Hauteurs et équations diophantiennes

Table des matières

1 Introduction

Lorsque l'on étudie une équation diophantienne, ou une classe d'équations, on commence par étudier la *finitude* du nombre de solutions. Ainsi, il existe quelques résultats célèbres donnant la finitude du nombre de solutions de certaines classes :

Theorem 1.1 (Thue, 1909)
Soit $\Phi(X,Y) \in \mathbb{Z}[X,Y]$, homogène de degré $n \geqslant 3$ et soit $A \in \mathbb{Z}\backslash\{0\}$. L'équation $\Phi(x,y) = A$ n'admet qu'un nombre fini de solutions entières

Theorem 1.2 (Skolem, 1929)
Soit $F(X,Y) \in \mathbb{Z}[X,Y]$ un irréductible tel que $F(0,0) = 0$. L'équation $F(x,y) = 0$ n'admet qu'un nombre fini de solutions entières à pgcd fixé.

Theorem 1.3 (Faltings, 1983)
Une courbe algébrique de genre supérieur ou égal à 2 n'a qu'un nombre fini de points K-rationnels pour tout corps de nombres K.

Mais un gros inconvénient de ces résultats vient du fait qu'ils ne sont pas explicites. On n'a aucune idée de la taille maximale des solutions. Une seconde étape dans la résolution de ces classes d'équations est donc de chercher une borne maximale pour la taille des solutions. Ainsi, on a des résultats explicites pour les problèmes de Thue dûs à Baker (1969), puis à Feldman (1971) et de Skolem, dû à Walsh (1992). Ces résultats sont sous la forme

Les solutions entières (x,y) vérifient

$$\max\{|x|,|y|\} \leqslant C$$

où C est une constante explicite qui dépend du polynôme donnant l'équation.

L'étape suivant dans la résolution de ces classes d'équations est la généralisation de ces résultats aux nombres algébriques. Mais on voit immédiatement un gros problème pour cette généralisation : on a besoin d'une bonne notion de "taille" pour les solutions. Sur $\mathbb{Z}$, on a une bonne notion de taille, c'est la valeur absolue usuelle. Ce qu'on demande à une *bonne notion*, c'est deux choses :

- D'abord une propriété de finitude. C'est à dire que l'on veut qu'il n'y ait qu'un nombre fini de solutions de taille bornée.
- On veut aussi un outil qui respecte l'arithmétique des solutions (comme la valeur absolue est multiplicative, elle conserva la décomposition en premiers).

Quand on se place sur un corps de nombre, la valeur absolue usuelle ne suffit plus. Déjà sur $\mathbb{Q}$, ça ne marche plus. On n'a plus la finitude. Par contre, un truc qui marche bien,

c'est prendre le max des valeurs absolues des numérateur et dénominateur. Si l'on se donne une borne B,

$$\{a/b \in \mathbb{Q} \text{ tq } \max\{|a|, |b|\} \leqslant B\}$$

est bien fini. Et en fait, c'est une généralisation de cette quantité aux nombres algébriques qui donne la bonne notion de taille pour les nombres algébriques.

2 Rappels sur les corps de nombres

2.1 Définitions

Definition 2.1 *1. Un corps de nombres est une extension finie de $\mathbb{Q}$.*

2. Soit K un corps de nombres. Un élément x de K est dit entier s'il est racine d'un polynôme unitaire de $\mathbb{Z}[X]$.

Proposition 2.2 *L'ensemble $\mathcal{O}_K$ des entiers d'un corps de nombres K est un anneau de Dedekind. C'est-à-dire qu'il est nœtherien, intègre et que tout idéal premier de $\mathcal{O}_K$ est maximal. D'autre part, tout idéal $\mathfrak{I}$ de $\mathcal{O}_K$ se décompose de façon unique en un produit d'idéaux premiers.*

2.2 Décomposition d'un idéal premier dans une extension

Soit K un corps de nombres et soit L une extension finie de K de degré $[L : K] = n$. Comme ce sont des anneaux de Dedekind, $\mathcal{O}_L$ est un $\mathcal{O}_K$-module de rang $\leqslant n$. D'autre part, tout idéal premier $\mathfrak{p}$ de $\mathcal{O}_K$ donne un idéal $\mathfrak{p}\mathcal{O}_L$ de $\mathcal{O}_L$. Ce dernier se décompose donc

$$\mathfrak{p}\mathcal{O}_L = \prod_{i=1}^{r} \mathfrak{P}_i^{e_i}$$

où les $\mathfrak{P}_i$ sont des idéaux premiers de $\mathcal{O}_L$ et les e_i sont des entiers $\geqslant 1$. Pour tout i, l'entier e_i est l'*indice de ramification* de $\mathfrak{p}$ en $\mathfrak{P}_i$.

