Quelques théorèmes classiques en Approximation diophantienne

Yamin Ould Ahmed

Mai-Juillet 2026

Ce stage de 3ième année de Licence ENS a été réalisé à l'Institut Fourier à Grenoble sous la direction de Tanguy Rivoal.

Résumé

Notre objectif est d'étudier quelques résultats importants de la première moitié du XX$^{\text{ème}}$ siècle en Approximation diophantienne. Parmi les plus célèbres nous avons le théorème de Gelfond–Schneider qui constitue le chapitre 3. Pour ensuite le généraliser avec le critère de Schneider–Lang qui occupe le chapitre 4. Mais avant nous étudierons un lemme arithmétique dû à Siegel. Enfin, nous verrons dans le chapitre 5 un résultat d'approximation de réels algébriques par des rationnels dû à Thue. Un théorème qui a reçu des améliorations par la suite, dont une due à une Siegel que nous étudierons dans le dernier chapitre.

Table des matières

1 Introduction

1.1 Rappels d'analyse complexe

Nous allons énoncer quelques notions classiques d'analyse complexe que nous utiliserons plus tard. Nous rappelons qu'on dit qu'une fonction f est *holomorphe* sur un ouvert $\mathcal{U}$ de $\mathbb{C}$ si f est dérivable au sens complexe en tout point de $\mathcal{U}$. Dans le cas où $\mathcal{U} = \mathbb{C}$ on dit que f est entière sur $\mathbb{C}$. Ensuite, on dit qu'une fonction f est *méromorphe* sur $\mathcal{U}$ s'il existe une partie discrète fermée D de $\mathcal{U}$ telle que f soit holomorphe sur $\mathcal{U} \setminus D$ et admet un pôle en tout point de D. On peut montrer qu'une fonction méromorphe sur $\mathcal{U}$ s'écrit comme le quotient de deux fonctions holomorphes sur $\mathcal{U}$. Enfin nous rappelons le théorème et la formule de Cauchy ainsi que le principe du maximum.

Théorème 1. (Cauchy) *Soit $\mathcal{U}$ un ouvert de $\mathbb{C}$ et $f : \mathcal{U} \to \mathbb{C}$. Les conditions suivantes sont équivalentes :*

1. *f est holomorphe sur $\mathcal{U}$;*
2. *f est analytique sur $\mathcal{U}$.*

Plus précisemment, lorsque ces conditions sont réalisées, pour tout disque ouvert $D(z_0, R) \subset \mathcal{U}$, la série de Taylor

$$\sum_{n=0}^{+\infty} \frac{f^{(n)}(z_0)}{n!}(z - z_0)^n$$

converge vers $f(z)$ pour tout $z \in D(z_0, R)$.

Théorème 2. (Formule de Cauchy)

Soit $\mathcal{U}$ un ouvert de $\mathbb{C}$ et f une fonction holomorphe sur $\mathcal{U}$. Soit $z_0 \in \mathcal{U}$ et $R \in \mathbb{R}_+^*$ tels que $\overline{D(z_0, R)} \subset \mathcal{U}$. Alors pour tout $z \in D(z_0, R)$ on a

$$f(z) = \frac{1}{2i\pi} \int_{C(z_0,R)} \frac{f(w)}{w - z} dw.$$

Théorème 3. (Principe du maximum)

Soit $\mathcal{U}$ un ouvert connexe de $\mathbb{C}$ et f une fonction holomorphe sur $\mathcal{U}$. Si $|f|$ admet un maximum local en $z_0 \in \mathcal{U}$, alors f est constante sur $\mathcal{U}$. En particulier si K est un compact inclu dans $\mathcal{U}$, alors

$$\max_{z \in K} |f(z)| = \max_{z \in \partial K} |f(z)|$$

où ∂K est la frontière de K.

1.2 Rappels algébriques et notations

Pour la suite, $\overline{\mathbb{Q}}$ désignera le sous corps de $\mathbb{C}$ des nombres complexes algébriques sur $\mathbb{Q}$. Nous rappelons qu'un nombre complexe x est dit *algébrique* sur $\mathbb{Q}$ s'il est racine d'un polynôme à coefficients rationnels, sinon on dit qu'il est *transcendant* sur $\mathbb{Q}$. Nous notons alors P_x le polynôme unitaire de degré minimal à coefficients rationnels qui annule x. On dit que le *degré de x* et les *conjugués de x* correspondent respectivement au degré de P_x et aux racines de P_x. De plus si son polynôme minimal est à coefficients dans $\mathbb{Z}$, alors nous disons que x est un *entier algébrique* (ou juste entier). Enfin nous désignons par *corps de nombres* toute extension de corps $\mathbb{K}$ de $\mathbb{Q}$ qui est un $\mathbb{Q}$-espace vectoriel de dimension finie (toutes les extensions de corps seront vues comme des inclusions de corps). Dans ce cas nous notons par $\mathcal{O}_{\mathbb{K}}$ l'ensemble des éléments de $\mathbb{K}$ qui sont des entiers algébriques. Maintenant nous allons rappeler quelques propriétés sur les entiers algébriques.

Propriété 1. *Soit $\mathbb{K}$ un corps de nombres. Alors :*

1. *$\mathcal{O}_{\mathbb{K}}$ est un sous anneau de $\mathbb{K}$;*
2. *pour $x \in \mathbb{K}$, il existe $d \in \mathbb{N}^*$ tel que $dx \in \mathcal{O}_{\mathbb{K}}$.*

Ainsi cette propriété met en évidence que les entiers algébriques de $\mathbb{K}$ se comportent comme les entiers relatifs sur $\mathbb{Q}$. De plus ceci nous amène naturellement à considérer la définition suivante.

Définition 1. *Soit $x \in \overline{\mathbb{Q}}$. On définit le* dénominateur *de x comme étant le plus petit entier naturel non nul d tel que dx soit un entier algébrique. On note cet entier* $\mathrm{den}(x)$.

Enfin, nous avons également le théorème de structure pour l'anneau des entiers algébriques d'un corps de nombre suivant (voir [5] pour la démonstration).

Théorème 4. *Soit $\mathbb{K}$ un corps de nombre de dimension h. Il existe $b_1, ..., b_h$ des entiers algébriques de $\mathbb{K}$ $\mathbb{Z}$-libres tels que :*

$$\mathcal{O}_{\mathbb{K}} = \mathbb{Z}b_1 + \mathbb{Z}b_2 + ... + \mathbb{Z}b_h.$$

Par ailleurs lors des démonstrations nous devrons régulièrement majorer $|x|$ pour $x \in \overline{\mathbb{Q}}$, mais aussi la valeur absolue de ses conjugués. Dans cette optique, nous introduisons la notion suivante.

Définition 2. *Si $x \in \overline{\mathbb{Q}}$, en notant $x_1 = x$, $x_2, ..., x_d$ ses conjugués, on définit* la maison *de x notée $\boxed{x}$ comme :*

$$\boxed{x} = \max_{1 \leq i \leq d} |x_i|$$

avec $|.|$ la valeur absolue canonique sur $\mathbb{C}$.

Afin de pouvoir la manipuler facilement dans les inégalités, nous donnons les propriétés suivantes de la maison.

Propriété 2. *Soit $x, y \in \overline{\mathbb{Q}}$. On a les assertions suivantes :*

1. $\boxed{xy} \leq \boxed{x}\boxed{y}$;
2. $\boxed{x+y} \leq \boxed{x} + \boxed{y}$;
3. *Si $x \in \mathbb{Q}$ alors $\boxed{x} = |x|$.*

Démonstration. Cela vient du fait que si $P_x(X) = \prod_{k=1}^{n}(X - x_k)$ et $P_y(X) = \prod_{l=1}^{m}(X - y_l)$, (avec $x_1 = x$ et $y_1 = y$), alors les polynômes

$$P(X) = \prod_{k=1}^{n}\prod_{l=1}^{m}(X - x_k y_l) \quad et \quad Q(X) = \prod_{k=1}^{n}\prod_{l=1}^{m}(X - (x_k + y_l))$$

sont, par un corollaire du théorème des polynômes symétriques, des polynômes à coefficients rationnels. Ainsi $P_{xy}(X)$ divise $P(X)$ et $P_{x+y}(X)$ divise $Q(X)$. On en déduit alors que toute racine de $P_{xy}(X)$ s'écrit comme le produit d'un conjugué de x et d'un conjugué de y, qui est donc inférieure en module à $\boxed{x}\,\boxed{y}$. De même, toute racine de $P_{x+y,\mathbb{Q}}(X)$ s'écrit comme la somme d'un conjugué de x et d'un conjugué y, qui est donc inférieur en module à $\boxed{x}+\boxed{y}$.

Enfin si $x \in \mathbb{Q}$, alors $P_x(X) = X - x$. Donc x n'a pas d'autres conjugués que lui même d'où $\boxed{x} = |x|$. □

Définition 3. *Soit $x \in \overline{\mathbb{Q}}$, on note $x_1 = x$, $x_2, ..., x_d$ ses conjugués sur $\mathbb{Q}$. On définit alors :*

1. la norme *de x notée $\mathcal{N}(x)$ par*

$$\mathcal{N}(x) = x_1 x_2 ..., x_d;$$

2. la trace *de x notée* $\mathrm{Tr}(x)$ *par*

$$\mathrm{Tr}(x) = x_1 + x_2 + ... + x_d.$$

Remarque 1. *En notant $P_x(X) = X^d + a_{d-1}X^{d-1} + ... + a_0$, les formules de Viète nous donnent les égalités suivantes :*

$$\mathcal{N}(x) = (-1)^d a_0 \quad et \quad \mathrm{Tr}(x) = -a_{d-1}.$$

En particulier, si x est un entier algébrique non nul alors P_x est un polynôme à coefficients entiers qui n'annule pas 0, d'où $|\mathcal{N}(x)| \geq 1$. Ceci sera important pour la suite, car les démonstrations que nous ferons reposeront sur un raisonnement par l'absurde où on trouve un entier algébrique non nul dont on prouve que le module de sa norme est strictement inférieur à 1.

2 Le lemme de Siegel

Pour montrer un critère de transcendance, les démonstrations reposent généralement sur la construction de fonctions de auxiliaires qui sont astreintes à satisfaire un certain nombre de conditions (devant s'annuler en plusieurs points par exemple). Puis en prenant l'image de la fonction en un point bien choisi, nous pourrons ainsi formuler notre contradiction. Cependant, si Hermite pour démontrer le théorème d'Hermite-Lindemann (avec $\alpha = 1$ trouva des formules explicites, le lemme de Siegel nous permet d'éviter cette étape et de construire la fonction auxiliaire avec les conditions que nous voulons sans avoir besoin d'en donner l'expression. Cette partie prend sa source dans le $2^{\text{ème}}$ paragraphe du chapitre 2 de [11].

2.1 Le cas $\mathbb{K} = \mathbb{Q}$

Nous commençons par démontrer le cas dans lequel $\mathbb{K}$ est le corps le plus simple à savoir $\mathbb{Q}$. Puis, nous utiliserons le théorème de structure de l'anneau des entiers algébriques d'un corps de nombres.

Lemme 1. *Soit $A \in \mathbb{N}$ et $M \in \mathcal{M}_{p,q}(\mathbb{Q})$ avec $0 < p < q$. Si pour $(i,j) \in \{1,2,3,...,p\} \times \{1,2,3,...,q\}$ on a $|M_{i,j}| \leq A$, alors il existe $X = (x_1, x_2, ..., x_q)^T \in \mathbb{Z}^q \setminus \{0\}$ tel que*

1. $MX = 0$;
2. $\forall i \in \{1,2,...,q\}, \quad |x_i| < 1 + (qA)^{\frac{p}{q-p}}$.

Démonstration. Prenons $M \in \mathcal{M}_{p,q}(\mathbb{Q})$ qui respecte les hypothèses du lemme. Comme pour $d \in \mathbb{N}^*$ on a $MX = 0 \iff dMX = 0$, quitte à remplacer M par dM, on peut supposer que $M \in \mathcal{M}_{p,q}(\mathbb{Z})$.

Soit $H \in \mathbb{N}$ et $X \in \{0, \pm 1, \pm 2, ..., \pm H\}^q$. Si $Y = MX$, alors

$$|y_k| \leq \sum_{l=1}^{q} |M_{k,l} x_l| \leq HAq.$$

Ensuite M étant à coefficients entiers, on en déduit donc que Y aussi et plus précisément $Y \in \{0, \pm 1, \pm 2, ..., \pm qAH\}$. Ainsi si on a H tel que $(2qAH+1)^p < (2H+1)^q$, par le principe des tiroirs il existe $X, X' \in \{0, \pm 1, \pm 2, ..., \pm H\}$ distincts tels que $M(X - X') = 0$ et donc chaque coordonnée de $X - X'$ est en module inférieure à $2H$ et $X - X' \neq 0$.

Aussi posons $2H$ comme l'unique entier pair tel que $(qA)^{\frac{p}{q-p}} - 1 \leq 2H < (qA)^{\frac{p}{q-p}} + 1$ (ce qui est possible car l'intervalle est de longueur 2 et contient l'une de ses extrémités). Comme

$$(2HqA+1)^p < (qA)^p (2H+1)^p \leq (2H+1)^{q-p}(2H+1)^p = (2H+1)^q,$$

par ce qui a été fait précédemment nous aurons $X \in \{0, \pm 1, \pm 2, ..., \pm 2H\}^q$ non nul tel que $MX = 0$. Ceci conclut donc la preuve du lemme. □

2.2 Le cas général

Maintenant, nous allons montrer le cas général où $\mathbb{K}$ est un corps de nombres.

Lemme 2. *Soit $\mathbb{K}$ un corps de nombres de degré h sur $\mathbb{Q}$. Soit $A \in \mathbb{N}$ et $M \in \mathcal{M}_{p,q}(\mathbb{K})$ avec $0 < p < q$. Si pour $(i,j) \in \{1,2,3,...,p\} \times \{1,2,3,...,q\}$ on a $\overline{|M_{i,j}|} \leq A$, alors il existe $X = (x_1, x_2, ..., x_q)^T \in \mathcal{O}_{\mathbb{K}}^q \setminus \{0\}$ tel que*

1. *$MX = 0$;*
2. *$\forall i \in \{1,2,...,q\}, \quad \overline{|x_i|} < c + c(cqA)^{\frac{p}{q-p}}$ (avec $c > 0$ une constante ne dépendant que de $\mathbb{K}$).*

Démonstration. Tout d'abord, on note $b = (b_1, b_2, ..., b_h)$ une $\mathbb{Z}$-base de $\mathcal{O}_{\mathbb{K}}$ qui est donc aussi une $\mathbb{Q}$-base de $\mathbb{K}$.

Étape 1 : Existence de γ_1 tel que $|g_k| \leq \overline{|a|}\gamma_1$.

Soit $a = g_1 b_1 + ... + g_h b_h \in \mathbb{K}$. On note $b^* = (b_1^*, b_2^*, ..., b_h^*)$ une base de $\mathbb{K}$ telle que $\mathrm{Tr}(b_j b_k^*) = \delta_{j,k}$. L'existence d'une telle base vient de l'injectivité de l'application linéaire $x \in \mathbb{K} \mapsto (\mathrm{Tr}(b_1 x), ..., \mathrm{Tr}(b_h x))^T \in \mathbb{Q}^h$. Ainsi par définition on a $g_k = \mathrm{Tr}(a b_k^*)$. Puis $\mathrm{Tr}(a b_k^*)$ étant la somme des conjugués de $a b_k^*$, alors

$$|g_k| \leq h\overline{|ab_k^*|} \leq h\overline{|a|} \max_{1 \leq \ell \leq h} \left(\overline{|b_l^*|}\right).$$

En posant $\gamma_1 := h \max\left(\overline{|b_l^*|}\right)$, on a le résultat et γ_1 ne dépend que de $\mathbb{K}$.

Étape 2 : Se ramener au lemme 1.

On souhaite résoudre sur $\mathcal{O}_{\mathbb{K}}$ le système suivant :

$$\forall k \in \{1,2,...,p\}, \quad M_{k,1}x_1 + ... + M_{k,q}x_q = 0$$

constitué de p équations et d'inconnues $x_1, ..., x_q$. Comme $\mathcal{O}_{\mathbb{K}}$ est engendré par b, résoudre le système précédent revient à résoudre celui-ci sur $\mathbb{Z}$ (en posant $x_k = x_{k,1}b_1 + ... + x_{k,h}b_h$) :

$$\forall k \in \{1,2,...,p\}, \quad \sum_{i=1}^{q} \sum_{j=1}^{h} M_{k,i} b_j x_{k,j} = 0$$

où les inconnues sont les $x_{k,j}$.
Ensuite (comme b est également une $\mathbb{Q}$-base de $\mathbb{K}$) en développant les $M_{k,i}b_j$ dans la base b, on obtient un système de ph équations linéaires à coefficients dans $\mathbb{Q}$ avec qh inconnues. Puis, par ce qu'on a fait à l'étape 1, nous savons que les coefficients sont en module $\leq \gamma_1 \max(\overline{|M_{k,\ell}b_r|})$ (car les coefficients du nouveau système sont tous des projections dans la base b d'un des $M_{k,\ell}b_r$ et on a une majoration par l'étape 1). Puis

$$\overline{|M_{k,\ell}b_r|} \leq \overline{|M_{k,\ell}|}\overline{|b_r|} \leq A \max(\overline{|b_\ell|}).$$

D'où $\gamma_1 \max(\overline{|M_{k,\ell} b_r|}) \leq \gamma_2 A$ avec $\gamma_2 := \gamma_1 \max(\overline{|b_\ell|})$ qui ne dépend que de la base b. Comme $ph < qh$, par le lemme 1, il existe une solution non triviale $(x_{k,\ell})$ à coefficients entiers du système telle que

$$|x_{k,\ell}| < 1 + (\gamma_2 hqA)^{\frac{p}{q-p}}.$$

Étape 3 : Conclusion.

Comme $(x_{k,\ell})$ est une solution non triviale à coefficients dans $\mathbb{Z}$, alors la $\mathbb{Z}$-liberté de b nous indique, qu'en posant pour $k \in \{1, ..., q\}$, $x_k := x_{k,1} b_1 + ... + x_{k,h} b_h$, (x_k) est une solution non triviale à coefficients dans $\mathcal{O}_\mathbb{K}$. De plus, elle est solution du premier système sur $\mathbb{K}$. Finalement,

$$\overline{|x_k|} \leq \sum_{\ell=1}^{h} (\overline{|b_\ell|} |x_{k,\ell}|) < h \max(\overline{|b_\ell|})(1 + (\gamma_2 hqA)^{\frac{p}{q-p}}).$$

Quitte à prendre $\max(\gamma_1, 1)$ on peut supposer que $\gamma_1 \geq 1$ ce qui donne l'inégalité voulue avec $c = \gamma_2 h$ qui ne dépend que de $\mathbb{K}$. □

3 Le théorème de Gelfond–Schneider

Nous commençons par un résultat sur la transcendance d'un certain type de nombre complexe que l'on peut retrouver dans le chapitre III de [11] et dans le premier chapitre II de [9]. La démonstration de ce théorème constituait le septième des vingt-trois problèmes posés par Hilbert au début du XX$^{\text{ème}}$ siècle. Il fut alors démontré indépendamment et simultanément par Gelfond et Schneider en 1934. Dans cette partie nous allons étudier ces deux démonstrations différentes.

