

[10] D. Roy – J.L. Thunder, *A note on Siegel's lemma over number fields*, Monatsh. Math. **120** (1995), 304-344.

[11] W.M. Schmidt, *Diophantine approximation and diophantine equations*, Springer LNM **1467**, 1991.

[12] C.L. Siegel, *Über Anwendungen diophantischer Approximationen*, Abh. Preuß. Akad. d. Wiss. Phys-Math. Kl., **1** (1929)(=Ges.Abh., **I**, 209-266).

[13] A. Thue, *Über Annäherungswerte algebraischer Zhalen*, J. Reine Angew. Math. **135** (1909), 284-305.

[14] J.L. Thunder, *Siegel's lemma for function fields*, Michigan Math. J. **42** (1995), 147-162.

[15] J.L. Thunder, *Asymptotic estimates for rational points of bounded height on flag varieties*, Comp. Math. **88** (1993), 155-186.

[16] A. Weil, *Basic number theory*, Springer CIM.