D'autre part, les idéaux premiers $\mathfrak{P}_i$ qui apparaissent dans la décomposition de $\mathfrak{p}\mathcal{O}_L$ sont exactement les idéaux de $\mathcal{O}_L$ dont l'intersection avec $\mathcal{O}_K$ donnent $\mathfrak{p}$ (clair par maximalité de $\mathfrak{p}$). Ainsi, comme $\mathcal{O}_L$ est un $\mathcal{O}_K$-module libre de type fini, le corps résiduel $\mathcal{O}_L/\mathfrak{P}_i$ est une extension finie de $\mathcal{O}_K/\mathfrak{p}$ pour tout indice i. On note f_i le degré de cette extension (le *degré résiduel*).

De même, toujours du fait de la structure de $\mathcal{O}_K$-module libre de type fini de $\mathcal{O}_L$, le quotient $\mathcal{O}_L/\mathfrak{p}\mathcal{O}_L$ est un $\mathcal{O}_K/\mathfrak{p}$-espace vectoriel de dimension finie.

Theorem 2.3 *Avec les notations précédentes, on a*

$$\sum_{i=1}^{r} e_i f_i = [\mathcal{O}_L/\mathfrak{p}\mathcal{O}_L : \mathcal{O}_K/\mathfrak{p}] = n.$$

Preuve : Considérons la suite strictement décroissante

$$\mathcal{O}_L \supset \mathfrak{P}_1\mathcal{O}_L \supset \mathfrak{P}_1^2\mathcal{O}_L \supset \ldots \supset \mathfrak{P}_1^{e_1} \supset \mathfrak{P}_1^{e_1}\mathfrak{P}_2\mathcal{O}_L \supset \ldots \supset \mathfrak{p}\mathcal{O}_L.$$

Par maximalité, il n'existe pas d'idéal compris entre deux termes consécutifs $\mathfrak{B}$ et $\mathfrak{B}\mathfrak{P}_i$ de cette suite. Ainsi, chacun des quotients $\mathfrak{B}/\mathfrak{B}\mathfrak{P}_i$ est un $\mathcal{O}_L/\mathfrak{P}_i$-module libre de type fini (i.e. un $\mathcal{O}_L/\mathfrak{P}_i$-ev) de dimension 1 (car tout sev du quotient $\mathfrak{B}/\mathfrak{B}\mathfrak{P}_i$ produit un idéal de $\mathfrak{B}$ contentant $\mathfrak{B}\mathfrak{P}_i$). C'est donc un $\mathcal{O}_K/\mathfrak{p}$-ev de dimension f_i. Or pour chaque indice i, il y a e_i inclusions du type $\mathfrak{B} \supset \mathfrak{B}\mathfrak{P}_i$. Il vient donc

$$[\mathcal{O}_L/\mathfrak{p}\mathcal{O}_L : \mathcal{O}_K/\mathfrak{p}] = \sum_{i=1}^{r} e_i f_i.$$

D'autre part, si $\mathcal{O}_K$ est principal, $\mathcal{O}_L$ apparaît comme un comme un $\mathcal{O}_K$-module libre de rang exactement n. Mais en localisant $\mathcal{O}_K$ en $\mathfrak{p}$ (i.e. en multipliant par $S_\mathfrak{p}^{-1}$ où $S_\mathfrak{p} = \mathcal{O}_K \setminus \mathfrak{p}$), on obtient un anneau principal, donc $S_\mathfrak{p}^{-1}\mathcal{O}_L$ est un $S_\mathfrak{p}^{-1}\mathcal{O}_K$-module libre de rang n. En passant au quotient, comme $\mathcal{O}_K/\mathfrak{p}$ est un corps, $\mathcal{O}_L/\mathfrak{p}\mathcal{O}_L$ est un $\mathcal{O}_K/\mathfrak{p}$-ev de dimension n. □

3 Valeurs absolues

Pour généraliser cela aux nombres algébriques, il faut d'abord d'autres valeurs absolues sur $\overline{\mathbb{Q}}$. Rappelons que

Definition 3.1 *Une valeur absolue sur un corps k, c'est une fonction $k \to [0, \infty)$ tq*

1. $|x| = 0 \Leftrightarrow x = 0$.
2. $|xy| = |x|.|y|$.
3. $|x + y| \leqslant |x| + |y|$ *(inégalité triangulaire).*

On peut remplacer la condition 3. *par la condition (plus forte)*

$$3'.\ |x + y| \leqslant \max\{|x|, |y|\}\ \textit{(inégalité ultramétrique).}$$

Sur un corps de nombres, on peut construire pas mal de valeurs absolues. Pour se donner une idée, on peut déjà regarder sur $\mathbb{Q}$.

3.1 Sur $\mathbb{Q}$.

Sur $\mathbb{Q}$, on a la valeur absolue usuelle (qui vérifie 3. mais pas 3'.. On dit qu'elle est archimédienne) :

$$|x| = \max\{x, -x\}.$$

On la note $|.|_\infty$.

D'autre part, pour tout nombre premier p, on peut définir une valeur absolue dite p-adique, que l'on note $|.|_p$.