3.1 Énoncé et ses applications

Théorème 5. *Soit $\alpha \in \overline{\mathbb{Q}} \setminus \{0\}$ et* log *une détermination du logarithme sur $\mathbb{C}$ définie en α. Si $\log(\alpha) \neq 0$ et β un irrationnel algébrique, alors $\alpha^\beta := e^{\beta \log(\alpha)}$ est transcendant.*

Ce théorème nous permet de démontrer la transcendance des nombres $e^\pi = i^{(-2i)}$ et $2^{\sqrt{2}}$ par exemple.

Prenons deux complexes algébriques $\mathbb{Q}$ non nuls α et β ainsi qu'une détermination du logarithme qui est définie en ces deux points telle que $\log(\alpha) \neq 0$. Nous pouvons remarquer que $\beta = \exp(\log(\beta)) = \exp(\frac{\log(\beta)}{\log(\alpha)} \log(\alpha))$ alors d'après le théorème de Gelfond–Schneider $\frac{\log(\beta)}{\log(\alpha)}$ est nécéssairement rationnel ou transcendant. Ce qui nous donne le théorème suivant.

Théorème 6. *Soit α, β des complexes algébriques sur $\mathbb{Q}$ non nuls. Pour toute détermination* log *du logarithme sur $\mathbb{C}$ définie en α et β, si $\log(\alpha)$ et $\log(\beta)$ sont $\mathbb{Q}-$linéairement indépendants, alors ils sont $\overline{\mathbb{Q}}-$linéairement indépendants.*

Remarque 2. *En fait ce théorème est équivalent à celui de Gelfond–Schneider. En effet, si on suppose le théorème 6 et qu'on pose* $\gamma = \alpha^\beta$*, comme* β *est irrationnel alors* $\log(\gamma)$ *et* $\log(\alpha)$ *sont* $\mathbb{Q}-$*linéairement indépendants. Or* β *est non nul algébrique sur* $\overline{\mathbb{Q}}$ *donc ils ne sont pas* $\overline{\mathbb{Q}}-$*linéairement indépendants. D'après le théorème 6,* γ *est donc nécessairement transcendant sur* $\mathbb{Q}$*.*

3.2 Démonstration de Schneider

Prenons log une détermination complexe du logarithme. Soit α un nombre complexe algébrique tel que $\log(\alpha) \neq 0$ et β un irrationnel algébrique et posons $\gamma = \alpha^\beta = e^{\beta \log(\alpha)}$.

Étape 0 : Hypothèses et choix des paramètres.

Dans la suite, supposons par l'absurde que α, β, et γ sont dans un corps de nombres $\mathbb{K}$ et posons $h = [\mathbb{K} : \mathbb{Q}]$. Si α est une racine de l'unité, alors l'irrationalité de β implique que γ ne l'est pas. Ainsi quitte à remplacer α, β, γ par γ, β^{-1}, α, nous pouvons supposer sans perte de généralité pour la suite que α *n'est pas une racine de l'unité.*

Définissons des entiers m et n par :

$$m = 4h + 3 \quad et \quad n = \frac{q^2}{m},$$

où q est un entier positif non nul divisible par m.

Notre objectif est de construire un entier algébrique non nul dont la norme en module est strictement inférieure à 1 ce qui donnera une contradiction. Pour cela, nous allons d'abord construire une fonction auxiliaire R qui respecte certaines conditions. En considérant l'image en un point bien choisi b_n tel que $R(b_n) \in \mathbb{K}$, nous pourrons avoir grâce à l'analyse complexe une majoration assez fine pour construire la contradiction en faisant tendre n vers l'infini. Dans la suite $c_1, c_2, ...$ correspondront à des constantes > 0 indépendantes de n et les O seront toujours considérés par rapport à la variable n.

Étape 1 : Construction d'une fonction auxiliaire R.

Dans la construction de la fonction auxiliaire nous allons vouloir exploiter la relation fonctionnelle $\alpha^{x+y} = \alpha^x \alpha^y$. Celle-ci avec l'hypothèse faite au départ donnent que si k, ℓ sont des entiers, alors $\alpha^{k+\ell\beta} = \alpha^k \gamma^\ell \in \mathbb{K}$ et c'est cela qui est au centre de la preuve. En effet, comme β est irrationnel nous avons alors "trop de nombres algébriques différents dans le corps de nombres $\mathbb{K}$" et c'est d'ici que va venir la contradiction.

Lemme 3. *Il existe une fonction* R *de la forme* $R(x) = P_1(x)a^{(m-1)x} + P_2(x)a^{(m-2)x} + ... + P_m(x)$ *telle que :*

1. *pour* $i \in \{1, ..., m\}$*,* $P_i \in \mathcal{O}_\mathbb{K}[X]$ *et* $\deg(P_i) \leq 2n - 1$ *;*
2. *les* P_i *ne sont pas tous nuls ;*
3. *pour* $\lambda, \mu \in \{1, ..., q\}$*,* $R(\lambda + \mu b) = 0$*;*
4. *la maison des coefficients des* P_i *est* $O(c_2^n n^n)$ *où* $c_2 > 0$ *est une constante indépendante de* n*.*

Démonstration. Trouver ces coefficients revient donc à résoudre un système linéaire de q^2 équations à $2nm = 2q^2 > q^2$ inconnues où les coefficients sont les $(\lambda + \mu b)^j \alpha^{(m-\ell)(\lambda+\mu\beta)} = (\lambda + \mu\beta)^j \ \alpha^{(m-\ell)\lambda} \ \gamma^{(m-\ell)\mu} \in \mathbb{K}$ avec $\lambda, \mu \in \{1, 2, ..., q\}$, $j \in \{0, 1, ..., 2n-1\}$ et $\ell \in \{1, 2, ..., m\}$. De plus nous avons la majoration suivante de leur maison

$$\begin{aligned}\overline{|(\lambda+\mu\beta)^j \alpha^{(m-\ell)\lambda}|} &\leq (\overline{|\lambda|} + \overline{|\mu\beta|})^j \overline{|\alpha|}^{\lambda(m-\ell)} \overline{|\gamma|}^{\mu(m-\ell)} \\ &\leq (q(1+\overline{|\beta|}))^{2n-1} [\max(1, \overline{|\alpha|}) \max(1, \overline{|\gamma|})]^n \\ &\leq c_1^n q^{2n-1}\end{aligned}$$

avec $c_1 = (1 + \overline{|\beta|})^2 \max(1, \overline{|\alpha|}) \max(1, \overline{|\gamma|})$.
Ainsi par le lemme 2, il existe de tels polynômes non tous nuls qui respectent les conditions 1. et 3. du lemme 3. De plus comme $\frac{q^2}{2q^2-q^2} = 1$, la maison de chacun de leurs coefficients est donc majorée par

$$c(1 + (c2q^2 q^{2n-1} c_1^n)) \leq 2c(c2q^2 q^{2n-1} c_1^n) = O(c_2^n n^n) \qquad (c_2 = 2mc_1)$$

où c est une constante ne dépendant que de $\mathbb{K}$ qu'on peut prendre ≥ 1. Ce qui conclut le théorème. □

Dans la suite nous considérons la fonction R définie comme dans le lemme 3.

Étape 2 : Construction d'un nombre algébrique non nul b_n**.**

Pour construire notre nombre algébrique non nul en utilisant la relation $\alpha^x \alpha^y = \alpha^{x+y}$, nous allons considérer une matrice de Vandermonde afin de démontrer l'existence d'entiers λ et de μ compris entre 1 et $2q$ tels que $R(\lambda + \mu\beta) \neq 0$.

Pour cela, définissons pour $k \in \{1, ..., m\}$ le $k^{\text{ème}}$ translaté du polynôme P_ℓ par

$$P_{k,\ell}(x) := \alpha^{(m-\ell)x} P_\ell(x + k).$$

Si bien que nous avons

$$\forall k \in \{1, ..., m\}, \quad R(x+k) = P_{k,1}(x) a^{(m-1)x} + P_{k,2}(x) a^{(m-2)x} + ... + P_{k,m}(x). \qquad (1)$$

Ensuite considérons seulement les polynômes non identiquement nuls. Notons donc $0 < \nu_1 < \nu_2 < ... < \nu_g < m$ tels que $P_{\nu_1}, ..., P_{\nu_g}$ soient les seuls P_k non nuls. Nous désignons alors respectivement par r_ℓ et $a_\ell \neq 0$ le degré et le coefficient dominant de P_{ν_ℓ}. Puis posons $\Delta(x)$ le déterminant associé à la matrice $N(x) := (P_{k,\nu_\ell}(x))_{k,\ell=1,...,g}$. Nous avons alors

$$\Delta(x) = a_1 a_2 ... a_g v x^{r_1 + r_2 + ... + r_g} + ...$$

où

$$\begin{aligned} v &= \det((\alpha^{(m-\nu_\ell)k})_{k,\ell=1,...,g}) \\ &= \prod_{1 \leq \ell < k \leq g} (\alpha^{m-\nu_k} - \alpha^{m-\nu_\ell}). \quad (Vandermonde) \end{aligned}$$

Nous savons que $v \neq 0$ car α n'est pas une racine de l'unité. Ceci nous indique que le degré de Δ est

$$0 \leq r_1 + r_2 + ... + r_g \leq m(2n-1) < 2mn = 2q^2.$$

Ainsi Δ n'est pas un polynôme identiquement nul. En utilisant l'équation (1) nous allons donc pouvoir trouver nos entiers λ et μ tels que $R(\lambda + \mu\beta) \neq 0$.

Lemme 4. *Il existe $\xi_n := \lambda + \mu\beta$ avec $\lambda \in \{1, ..., q\}$ et $\mu \in \{1, ..., 2q\}$, et $k_n \in \{1, ..., m\}$ tels que*

$$R(\xi_n + k_n) \neq 0.$$

Démonstration. Comme Δ est un polynôme non nul de degré $< 2q^2$, alors il existe $\xi_n = \lambda + \mu b$ ($\lambda = 1, .., q$; $\mu = 1, ..., 2q$) parmi les $2q^2$ possibles qui n'est pas une racine de Δ. Par ailleurs si on écrit l'égalité (1) sous forme matricielle, on a :

$$(R(x+1), ..., R(x+g))^T = N(x)(a^{(m-\nu_1)x}, ..., a^{(m-\nu_g)x})^T.$$

Comme $\Delta(\xi) \neq 0$, on sait que la matrice $N(\xi)$ est inversible. De plus le vecteur $(a^{(m-\nu_1)\xi_n}, ..., a^{(m-\nu_g)\xi_n})^T$ est non nul. Ceci implique que $(R(\xi_n+1), ..., R(\xi_n+g))^T$ est non nul. D'où l'existence d'un entier $0 < k_n \leq m$ tel que $R(\xi_n + k_n) \neq 0$. □

Dans la suite, nous posons :

$$b_n := R(\xi_n + k_n) \neq 0.$$

Étape 3 : Minoration de $|\mathcal{N}(b_n)|$.

Tout d'abord, comme b_n est algébrique, nous savons que si $d_n = \text{den}(b_n)$ alors $d_n b_n \in \mathcal{O}_{\mathbb{K}}$. De plus $b_n \neq 0$ d'où

$$|d_n|^h |\mathcal{N}(b_n)| \geq |\mathcal{N}(d_n b_n)| \geq 1.$$

Ainsi trouver une minoration de $|\mathcal{N}(b_n)|$ revient à trouver une majoration de d_n.

Lemme 5. *Si $c_3 = \text{den}(\alpha)\,\text{den}(\beta)\,\text{den}(\gamma)$ alors*

$$c_3^{2n-1+(m-1)(3q+m)} b_n \in \mathcal{O}_{\mathbb{K}}.$$

En particulier on a $d_n \leq c_3^{2n-1+(m-1)(3q+m)}$.

Démonstration. Posons $d = c_3^{2n-1+(m-1)(3q+m)}$. Comme les coefficients des polynômes $P_k(x)$ sont dans $\mathcal{O}_{\mathbb{K}}$, pour que $dR(\xi_n + k_n)$ soit entier sur $\mathbb{K}$, il suffit juste que les $d(\xi_n + k_n)^j a^{(m-l)(\xi_n+k_n)}$ soient entiers sur $\mathbb{K}$ pour $j = 0, 1, ..., 2n-1$ et $l = 1, ..., m$ (car $\mathcal{O}_{\mathbb{K}}$ est un anneau). Alors

$$\begin{aligned} &d(\xi_n + k_n)^j \alpha^{(m-l)(\xi+k)} \\ &\qquad = d(\lambda + k_n + \mu\beta)^j (\alpha^{\lambda+k_n}\gamma^{\mu})^{(m-l)} \\ &\qquad = c_3^{2n-1-j+(m-1)(3q+m)-(\lambda+k_n+\mu)(m-l)} \big((\lambda + k_n + \mu\beta)c_3\big)^j \big((c_3\alpha)^{\lambda+k_n}(c_3\gamma)^{\mu}\big)^{(m-l)}. \end{aligned}$$

Or $0 \leq j \leq 2n-1$ et $1 \leq l \leq m$ et $0 \leq \lambda + k_n + \mu \leq 3q + m$ d'où $d(\xi_n + k_n)^j \alpha^{(m-l)(\xi_n+k)}$ est un entier sur $\mathbb{K}$. □

En conclusion, $\mathcal{N}(b_n) \geq (d_n)^{-h} \geq c_3^{-h\left(2n-1+(m-1)(3q+m)\right)} \geq c_4^{-n}$, où $c_4 := c_3^{2h+4mh}$.

Étape 4 : Majoration de $|\mathcal{N}(b_n)|$.

Tout d'abord par le lemme 3 les coefficients de $P_1, P_2, ..., P_m$ sont en module $O(c_2^n n^n)$. De plus pour $j = 0, 1, ..., 2n-1$ et $l = 1, 2, ..., m$ nous avons les majorations suivantes :

$$\overline{|(\xi + k_n)^j|} \leq (\overline{|\lambda|} + \overline{|k_n|} + \overline{|\mu|}\overline{|\beta|})^j \leq (qc_5)^{2n-1}$$

et

$$\overline{|\alpha^{(m-l)(\lambda+k_n+\mu\beta)}|} \leq (\overline{|\alpha|}^{\lambda+k_n}\overline{|\gamma|}^{\mu})^{m-l} \leq (c_6)^n c_7$$

avec $c_5 := (2 + m + \overline{|\beta|})$, $c_6 := \overline{|\alpha|}^2\overline{|\gamma|}$, $c_7 := \overline{|\alpha|}^{m^2}$. Ainsi

$$\overline{|b_n|} = O(2nm(qc_5)^{2n-1}c_6^n c_2^n n^n) = O(c_8^n n^n) \quad \text{avec} \quad c_8 = 2c_2c_5^2c_6.$$

Nous avons donc une première majoration de $|\mathcal{N}(b_n)|$ en $O(c_8^{hn}n^{hn})$. Cependant cette majoration n'est pas assez fine pour rencontrer une contradiction avec notre minoration de $|\mathcal{N}(b_n)|$. Pour l'améliorer, nous allons chercher une bonne estimation $|b_n|$. Pour cela, nous allons appliquer la formule de Cauchy à la fonction suivante :

$$S(x) = R(x) \prod_{f,g=1}^{q} \left(\frac{\xi_n + k_n - f - g\beta}{x - f - g\beta} \right),$$

où $S(\xi_n + k_n) = b_n$.

En effet, quand n devient grand la fonction R s'annule en un grand nombre de points et donc la fonction S associée où nous avons enlevé ces zéros de R reste "petite". Ceci va nous donner la bonne majoration.

Lemme 6. *Il existe une constante $c_{14} > 0$ indépendante de n telle que*

$$|b_n| = O(c_{14}^n n^{(3-\frac{m}{2})n}).$$

Démonstration. D'abord, la fonction entière R s'annule en les $f + g\beta$ pour $f, g \in \{1, ..., q\}$ alors la fonction S est également entière sur $\mathbb{C}$. Ensuite

$$|\xi_n + k_n| \leq (2 + |\beta|)q + m \leq n + q(2 + |\beta|) = n + O(\sqrt{n}),$$

d'où en prenant le cercle $\mathcal{C}$ de centre 0 et de rayon $n + m(q(2 + |\beta|)$ d'après la formule de Cauchy on a :

$$b_n = \frac{1}{2\pi i} \int_{\mathcal{C}} \frac{S(x)}{x - \xi_n - k_n} dx.$$

Il faut donc maintenant estimer les différents termes de l'intégrale.

— $|R(x)|$ pour $x \in \mathcal{C}$:

$$|R(x)| = O(2nm(nc_9)^{2n-1}c_{10}^n c_2^n n^n) = O(c_{11}^n n^{3n}),$$

où c_9 et c_{10} sont des constantes indépendantes de n et $c_{11} := 2c_2c_9^2c_{10}$.

— Pour $x \in \mathcal{C}$: $|x - \xi_n - k_n| \geq |x| - |\xi_n + k_n| \geq n$.
— Pour $x \in \mathcal{C}$ et $f, g = 1, ..., q$:

$$|x - f - g\beta| \geq |x| - |f + g\beta| \geq |x| - q(1 + |\beta|) \geq n.$$

— Pour $f, g = 1, ..., q$:

$$|\xi_n + k_n - f - g\beta| \leq |\xi_n + k_n| + \lambda + \mu|\beta| \leq m + q(3 + 2|\beta|) \leq qc_{12}$$

où $c_{12} = 3 + 2|\beta| + m$.
— Pour $f, g = 1, ..., q$:

$$\prod_{f,g=1}^{q} \left| \frac{\xi_n + k_n - f - gb}{x - f - gb} \right| \leq (qc_{12})^{q^2} n^{-q^2} = O(c_{13}^n n^{-\frac{mn}{2}})$$

où $c_{13} = c_{12}^m$.
Ainsi, on obtient que

$$\begin{aligned} |b_n| &\leq \frac{1}{2\pi} \int_{\mathcal{C}} \frac{|S(x)|}{|x - \xi_n - k_n|} dx \\ &= \frac{1}{2\pi n}(n + O(\sqrt{n}))O(c_{13}^n n^{-\frac{mn}{2}})O(n^{3n} c_{11}^n) \\ &= O(c_{14}^n n^{(3-\frac{m}{2})n}), \end{aligned}$$

avec $c_{14} = c_{11}c_{13}$. □

Enfin, on sait que b_n a au plus $h - 1$ conjugués différents de lui même d'où :

$$|\mathcal{N}(b_n)| \leq |b_n| \max(1, \overline{|b_n|})^{h-1} = O(c_{15}^n n^{(3-\frac{m}{2})n + 2(h-1)n})$$

avec $c_{15} := c_8 c_{14}$.

Étape 5 : Conclusion.

Par définition de m, on a

$$(3 - \frac{m}{2})n + 2(h-1)n = -\frac{mn}{2}.$$

On a donc démontré que :
— $|\mathcal{N}(b_n)| = O(c_{15}^n n^{-\frac{mn}{2}})$
— $|\mathcal{N}(b_n)| \geq c_4^{-n}$.
Or ces deux hypothèses sont contradictoires par croissance comparée lorsque $n \longrightarrow +\infty$. En conclusion l'hypothèse initiale est absurde, donc $\gamma = \alpha^\beta = e^{\beta \log(a)}$ est transcendant sur $\mathbb{Q}$.

3.3 Démonstration de Gelfond

Nous nous proposons de voir une autre démonstration du théorème, due à Gelfond que nous pouvons retrouver dans le chapitre III de [11]. Bien que le raisonnement soit similaire à celui précédemment, il est néanmoins intéressant car il utilise une autre fonction auxiliaire. En effet, on demande cette fois ci à la fonction de s'annuler à un grand ordre en des points bien choisis. Heuristiquement cela empêche la fonction de prendre des grandes valeurs. Nous allons pouvoir trouver notre nombre algébrique dont la majoration du module de sa norme donnera une contradiction.

Dans la suite α, β, γ et $\mathbb{K}$ sont utilisés comme précédemment à l'exception du fait qu'on ne suppose plus que α n'est pas une racine de l'unité. Comme la démonstration est similaire à celle de Schneider et qu'on la généralisera dans la prochaine partie, on ne détaillera pas les calculs. On définit des entiers m et n par

$$m = 2h + 2 \quad et \quad n = \frac{q^2}{2m},$$

où q est un entier divisible par $2m$.

A partir de maintenant $c_1, c_2, ...$ qui seront introduites dans cette partie désigneront des constantes > 0 indépendantes de q.

Étape 1 : Construction d'une fonction auxiliaire R et de b_n**.**

Posons

$$\rho_1, \rho_2, ..., \rho_t = (\lambda + \beta\mu)\log(\alpha) \quad \lambda, \mu = 1, 2, ..., q.$$

Nous souhaitons construire une fonction non nulle R telle que

1. $R(x) = \eta_1 e^{\rho_1 x} + ... + \eta_t e^{\rho_t x}$ avec $\eta_1, ..., \eta_t \in \mathcal{O}_{\mathbb{K}}$;
2. $R^{(k)}(\ell) = 0$ pour $k = 0, 1, ..., n-1$ et $\ell = 1, ..., m$;
3. Il existe une constante c_2 telle que $\max\limits_{1 \leq i \leq t}(|\overline{\eta_i}|) = O(c_2^n n^{\frac{n}{2}})$.

En considérant les mn équations $(\log a)^{-k} R^{(k)}(\ell) = 0$, il nous suffit juste de résoudre un système linéaire sur $\mathcal{O}_{\mathbb{K}}$ où les $2mn$ inconnues sont les η_i et à mn équations dont les coefficients sont les $(\lambda + \mu\beta)^k (\alpha^\lambda \gamma^\mu)^\ell \in \mathbb{K}$. De plus la maison des coefficients est en $O(c_1^n n^{\frac{n}{2}})$. D'après le lemme de Siegel, il existe donc $(\eta_1, ..., \eta_t)$ non trivial dans $\mathcal{O}_{\mathbb{K}}$ solution de ce système telle que

$$\max_{1 \leq i \leq t}(|\overline{\eta_i}|) = O(c_2^n n^{\frac{n}{2}}).$$

Il ne reste plus qu'à montrer que R est non nulle. Pour cela nous allons utiliser le lemme suivant :

Lemme 7. *Soit $r \in \mathbb{N}^*$ et soient $\mu_1, ..., \mu_r$ des complexes distincts. Si $f(x) = \lambda_1 e^{\mu_1 x} + ... + \lambda_r e^{\mu_r x}$ avec $\lambda_1, ..., \lambda_r$ des complexes alors*

$$f = 0 \Longleftrightarrow (\lambda_1, ..., \lambda_r) = 0.$$

Démonstration. Le sens réciproque est immédiat, montrons le sens direct. Si f est identiquement nulle alors en particulier

$$\forall k \in \mathbb{N}, \quad 0 = f^{(k)}(0) = \lambda_1 \mu_1^k + ... + \lambda_r \mu_r^k.$$

Soit $j \in \{1, ..., r\}$,, comme les μ_i sont distincts on peut poser le polynôme

$$L_j(X) = \prod_{\substack{1 \leq i \leq r \\ i \neq j}} \frac{X - \mu_i}{\mu_j - \mu_i}.$$

Ainsi,

$$0 = \lambda_1 L_j(\mu_1) + ... + \lambda_r L_j(\mu_r) = \lambda_j.$$

□

Comme $\log(\alpha) \neq 0$ et β est irrationnel, alors les ρ_i sont distincts et donc par le lemme 7 : R est non identiquement nulle. Ceci nous permet d'affirmer :

$$\forall \ell \in \{1, ..., m\}, \exists k \in \mathbb{N}^*, \quad R^{(k)}(\ell) \neq 0.$$

Finalement, cela justifie la définition de p, un entier tel que :

1. $\forall k \in \{0, ..., p-1\}, \forall \ell \in \{1, .., m\} \quad R^{(k)}(\ell) = 0$;
2. $\exists\, \ell_0 \in \{1, .., m\}, \quad R^{(p)}(\ell_0) \neq 0.$

Remarque 3. *Par construction $p \geq n$ donc quitte à prendre n grand nous pourrons prendre p arbitrairement grand.*

Enfin posons

$$b_p := (\log \alpha)^{-p} R^{(p)}(\ell_0) \in \mathbb{K}^*.$$

Étape 2 : Premières majorations.

Comme auparavant nous cherchons un entier c tel que cb_p soit un entier algébrique. Pour cela en posant $c_3 = \text{den}(\alpha)\,\text{den}(\gamma)$, nous obtenons que $c_3^{p+2mq} b_p$ est un entier algébrique et on note $k \leq h$ son degré. D'où les inégalités suivantes :

$$\left|\overline{c_3^{p+2mq} b_p}\right| \geq 1$$

et

$$1 \leq |\mathcal{N}(c_3^{p+2mq} b_p)| \leq c_3^{p+2mq} |b_p| \left|\overline{c_3^{p+2mq} b_p}\right|^{k-1} \leq (c_3^{p+2mq})^{h-1} |b_p| \left|\overline{b_p}\right|^{h-1}.$$

Puis, à l'aide du lemme de Siegel, nous avons une majoration de la maison des η_j en $O(c_2^n n^{\frac{n}{2}})$ et donc cela nous donne :

$$\left|\overline{b_p}\right| = O(c_5^p p^{\frac{p}{2}}).$$

Étape 3 : Estimation de $|b_p|$.

Comme précédemment, nous allons poser une fonction entière S où l'on enlève tous les zéros connus de R tout en renormalisant de sorte à avoir $S(\ell_0) = b_p$. Pour cela, prenons

$$S(x) = \frac{p!}{(\log(\alpha))^p} \frac{R(x)}{(x-\ell_0)^p} \prod_{\substack{k=1,\\ k\neq \ell_0}}^{m} \left(\frac{\ell_0 - k}{x-k}\right)^p.$$

Puis utilisons que

$$b_p = S(\ell_0) = \frac{1}{2\pi i} \int_{\mathcal{C}} \frac{S(x)}{x-\ell_0} dx,$$

où $\mathcal{C}$ est le cercle de centre 0 et de rayon $m(1+\frac{p}{q})$.
Enfin après toutes les majorations des termes de l'intégrale (où les raisonnements sont les mêmes que ceux utilisés pour la preuve de Schneider), on obtient

$$|b_p| = O(c_6^p p^{\frac{3-m}{2}}).$$

Étape 4 : Conclusion.

En regroupant toutes les inégalités, nous avons

$$1 \leq |\mathcal{N}(c_3^{p+2mq} b_p)| = O(c_6^p p^{(h-1)p+\frac{3-m}{2}p}).$$

Or l'exposant de p est égal

$$(h-1)p + \frac{3-m}{2}p = -\frac{p}{2},$$

donc en prenant p suffisament grand on aurait

$$1 \leq |\mathcal{N}(c_3^{p+2mq} b_p)| < 1,$$

ce qui nous donne la contradiction.

4 Le critère de Schneider–Lang

Dans ce chapitre allons démontrer un résultat important démontré par Schneider en 1949 [9] et raffiné par Lang en 1966 (dans l'appendice de [5]) qui est une généralisation du théorème de Gelfond–Schneider. Ce critère de transcendance est parmi les plus généraux connus et permet de prouver la transcendance de nombres classiques comme π, e ou encore $2^{\sqrt{2}}$ et e^{π}. Cependant, avant de pouvoir énoncer le théorème nous avons besoin d'introduire quelques outils pour pouvoir le comprendre. Nous nous sommes grandement appuyé sur de la preuve fournie dans [5] et [9].

4.1 Quelques notions d'extensions transcendantes

Nous avons souvent auparavant associé la propriété de finitude d'une extension $\mathbb{K}/\mathbb{Q}$ avec la dimension finie de $\mathbb{K}$ vu comme un $\mathbb{Q}$-espace vectoriel. Cependant, on peut aussi voir la finitude avec le nombre fini d'éléments qui engendre la $\mathbb{Q}$-algèbre $\mathbb{K}$ qui n'est pas équivalente à la première vision. Ceci amène à élaborer une théorie sur les extensions transcendantes qu'on retrouve dans le chapitre VIII de [5].

Définition 4. *On dit que l'extension de corps $\mathbb{L}/\mathbb{K}$ est de* type fini *si $\mathbb{L} = \mathbb{K}(x_1, x_2, ..., x_n)$ où les x_i sont des éléments de $\mathbb{L}$ et $n \in \mathbb{N}$.*

Remarque 4. *$\mathbb{L}/\mathbb{K}$ peut être de type fini sans être algébrique. Par exemple $\mathbb{Q}(\pi)/\mathbb{Q}$ est de type fini mais n'est pas algébrique (π étant transcendant sur $\mathbb{Q}$).*

En revanche si $\mathbb{L}/\mathbb{K}$ est de dimension finie, alors elle est de type fini ($\mathbb{L}$ est engendré par sa base). De plus dans notre cas où $\mathbb{K} = \mathbb{Q}$, le théorème de l'élément de primitif ajoute que plus précisémment $\mathbb{L}$ est engendré par un seul élément.

Maintenant, nous voudrions un moyen de classer les extensions de type fini. En effet, par exemple, on voit bien que les extensions $\mathbb{Q}(\pi)/\mathbb{Q}$ et $\mathbb{Q}(\sqrt{2})/\mathbb{Q}$ ne sont pas "de même nature" (l'une contenant un nombre transcendant et l'autre étant une extension algébrique). Pour cela, comme ces extensions peuvent contenir des éléments qui ne sont pas algébriques, nous allons devoir donner une notion plus général que l'algébricité.

On va donc s'inspirer de l'algébre linéaire et définir une notion sur les algèbres similaire à la liberté sur les espaces vectoriels : l'indépendance algébrique.

Définition 5. *Si $\mathbb{L}/\mathbb{K}$ est une extension, $d \geq 1$ et $y_1, ..., y_d$ des éléments de $\mathbb{L}$. On dit que $(y_i)_{1 \leq i \leq d}$ est* algébriquement indépendante sur $\mathbb{K}$ *si pour $P \in \mathbb{K}[X_1, ..., X_d]$*

$$P(y_1, ..., y_d) = 0 \iff P = 0.$$

Remarque 5. *En particulier, si y est algébriquement indépendant sur $\mathbb{K}$ alors il est transcendant sur $\mathbb{K}$.*

De plus, comme la liberté une famille finie d'éléments algébriquement indépendante si, et seulement si, toute sous-famille est algébriquement indépendante ce qui nous permet d'étendre la définition aux familles infinies. Une famille est donc algébriquement indépendante si, et seulement si, toute sous famille finie est algébriquement indépendante.

Nous pouvons également définir une notion similaire à celle de famille génératrice.

Définition 6. *Soit $\mathbb{L}/\mathbb{K}$ une extension et S une famille d'éléments de $\mathbb{L}$. On dit que la famille S* engendre algébriquement $\mathbb{L}$ *sur $\mathbb{K}$ si l'extension $\mathbb{L}/\mathbb{K}(S)$ est algébrique.*

Remarque 6. *Cette définition est cohérente avec la notion d'indépendance algébrique car si S engendre algébriquement $\mathbb{L}$, alors l'algébricité de l'extension $\mathbb{L}/\mathbb{K}(S)$ empêche un élément de $\mathbb{L}$ d'être algébriquement indépendant de S.*

Définition 7. *Soit $\mathbb{L}/\mathbb{K}$ une extension et B une famille d'éléments de $\mathbb{L}$. On dit que B est une* base de transcendance *de $\mathbb{L}$ sur $\mathbb{K}$, si c'est une famille algébriquement indépendante sur $\mathbb{K}$ qui engendre algébriquement $\mathbb{L}$. Autrement dit si B est une famille algébriquement indépendante maximale.*

Ainsi l'analogie avec l'algèbre linéaire nous amène à formuler le théorème suivant sur la base incomplète.

Théorème 7. *Soit $\mathbb{L}/\mathbb{K}$ une extension.*

1. *Soit T une famille finie d'éléments de $\mathbb{L}$ algébriquement génératrice et $S \subseteq T$, une famille algébriquement indépendante. Alors il existe $S \subseteq B \subseteq T$ une base de transcendance de $\mathbb{L}$ sur $\mathbb{K}$;*
2. *De plus si $\mathbb{L}/\mathbb{K}$ est de type fini, alors il existe une base de transcendance finie sur $\mathbb{K}$;*
3. *Si on dispose d'une base de transcendance finie de $\mathbb{L}$ sur $\mathbb{K}$, alors toutes les autres bases de transcendance $\mathbb{L}$ ont le même cardinal que cette dernière. On nomme ce cardinal* le degré de transcendance *de $\mathbb{L}$ sur $\mathbb{K}$.*

Remarque 7. *Ce théorème nous dit donc que dans le cas où l'extension est de type fini, on peut toujours compléter une famille algébriquement indépendante ou extraire d'une famille qui engendre algébriquement l'extension en une base de transcendance de celle-ci.*

Remarque 8. *La réciproque du point 2 du théorème est fausse. En effet, par exemple $\overline{\mathbb{Q}}/\mathbb{Q}$ est de degré de transcendance 0 mais n'est pas de type fini.*

Démonstration. • Prenons T et S comme dans le théorème. Nous allons construire par récurrence une famille algébriquement indépendante $S \subseteq B \subseteq T$ telle que si $x \in T \setminus B$ alors x est algébrique sur $\mathbb{K}(B)$. Ceci suffit à conclure que B est une base de transcendance car on a alors que

$$\mathbb{L}/\mathbb{K}(T) \quad \text{et} \quad \mathbb{K}(T)/\mathbb{K}(B)$$

sont algébriques et donc $\mathbb{L}/\mathbb{K}(B)$ aussi.

On écrit $T = \{x_1, ..., x_n\}$ et $S = \{x_1, ..., x_d\}$ et on applique l'algorithme suivant :

1. $S_0 := S$;
2. pour $i \in \{1, ..., n-d\}$, $S_i := S_{i-1} \cup \{x_{d+i}\}$ si $S_{i-1} \cup \{x_{d+i}\}$ est algébriquement indépendante, sinon $S_i := S_{i-1}$.

On pose $B := S_{n-d} = \{y_1, ...y_m\}$. Si $x \in T \setminus B$, alors par construction $B \cup \{x\}$ n'est pas algébriquement indépendante. Il existe donc $P \in \mathbb{K}[X_1, ..., X_{m+1}]$ non nul tel que :

$$P(y_1, ..., y_m, x) = 0.$$

On écrit

$$P(X_1, ..., X_{m+1}) = \sum_{j=0}^{h} Q_j(X_1, ..., X_m) X_{m+1}^j$$

avec les Q_j dans $\mathbb{K}[X_1, ..., X_m]$. Enfin $P \neq 0$, d'où

$$\exists \ell \in \{0, ..., h\}, \quad Q_l \neq 0.$$

et donc par indépendance algébrique de B : $Q_l(y_1, ..., y_m) \neq 0$. Ceci conclut sur l'algébricité de x sur $\mathbb{K}(B)$ car en posant $R(X) = \sum_{j=0}^{h} Q_j(y_1, ..., y_m)X^j \in \mathbb{K}(B)[X]$, on a $R \neq 0$ et $R(x) = 0$.

• Supposons que $\mathbb{L}/\mathbb{K}$ est finie. Il existe donc une famille finie T telle que $\mathbb{L} = \mathbb{K}(T)$. En particulier T engendre algébriquement $\mathbb{L}$. En prenant $S = \emptyset \subset T$, on a bien les hypothèses du premier point qui nous donne l'existence d'une base de transcendance.

• Considérons

$$\mathcal{A} = \{\text{card}(B)/\text{B base de transcendance finie de } \mathbb{L}\},$$

par hypothèse c'est une partie non vide de $\mathbb{N}$. On note n son minimum. On pose également

$$\mathcal{B} = \{\text{bases de transcendance de cardinal } n\}.$$

Nous allons montrer que toute famille algébriquement indépendante est de cardinal $\leq n$ ce qui conclura. Pour cela montrons par récurrence pour $t \in \{0, .., n\}$ l'assertion $\mathcal{P}(t)$ suivante :

Si S famille finie d'éléments de $\mathbb{L}$ algébriquement indépendante telle qu'il existe $B \in \mathcal{B}$ avec $\text{card}(B \cap S) = n - t$, alors $\text{card}(S) \leq n$.

Si $\text{card}(B \cap S) = n$ alors comme $\text{card}(B) = n$, cela implique que $B \subseteq S$ et donc par maximalité de B on a l'égalité d'où $\mathcal{P}(0)$.