Definition 3.2 *Pour tout $x \in \mathbb{Q}$ il existe un entier $v_p(x)$ et des entiers a et b tels que $x = p^{v_p(x)}.\frac{a}{b}$ et $p \nmid ab$ (pour $x = 0$, on pose $v_p(0) = +\infty$). On appelle cela la valuation en p et l'on définit*

$$|x|_p = p^{-v_p(x)}.$$

$|.|_p$ est bien une valeur absolue, et elle vérifie même l'inégalité ultramétrique. On dira qu'elle est non archimédienne. Et on peut définir une notion de taille si l'on dit qu'un entier est "petit" s'il est divisible par une grande puissance de p.

Remarque : avec une valeur absolue, on peut définir une métrique, et si la valeur absolue est ultramétrique, on à des résultats rigolos. Par exemple, dans une boule, tout point à l'intérieur de la boule est centre de la boule. Ou encore, pour qu'une série converge, il suffit que son terme général tende vers 0.

On appelle valeurs absolues standards sur $\mathbb{Q}$ l'ensemble de ces valeurs absolues et on le note $\mathcal{M}_\mathbb{Q}$. (En fait, on note $\mathcal{M}_\mathbb{Q}$ l'ensemble des premiers, union $\{\infty\}$ pour pouvoir indicer).
Une première propriété de cet ensemble, c'est la formule du produit :

$$\forall\, x \in \mathbb{Q}^*, \qquad \prod_{v \in \mathcal{M}_\mathbb{Q}} |x|_v \;=\; 1.$$

Preuve : Si l'on isole les valeurs absolues p-adiques, et si l'on écrit $x = a/b$, ne vont intervenir dans le produit que les valeurs absolues . correspondant à un premier qui divise a ou b. Et avec le "$-$" qu'on a dans la définition des valeurs absolues , on a

$$\prod_{p|ab} |a/b| = |b/a|_\infty.$$

□

3.2 Sur un corps de nombres K.

Soit donc $K/\mathbb{Q}$ un corps de nombres de degré $[K : \mathbb{Q}] = n$. Il existe n plongements du corps K dans $\mathbb{C}$ et chaque plongement σ donne une valeur absolue archimédienne $|.|_\sigma$:

$$|x|_\sigma = |\sigma(x)|_\infty.$$

Rq : comme les plongements complexes sont conjugués deux à deux, en notant $n = r_1 + 2r_2$ où r_1 est le nombre de plongements réels et r_2 le nombre de couples de plongements complexes conjugués, on peut montrer qu'il y a exactement $r_1 + r_2$ telles valeurs absolues distinctes.

D'autre part, l'anneau des entiers $\mathcal{O}_K$ de K étant un anneau de Dedekind, on peut généraliser la notion de valeur absolue non archimédienne. Étant donné un idéal premier $\mathfrak{p}$ de $\mathcal{O}_K$, on peut définir la notion de valuation en $\mathfrak{p}$ pour les éléments de K par

$$\forall\, x \in K, \qquad x\mathcal{O}_K = \prod_{\mathfrak{p}} \mathfrak{p}^{v_\mathfrak{p}(x)}.$$

D'autre part, tout idéal premier $\mathfrak{p}$ de K vit au dessus d'un nombre premier p, et si l'on note alors $e_\mathfrak{p} = v_\mathfrak{p}(p)$ (c'est l'indice de ramification de p en $\mathfrak{p}$), on peut définir la valeur absolue $\mathfrak{p}$-adique par

$$\forall x \in K, \qquad |x|_\mathfrak{p} = p^{-v_\mathfrak{p}(x)/e_\mathfrak{p}}.$$

On note encore $\mathcal{M}_K$ l'ensemble de ces valeurs absolues (avec bien sur $\mathcal{M}_K^\infty$ les valeurs absolues archimédiennes et $\mathcal{M}_K^0$ les non archimédiennes). Les éléments de $\mathcal{M}_K$ sont les valeurs absolues qui prolongent les valeurs absolues standards de $\mathbb{Q}$.

3.2.1 Complétion

Sur K, on a encore la formule du produit mais il faut pondérer les valeurs absolues de $\mathcal{M}_K$. Pour cela, étant donnée une place $v \in \mathcal{M}_K$, on on considère le complété K_v de K pour la métrique induite pas v. Si v est une place archimédienne, K_v est égale à $\mathbb{R}$ ou $\mathcal{C}$ selon que v soit issue d'un plongement réel ou complexe. Si v est non-archimédienne, elle est issue d'un idéal premier $\mathfrak{p}$ de $\mathcal{O}_K$, et l'on peut dire plus de choses.

Proposition 3.3 *Soit R un système de représentant de $\mathcal{O}_K/\mathfrak{p}$ dans $\mathcal{O}_K$ et soit $(\pi_m)_{m\in\mathbb{Z}}$ une suite d'éléments de $\mathcal{O}_K$ telle que pour tout $m \in \mathbb{Z}$ on ait $\pi_m \in \mathfrak{p}^m \setminus \mathfrak{p}^{m+1}$ (autrement dit, π_m est exactement de valuation m). Alors tout élément $x \in K_v$ s'écrit de façon unique sous la forme d'une série de Laurent*

$$x = \sum_{m \geqslant v_\mathfrak{p}(x)} a_m \pi_m$$

avec $a_m \in R$ et $a_{v_\mathfrak{p}(m)} \neq 0$.