Supposons maintenant $\mathcal{P}(t-1)$ et prenons une famille S comme définie dans $\mathcal{P}(t)$. Si $S \subseteq B$, alors S vérifie $\mathcal{P}(t)$. Nous pouvons donc supposer que $S \not\subset B$ et posons

$$B = \{b_1, .., b_n\} \quad \text{et} \quad B \cap S = \{b_{t+1}, ..., b_n\}.$$

Prenons $s \in S$ algébriquement indépendant de $B \cap S$. Par maximalité de B, on sait que $B \cup \{s\}$ n'est pas algébriquement indépendante. D'où

$$\exists P \in \mathbb{K}[X_1, ..., X_n, X] \setminus \{0\}, \quad P(b_1, ..., b_n, s) = 0.$$

L'indépendance algébrique de S assure l'apparition d'une indéterminée X_i pour $1 \leq i \leq t$ dans P, sans perte de généralité supposons que c'est X_1. Montrons que

$$C := \{s, b_2, ..., b_n\} \in \mathcal{B}.$$

On sait déjà que C est une famille algébriquement indépendante de cardinal n. Puis $P(X, b_2, ..., b_n, s) \in \mathbb{K}(C)[X] \setminus \{0\}$ annule b_1 ce qui justifie l'algébricité de $\mathbb{L}/\mathbb{K}(C)$. On a ainsi $C \in \mathcal{B}$. Comme $\text{card}(S \cap C) = n - (t-1)$, par hypothèse de récurrence on a $\text{card}(S) \leq n$. D'où $\mathcal{P}(t)$.

En conclusion, nous avons montré que toute famille algébriquement indépendante est de cardinal $\leq n$. En particulier, toutes les bases de transcendance sont de cardinal n par minimalité de n. □

Enfin nous allons démontré une propriété qui nous sera utile pour montrer le critère de Schneider–Lang nous donnant une caractérisation des extensions de type fini de degré de transcendance ≥ 2.

Propriété 3. *Soit $\mathbb{K}$ un corps et $x_1, ..., x_n$ des éléments dans une extension de $\mathbb{K}$, alors les assertions suivantes sont équivalentes.*

1. *$\mathbb{K}(x_1, ..., x_n)$ est de de degré de transcendance ≥ 2.*
2. *Il existe $1 \leq i < j \leq n$ tel que x_i et x_j sont algébriquement indépendants.*

Démonstration. 2. $\Longrightarrow$ 1..

Supposons que l'on a x_k et x_ℓ algébriquement indépendants. Comme $\mathbb{K}(x_1, ..., x_n)/\mathbb{K}$ est de type fini, on peut compléter $\{x_k, x_\ell\}$ en une base de transcendance de $\mathbb{K}(x_1, ..., x_n)$ qui est de cardinal ≥ 2.

1. $\Longrightarrow$ 2..

Comme $x_1, ..., x_n$ engendrent l'extension de type fini $\mathbb{K}(x_1, ..., x_n)$, on peut en extraire une base de transcendance B qui est donc de cardinal ≥ 2. En prenant deux éléments distincts de B, on a le résultat. □

4.2 Le critère et ses applications

Avant d'énoncer le théorème, nous allons définir la dernière notion d'analyse complexe qui nous manque.

Définition 8. *Soit f une fonction entière sur $\mathbb{C}$ et $\rho \in \mathbb{R}_+$. On dit que f est **d'ordre** $\leq \rho$ lorsque*

$$\exists M(f) > 1, \exists R(f) > 0, \forall R > R(f), \forall z \in \overline{D(0,R)}, \quad |f(z)| \leq (M(f))^{R^\rho}.$$

On dit q'une fonction méromorphe sur $\mathbb{C}$ est d'ordre $\leq \rho$ si elle est le quotient de fonctions entières d'ordre $\leq \rho$.

Demander qu'une fonction entière soit d'ordre $\leq \rho$ peut sembler assez faible comme contrainte car on a seulement une majoration exponentielle d'une puissance du rayon. Cependant pour des fonctions entières, il est difficile de demander un comportement asymptotique plus fort sans éliminer un nombre important de fonctions entières. Par exemple si on avait imposé une majoration en R^ρ les fonctions d'ordre $\leq \rho$ seraient les fonctions polynômiales de degré inférieur ou égal à la partie entière de ρ ce qui n'est pas intéressant ici. On peut néanmoins remarquer immédiatement certaines propriétés.

Propriété 4. *Si f, g sont deux fonctions entières d'ordre $\leq \rho$ alors :*

1. *$f+g$ est d'ordre $\leq \rho$ et $R(f+g) = \max(R(f), R(g), 2^{\frac{1}{\rho}})$ et $M(f+g) = M(f)+M(g)$ conviennent ;*
2. *fg est d'ordre $\leq \rho$ et $R(fg) = \max(R(f), R(g))$ et $M(fg) = M(f)M(g)$ conviennent.*

Démonstration. • Pour $f+g$. Si $R > \max(R(f), R(g), 2^{\frac{1}{\rho}})$, pour $z \in \overline{D(0,R)}$

$$\begin{aligned}|f(z)+g(z)| &\leq |f(z)|+|g(z)| \\ &\leq (M(f))^{R^\rho} + (M(g))^{R^\rho} \\ &\leq 2(\frac{M(f)+M(g)}{2})^{R^\rho} \quad \text{(convexité)} \\ &\leq (M(f)+M(g))^{R^\rho}.\end{aligned}$$

• Pour fg. Si $R > \max(R(f), R(g))$, pour $z \in \overline{D(0,R)}$

$$|f(z)g(z)| \leq |f(z)||g(z)| \leq (M(f)M(g))^{R^\rho}.$$

□

Nous avons maintenant tous les outils nécessaires pour énoncer le critère.

Théorème 8. (Schneider–Lang) *Soit $\mathbb{K}$ un corps de nombres. Soit $f_1, ..., f_N$ des fonctions méromorphes sur $\mathbb{C}$ d'ordre $\leq \rho$ telles que le corps $\mathbb{K}(f_1, ..., f_N)$ est de degré de transcendance ≥ 2 sur $\mathbb{K}$ et $\mathbb{K}[f_1, ..., f_N]$ soit stable par la dérivation complexe notée D. Soit $\omega_1, ..., \omega_m$ des nombres complexes distincts qui ne sont pas des pôles des f_i. Supposons que $f_i(\omega_j) \in \mathbb{K}$ pour $i = 1, ..., N$ et $j = 1, .., m$. Alors nous avons*

$$m \leq 8\rho\,[\mathbb{K} : \mathbb{Q}].$$

La propriété 3 nous dit qu'on a au moins deux des f_i qui sont algébriquement indépendantes. Ce critère est un résultat important car il est une généralisation d'autres théorèmes de transcendance importants, comme celui de Hermite-Lindemann ou encore de Gelfond-Schneider qu'on a étudié dans la partie précédente, en prenant des fonctions particulières.

Corollaire 1. ***Hermite-Lindemann***

Si α est algébrique sur $\mathbb{Q}$ et non nul, alors e^α est transcendant. En particulier si $\alpha \in \overline{\mathbb{Q}} \setminus \{0, 1\}$, alors $\log(\alpha) \notin \overline{\mathbb{Q}}$ où $\log$ est n'importe quelle fonction telle que $e^{\log(\alpha)} = \alpha$.

Ce théorème nous donne la transcendance de π car sinon cela impliquerait que $e^{2i\pi} = 1$ est transcendant ce qui est absurde.

Démonstration. Supposons par l'absurde que α et e^α sont algébriques.

Posons $\mathbb{K} = \mathbb{Q}(\alpha, e^\alpha)$ et les fonctions $f(z) = z$ et $g(z) = e^z$. Ces fonctions sont entières et d'ordre ≤ 1 telles que $\mathbb{K}[f, g]$ est stable par la dérivation complexe. Par ailleurs, pour $n \in \mathbb{N}$,

$$f(n\alpha) = n\alpha \in \mathbb{K}, \quad g(n\alpha) = (e^\alpha)^n \in \mathbb{K}.$$

Comme les $n\alpha$ sont distincts, si nous montrons que $\mathbb{K}(f, g)$ est de degré de transcendance ≥ 2 autrement dit que f, g sont algébriquement indépendantes alors le critère de Schneider–Lang nous dit que les fonctions f, g ne peuvent prendre simultanément des valeurs dans $\mathbb{K}$ qu'en au plus $8[\mathbb{K} : \mathbb{Q}]$ points distincts. Or ici on a en une infinité ce qui est absurde.

Supposons par l'absurde que f et g ne sont pas algébriquement indépendantes, alors on a

$$P(X,Y) = \sum_{0\leq k,\ell\leq m} a_{k,l}X^\ell Y^k = \sum_{0\leq k\leq m} P_k(X)Y^k \in \mathbb{K}[X,Y],$$

avec $P \neq 0$ et $P(f,g) = 0$.
Comme $P \neq 0$, il existe $k \in \{0,..,m\}$ tel que $P_k \neq 0$. On note donc j le max de ces k, et $\ell \geq 0$ le degré de P_j. Ainsi en étudiant le comportement asymptotique de $P(f,g)$ sur $\mathbb{R}$ à l'infini on a

$$0 = P(f(x),g(x)) = \sum_{0\leq k\leq m} P_k(x)e^{kx} \underset{+\infty}{\sim} a_{\ell,j}x^\ell e^{jx} \neq 0.$$

Or ceci est absurde. D'où f, g sont algébriquement indépendantes. □

On rappelle également le théorème de Gelfond–Schneider.

Corollaire 2. ***Gelfond–Schneider***
Soit $\alpha \in \mathbb{C}^*$ algébrique tel que $\log(\alpha) \neq 0$ et β irrationnel algébrique. Alors $\alpha^\beta := e^{\beta\log(\alpha)}$ est transcendant.

Démonstration. De manière analogue, on raisonne par l'absurde en supposant que α^β est algébrique et on pose $\mathbb{K} = \mathbb{Q}(\alpha, \beta, \alpha^\beta)$. Puis, on considère les fonctions d'ordre ≤ 1 $f(z) = e^z$ et $g(z) = e^{z\beta}$. Ensuite, on a bien $\mathbb{K}[f,g]$ stable pour la dérivation complexe et pour $k \in \mathbb{N}$

$$f(\log(\alpha)k) = \alpha^k \in \mathbb{K}, \quad g(\log(\alpha)k) = (\alpha^\beta)^k \in \mathbb{K}.$$

De plus les $k\log(\alpha)$ sont distincts.

Enfin, en utilisant l'irrationalité de β et le lemme 7, on montre que f, g sont algébriquement indépendantes. Ainsi, f, g prennent simultanément une valeur algébrique sur une infinité de points ce qui contredit le critère de Schneider-Lang. □

4.3 Dérivation de polynômes à plusieurs indéterminées et majorations

Lors de la démonstration du théorème, on sera amené à majorer les dérivées d'une fonction $f = P(f_1,...,f_N)$, avec P un polynôme à plusieurs indeterminées à coefficients dans un corps de nombres, évaluées en un point où elles prennent des valeurs algébriques. Le but de cette partie est d'établir une telle majoration du module de ces fonctions évaluées en ce point mais aussi une bonne majoration du dénominateur de ces nombres.

Dans toute la suite de cette partie, $\mathbb{K}$ est un corps de nombres et on travaille sur l'anneau $\mathbb{K}[X_1,...,X_N]$ où les X_i sont des indéterminées. On introduit avant plusieurs notations.

Si $\alpha = (\alpha_1,...,\alpha_N) \in \mathbb{N}^N$, on pose

$$X^\alpha = X_1^{\alpha_1}...X_N^{\alpha_N}.$$

On dit que X^α est un monôme de multi-indice α et on pose $|\alpha| = \alpha_1 + ... + \alpha_N$.

Ainsi si $P \in \mathbb{K}[X_1, ..., X_N]$, on a une famille presque nulle $(p_\alpha)_{\alpha \in \mathbb{N}^N}$ d'éléments de $\mathbb{K}$ telle que :

$$P(X_1, .., X_N) = \sum_{\alpha \in \mathbb{N}^N} p_\alpha X^\alpha.$$

On note $\deg(P) = \max\{|\alpha| / p_\alpha \neq 0\}$ le degré de P.

On rappelle également que $\frac{\partial}{\partial X_i}$ désigne la dérivée partielle selon la i-ème coordonnée et est définie sur les monômes par

$$\frac{\partial X^\alpha}{\partial X_i} = \alpha_i X_1^{\alpha_1} ... X_{i-1}^{\alpha_{i-1}} X_i^{\alpha_i - 1} X_{i+1}^{\alpha_{i+1}} ... X_N^{\alpha_N}.$$

Puis on l'étend par $\mathbb{K}$-linéarité sur tout $\mathbb{K}[X_1, ..., X_N]$.

Enfin on étend les notations données sur $\mathbb{K}$ à $\mathbb{K}[X_1, ..., X_N]$. Ainsi le dénominateur de P noté $\text{den}(P)$ désigne le ppcm des dénominateurs des coefficients de P. Par ailleurs la maison de P, notée $\overline{|P|}$ désigne le maximum des maisons de ses coefficients.

Définition 9. ***Dérivation par rapport à une famille de polynômes.***

Si $R = (R_1, ..., R_N)$ est une famille d'éléments de $\mathbb{K}[X_1, ..., X_N]$, alors on définit la dérivation par rapport à la famille R par une application D_R définie de $\mathbb{K}[X_1, ..., X_N]$ dans $\mathbb{K}[X_1, ..., X_N]$ telle que

$$D_R(P) = \sum_{j=1}^{N} R_j \frac{\partial P}{\partial X_j}.$$

Remarque 9. *La $\mathbb{K}$-linéarité des dérivées partielles donnent que l'application D_R est $\mathbb{K}$-linéaire. De plus D_R vérifie l'identité $D_R(PQ) = D_R(P)Q + PD_R(Q)$ d'où la dénomination dérivation pour la qualifier.*

Maintenant, on va définir une relation d'ordre sur les polynômes qui est compatible avec cette définition.

Définition 10. ***Domination de polynômes***

Soit $M \in \mathbb{R}_+[X_1, ..., X_N]$ et $P \in \mathbb{K}[X_1, ..., X_N]$. On dit que M domine P et on le note $P \leq M$ si

$$\forall \alpha \in \mathbb{N}^N, \quad \overline{|p_\alpha|} \leq m_\alpha.$$

Les propriétés suivantes montrent que cette notion est compatible avec la dérivation.

Propriété 5. *Si $P_1 \leq M_1$ et $P_2 \leq M_2$, alors :*

1. *$P_1 + P_2 \leq M_1 + M_2$;*
2. *$P_1 P_2 \leq M_1 M_2$;*
3. *pour tout $j \in \{1, ..., N\}$, $\quad \frac{\partial P_1}{\partial X_j} \leq \frac{\partial M_1}{\partial X_j}$.*

Puis dans nos majorations, on utilisera les propriétés suivantes.

Propriété 6. *Si $P \in \mathbb{K}[X_1, ..., X_N]$ de degré r alors*

$$P \leq \overline{|P|}\,(1 + X_1 + ... + X_N)^r.$$

Démonstration. En effet d'après la formule du multinôme

$$(1 + X_1 + ... + X_N)^r = \sum_{\alpha \in \mathbb{N}^N, |\alpha| \leq r} \frac{r!}{(r - |\alpha|)!\alpha_1!...\alpha_N!} X^\alpha.$$

De plus comme les $\frac{r!}{(r-|\alpha|)!\alpha_1!...\alpha_N!}$ sont des entiers ≥ 1, alors comme P est de degré $\leq r$, on sait qu'il est une somme de monôme $p_\alpha X^\alpha$ avec $|\alpha| \leq r$. Alors nécessairement $\overline{|p_\alpha|}$ est inférieur ou égal au produit de $\overline{|P|}$ avec le α-ième coefficient de $(1 + X_1 + ... + X_N)$. D'où le résultat. □

Propriété 7. *Si $M \in \mathbb{R}_+[X_1, ..., X_N]$ domine $P \in \mathbb{K}[X_1, ..., X_N]$ et $x_1, ..., x_N \in \mathbb{K}$ alors*

$$\overline{|P(x_1, ..., x_N)|} \leq M(\overline{|x_1|}, ..., \overline{|x_N|}).$$

Démonstration. On reprend les notations et les hypothèses de la propriété, alors

$$\begin{aligned}
\overline{|P(x_1, ..., x_N)|} &= \overline{\left| \sum_{\alpha \in \mathbb{N}^N} p_\alpha x_1^{\alpha_1} ... x_N^{\alpha_N} \right|} \\
&\leq \sum_{\alpha \in \mathbb{N}^N} \overline{|p_\alpha|} \overline{|x_1|}^{\alpha_1} ... \overline{|x_N|}^{\alpha_N} \\
&\leq \sum_{\alpha \in \mathbb{N}^N} m_\alpha \overline{|x_1|}^{\alpha_1} ... \overline{|x_N|}^{\alpha_N} = M(\overline{|x_1|}, ..., \overline{|x_N|}).
\end{aligned}$$

□

Nous avons tous les outils nécessaires pour énoncer le lemme qui sera important pour la suite.

Lemme 8. *Soit $\mathbb{K}$ un corps de nombres et $f_1, ..., f_N$ des fonctions holomorphes au voisinage d'un point $\omega \in \mathbb{C}$. De plus supposons que $f_i(\omega) \in \mathbb{K}$ pour $i = 1, .., N$ et que $\mathbb{K}[f_1, ..., f_N]$ est stable pour la dérivation complexe notée D. Enfin on pose $f = P(f_1, ..., f_N) \in \mathbb{K}[f_1, ..., f_N]$ avec P de degré inférieur ou égal r. Alors il existe un entier $C > 0$ ne dépendant que de $\mathbb{K}$ des f_j et de ω tel que, pour tout $k \in \mathbb{N}$:*

1. $\overline{|D^k(f)(\omega)|} \leq \overline{|P|} r^k k! C^{r+k}$;
2. $\mathrm{den}(P) C^{k+r} D^k(f)(\omega) \in \mathcal{O}_\mathbb{K}$.

Démonstration. Étape 1 : Transformation de la dérivée.

Tout d'abord pour pouvoir se ramener aux propriétés précédentes, on va transformer la dérivée complexe en une dérivée sur l'anneau $\mathbb{K}[X_1, ..., X_N]$ plus facile à manipuler. Pour cela on utilise que $D(f_j) = R_j(f_1, ..., f_N) \in \mathbb{K}[f_1, ..., f_N]$.

On pose donc D_R la dérivation par rapport à la famille de polynômes $R = (R_1, ..., R_N)$. Montrons que pour toute fonction $g = Q(f_1, .., f_N) \in \mathbb{K}[f_1, ..., f_N]$, on a $D(g)(\omega) = D_R(Q)(f_1(\omega), ..., f_N(\omega))$. On obtient

$$\begin{aligned} D(g)(\omega) &= \sum_{j=1}^{N} D(f_j)(\omega) \frac{\partial Q}{\partial X_j}(f_1(\omega), ..., f_N(\omega)) \\ &= \sum_{j=1}^{N} R_j(f_1(\omega), ..., f_N(\omega)) \frac{\partial Q}{\partial X_j}(f_1(\omega), ..., f_N(\omega)) \\ &= D_R(Q)(f_1(\omega), ..., f_N(\omega)). \end{aligned}$$

Ainsi par récurrence on montre en particulier que

$$\forall k \in \mathbb{N}, \quad D^k(f) = D_R^k(P)(f_1(\omega), ..., f_N(\omega)).$$

Étape 2 : Majoration de $\overline{|D^k(f)(\omega)|}$.

Après avoir transformé la dérivée complexe en une dérivée de polynôme, on va pouvoir utiliser les outils introduits précédemment pour notre majoration. Tout d'abord notons $Q = 1 + X_1 + ... + X_N$ et h le degré maximal des R_j qui ne dépend que des f_j. La propriété 6 nous donne

$$\forall j \in \{1, ..., N\}, \quad R_j \leq \overline{|R_j|} Q^h. \tag{2}$$

A partir de cela, montrons par récurrence sur k l'assertion suivante

$$\mathcal{P}(k) : D_R^k(P) \leq \overline{|P|}(rc_1h)^k k! Q^{r+k(h-1)} \quad \big(\text{avec } c_1 := 1 + \overline{|R_1|} + ... + \overline{|R_N|}\big).$$

Le cas $k = 0$ vient du fait que $\deg(P) \leq r$, donc par la propriété 6 : $D_r^0(P) = P \leq Q^r$. Maintenant prenons $k \in \mathbb{N}$ et supposons $\mathcal{P}(k)$. On sait que

$$D_R^{k+1}(P) = D_R(D_R^k(P)) = \sum_{j=1}^{N} R_j \frac{\partial D_R^k(P)}{\partial X_j}.$$

Puis par hypothèse de récurrence, par (2) en utilisant les trois points de la propriété 5

$$\begin{aligned} D_R^{k+1}(P) &\leq \sum_{j=1}^{N} \overline{|R_j|} Q^h \frac{\partial \overline{|P|}(rhc_1)^k k! Q^{r+k(h-1)}}{\partial X_j} \\ &\leq \overline{|P|}(rhc_1)^k k!(r + k(h-1)) Q^{r+h-1+k(h-1)} \sum_{j=1}^{N} \overline{|R_j|} \\ &\leq \overline{|P|}(rh)^k k!(r + k(h-1)) Q^{r+(k+1)(h-1)} c_1^{k+1} \qquad (car \sum_{j=1}^{N} \overline{|R_j|} \leq c_1) \\ &\leq \overline{|P|}(rh)^{k+1}(k+1)! Q^{r+(k+1)(h-1)} c_1^{k+1}. \end{aligned}$$

La dernière inégalité vient de

$$r + k(h-1) \leq r + kh \leq rh + rhk = rh(k+1).$$

On a donc montré $\mathcal{P}(k+1)$.

Enfin, comme les $f_i(\omega)$ sont dans $\mathbb{K}$, $D^k(f)(\omega) \in \mathbb{K}$ et par la propriété 7

$$\begin{aligned}\overline{|D^k(f)(\omega)|} &= \overline{|D_R(P)(f_1(\omega), ..., f_N(\omega))|} \\ &\leq (rc_1h)^k k!(Q(\overline{|f_1(\omega)|}, ..., \overline{|f_N(\omega)|}))^{r+k(h-1)} \\ &\leq \overline{|P|} r^k k! c_2^{k+r}\end{aligned}$$

avec $c_2 := [c_1 h Q(\overline{|f_1(\omega)|}, ..., \overline{|f_N(\omega)|})]^h$ qui ne dépend que des f_i et de ω. Ainsi tout entier $C \geq c_2$ convient pour satisfaire le premier point du lemme. Maintenant nous allons chercher d'autres conditions pour que C respecte le second point aussi.

Étape 3 : Majoration du dénominateur de $D^k(f)(\omega)$.

Tout d'abord pour $k \in \mathbb{N}$, on a montré dans l'étape 1 la domination de $D_R^k(P)$ par $(rc_1h)^k k! Q^{r+k(h-1)}$ et comme $\deg(Q^{k(h-1)+r}) = k(h-1)+r$, on en déduit que $\deg(D_R^k(P)) \leq k(h-1)+r$. Ensuite on pose

$$c_3 := \operatorname{ppcm}_{1 \leq j \leq N} \operatorname{den}(R_j), \quad c_4 := \operatorname{ppcm}_{1 \leq j \leq N} \big(\operatorname{den}(f_j(\omega))\big).$$

Montrons alors par récurrence pour $k \in \mathbb{N}$ que

$$c_3^k \operatorname{den}(P) D_R^k(P) \in \mathcal{O}_{\mathbb{K}}[X_1, ..., X_N].$$

Le cas $k = 0$ vient de la définition de $\operatorname{den}(P)$. Puis si la propriété est vraie au rang k alors

$$c_3^{k+1} \operatorname{den}(P) D_R^{k+1}(P) = \sum_{j=1}^{N} c_3 R_j \frac{\partial(\operatorname{den}(P) c_3^k D_R^k(P))}{\partial X_j}.$$

Comme $\mathcal{O}_{\mathbb{K}}[X_1, ..., X_N]$ est un anneau stable par les dérivées partielles, par définition de c_3 et par hypothèse de récurrence on a le résultat au rang $k+1$.

Ainsi, $c_3^k \operatorname{den}(P) D_R^k(P)$ est un polynôme à coefficients dans $\mathcal{O}_{\mathbb{K}}$ de degré $\leq r + k(h-1)$. Comme les $c_4 f_j(\omega)$ sont des entiers algébriques, on en déduit que

$$c_3^k c_4^{r+k(h-1)} \operatorname{den}(P) D_R^k(P)(f_1(\omega), ..., f_N(\omega)) \in \mathcal{O}_{\mathbb{K}}.$$

En posant $c_5 := c_3 c_4^h$, c_5 ne dépend que des f_j et de ω et

$$c_5^{k+r} \operatorname{den}(P) D^k(f)(\omega) = c_5^{k+r} \operatorname{den}(P) D_R^k(P)(f_1(\omega), ..., f_N(\omega)) \in \mathcal{O}_{\mathbb{K}}.$$

En conclusion en prenant C un multiple de c_5 plus grand que c_2 on finit la démonstration. □

4.4 Preuve du théorème

Maintenant, reprenons les hypothèses et les notations du théorème. Comme pour le théorème de Gelfond–Schneider la preuve repose sur la construction d'un entier algébrique γ non nul via une fonction auxiliaire. Grâce à cette dernière on pourra estimer le module de la norme en $O(s^{sC})$ où s une variable qu'on fera tendre à l'infini et C est une fonction de ρ, m, $d = [\mathbb{K} : \mathbb{Q}]$. Ainsi ceci imposera que C soit positive ce qui nous donnera l'inégalité du théorème. Dans la suite les c_i désigneront des constantes > 0 dépendantes éventuellement de m, des f_i, des ω_k et de $\mathbb{K}$.

Étape 1 : Construction d'une fonction auxiliaire.

Le degré de transcendance de $\mathbb{K}(f_1, ..., f_N)$ étant ≥ 2 par la propriété 3 on sait qu'au moins 2 des f_i sont algébriquement indépendantes. Supposons sans perte de généralité que ce sont f_1 et f_2 et posons $f = f_1$ et $g = f_2$. On définit un entier r qu'on pourra prendre arbitrairement grand tel qu'il soit un multiple de $2m$ et on pose $n = \frac{r^2}{2m} \geq r$.

On souhaite construire une fonction F méromorphe sur $\mathbb{C}$ de la forme $F = \sum_{i,j=1}^{r} a_{i,j} f^i g^j$ où les $a_{i,j}$ vérifient le lemme suivant.

Lemme 9. *Il existe une famille $(a_{i,j})_{1\leq i,j\leq r}$ d'éléments de $\mathcal{O}_\mathbb{K}$ telle que :*

1. *F n'est pas identiquement nulle ;*
2. *$\max(\overline{|a_{i,j}|}) = O(n^{2n})$;*
3. *$\forall \ell \in \{0, ..., n\}$, $\forall k \in \{1, ..., m\}$, $\quad \sum_{i,j=1}^{r} a_{i,j} D^\ell(f^i g^j)(\omega_k) = D^\ell F(\omega_k) = 0.$*

Démonstration. Pour cela, remarquons que la 3e condition du lemme traduit un système linéaire faisant intervenir $2mn$ inconnues pour mn avec pour coefficients $D^\ell(f^i g^j)(\omega_k)$ dans $\mathbb{K}$. Ensuite pour appliquer le lemme de Siegel cherchons une majoration de la maison des coefficients.

Pour tout $i, j \in \{1, ..., r\}$, si $P(X_1, ..., X_N) = X_1^i X_2^j \in \mathbb{K}[X_1, ..X_N]$, alors on sait que : $f^i g^j = P(f_1, ..., f_N)$, $\mathrm{den}(P) = 1$, $\overline{|P|} = 1$ et $\deg(P) \leq 2r$. Comme toutes les hypothèses du lemme 8 sont respectées pour tous les ω_k, alors pour $k \in \{1, ..., m\}$ il existe donc un entier $C_{1,k} > 0$ dépendant de $\mathbb{K}$, des f_ℓ et de ω_k tel que :

$$\forall \ell \in \mathbb{N}, \forall (i, j) \in \{1, .., r\}^2, \ \overline{\left|D^\ell(f^i g^j)(\omega_k)\right|} \leq (2r)^\ell \ell! C_{1,k}^{2r+\ell}$$

et

$$\forall \ell \in \mathbb{N}, \forall (i, j) \in \{1, .., r\}^2, \ C_{1,k}^{2r+\ell} D^\ell(f^i g^j)(\omega_k) \in \mathcal{O}_\mathbb{K}.$$

En posant $c_1 := \mathrm{ppcm}(C_{1,k})$, dans le cas où $\ell \in \{0, ..., n\}$:

$$\forall (i, j) \in \{1, .., r\}^2, \quad \overline{\left|D^\ell(f^i g^j)(\omega_k)\right|} \leq c_1^{2r+\ell} n! (2\sqrt{2m})^n n^{n/2} \leq c_2^n n^{n/2} n!, \tag{3}$$

$$\forall (i, j) \in \{1, .., r\}^2, \quad c_1^{3n} D^\ell(f^i g^j)(\omega_k) = c_1^{3n-2r-\ell}(c_1^{2r+\ell} D^\ell(f^i g^j)(\omega_k)) \in \mathcal{O}_\mathbb{K}. \tag{4}$$

avec $c_2 := 2\sqrt{2m} c_1^3$.

Ainsi par le lemme de Siegel et la majoration (3), il existe une famille $(a_{i,j})$ d'éléments de $\mathcal{O}_K$ non nulle vérifiant le système définit par la $3^{\text{ème}}$ condition du lemme et telle que

$$\max_{1\leq i,j\leq r}(\overline{|a_{i,j}|}) = O((2mnc_3^n n! n^{n/2})^{\frac{mn}{2mn-mn}}) = O(nc_3^n n! n^{n/2})$$

où c_3 est une constante dépendant de c_2 et $\mathbb{K}$. Puis par la formule de Stirling, on a

$$nc_3^n n^{n/2} n! = O(n^{2n}).$$

On en déduit donc

$$\max(\overline{|a_{i,j}|}) = O(n^{2n}).$$

Enfin $F = P(f,g)$ avec $P = \sum_{i,j=1}^{r} a_{i,j}X^iY^j \in \mathbb{K}[X,Y]$, de plus comme f et g sont algébriquement indépendantes et que les $a_{i,j}$ ne sont pas tous nuls, alors F n'est pas identiquement nulle. □

Remarque 10. *Lors de la démonstration, on a plus généralement donné les majorations de la maison et du dénominateur de $D^\ell(f^ig^j)(\omega_k)$ pour tout $\ell \in \mathbb{N}$ suivantes :*

$$\overline{|D^\ell(f^ig^j)(\omega_k)|} \leq (2r)^\ell \ell! c_1^{2r+\ell}$$

et

$$c_1^{2r+\ell}D^\ell(f^ig^j)(\omega_k) \in \mathcal{O}_{\mathbb{K}}.$$

Maintenant nous allons utiliser cette fonction F pour construire un nombre algébrique dont on peut faire une estimation du dénominateur. Dans un premier temps remarquons que F étant une fonction méromorphe sur $\mathbb{C}$ non identiquement nulle, pour $k \in \{1,...,m\}$, il existe $\ell_k \in \mathbb{N}$ tel que $D^{\ell_k}(F)(\omega_k) \neq 0$ et $D^j(F)(\omega_k) = 0$ pour $j \in \{0,..,\ell_k - 1\}$. Puis on pose $s = \min(\ell_k)$ et quitte à réindicer sans perte de généralité on peut supposer que $s = \ell_1$. Par définition de F on sait que $s \geq n$. On définit ensuite

$$\gamma_s := D^s(F)(\omega_1) \neq 0.$$

γ_s est un élément de $\mathbb{K}$ car par stabilité de la dérivation complexe dans $\mathbb{K}[f_1,...,f_N]$ on sait que $D^s(F)(\omega_1)$ est un élément de celui-ci et comme les $f_i(\omega_1)$ sont dans $\mathbb{K}$ cela conclut.

Étape 2 : Premières estimations de $|\mathcal{N}(\gamma_s)|$.

Toute la suite reposera sur ce nombre algébrique γ_s. Pour cela on va donner quelques propriétés sur γ_s.

Propriété 8. Propriétés de γ_s. *Il existe une constante entière $c_4 > 0$ telle que :*

1. $c_4^s\gamma_s \in \mathcal{O}_{\mathbb{K}}$;
2. $\overline{|c_4^s\gamma_s|} = O(s^{4s})$.

Démonstration. Par la remarque 10, comme $r \leq n \leq s$, on sait que

$$\forall i, j \in \{1, ..., r\}, \quad c_1^{3s} D^s(f^i g^j)(\omega_1) \in \mathcal{O}_{\mathbb{K}}.$$

On pose $c_4 = c_1^3$. Ainsi comme les $a_{i,j}$ sont dans $\mathcal{O}_{\mathbb{K}}$, on en déduit

$$c_4^s \gamma_s = \sum_{i,j=1}^{r} a_{i,j}(c_4^s D^s(f^i g^j)(\omega_k)) \in \mathcal{O}_{\mathbb{K}}.$$

Ensuite toujours par la remarque 10 et le point 2 du lemme 9, on a

$$\begin{aligned}
\overline{|c_4^s \gamma_s|} &\leq \sum_{i,j=1}^{r} \overline{|a_{i,j}|}(\overline{|c_4|})^s \overline{|D^s(f^i g^j)(\omega_k)|} \\
&= O(2mnn^{2n} c_4^s (2r)^s s! c_1^{2r+s}) \\
&= O(2mss^{2s}(c_4^2 2\sqrt{2m})^s s^{s/2} s!) \\
&= O(s^{4s}).
\end{aligned}$$

□

Cela nous permet d'avoir les estimations suivantes.

Lemme 10. *On rappelle que $d = [\mathbb{K} : \mathbb{Q}]$. Alors :*

1. $|\mathcal{N}(c_4^s \gamma_s)| \geq 1$;
2. $\overline{|\gamma_s c_4^s|} \geq 1$;
3. $|\mathcal{N}(c_4^s \gamma_s)| = O(c_4^s |\gamma_s| s^{4(d-1)s})$.

Démonstration. On sait déjà que $c_4^s \gamma_s$ est un entier algébrique, donc le module de sa norme est ≥ 1. Ensuite en notant k le degré de $c_4^s \gamma_s$ on a

$$(\overline{|\gamma_s c_4^s|})^k \geq |\mathcal{N}(c_4^s \gamma_s)| \geq 1.$$

On en déduit donc $\overline{|\gamma_s c_4^s|} \geq 1$. Enfin comme $\gamma_s c_4^s \in \mathbb{K}$, on sait que $k \leq d$ et donc par la propriété 8

$$|\mathcal{N}(c_4^s \gamma_s)| \leq c_4^s |\gamma_s| (\overline{|\gamma_s c_4^s|})^{k-1} \leq c_4^s |\gamma_s| (\overline{|\gamma_s c_4^s|})^{d-1} = c_4^s |\gamma_s| O(s^{4(d-1)s}).$$

□

Étape 3 : Majoration de $|\gamma_s|$.

Afin de démontrer le théorème il nous suffit de montrer que $|\gamma_s| = O(s^{Cs})$ avec C une constante indépendante de s. Pour se faire, nous allons d'abord régulariser F pour en faire une fonction entière F_1 qui sera toujours d'ordre $\leq \rho$. Puis on appliquera le principe du maximum à la fonction F_1 divisée par ses zéros connus. Les deux prochains lemmes servent à la régularisation de la fonction F.