Rq : en pratique, on choisit un élément $\pi \in \mathfrak{p} \setminus \mathfrak{p}^2$ (un tel élément est appelé uniformisante) et l'on prend pour tout $m \in \mathbb{Z}$, $\pi_m = \pi^m$.

Corollary 3.4 *Soient K un corps de nombres et L une extension finie de K. Soient également $\mathfrak{P}$ un idéal premier de $\mathcal{O}_L$ et $\mathfrak{p}$ l'idéal de $\mathcal{O}_K$ qui vit sous $\mathfrak{P}$. En notant e l'indice de ramification de $\mathfrak{p}$ en $\mathfrak{P}$ et f le degré résiduel $[\mathcal{O}_L/\mathfrak{P} : \mathcal{O}_K/\mathfrak{p}]$, on a*

$$[L_\mathfrak{P} : K_\mathfrak{p}] = ef.$$

Preuve : Soit Π un élément de $\mathfrak{P} \setminus \mathfrak{P}^2$ et soit $\pi \in \mathfrak{p} \setminus \mathfrak{p}^2$. Par construction, on a $v_\mathfrak{P}(\Pi) = v_\mathfrak{p}(\pi) = 1$ (un tel élément est appelé *uniformisante*) et en notant

$$\Pi_m = \pi^{\lfloor m/e \rfloor} \Pi^{m - e\lfloor m/e \rfloor},$$

on a $v_\mathfrak{P}(\Pi_m) = m$ pour tout $m \in \mathbb{Z}$, et d'après la proposition précédente, tout élément $x \in L_\mathfrak{P}$ s'écrit de façon unique sous la forme

$$x = \sum_{m \geqslant v_\mathfrak{P}(x)} a_m \Pi_m \tag{1}$$

où les a_m sont pris dans un système de représentant de $\mathcal{O}_L/\mathfrak{P}$ dans $\mathcal{O}_L$. Mais $\mathcal{O}_L/\mathfrak{P}$ est un $\mathcal{O}_K/\mathfrak{p}$-ev de dimension f, donc si $\{\overline{\varepsilon_1}, \dots, \overline{\varepsilon_f}\}$, $\varepsilon_i \in \mathcal{O}_L$ en est une base, chacun des a_m peut s'écrire

$$a_m = \sum_{j=1}^{f} a_{j,m}\varepsilon_j$$

où les $a_{j,m}$ sont pris dans un système de représentants de $\mathcal{O}_K/\mathfrak{p}$ dans $\mathcal{O}_K$. En injectant tout ça dans (1), la famille $\{\Pi^i\varepsilon_j,\ 0 \leqslant i \leqslant e-1,\ 1 \leqslant j \leqslant f\}$ apparaît comme une $K_\mathfrak{p}$-base de $L_\mathfrak{P}$. □

Avant de montrer la formule du produit dans le cas général, donnons un corollaire du corollaire qui nous servira plus tard

Corollary 3.5 *Soit L une extension finie de K de degré n. Alors pour tout $v \in \mathcal{M}_K$,*

$$\sum_{w \in \mathcal{M}_L,\ w|v} [L_w : K_v] = n.$$

Preuve : Si w est non archimédienne, c'est immédiat avec le corollaire précédent et le théorème 2.3.

Supposons donc que v soit archimédienne. Il existe un plongement $\sigma : K \hookrightarrow \mathbb{C}$ qui définit v. Or, si σ est réel (i.e. $K_v = \mathbb{R}$), il existe $n = r_1 + 2r_2$ plongements de $L \hookrightarrow \mathbb{C}$ qui prolongent σ, dont r_1 plongements réels et r_2 couples de plongements complexes conjugués. Ainsi,

$$\sum_{w \in \mathcal{M}_L,\ w|v} [L_w : K_v] = \sum_{i=1}^{r_1} [\mathbb{R} : \mathbb{R}] + \sum_{i=1}^{r_2} [\mathbb{C} : \mathbb{R}] = n.$$

Par ailleurs, si σ est un plongement complexe (i.e. $K_v = \mathbb{C}$), il existe exactement n plongements indépendants (i.e. non conjugués) de $L \hookrightarrow \mathbb{C}$ qui prolongent σ, et chacun donne une place $w \in \mathcal{M}_L$ telle que $w \mid v$. Ainsi,

$$\sum_{w \in \mathcal{M}_L,\ w|v} [L_w : K_v] = \sum_{i=1}^{n} [\mathbb{C} : \mathbb{C}] = n.$$