Lemme 11. *Soit ϕ une fonction entière d'ordre $\leq \rho$. Soit $k_1, ...k_t$ des entiers > 0 et $z_1, ..., z_t$ des complexes distincts. Si la fonction $\psi(z) = \frac{\phi(z)}{(z-z_1)^{k_1}...(z-z_t)^{k_t}}$ est entière, alors ψ est d'ordre $\leq \rho$.*

Démonstration. Soit $R \geq \max(R(\phi), |z_1| + 1, ..., |z_t| + 1)$, par le principe du maximum on a

$$\max_{|z|\leq R} |\psi(z)| = \max_{|z|=R} |\psi(z)|.$$

Prenons $z \in \mathbb{C}$ tel que $|z| = R$ et prenons $j \in \{1, ..., t\}$. Alors

$$|z - z_j| \geq |z| - |z_j| \geq 1 \qquad (R \geq |z_j| + 1)$$

et

$$|\phi(z)| \leq M(\phi)^{R^\rho} \qquad (R \geq R(\phi)).$$

On a donc

$$|\psi(z)| \leq (M(f))^{R^\rho}.$$

Ainsi en prenant

$$R(\psi) = \max(R(\phi), |z_1| + 1, ..., |z_t| + 1), \quad M(\psi) = M(\phi).$$

on obtient bien que ψ est d'ordre $\leq \rho$. □

Lemme 12. *Il existe une fonction entière θ d'ordre $\leq \rho$ telle que θf et θg soient des fonctions entières d'ordre $\leq \rho$ et $\theta(\omega_1) \neq 0$.*

Démonstration. f et g sont des fonctions méromorphes d'ordre $\leq \rho$, donc par définition il existe $\phi_1, \phi_2, \phi_3, \phi_4$ des fonctions entières d'ordre $\leq \rho$ telles que

$$f = \frac{\phi_1}{\phi_2}, \qquad g = \frac{\phi_3}{\phi_4}.$$

Ensuite il existe k_1 (respectivement $k_{2,}, k_3, k_4$) tel que $\psi_1(z) = \frac{\phi_1}{(z-\omega_1)^{k_1}}$ soit entière et $\psi_1(\omega_1) \neq 0$ (de même pour ϕ_2, ϕ_3, ϕ_4). Ainsi

$$f(z) = (z - \omega_1)^{k_1 - k_2} \frac{\psi_1(z)}{\psi_2(z)}, \qquad g(z) = (z - \omega_1)^{k_3 - k_4} \frac{\psi_3(z)}{\psi_4(z)}.$$

Posons $\theta = \psi_2\psi_4$. Alors θ est bien une fonction entière d'ordre $\leq \rho$ par produit de fonctions entières d'ordre $\leq \rho$. De plus $\theta(\omega_1) = \psi_2(\omega_1)\psi_4(\omega_1) \neq 0$. Enfin

$$(\theta f)(z) = (z - \omega_1)^{k_1 - k_2} \psi_1(z)\psi_4(z), \qquad (\theta g)(z) = (z - \omega_1)^{k_3 - k_4} \psi_3(z)\psi_2(z).$$

ω_1 n'étant pas un pôle de f ou de g, on sait que $k_1 - k_2 \geq 0$ et $k_3 - k_4 \geq 0$. Alors θf et θg sont des fonctions entières d'ordre $\leq \rho$ comme produit de fonction entières d'ordre $\leq \rho$. □

Remarque 11. *La fonction entière θ dépend uniquement de f, g et ω_1. C'est donc aussi le cas pour $R(\theta)$ et $M(\theta)$.*

Dans la suite, on pose

$$F_1 = (\theta)^{2r} F \quad \text{et} \quad H(z) = \frac{F_1(z)}{(z-\omega_1)^s ... (z-\omega_m)^s}.$$

Maintenant énonçons quelques propriétés sur ces deux fonctions.

Propriété 9. *On définit $R_0 = \max(R(\theta), R(\theta f), R(\theta g), 2|\omega_1|, ..., 2|\omega_m|)$ qui ne dépend que de f, g et des ω_i, alors :*

1. *F_1 et H sont des fonctions entières ;*
2. *$H(\omega_1) = \frac{\gamma_s}{s!(\omega_1-\omega_2)^s ....(\omega_1-\omega_m)^s}(\theta(\omega_1))^{2r}$;*
3. *pour $R \geq R_0$, et $z \in \mathbb{C}$ tel que $|z| = R$*

$$|F_1(z)| \leq c_6^{\sqrt{s}R^\rho} s O(s^{2s}) \quad et \quad |H(z)| \leq \frac{2^{sm} s c_6^{\sqrt{s}R^\rho} O(s^{2s})}{R^{sm}}.$$

Démonstration. • Tout d'abord

$$\theta^{2r} F = \theta^{2r} \sum_{i,j=1}^{r} a_{i,j} f^i g^j = \sum_{i,j=1}^{r} \theta^{2r-i-j} a_{i,j} (\theta f)^i (\theta g)^j.$$

D'où $\theta^{2r} F$ est une fonction entière par sommes et produits de fonctions entières. Ensuite θ étant une fonction entière et F annulant tous les ω_j jusqu'au moins l'ordre s, on en déduit que c'est aussi le cas pour F_1. Ainsi H est également une fonction entière.

• Ensuite, on sait que $\theta(\omega_1) \neq 0$, donc F_1 annule ω_1 à exactement l'ordre s, et donc on en déduit le résultat.

• Soit $R \geq R_0$, et $z \in \mathbb{C}$ avec $|z| = R$, alors

$$\begin{aligned}
|F_1(z)| &\leq \sum_{i,j=1}^{r} |\theta^{2r-i-j}(z) a_{i,j} (\theta f(z))^i (\theta g(z))^j| \\
&\leq \sum_{i,j=1}^{r} |\overline{a_{i,j}}| (M(\theta)^{2r-i-j} M(\theta f)^i M(\theta g)^j)^{R^\rho} && (\theta f, \theta g, \theta \quad d'ordre \leq \rho) \\
&\leq 2mn (M(\theta)^2 M(\theta f) M(\theta g))^{r R^\rho} \max(|\overline{a_{i,j}}|) \\
&\leq s c_6^{\sqrt{s}R^\rho} O(s^{2s}) && (c_6 = M(\theta)^2 M(\theta f) M(\theta g)).
\end{aligned}$$

Par ailleurs on a $R/2 \leq R - |\omega_k| \leq |z - \omega_k|$ pour $k \in \{1, ..., m\}$. Ceci nous permet d'en déduire l'inégalité suivante pour $|H(z)|$:

$$|H(z)| = \frac{|F_1(z)|}{|z-\omega_1|^s ....|z-\omega_m|^s} \leq \frac{2^{sm} s c_6^{\sqrt{s}R^\rho} O(s^{2s})}{R^{sm}}.$$

□

A partir de cette propriété, on obtient la majoration suivante de $|\gamma_s|$.

Lemme 13. *On dispose d'une constante C ne dépendant que des paramètres du théorème de Schneider-Lang telle que*

$$|\gamma_s| = O(C^s s^{s(3-m/2\rho)}).$$

Démonstration. Tout d'abord, d'après la propriété 9 on sait que H est une fonction entière et donc par le principe du maximum

$$\forall R \geq |\omega_1|, \quad |H(\omega_1)| \leq \max_{|z|=R} |H(z)|.$$

En particulier

$$\forall R \geq R_0 \geq |\omega_1|, \quad |H(\omega_1)| \leq \frac{2^{sm} s c_6^{\sqrt{s}R^\rho} O(s^{2s})}{R^{sm}}.$$

Maintenant on va choisir R en fonction de s afin d'optimiser la majoration. Comme $s \geq r$ pour r suffisament grand, en posant $R = s^{1/2\rho}$, on a que $R \geq R_0$ d'où

$$|H(\omega_1)| \leq \frac{2^{sm} s c_6^s O(s^{2s})}{s^{sm/2\rho}}.$$

Finalement pour $|\gamma_s|$, on a

$$\begin{aligned}
|\gamma_s| &= (1/|\theta(\omega_1)|)^{2r} |H(\omega_1)| s! \prod_{k=2}^{m} |\omega_1 - \omega_k|^s \\
&\leq \frac{(c_7^2 c_8^{(m-1)} 2^m)^s s c_6^s \, s! \, O(s^{2s})}{s^{sm/2\rho}} && (c_7 = \max(1, 1/|(\theta(\omega_1)|), \ c_8 = \max_{2\leq k\leq m} |\omega_1 - \omega_k|) \\
&= O(c_9^s s^{s(3-m/2\rho)}) && (s\, s! = O(s^s), \ c_9 = 2^m c_6 c_7^2 c_8^{m-1}).
\end{aligned}$$

□

Étape 4 : Conclusion.

A l'aide des lemmes 10 et 13, on a les estimations suivantes

$$|\mathcal{N}(c_4^s \gamma_s)| \geq 1 \quad \text{et} \quad |\mathcal{N}(c_4^s \gamma_s)| = O((c_4 c_9)^s s^{s(4(d-1)+3-m/2\rho)}).$$

Il faut donc $4(d-1) + 3 - m/2\rho) \geq 0$ car sinon par croissance comparée en prenant r suffisamment grand on aurait ($s \geq r$) :

$$O((c_4 c_9)^s s^{s(4(d-1)+3-m/2\rho)}) < 1$$

ce qui est absurde.

En conclusion on a l'inégalité suivante

$$4(d-1) + 3 - m/2\rho) \geq 0.$$

D'où

$$m \leq 2\rho(4d-1) \leq 8\rho[\mathbb{K} : \mathbb{Q}].$$

5 Le théorème de Thue

Durant ce chapitre nous allons démontrer un résultat d'approximation rationnelle des nombres algébriques démontré par Thue en 1908. Ce résultat est une généralisation d'un résultat d'approximation due à Liouville rappelé plus loin et il sera à son tour généralisé en 1955 par Roth. Dans cette partie nous nous sommes majoritairement appuyé des cours de Mignotte [6] et de Guth [4].

Théorème 9. (Thue) *Soit α est un réel irrationnel algébrique sur $\mathbb{Q}$, et on note d son degré. Si $\rho > \frac{d}{2} + 1$, alors il n'existe qu'un nombre fini de rationnels $\frac{p}{q}$ solutions de l'inéquation :*

$$\left|\alpha - \frac{p}{q}\right| < \frac{1}{q^\rho}.$$

Dans cette partie, on utilisera les notations suivantes.

1. Pour un réel α, on note $\lfloor \alpha \rfloor$ sa *partie entière* et $\{\alpha\} = \alpha - \lfloor \alpha \rfloor$ sa *partie fractionnaire*.
2. Pour $z \in \mathbb{C}$
$$z^* = \max(|z|, 1).$$
3. Pour $P = \sum_{\alpha \in \mathbb{N}^N} p_\alpha X^\alpha \in \mathbb{C}[X_1, ..., X_N]$, on définit sa *hauteur* par
$$\mathrm{H}(P) = \max_{\alpha \in \mathbb{N}^N} |p_\alpha|.$$
4. On définit également la hauteur d'un nombre rationnel p/q écrit sous forme irréductible par
$$\mathrm{H}\left(\frac{p}{q}\right) = \max(|p|, |q|)$$
(A noter que $\mathrm{H}\left(\frac{p}{q}\right) = \mathrm{H}(qX - p)$).

5.1 Premiers résultats

Dans cette partie nous allons nous intéresser à des résultats antérieurs d'approximations de nombres algébriques par des rationnels car même s'ils sont moins généraux, la méthode utilisée néanmoins sera la même quoiqu'un peu plus générale nous permettra de démontrer le théorème de Thue. Nous commençons d'abord par un résultat classique d'approximation de réels irrationnels par des rationnels reposant sur le principe des tiroirs.

Théorème 10. (Dirichlet) *Soit $\alpha \in \mathbb{R} \setminus \mathbb{Q}$ et Q un réel ≥ 2. Alors il existe des entiers p et q premiers entre eux tels que*

$$1 \leq q < Q \text{ et } |\alpha q - p| \leq \frac{1}{Q} < \frac{1}{q}.$$

Démonstration. Quitte à considérer $\lfloor Q \rfloor \geq 2$, il suffit de le montrer dans le cas où Q est entier ≥ 2 pour avoir le résultat général. Supposons donc Q entier.

Ensuite nous découpons l'intervalle $[0;1]$ en Q intervalles $[\frac{i}{Q}; \frac{i+1}{Q}]$ avec $i = 0, 1, ..., Q-1$. Puis remarquons due à l'irrationalité de α que les $Q+1$ nombres $0, 1, \{\alpha\}, ..., \{(Q-1)\alpha\}$ sont distincts. En effet, tout d'abord l'irrationalité de α indique que $k\alpha$ n'est jamais un entier et donc $0 < \lfloor k\alpha \rfloor < 1$ pour $k = 1, ..., Q-1$. De plus si nous supposons par l'absurde que l'on a de $1 \leq k < l \leq Q-1$ tels que $\{k\alpha\} = \{l\alpha\}$, alors on obtient l'égalité suivante

$$\alpha = \frac{\lfloor l\alpha \rfloor - \lfloor k\alpha \rfloor}{k-l} \in \mathbb{Q}\,.$$

Ceci contredit l'irrationalité de α. Les $\{k\alpha\}$ sont donc bien distincts.

Comme $Q \geq 2$, 0 et 1 ne peuvent être dans le même intervalle. Ainsi d'après le principe des tiroirs, il existe $0 \leq q_1 < q_2 \leq Q-1$ et des entiers p_1, p_2 tels que

$$|q_1\alpha - p_1 - (q_2\alpha - p_2)| \leq Q^{-1}.$$

D'où en prenant $q = q_2 - q_1$ et $p = p_1 - p_2$, on a bien

$$0 < q < Q \quad \text{et} \quad |q\alpha - p| \leq Q^{-1}.$$

Enfin en notant $d = \text{pgcd}(p, q) \geq 1$ et en divisant l'inégalité précédente, on a

$$0 < \frac{q}{d} \leq q < Q \quad et \quad \left|\frac{q}{d}\alpha - \frac{p}{d}\right| \leq \frac{1}{dQ} \leq \frac{1}{Q}$$

avec p/d et q/d des entiers premiers entre eux. □

Corollaire 3. *Soit $\alpha \in \mathbb{R}$, alors :*

1. *si α est rationnel alors il existe une constante $c = c(\alpha) > 0$ telle que pour tout rationnel $\frac{p}{q}$ distinct de α*
$$\left|\alpha - \frac{p}{q}\right| \geq \frac{c(\alpha)}{q};$$
2. *si α est irrationnel alors il existe une infinité de rationnels $\frac{p}{q}$ tels que*
$$\left|\alpha - \frac{p}{q}\right| < \frac{1}{q^2}.$$

Démonstration. • Pour le premier point, posons $\alpha = \frac{a}{b} \in \mathbb{Q}$ et prenons un rationnel $\frac{p}{q} \neq \alpha$ alors

$$\left|\frac{p}{q} - \alpha\right| = \frac{|pb - aq|}{qb} > 0.$$

$|pb - aq|$ est donc un entier positif non nul d'où

$$\left|\frac{p}{q} - \alpha\right| \geq \frac{1}{qb}.$$

Ainsi la constante $c(\alpha) = b^{-1}$ convient.
• Le point 2 est une conséquence immédiate du théorème 10 □

Avant d'effectuer la démonstration du théorème de Thue, nous allons d'abord nous intéresser l'approximation de Liouville pour comprendre l'idée du raisonnement que nous allons utiliser.

Théorème 11. (Liouville) *Soit α est un réel algébrique sur $\mathbb{Q}$ de degré $d \geq 2$ sur $\mathbb{Q}$. Alors il existe une constante strictement positive $c = c(\alpha)$ (calculable) telle que pour tout rationnel $\frac{p}{q}$ on a :*

$$\left|\alpha - \frac{p}{q}\right| \geq \frac{c}{q^d}.$$

Démonstration. On note P_α le polynôme minimal de α sur $\mathbb{Q}$. Prenons un rationnel $r = \frac{p}{q}$, comme $d \geq 2$ on sait que $\frac{p}{q} \neq \alpha$. Ainsi l'irréductibilité de P_α sur $\mathbb{Q}$ entraine que

$$P_\alpha(r) \neq 0.$$

Par ailleurs si nous prenons $a = a(\alpha) \in \mathbb{N}^*$ tel que $aP_\alpha \in \mathbb{Z}[X]$, alors

$$aq^d P_\alpha(r) \in \mathbb{Z}^*.$$

Ceci nous fournit l'inégalité

$$|P_\alpha(r)| \geq \frac{1}{aq^d}.$$

Ensuite par l'inégalité des accroissements finis, il existe un réel $c \in [\min(\alpha, r); \max(\alpha, r)]$ tel que

$$|P_\alpha(r)| = |P_\alpha(r) - P_\alpha(\alpha)| \leq |P'(c)|\,|(\alpha - r)|.$$

Si r vérifie $|\alpha - r| \leq 1$, alors

$$|P_\alpha(r)| \leq d^2\,\mathrm{H}(P_\alpha)(|\alpha| + 1)^{d-1}|\alpha - r|.$$

D'où

$$\left|\alpha - \frac{p}{q}\right| \geq \frac{1}{d^2\,\mathrm{H}(aP_\alpha)(|\alpha| + 1)^{d-1}q^d}.$$

Enfin comme aP_α est à coefficients entiers alors $\mathrm{H}(aP_\alpha) \geq 1$. Ainsi cette inégalité reste vraie lorsque $|\alpha - r| > 1$. En conclusion la constante

$$c(\alpha) = \frac{1}{d^2\,\mathrm{H}(aP_\alpha)(|\alpha| + 1)^{d-1}}$$

convient. □

Ainsi ce théorème montre que d'une certaine manière les nombres algébriques s'approximent mal par des rationnels. Ceci nous permet de démontrer la transcendance de nombres réels qu'on peut bien approcher par des rationnels comme le montre le corollaire suivant.

Corollaire 4. *Le nombre $x = \sum_{n=0}^{+\infty} 2^{-n!}$ est transcendant sur $\mathbb{Q}$.*

Démonstration. Tout d'abord on sait que x est irrationnel car son développement en base 2 n'est pas périodique. Ensuite prenons $n \in \mathbb{N}$. On écrit

$$\sum_{k=0}^{n} \frac{1}{2^{k!}} = \frac{\sum_{k=0}^{n} 2^{n!-k!}}{2^{n!}} = \frac{p_n}{q_n}.$$

Ainsi on a $q_n \leq 2^{n!}$. Par ailleurs nous savons également que

$$\sum_{k=n+1}^{+\infty} 2^{-k!} \leq \frac{2}{2^{(n+1)!}} \leq \frac{2}{q_n^{n+1}}.$$

Supposons par l'absurde que x est algébrique de degré d sur $\mathbb{Q}$. Par l'inégalité précédente on a

$$\forall n \geq d, \qquad \left| x - \frac{p_n}{q_n} \right| \leq \frac{2}{q_n^{d+1}}.$$

Par ailleurs comme x est algébrique de degré d, par le théorème de Liouville il existe une constante $c = c(\alpha) > 0$ telle que

$$\forall n \geq d, \quad \frac{c}{q_n^d} \leq \left| x - \frac{p_n}{q_n} \right| \leq \frac{2}{q_n^{d+1}}.$$

Or cela est impossible car c est strictement positif et $\lim\limits_{n \to +\infty} q_n = +\infty$. En conclusion x est bien transcendant sur $\mathbb{Q}$ □

Dans cette démonstration on retrouve la même méthode que celle utilisée lors de la preuve du critère de Schneider-Lang qui repose toujours sur les mêmes étapes :

1. construction d'une fonction auxiliaire vérifiant des conditions d'annulation en des points bien choisis (ici P_α) ;
2. construction d'un nombre algébrique γ (dans cette partie il sera rationnel ce qui nous permet d'étudier $|\gamma|$ au lieu de $|\mathcal{N}(\gamma)|$) ;
3. minoration de $|\gamma|$ résultant de propriétés arithmétiques ;
4. majoration de $|\gamma|$ résultant de propriétés analytiques ;
5. enfin l'incompatibilité entre la majoration et la minoration permet de conclure.

5.2 Preuve du théorème

Pour pouvoir améliorer l'exposant d en $d/2+1$, Thue eut l'idée d'utiliser deux approximations rationnelles de α. Cependant cela implique d'avoir recours à des polynômes à deux variables ce qui complique la preuve. Dans la suite α désignera un réel algébrique de degré $d \geq 2$ et ρ un réel vérifiant $\frac{d}{2} + 1 < \rho < d$ (cela suffit pour démontrer le théorème car le théorème de Liouville nous fournit la preuve dans le cas $\rho \geq d$).

Étape 0 : Choix des paramètres.

Nous allons démontrer le théorème en raisonnant par l'absurde. Supposons donc qu'il existe une infinité de rationnels p/q vérifiant l'inégalité suivante :

$$\left|\alpha - \frac{p}{q}\right| < \frac{1}{q^\rho}. \tag{5}$$

Considérons deux approximations rationnelles irréductibles (p_1/q_1) et (p_2/q_2) vérifiant (5). De plus comme l'inéquation (5) admet une infinité de solutions rationnelles, on peut prendre q_1, q_2 et $(\log(q_2)/\log(q_1))$ arbitrairement grands. Pour l'instant on les prendra tous les trois ≥ 2. Ensuite prenons $0 < \varepsilon < \frac{1}{3}$ suffisamment petit pour que $\rho > 1 + (1+3\,\varepsilon)\frac{d}{2}$. Enfin nous définissons les entiers r et ℓ suivants

$$\ell = \left\lfloor \frac{\log(q_2)}{\log(q_1)} \right\rfloor + 1, \qquad r = \left\lfloor \frac{(1+\varepsilon)\,d}{2}\ell \right\rfloor.$$

Dans la suite $c_1, c_2, ...$ désigneront des constantes ne dépendant que de α, d, ρ et ε.

Étape 1 : Construction de la fonction auxiliaire.

Par analogie avec la preuve de Liouville, nous allons construire un polynôme à deux variables $A(X,Y)$ d'une certaine forme vérifiant des conditions d'annulation au point (α,α).

Lemme 14. *Il existe $P, Q \in \mathbb{Z}[X]$ non tous les deux nuls et de degré $\leq r$ tels que si l'on pose $A(X,Y) = P(X) - YQ(X) \in \mathbb{Z}[X,Y]$, alors :*

1. $\forall j \in \{0,...,\ell-1\}, \quad \frac{1}{j!}\partial_1^j A(\alpha,\alpha) = \frac{1}{j!}\frac{\partial^j A}{\partial X^j}(\alpha,\alpha) = 0$;
2. $\mathrm{H}(A) \leq e^{C\ell}$ *avec C une constante indépendante de ℓ.*

Démonstration. Écrivons $P(X) = \sum_{i=0}^{r} a_i X^i$ et $Q(X) = \sum_{i=0}^{r} b_i X^i$ où les a_i et les b_i sont nos inconnues dans $\mathbb{Z}$. Nous voulons résoudre le système suivant

$$\forall j \in \{0,...,\ell-1\}, \quad \sum_{k=j}^{r} \binom{k}{j} \alpha^{k-j} a_k \; + \sum_{k=j}^{r} \binom{k}{j} \alpha^{k-j+1} b_k = 0.$$

Or ici les coefficients sont à valeurs dans $\mathbb{Z}[\alpha]$. Cependant pour pouvoir utiliser le lemme de Siegel nous devons avoir un système d'équations linéaires avec des coefficients dans $\mathbb{Q}$.

Posons $\mathbb{K} = \mathbb{Q}(\alpha)$, on sait que $1, \alpha, ..., \alpha^{d-1}$ est une $\mathbb{Q}$-base de $\mathbb{K}$. Puis en utilisant le même raisonnement que celui fait à l'étape 1 de la démonstration du lemme 2, on trouve une constante c_1 ne dépendant que de $\mathbb{K}$ telle que si $x = x_1 + ... + x_d\alpha^{d-1} \in \mathbb{K}$, alors

$$\forall k \in \{1,...,d\}, \quad |x_k| \leq c_1 \overline{|x|}.$$

Ensuite pour $i \in \mathbb{N}$, on pose

$$\alpha^i = \sum_{k=0}^{d-1} A_{i,k}\alpha^k, \quad A_{i,k} \in \mathbb{Q}\,.$$

Ainsi le système initial se transforme en

$$\forall j \in \{0, ..., \ell - 1\}, \quad \sum_{k=0}^{d-1} \alpha^k \left(\sum_{i=j}^{r} \binom{i}{j} A_{i-j,k} a_i + \binom{i}{j} A_{i-j+1,k} b_i \right) = 0.$$

Par $\mathbb{Q}$ liberté de $1, \alpha, ..., \alpha^{d-1}$ on obtient un système de $d\ell$ équations et $2(r+1) \geq (1+\varepsilon)d\ell > dl$ inconnues. Par ailleurs les coefficients sont les $\binom{i}{j}A_{i-j,k}$ et $\binom{i}{j}A_{i-j+1,k}$ avec $j = 0, 1, ..., \ell - 1$, $i = j, ..., r$ et $k = 0, ..., d-1$ qui sont rationnels. De plus on a les majorations suivantes :

$$\left| \binom{i}{j} A_{i-j,k} \right| \leq 2^i |A_{i-j,k}| \leq 2^r c_1 |\overline{\alpha}|^{i-j} \leq 2^r c_1 (|\overline{\alpha}|^*)^{r+1}$$

et de manière analogue

$$\left| \binom{i}{j} A_{i-j+1,k} \right| \leq 2^r c_1 (|\overline{\alpha}|^*)^{r+1}.$$

Le lemme de Siegel nous permet donc d'avoir de tels polynômes $P(X)$, $Q(X)$ non tous les deux nuls et on a également

$$\begin{aligned}
\log(\mathrm{H}(A)) &\leq \log\left(1 + \left(2(r+1)2^r c_1 (|\overline{\alpha}|^*)^{r+1}\right)^{\frac{d\ell}{2(r+1)-d\ell}}\right) \\
&\leq \frac{d\ell}{2(r+1)-d\ell} \log((c_2^r(r+1)) + \log(2) \qquad\qquad (c_2 = 4c_1^*(|\overline{\alpha}|^*)^2 > 1) \\
&\leq \varepsilon^{-1} r(\log(c_2) + \frac{\log(r+1)}{r} + \varepsilon \frac{\log(2)}{r}) \\
&\leq c_3 \ell
\end{aligned}$$

avec $c_3 = (\log(c_2) + 1 + \varepsilon)(1+\varepsilon)\frac{d}{2\varepsilon}$. D'où $\mathrm{H}(A) \leq e^{c_3 \ell}$. □

On considérera ce polynôme avec $C = c_3$ dans la suite, même si l'on sera amené aussi à supposer q_1 et q_2 grands relativement à c_3. Enfin on écrira

$$P(X) = \sum_{k=0}^{r} a_k X^k \quad et \quad Q(X) = \sum_{k=0}^{r} b_k X^k.$$

Étape 2 : Construction de γ.

Pour construire un γ convenable, on aimerait prendre $\gamma = (j!)^{-1}\partial_1^j A(p_1/q_1, p_2/q_2)$ avec $j \leq \delta\ell + 1$ où δ dépend de q_1, et qui est petit lorsque q_1 est grand. En effet lorsque $\delta\ell < \ell$ on pourra, comme dans la démonstration de Liouville, utiliser un théorème d'analyse pour majorer $|\gamma|$. Le facteur $(j!)^{-1}$ permet de diminuer la taille des coefficients de $\partial_1^j A$ tout en les gardant entiers.

Pour cela il nous faut d'abord étudier le lien entre $\mathrm{H}(A)$ et l'ordre auquel il s'annule en un point de coordonnées rationnels. Au vu de la ressemblance de $A(X,Y)$ avec un polynôme à une variable, dans la suite nous allons nous intéresser aux polynômes de $\mathbb{Z}[X]$, puis nous adapterons au cas particulier de $A(X,Y)$.

Prenons un rationnel $r = p/q$ sous forme irréductible et un entier n. Le polynôme P le plus simple de $\mathbb{Z}[X]$ qui annule r à l'ordre n est $P(X) = (qX - p)^n$. Ainsi, le coefficient dominant de P est q^n tandis que le coefficient constant est p^n. On en déduit que

$$\mathrm{H}(r)^n \leq \mathrm{H}(P).$$

Pouvons nous alors trouver un polynôme non nul $P \in \mathbb{Z}[X]$ avec des coefficients plus petits ? La proriété qui suit nous en prouve l'impossibilité.

Propriété 10. *Soit $P \in \mathbb{Z}[X]$. Si $P^{(j)}(r) = 0$ pour $j \in \{0, ..., n-1\}$, alors*

$$P(X) = (qX - p)^n P_1(X)$$

avec $P_1 \in \mathbb{Z}[X]$.

Démonstration. Tout d'abord pour un polynôme $Q \in \mathbb{Z}[X]$ on définit son contenu noté $\mathrm{C}(Q)$ comme le pgcd de ses coefficients. Par le lemme de Gauss nous savons que

$$\forall P, Q \in \mathbb{Z}[X], \quad \mathrm{C}(PQ) = \mathrm{C}(P)\,\mathrm{C}(Q).$$

Étant donné $Q \in \mathbb{Z}[X]$ qui annule r à l'ordre n, on sait alors qu'il existe $Q_1 \in \mathbb{Q}[X]$ tel que

$$Q(X) = (qX - p)^n Q_1(X).$$

Par ailleurs prenons $a \in \mathbb{N}^*$ tel que $aQ_1(X) \in \mathbb{Z}[X]$. Nous avons alors l'égalité de contenus suivante :

$$a\,\mathrm{C}(Q) = \mathrm{C}\left((qX - p)^n\right)\mathrm{C}(aQ_1).$$

Or comme q^n et p^n sont premiers entre eux et sont respectivement le coefficient dominant et constant de $(qX-p)^n$, alors $\mathrm{C}\left((qX-p)^n\right) = 1$ et donc a divise $\mathrm{C}(aQ_1)$ d'où $Q_1 \in \mathbb{Z}[X]$. □

Nous pouvons alors en déduire immédiatement le corollaire suivant.

Corollaire 5. *Soit $P \in \mathbb{Z}[X]$ non nul. Si $P^{(j)}(r) = 0$ pour $j \in \{0, ..., n-1\}$, alors*

$$\mathrm{H}(r)^n \leq \mathrm{H}(P).$$

Ainsi pour annuler un nombre rationnel à un grand ordre cela demande un polynôme avec des coefficients entiers grands. Maintenant adaptons ce résultat pour $A(X, Y)$ afin de construire γ.

Lemme 15. *Si $\log(q_1) > 2c_3$, alors il existe un entier $j \leq \frac{c_4}{\log(q_1)}\ell + 1$ où $c_4 := 2\big((1+\varepsilon)d + c_3\big)$ tel que*

$$\partial_1^j A\left(\frac{p_1}{q_1}, \frac{p_2}{q_2}\right) \neq 0.$$

Démonstration. Pour pouvoir se ramener à un polynôme à une variable, on considère le wronskien $W(X)$ associé à $P(X)$ et $Q(X)$:

$$W(X) = \begin{vmatrix} P(X) & Q(X) \\ P'(X) & Q'(X) \end{vmatrix} = P(X)Q'(X) - Q(X)P'(X)$$

Ensuite notons j l'ordre de p_1/q_1 par rapport au polynôme $A(X, p_2/q_2)$. Nous pouvons donc écrire

$$A\left(X, \frac{p_2}{q_2}\right) = R(X)(X - \frac{p_1}{q_1})^j \qquad (R(X) \in \mathbb{Q}[X]).$$

D'où

$$\begin{aligned} P(X) &= \frac{p_2}{q_2}Q(X) + R(X)(X - \frac{p_1}{q_1})^j, \\ P'(X) &= \frac{p_2}{q_2}Q'(X) + \left(jR(X) + R'(X)(X - \frac{p_1}{q_1})\right)(X - \frac{p_1}{q_1})^{j-1}. \end{aligned}$$

En injectant ces égalités dans $W(X)$ on obtient

$$W(X) = \left(R(X)(X - \frac{p_1}{q_1}) - jR(X) + R'(X)(X - \frac{p_1}{q_1})\right)(X - \frac{p_1}{q_1})^{j-1}.$$

Ceci nous permet d'affirmer que p_1/q_1 est au moins d'ordre $j - 1$ par rapport à W. Par ailleurs

$$PQ' = \sum_{n=0}^{2r} \left(\sum_{k=0}^{n} a_{n-k}\, k b_k \right) X^n \quad (a_k, b_k = 0 \ pour \ k \geq r + 1)$$

et

$$\begin{aligned} \left| \sum_{k=0}^{n} a_{n-k}\, k b_k \right| &\leq \sum_{k=0}^{n} |a_{n-k}\, k b_k| \\ &\leq \sum_{k=0}^{r} \mathrm{H}(P)\, \mathrm{H}(Q) r \\ &\leq 2r^2\, \mathrm{H}(A)^2. \end{aligned}$$

En raisonnant de manière analogue pour $P'Q$, on a

$$\begin{aligned} \mathrm{H}(W) &\leq \mathrm{H}(PQ') + \mathrm{H}(P'Q) \\ &\leq 4r^2\, \mathrm{H}(A). \end{aligned}$$

Maintenant nous allons distinguer deux cas.

• Premier cas : $W \neq 0$.

D'après le corollaire 5 W étant un polynôme non nul à coefficients entiers qui annule p_1/q_1 jusqu'à l'ordre $j-1$, alors on a l'inégalité

$$q_1^{j-1} \leq \mathrm{H}(p_1/q_1)^{j-1} \leq \mathrm{H}(W) \leq 4r^2\,\mathrm{H}(A)^2.$$

En passant au logarithme on obtient

$$\begin{aligned} j &\leq 1 + \frac{2}{\log(q_1)}(\log(2r) + \log(\mathrm{H}(A))) \\ &\leq 1 + \frac{2}{\log(q_1)}(2r + c_3\ell) \\ &\leq 1 + \frac{c_4\ell}{\log(q_1)} \end{aligned}$$

avec $c_4 = 2\big((1+\varepsilon)d + c_3\big)$. Ceci conclut le lemme dans ce cas.

• Second cas : $W = 0$.

Si $Q = 0$, alors on a l'égalité $A(X,Y) = P(X)$. De plus comme A est non nul, on en déduit que P est un polynôme non nul à coefficients entiers qui annule p_1/q_1 jusqu'à l'ordre j et donc par le corollaire 5

$$j \leq \frac{\log(\mathrm{H}(A))}{\log(q_1)} \leq \frac{c_3\ell}{\log(q_1)} \leq \frac{2((1+\varepsilon)d + c_3)\ell}{\log(q_1)} \leq 1 + \frac{c_4\ell}{\log(q_1)}.$$

Si $Q \neq 0$, le wronskien étant le numérateur de la dérivée de P/Q, les deux polynômes sont colinéaires dans $\mathbb{Q}$. Il existe donc $a \in \mathbb{Z}$ et $b \in \mathbb{N}^*$ premiers entre eux tels que $aQ(X) = bP(X)$. On peut donc réécrire $bA(X,Y) = (a - bY)Q(X)$. Si $a/b \neq p_2/q_2$, alors $a - (p_2/q_2)b \neq 0$ et donc Q est un polynôme non nul à coefficients entiers annulant p_1/q_1 à l'ordre j. Par le corollaire 5 on a donc

$$j \leq \frac{\log(\mathrm{H}(A))}{\log(q_1)} \leq 1 + \frac{c_4\ell}{\log(q_1)}.$$

En revanche si $a = p_2$ et $b = q_2$, Q étant non nul à coefficients entiers, on a

$$q_2 \leq \mathrm{H}(p_2/q_2) \leq \mathrm{H}(A)$$

et donc

$$\log(q_2) \leq \log(\mathrm{H}(A)) \leq c_3\left(\frac{\log(q_2)}{\log(q_1)} + 1\right) \leq 2c_3\frac{\log(q_2)}{\log(q_1)}$$

car $(\log(q_2)/\log(q_1)) \geq 2$. Ainsi comme $\log(q_1) > 2c_3$ on a une contradiction, d'où $a/b \neq p_2/q_2$. En conclusion dans tous les cas on obtient que

$$j \leq 1 + \frac{c_4\ell}{\log(q_1)}.$$

Puis par définition de j

$$\partial_1^j A\left(\frac{p_1}{q_1}, \frac{p_2}{q_2}\right) \neq 0.$$

□

Remarque 12. *Plus précisément dans ce lemme nous avons en fait donné une condition nécessaire pour qu'un polynôme non nul de la forme $R(X,Y) = R_0(X) - YR_1(X) \in \mathbb{Z}[X,Y]$ s'annule en un point $x = (x_1, x_2) \in \mathbb{Q}^2$ à l'ordre m qui est*

$$\mathrm{H}(R) \geq \min([\sqrt{2}\max(\deg(P), \deg(Q))]^{-1}\,\mathrm{H}(x_1)^{\frac{m-1}{2}},\, \mathrm{H}(x_2)).$$

Comme nous pouvons prendre q_1 aussi grand qu'on veut, nous pouvons supposer vérifiées les hypothèses du lemme et avoir l'existence d'un tel j. Nous imposons également que $\log(q_1) > 2c_4$ et comme $\ell \geq 2$ on a $j < \ell$. Nous définissons ainsi

$$\gamma := \frac{1}{j!}\partial_1^j A\left(\frac{p_1}{q_1}, \frac{p_2}{q_2}\right) \neq 0.$$

Étape 3 : Minoration de $|\gamma|$***.***

Écrivons explicitement γ

$$\gamma = P^{(j)}\left(\frac{p_1}{q_1}\right) - \frac{p_2}{q_2}Q^{(j)}\left(\frac{p_1}{q_1}\right) = \sum_{k\geq j}^{r} a_k \binom{k}{j}\left(\frac{p_1}{q_1}\right)^{k-j} + \frac{p_2}{q_2}\sum_{k\geq j}^{r} b_k\binom{k}{j}\left(\frac{p_1}{q_1}\right)^{k-j}.$$

Les $a_k\binom{k}{j}$ et les $b_k\binom{k}{j}$ étant des entiers, on en déduit donc que $\gamma q_1^{r-j} q_2$ est un entier non nul d'où

$$|\gamma| \geq q_1^{j-r} q_2^{-1}.$$

Étape 4 : Majoration de $|\gamma|$***.***

Pour simplifier les notations posons

$$R(X,Y) = \frac{1}{j!}\partial_1^j A(X,Y).$$

Comme $j < \ell$ alors nous savons que $\partial_1^k R(\alpha,\alpha) = 0$ pour $k = 0, ..., \ell - j - 1$. Ainsi d'après la formule de Taylor et comme $R(X,\alpha)$ est un polynôme de degré $\leq r - j$, on a

$$\begin{aligned}
|R(p_1/q_1,\alpha) - R(\alpha,\alpha)| &= \left|\sum_{k=\ell-j}^{r-j} \frac{(p_1/q_1 - \alpha)^k}{k!}\partial_1^k R(\alpha,\alpha)\right| \\
&= \left|\sum_{k=\ell}^{r} \frac{(p_1/q_1 - \alpha)^{k-j}}{(k-j)!j!}\partial_1^k A(\alpha,\alpha)\right| \\
&= \left|\sum_{k=\ell}^{r} (p_1/q_1 - \alpha)^{k-j}\binom{k}{j}\frac{1}{k!}\partial_1^k A(\alpha,\alpha)\right| \\
&\leq q_1^{-\rho(\ell-j)}\sum_{k=\ell}^{r} 2^k\left|\frac{1}{k!}\partial_1^k A(\alpha,\alpha)\right| \qquad (|p_1/q_1 - \alpha| < q_1^{-\rho} \leq 1).
\end{aligned}$$

Par ailleurs en explicitant $|\frac{1}{k!}\partial_1^k A(\alpha,\alpha)|$, on obtient

$$\left|\frac{1}{k!}\partial_1^k A(\alpha,\alpha)\right| \leq 2(\alpha^*)^{r+1}2^{r+1}\,\mathrm{H}(A).$$

D'où

$$|R(p_1/q_1,\alpha) - R(\alpha,\alpha)| \leq q_1^{-\rho(\ell-j)}\,\mathrm{H}(A)(\alpha^*)^{r+1}2^{2r+3} \leq q_1^{-\rho(\ell-j)}\,\mathrm{H}(A)c_5^r,$$

avec $c_5 := 2^5(|\alpha|^*)^2$.