□

3.2.2 Formule du produit

On peut maintenant donner la formule du produit dans un corps de nombres K quelconque :

$$\forall x \in K^*, \qquad \prod_{v \in \mathcal{M}_K} |x|_v^{n_v} = 1.$$

Preuve : Encore une fois, on sépare les cas archimédien et non archimédien. Ainsi, pour $x \in K^*$,

$$\begin{aligned}\prod_{v\in\mathcal{M}_K^\infty} |x|_v^{n_v} &= \prod_{\sigma \text{ réel}} |\sigma(x)|_\infty \times \prod_{\sigma \text{ complexe}} |\sigma(x)|_\infty^2 \\ &= \prod_{\sigma} |\sigma(x)|_\infty \\ &= |\mathcal{N}_{K/\mathbb{Q}}(x)|_\infty.\end{aligned}$$

D'autre part, pour tout idéal premier $\mathfrak{p} \subset \mathcal{O}_K$, si l'on note p le nombre premier qui vit sous $\mathfrak{p}$, on a $n_\mathfrak{p} = e_\mathfrak{p} f_\mathfrak{p}$ où $f_\mathfrak{p}$ est le degré du corps résiduel : $f_\mathfrak{p} = [\mathcal{O}_K/\mathfrak{p} : \mathbb{Z}/p\mathbb{Z}]$. Ainsi,

$$\prod_{v\in\mathcal{M}_K^0} |x|_v^{n_v} = \prod_p \prod_{\mathfrak{p}|p} (p^{f_\mathfrak{p}})^{-v_\mathfrak{p}(x)}.$$

Or $p^{f_\mathfrak{p}}$ est la norme de l'idéal $\mathfrak{p}$. Ainsi,

$$\begin{aligned}\prod_{v\in\mathcal{M}_K^0} |x|_v^{n_v} &= \prod_p \prod_{\mathfrak{p}|p} (\mathcal{N}_{K/\mathbb{Q}}\mathfrak{p})^{-v_\mathfrak{p}(x)} \\ &= \left| \mathcal{N}_{K/\mathbb{Q}} \left(\prod_\mathfrak{p} \mathfrak{p}^{-v_\mathfrak{p}(x)} \right) \right|_\infty \\ &= |\mathcal{N}_{K/\mathbb{Q}}(x)|_\infty^{-1}\end{aligned}$$

□

Rq : l'anneau des entiers de K peut être défini à partir des valeurs absolues standards de K :

$$\mathcal{O}_K = \{x \in K,\ |x|_v \leqslant 1\ \forall v \in \mathcal{M}_K^0\}.$$

De même, si $S \subset \mathcal{M}_K$ est un ensemble fini contenant les places archimédiennes, l'anneau des S-entiers de K est

$$\mathcal{O}_{S,K} = \{x \in K,\ |x|_v \leqslant 1\ \forall v \in \mathcal{M}_K,\ v \notin S\}.$$

4 Hauteurs d'un nombre algébrique

Definition 4.1 *Soit K un corps de nombres et soit $\mathcal{M}_K$ l'ensemble de ses valeurs absolues standards. La hauteur relative (i.e. sur K) d'un nombre $\alpha \in K$ est*

$$H_K(\alpha) = \prod_{v\in\mathcal{M}_K} \max\{1, |\alpha|_v^{n_v}\}.$$

D'autre part, pour des raisons pratiques, on préfère souvent travailler dans un univers additif. On définit donc également la hauteur logarithmique :

$$h_K(\alpha) = \log H_K(\alpha) = \sum_{v\in\mathcal{M}_K} n_v \log \max\{1, |\alpha|_v\}.$$

Rq : pour $\alpha = a/b \in \mathbb{Q}$, on retrouve bien $H_\mathbb{Q}(a/b) = \max\{|a|, |b|\}$.

4.1 Hauteur absolue

Ainsi définie, la hauteur d'un nombre algébrique dépend du corps de nombres dans lequel on travaille. Précisément,

Proposition 4.2 *Soit L une extension finie de K. alors pour tout $\alpha \in K$, on a*

$$h_L(\alpha) = [L : K] h_K(\alpha).$$

Preuve :

$$\begin{aligned} h_L(\alpha) &= \sum_{w \in \mathcal{M}_L} n_w \log \max\{1, |\alpha|_w\} \\ &= \sum_{v \in K} \sum_{w \in \mathcal{M}_L,\, w|v} n_w \log \max\{1, |\alpha|_v\} \end{aligned}$$

Mais d'après la formule des degrés, on a $n_w = [L_w : \mathbb{Q}_w] = [L_w : K_w] n_v$ pour tout $w \mid v$. Ainsi,

$$\begin{aligned} h_L(\alpha) &= \sum_{v \in K} n_v \log \max\{1, |\alpha|_v\} \sum_{w \in \mathcal{M}_L,\, w|v} [L_w : K_w] \\ &= [L : K] h_K(\alpha) \qquad \text{d'après le corollaire 3.5.} \end{aligned}$$

□

Autrement dit, si L et K sont deux corps de nombres, pour tout $\alpha \in L \cap K$, on a

$$\frac{h_L(\alpha)}{[L : \mathbb{Q}]} = \frac{h_K(\alpha)}{[K : \mathbb{Q}]}.$$

Cela nous permet donc de définir une hauteur *absolue* (i.e. qui ne dépend pas du corps sur lequel on se place) :

Definition 4.3 *Soit K un corps de nombres. Pour tout $\alpha \in K$, on note*

$$h(\alpha) = \frac{1}{[K : \mathbb{Q}]} h_K(\alpha) = \sum_{v \in \mathcal{M}_K} \frac{[K_v : \mathbb{Q}_v]}{[K : \mathbb{Q}]} \log \max\{1, |\alpha|_v\}.$$

4.2 Propriétés élémentaire

Proposition 4.4 Règles de Calcul

Soient K un corps de nombres et $\alpha_1, \ldots, \alpha_r \in K$.

1. $h(\alpha_1 \cdots \alpha_r) \leqslant h(\alpha_1) + \ldots + h(\alpha_r)$.
2. $h(\alpha_1 + \ldots + \alpha_r) \leqslant h(\alpha_1) + \ldots + h(\alpha_r) + \log r$

Preuve :

1. Cela vient de la multiplicativité des valeurs absolues (combiné avec le log). À noter que l'inégalité vient du max. On ne peut pas espérer mieux.
2. Cela vient des inégalités triangulaire et ultramétrique. Encore une fois, le $\log r$ vient du max que l'on a dans la définition de la hauteur.

□

Proposition 4.5 Invariance
La notion de hauteur absolue est invariante par
– *changement de corps,*
– *changement de signe ($h(-\alpha) = h(\alpha)$),*
– *passage à l'inverse ($h(1/\alpha) = h(\alpha), \quad \forall \alpha \in K^*$),*
– *par l'action du groupe de Galois sur $\overline{\mathbb{Q}}$.*

Preuve :
– Pour le changement de corps, tout a été fait pour.
– Pour le changement de signe, c'est clair.
– Pour le passage à l'inverse, tout repose sur la formule du produit. En effet, celle-ci peut se lire

$$\prod_{v/|\alpha|_v \geqslant 1} |\alpha|_v = \left(\prod_{v/|\alpha|_v \leqslant 1} |\alpha|_v \right)^{-1} = \prod_{v/|\alpha^{-1}|_v \geqslant 1} |\alpha^{-1}|_v$$

– Pour l'action du groupe de Galois, cela vient du fait que tout automorphisme de Galois effectue une permutation des places. Comme elles sont prises dans leur ensemble, cela ne change rien.

□

4.3 Théorème de Northcott

Theorem 4.6 *Soient $B, D > 0$ deux bornes. Alors l'ensemble*

$$\{\alpha \in \overline{\mathbb{Q}} \,/\, [\mathbb{Q}(\alpha) : \mathbb{Q}] \leqslant D \text{ et } h(\alpha) \leqslant B\}$$

est fini. En particulier, sur un corps de nombres K fixé, l'ensemble

$$\{\alpha \in K \,/\, h(\alpha) \leqslant B\}$$

est fini.

Preuve : Tout d'abord, montrons que cela marche sur $\mathbb{Q}$. Soit donc $a/b \in \mathbb{Q}$ sous sa forme irréductible. On a

$$h(a/b) = \sum_{v \in \mathcal{M}_{\mathbb{Q}}} \log \max\{1, |a/b|_v\} = \log \max\{1, |a/b|_\infty\} + \sum_p \log \max\{1, |a/b|_p\}.$$

Or les seuls places archimédiennes v telles que $|a/b|_v > 1$ correspondent aux nombres premiers qui divisent b. Autrement dit,

$$\sum_p \log \max\{1, |a/b|_p\} = \log |b|_\infty.$$

Mais alors, soit $|a/b| < 1$ et $h(a/b) = \log |b|$, soit $|a/b| > 1$ et $h(a/b) = \log |b| + \log |a/b| = \log |a|$. Dans tous les cas, $h(a/b) = \log \max\{|a|, |b|\}$ et il n'y a qu'un nombre fini de rationnels qui sont de hauteur bornée.

Cela va nous servir pour montrer le cas général. En effet, soient $B, D > 0$ deux bornes et soit $\alpha \in \overline{\mathbb{Q}}$ tel que $[\mathbb{Q}(\alpha) : \mathbb{Q}] \leqslant D$ et $h(\alpha) \leqslant B$. Le polynôme minimal P_α de α est de degré au plus D, et ses coefficients (rationnels) sont tous fonctions symétriques de α et ses conjugués. Or la hauteur est invariante par l'action du groupe de Galois, donc tous les conjugués de α ont la même hauteur (majorée par B) et grâce aux règles de calcul sur les hauteurs, on peut majorer la hauteur de chacun des D coefficients de P_α en fonction de B. Comme il n'existe qu'un nombre fini de rationnels de hauteur bornée, il n'existe qu'un nombre fini de polynômes minimaux, donc qu'un nombre fini de racines. Autrement dit, il ne peut y avoir qu'un nombre fini de tels α. □.

5 Hauteurs et polynômes

Résoudre une équation diophantienne sur $\overline{\mathbb{Q}}$, c'est donc borner la hauteur des solutions algébriques. Sur $\mathbb{Z}$, on borne la valeur absolue des solutions en fonction des valeurs absolues des coefficients (entiers) du polynôme qui nous donne l'équation. Pour pouvoir généraliser cela à $\overline{\mathbb{Q}}$, il faut nous faut donc une notion de hauteur pour les polynômes à coefficients algébriques. Pour cela, on passe par l'espace projectif $\mathbb{P}^r(\overline{\mathbb{Q}})$. En effet, dans tout ce que l'on vient de faire, on peut remplacer K par $\mathbb{P}^r(K)$. Ainsi, pour un point $P = (\alpha_0, \ldots, \alpha_r) \in \mathbb{P}^r(K)$, on note

$$h(P) = \sum_{v \in \mathcal{M}_K} \frac{[K_v : \mathbb{Q}_v]}{[K : \mathbb{Q}]} \log \max\{|\alpha_0|_v, \ldots, |\alpha_r|_v\}$$

La formule du produit nous assure que cette fonction est bien définie sur $\mathbb{P}^r(K)$, et la hauteur d'un nombre algébrique α est la hauteur du point $(1, \alpha) \in \mathbb{P}^1(K)$. D'autre part, on a les mêmes règles de calcul, propriétés d'invariance, et bien sur le même résultat de finitude.

Par ailleurs, cette notion de hauteur sur l'espace projectif nous permet de définir la hauteur d'un polynôme à coefficients algébriques. En effet, étant donné un polynôme F à coefficients algébriques, il existe un corps de nombres K contenant tous ses coefficients, et l'on définit la hauteur de F comme étant la hauteur du vecteur de l'espace projectif dont les coordonnées sont les coefficients de F. On a bien sûr des règles de calcul pour la hauteur de la somme ou de produits de polynômes, ainsi que des propriétés propres aux polynômes. Parmi ces dernières, on peut en citer deux particulièrement utiles :

– Étant donnés deux polynômes F et G dans $\overline{\mathbb{Q}}[X_1, \ldots, X_n]$, on peut borner la hauteur du *résultant* $\mathrm{Res}_{X_i}(F, G)$ pour tout indice i. Cela nous permet d'utiliser la théorie de l'élimination.
– Comme on sait majorer la hauteur du produit de deux polynômes en fonction des hauteurs des facteurs, en remarquant que la hauteur d'un nombre algébrique α n'est rien d'autre que la hauteur du polynôme $X - \alpha$, on obtient rapidement une majoration de la hauteur d'un nombre en fonction de celle d'un polynôme qu'il annule.

Une première application de ces deux propriétés est la suivante. Étant donnée deux équations algébriques en 2 variables, multiplicativement indépendantes. Les solutions communes aux deux équations sont en nombre fini, et leur hauteur est bornée de façon explicite.

6 Hauteur et pgcd

Une classe importante de problèmes diophantiens est l'étude des solutions entières des équations du type $F(x, y) = 0$ à pgcd fixé (ici, F est un polynôme à coefficients entiers). Or la notion de hauteur pour les nombres algébriques nous permet d'avoir une généralisation de la notion de pgcd à $\overline{\mathbb{Q}}$, et l'on peut ainsi généraliser les résultats obtenus sur les entiers.

Pour cela, il nous faut une notion de *hauteur locale*. Toujours grâce à la formule du produit, on a

$$\begin{aligned} h(\alpha) &= \sum_{v \in \mathcal{M}_K} \frac{[K_v : \mathbb{Q}_v]}{[K : \mathbb{Q}]} \log \max\{1, |\alpha|_v\} \\ &= -\sum_{v \in \mathcal{M}_K} \frac{[K_v : \mathbb{Q}_v]}{[K : \mathbb{Q}]} \log \min\{1, |\alpha|_v\} \end{aligned}$$

Ainsi, pour toute place $v \in \mathcal{M}_K$, et pour tout $\alpha \in K$, on note

$$h_v(\alpha) = -\frac{[K_v : \mathbb{Q}_v]}{[K : \mathbb{Q}]} \log \min\{1, |\alpha|_v\}$$

de telle sorte que

$$h(\alpha) = \sum_{v \in \mathcal{M}_K} h_v(\alpha).$$

Definition 6.1 *Soient K un corps de nombres et α et β dans K. Le pgcd logarithmique de α et β est donné par*

$$\mathrm{lgcd}(\alpha, \beta) = \sum_{v \in \mathcal{M}_K} \min\{h_v(\alpha), h_v(\beta)\}.$$

On peut vérifier rapidement que si α et β sont dans $\mathbb{Z}$, alors $\mathrm{lgcd}(\alpha, \beta) = \log |\gcd(\alpha, \beta)|$.

7 Application au problème de Skolem

7.1 Une généralisation

Grace à cette notion de pgcd sur les nombres algébriques, on obtient une version explicite et généralisée du résultat de Skolem :

Theorem 7.1 *Soit $F(X,Y) \in \overline{\mathbb{Q}}[X,Y]$ un irréductible tel que $F(0,0) = 0$ et tel que $\partial_Y F(0,0) \neq 0$. Pour toute solution $(\alpha,\beta) \in \overline{\mathbb{Q}}^2$ de l'équation $F(x,y) = 0$ et pour tout $\varepsilon \in]0,1]$, on a soit*

$$h(\alpha) \ll \varepsilon^{-2}(h(F)+1)$$

soit

$$|h(\alpha) - \nu\, lgcd(\alpha,\beta)| \leqslant \varepsilon h(\alpha) + O(\varepsilon^{-1}(h(F)+1)).$$

7.2 Une idée de la preuve

On se place dans un corps de nombres K contenant tous les nombres algébriques de notre problème (il n'y en a qu'un nombre fini...). D'après les hypothèses, il existe une série formelle $\mathcal{Y}(X) = \sum_{k\geqslant 1} a_k X^k$ à coefficients dans K telle que $F(X,\mathcal{Y}(X)) = 0$. D'autre part, un théorème d'Eisenstein nous permet de contrôler les coefficients de $\mathcal{Y}$: il existe des réels $A_v \geqslant 1$, $v \in \mathcal{M}_K$ presque tous égaux à 1 tels que

$$|a_k|_v \leqslant A_v^k, \qquad v \in \mathcal{M}_K,\ k \geqslant 1,$$

et tels que

$$\sum_{v\in\mathcal{M}_K} \frac{[K_v:\mathbb{Q}_v]}{[K:\mathbb{Q}]} \log A_v \ll h(F)+1.$$

Grace à ca, on peut isoler les valuations de K qui vont être significatives :

$$S = \left\{ v \in \mathcal{M}_K \ : \ |\alpha|_v \begin{array}{c} \leqslant \\ < \end{array} \begin{array}{c} (2A_v)^{-1} \\ A_v^{-1} \end{array} \begin{array}{l} \text{si } v \text{ est infinie,} \\ \text{sinon.} \end{array} \right\}$$

Comme la "hauteur" des A_v est bornée, c'est principalement les valuations de S qui vont compter dans la hauteur de α. Autrement dit

$$h(\alpha) \approx h_S(\alpha)$$

où h_S est la somme sur les valautions de S. Par ailleurs, comme les A_v majorent les coefficients de $\mathcal{Y}$, pour toute place $v \in S$, $\mathcal{Y}$ v-converge pour $X = \alpha$. (On note $\mathcal{Y}_v(\alpha)$ sa somme). Comme $F(X,\mathcal{Y}(X)) = 0$, les $\mathcal{Y}_v(\alpha)$ sont des racines de $F(\alpha,Y)$ et l'on peut isoler les $v \in S$ pour lesquels cette série converge vers β. Ainsi

$$T = \left\{ v \in S \ : \ \mathcal{Y}_v(\alpha) = \sum_k a_k \alpha^k = \beta \right\}.$$

Dans T, on a les valuations pour lesquelles à la fois α et β sont petits. Autrement dit, on a

$$h_T(\alpha) \approx \mathrm{lgcd}(\alpha, \beta).$$

Il nous reste donc à relier $h_S(\alpha)$ et $h_T(\alpha)$. Pour cela, on construit un polynôme $G(X, Y)$ tel que

$$\begin{array}{l} \deg_X(G) = N,\ \deg_Y(G) < \deg_Y(F) \\ \mathrm{ord}_x G(x, Y(x)) \geqslant nN(1-\varepsilon), \\ h(G) \ll \varepsilon^{-1} N(h(F)+1) \end{array}$$

et l'on étudie $\gamma = G(\alpha, \beta)$. Si $\gamma = 0$, on applique le résultant et $h(\alpha)$ est bornée. Sinon, on borne $h(\gamma)$ en fonction $nh_T(\alpha)$ et $h_S(\alpha)$, ce qui nous permet de les relier et d'obtenir l'autre inégalité.

7.3 Quelques remarques finales

– En optimisant le paramètre ε, on obtient

$$h(\alpha) \ll h(F) + 1$$

ou

$$|h(\alpha) - \nu \mathrm{lgcd}(\alpha, \beta)| \ll \sqrt{h(\alpha)(h(F)+1)}.$$

– Ce théorème nous donne une version explicite d'un résultat connu : la quasi-équivalence des hauteurs : si (α, β) annule un polynôme F, de degrés m en X et n en Y, alors $h(\alpha)/n$ est approximativement égale à $h(\beta)/m$, ce qui nous permet entre autres d'obtenir une version symétrique du théorème principal.

Références

[1] M. Abouzaid, *Aspects effectifs d'analyse diophantienne*, thèses, Université Bordeaux 1 (2006).

[2] H. Cohen, *A course in algebraic number theory*, en préparation.

[3] M. Hindry, J. Silverman, *Diophantine Geometry, an introduction*, GTM **201**, Springer. (2000).

[4] S. Lang, *Algebraic Number Theory*, GTM **110**, Springer, Second Edition.

[5] P. Samuel, *Théorie algébrique des nombres*, Hermann, Seconde Edition (1971).