Ensuite majorons $|R(p_1/q_1,p_2/q_2) - R(p_1/q_1,\alpha)|$:

$$\begin{aligned}|R(p_1/q_1,p_2/q_2) - R(p_1/q_1,\alpha)| &= |p_2/q_2 - \alpha|\ \frac{1}{j!}|Q^{(j)}(p_1/q_1)| \\ &\leq q_2^{-\rho}2^{r+1}(|\alpha|+1)^r\,\mathrm{H}(A) \\ &\leq q_2^{-\rho}c_6^r\,\mathrm{H}(A).\end{aligned}$$

avec $c_6 := 2^2(|\alpha|+1)$. Nous pouvons obtenir une majoration sur γ :

$$\begin{aligned}|\gamma| &= |R(p_1/q_1,p_2/q_2)| \\ &= |R(p_1/q_1,p_2/q_2) - R(\alpha,\alpha)| && (R(\alpha,\alpha)=0) \\ &\leq |R(p_1/q_1,p_2/q_2) - R(p_1/q_1,\alpha)| + |R(p_1/q_1,\alpha) - R(\alpha,\alpha)| \\ &\leq \mathrm{H}(A)(c_6^r q_2^{-\rho} + c_5^r q_1^{-\rho(\ell-j)}) \\ &\leq e^{c_7\ell}(q_2^{-\rho} + q_1^{-\rho(\ell-j)})\end{aligned}$$

avec $c_7 := c_3 + \frac{(1+\varepsilon)d}{2}\max(\log(c_5),\log(c_6),0)$.

Étape 5 : Conclusion.

Grâce aux étapes 3 et 4 nous avons l'encadrement suivant pour $|\gamma|$:

$$q_2^{-1}q_1^{-(r-j)} \leq |\gamma| \leq e^{c_7\ell}(q_2^{-\rho} + q_1^{-\rho(\ell-j)}).$$

Ceci nous donne

$$e^{c_7\ell} \geq q_2^{\rho-1}q_1^{\rho\ell-r-j(\rho-1)} \geq q_2^{\rho-1}q_1^{-r-j(\rho-1)}.$$

Puis en passant au logarithme et en utilisant que $j \leq 1 + (c_4/\log(q_1))\ell$, on obtient

$$c_7\ell \geq (\rho-1)\log(q_2) - \log(q_1)\left((\rho-1)(1+\frac{c_4\ell}{\log(q_1)}) + r\right).$$

D'où

$$\frac{c_7 + (\rho-1)c_4}{\log(q_1)}\ell \geq (\rho-1)\frac{\log(q_2)}{\log(q_1)} - (\rho-1+r).$$

Enfin, en supposant (comme on le peut) $\log(q_1) > \frac{\varepsilon}{c_7+(\rho-1)c_4}$, on a

$$\begin{aligned}\rho - 1 &\geq (\rho-1)\frac{\log(q_2)}{\log(q_1)} - (1+\varepsilon)\frac{d\ell}{2} - \ell\,\varepsilon \\ &\geq (\rho-1)(\ell-1) - (1+2\,\varepsilon)\frac{d\ell}{2} \qquad\qquad (\ell = \left\lfloor \frac{\log(q_2)}{\log(q_1)} \right\rfloor + 1,\ d \geq 2).\end{aligned}$$

On en déduit l'inégalité

$$\ell \leq \frac{2(\rho-1)}{\rho - 1 - (1+2\,\varepsilon)(d/2)} \qquad (\rho > 1 + (1+3\,\varepsilon)(d/2) > 1 + (1+2\,\varepsilon)(d/2)).$$

Cela impose alors que $\log(q_2)/\log(q_1)$ soit borné ce qui contredit l'existence d'une infinité de solutions rationnelles à l'inéquation (5) car sinon on pourrait prendre $\log(q_2)/\log(q_1)$ aussi grand qu'on veut. En conclusion l'inéquation (5) admet un nombre fini de solutions rationnelles.

Remarque 13. *Lorsqu'on analyse la démonstration,on voit qu'on a également montré que si nous avons une solution p_1/q_1 de l'inéquation* (5) *telle que* $\log(q_1) > \max(\frac{\varepsilon}{c_7+(\rho-1)c_4}, 2c_3, 2c_4)$, *alors pour toute solution rationnelle p_2/q_2 de l'inéquation telle que $q_2 > q_1$*

$$\frac{\log(q_2)}{\log(q_1)} \leq \frac{2(\rho-1)}{\rho - 1 - (1+2\,\varepsilon)(d/2)}.$$

6 Amélioration de Siegel

Bien que le théorème de Thue soit une amélioration notable par rapport au théorème de Liouville, cependant il n'est toujours pas optimal. Ainsi, l'exposant sera amélioré dans le temps avec des démonstrations généralisant toujours plus la méthode de Thue. Nous avons notamment celle de Siegel démontrée dans [6].

Théorème 12. (Siegel) *Soit α un réel algébrique de degré $d \geq 2$. Si $\rho > \min_{1\leq s\leq d-1}\left(s + \frac{d}{s+1}\right)$, alors il existe un nombre fini de solutions rationnelles $\frac{p}{q}$ solutions de*

$$\left|\alpha - \frac{p}{q}\right| < \frac{1}{q^\rho}.$$

6.1 Preuve de Siegel

Dans la suite nous allons voir la démonstration de Siegel qui améliore donc l'exposant $d/2 + 1$ en $2\sqrt{d}$. Elle se différencie de celle de Thue par la construction de la fonction auxiliaire qui est plus générale. En effet, plutôt que de s'intéresser à un polynôme $A(X,Y)$ de degré 1 selon la variable Y, Siegel considère un polynôme $A(X,Y)$ de degré au plus $d-1$

selon Y. Cependant, cela va compliquer la construction d'un rationnel non nul convenable γ.

Étape 0 : Choix des paramètres.

Soit $\rho > \min_{1\leq s\leq d-1}\left(s+\frac{d}{s+1}\right)$. Nous allons raisonner par l'absurde. Supposons donc qu'il existe une infinité de rationnels p/q vérifiant l'inégalité suivante :

$$\left|\alpha - \frac{p}{q}\right| < \frac{1}{q^\rho}. \tag{6}$$

Considérons encore deux approximations rationnels irréductibles (p_1/q_1) et (p_2/q_2) vérifiant (6). Par hypothèse, on peut prendre q_1, q_2 et $\log(q_2)/\log(q_1)$ arbitrairement grands. Pour l'instant, on les prendra tous les trois $\geq d$. Puis considérons s comme l'entier entre 1 et $d-1$ tel que $s+\frac{d}{s+1} = \min_{1\leq k\leq d-1}(k+\frac{d}{k+1})$. Ensuite prenons un réel $\varepsilon > 0$ qu'on prendra arbitrairement petit. Enfin définissons les entiers :

$$\ell := \left\lfloor \frac{\log(q_2)}{\log(q_1)} \right\rfloor, \quad r := \left\lfloor \frac{(1+\varepsilon)d}{s+1}\ell \right\rfloor.$$

On remarque que

$$(r+1)(s+1) > (1+\varepsilon)d\ell > d\ell.$$

Étape 1 : Construction de la fonction auxiliaire.

La construction de la fonction auxiliaire de Siegel se fait de la même manière que celle Thue. On va donc encore construire un polynôme à deux variables $A(X,Y)$ à coefficients entiers tel que $A(X,\alpha)$ annule α jusqu'au moins l'ordre ℓ. Seulement cette fois on autorise $A(X,Y)$ à avoir un degré selon Y inférieur ou égal à s.

Lemme 16. *Il existe* $(a_{k,h})_{\substack{0\leq k\leq r\\ 0\leq h\leq s}}$ *des entiers non tous nuls tels que si* $A(X,Y) := \sum_{k=0}^{r}\sum_{h=0}^{s} a_{k,h}X^kY^h \in \mathbb{Z}[X,Y]$ *alors :*

1. $\forall j \in \{0,...,\ell-1\},\ \frac{1}{j!}\partial_1^j A(\alpha,\alpha) = 0$ *;*
2. $\mathrm{H}(A) \leq e^{C\ell}$ *avec* C *une constante indépendante de* ℓ*.*

Démonstration. On procède de la même manière que pour le théorème de Thue. Ainsi démontrer le lemme revient à trouver une solution dans $\mathbb{Z}$ non triviale du système linéaire suivant :

$$\forall j \in \{0,...,\ell-1\},\ \sum_{k=j}^{r}\sum_{h=0}^{s}\binom{k}{j}a_{k,h}\alpha^{k-j+h} = 0,$$

où les inconnues sont les $a_{k,h}$, et les coefficients sont dans $\mathbb{Q}(\alpha)$. Puis comme on l'avait fait précédemment, on se ramène à un autre système linéaire de $d\ell$ équations et $(r+1)(s+1) > d\ell$ inconnues où les coefficients sont des entiers majorés en module par c_1^r où c_1 est une constante > 1 ne dépendant que de α et s.

Le lemme de Siegel nous permet alors d'affirmer l'existence d'entiers $a_{k,h}$ non tous nuls qui vérifient le point 1. du lemme 16. De plus, le lemme de Siegel nous donne également la majoration

$$\begin{aligned}\log(\mathrm{H}(A)) &\leq \log\left(1+\left((s+1)(r+1)c_1^r\right)^{\frac{d\ell}{(s+1)(r+1)-d\ell}}\right)\\ &\leq \frac{d\ell}{(s+1)(r+1)-d\ell}\log\left((s+1)(r+1)c_1^r\right)+\log(2)\\ &\leq \varepsilon^{-1}\,r\left(\log(c_1)+\frac{\log(s+1)}{r}+\frac{\log(r+1)}{r}+\frac{\varepsilon\log(2)}{r}\right)\\ &\leq c_2\ell,\end{aligned}$$

où $c_2 := \frac{(1+\varepsilon)d}{\varepsilon(s+1)}\big(\log(c_1(s+1))+\varepsilon\log(2)\big)$ ce qui démontre le lemme. □

Dans la suite, on considérera le polynôme $A(X,Y) := \sum_{k=0}^{r}\sum_{h=0}^{s} a_{k,h}X^kY^h$ respectant les conditions du lemme et en prenant c_2 comme constante C, même si on sera amené à prendre $\log(q_2)/\log(q_1)$ grand.

Étape 2 : Construction de γ.

La construction de γ repose sur la même idée que pour Thue. On montre que l'ordre j auquel le polynôme $A(X,p_2/q_2)$ annule p_1/q_1 est petit par rapport à ℓ et puis on pose $\gamma := \frac{1}{j!}\partial_1^j A(p_1/q_1,p_2/q_2)$. Pour pouvoir majorer j, il faut se ramener à majorer l'ordre d'un polynôme non nul à coefficient entier en p_1/q_1 et conclure avec le corollaire 5. Pour cela on va considérer un wronskien et donc donnons tout d'abord une condition suffisante pour qu'un wronksien ne soit pas nul pour une famille finie de fractions rationnelles.

Lemme 17. *Soit $\mathbb{K}$ est un corps de caractéristique nulle et $f_1,...,f_n \in \mathbb{K}(X)$. Si $f_1,...,f_n$ sont libres sur $\mathbb{K}$ alors*

$$W(f_1,...,f_n) := \det\big(M[f_1,....,f_n]\big) \neq 0,$$

où $M[f_1,...,f_n] = (f_j^{(i-1)})_{1\leq i,j\leq n}$.

Démonstration. Montrons par récurrence que l'assertion $\mathcal{P}(n)$ suivante est vraie pour tout $n \in \mathbb{N}^*$:

$$\forall f_1,...,f_n \in \mathbb{K}(X), \quad (f_1,...,f_n \ \mathbb{K}\text{-libres}) \Longrightarrow W(f_1,...,f_n) \neq 0.$$

Tout d'abord l'assertion est immédiate pour $n=1$. Maintenant prenons un entier $n \geq 2$ et supposons que $\mathcal{P}(n-1)$ est vérifiée. Prenons donc $f_1,...,f_n \in \mathbb{K}(X)$ libres sur $\mathbb{K}$.

Dans un premier temps si pour tout $i \in \{1,...,n\}$ on a $f_i^{(n-1)} = 0$ alors tous les f_i sont des polynômes de degré $\leq n-2$. Ainsi $f_1,...,f_n$ est une famille de n polynômes de degré $n-2$ et donc par un argument de dimension elle n'est pas libre sur $\mathbb{K}$ ce qui contredit l'hypothèse initiale des f_i. On a donc i_0 tel que $f_{i_0}^{(n-1)} \neq 0$. Quitte à changer les indices, supposons sans perte de généralité que $i_0 = n$.

Ensuite, prenons $(R_1, ..., R_n)^T \in \mathbb{K}(X)^n$ tel que

$$M[f_1, ..., f_n](R_1, ..., R_n)^T = 0.$$

Alors comme $f_n^{(n-1)} \neq 0$ (donc $f_n^{(k)} \neq 0$ pour $k \leq n-1$) nous avons

$$\forall h \in \{0, ..., n-1\}, \quad \sum_{k=1}^{n-1} R_k \frac{f_k^{(h)}}{f_n^{(h)}} = R_n.$$

Maintenant, posons $g_k = \left(\frac{f_k}{f_n}\right)'$ pour $k = 1, ..., n-1$. Alors pour $h \in \{0, ..., n-2\}$

$$\begin{aligned}
A_h &:= \sum_{k=1}^{n-1} R_k g_k^{(h)} \\
&= \sum_{k=1}^{n-1} \sum_{j=0}^{h+1} \binom{h+1}{j} R_k \left(\frac{1}{f_n}\right)^{(h+1-j)} f_k^{(j)} && \text{(Formule de Leibniz)} \\
&= \sum_{j=0}^{h+1} \binom{h+1}{j} \left(\frac{1}{f_n}\right)^{(h+1-j)} f_n^{(j)} \sum_{k=1}^{n-1} R_k \frac{f_k^{(j)}}{f_n^{(j)}} \\
&= \left(\frac{f_n}{f_n}\right)^{(h+1)} R_n \\
&= 0 && (h+1 \geq 1).
\end{aligned}$$

D'où

$$M[g_1, ..., g_{n-1}](R_1, ..., R_{n-1})^T = (A_0, ..., A_{n-2})^T = 0.$$

Enfin, si $\lambda_1, ..., \lambda_{n-1} \in \mathbb{K}$ tels que

$$\lambda_1 g_1 + ... + \lambda_{n-1} g_{n-1} = 0$$

alors

$$\left(\frac{\lambda_1 f_1 + ... + \lambda_{n-1} f_{n-1}}{f_n}\right)' = 0.$$

Comme $\mathbb{K}$ est de caractéristique nulle alors $\frac{\lambda_1 f_1 + ... + \lambda_{n-1} f_{n-1}}{f_n}$ est un polynôme de degré 0. Par la $\mathbb{K}$ liberté des f_i on obtient la nullité des λ_i et donc de la $\mathbb{K}$ liberté des g_i. Ainsi par hypothèse de récurrence $M[g_1, ..., g_{n-1}]$ est inversible et donc

$$\forall k \in \{1, ..., n-1\}, \quad R_k = 0.$$

Puis la non nullité de f_n implique $R_n = 0$. En conclusion, $M[f_1, ..., f_n]$ est injective donc inversible. Ceci prouve que $\mathcal{P}(n)$ est vraie. □

Maintenant on va chercher une meilleur écriture de $A(X, Y)$ pour étudier plus facilement l'ordre de p_1/q_1.

Lemme 18. *Il existe un entier $t \leq s$ et $U_0(X), .., U_t(X) \in \mathbb{Q}[X]$ et $V_1(Y), ..., V_t(Y) \in \mathbb{Q}[Y]$ tels que :*

1. $A(X,Y) = \sum_{k=0}^{t} U_k(X)V_k(Y)$;
2. $U_0, ..., U_t$ *libres sur* $\mathbb{Q}$;
3. $V_0, ..., V_t$ *libres sur* $\mathbb{Q}$.

Démonstration. Considérons l'ensemble

$$\mathcal{P} = \{h \in \mathbb{N} / \exists U_0, ..., U_h \in \mathbb{Q}[X], \exists V_0, ..., V_h \in \mathbb{Q}[Y],\ A(X,Y) = \sum_{k=0}^{h} U_k(X)V_k(Y)\}$$

Comme

$$A(X,Y) = \sum_{h=0}^{s} \big(\sum_{k=0}^{r} a_{k,h}X^h\big)Y^h,$$

alors $s \in \mathcal{P}$ qui est donc non vide. Ainsi on peut poser $t := \min(\mathcal{P})$. Par définition $t \leq s$. Ensuite prenons les $U_0, ..., U_t$ et $V_0, ..., V_t$ associés à t. Supposons que les $U_0, ..., U_t$ ne sont pas libres sur $\mathbb{Q}$. Alors quitte à changer les indices on peut supposer sans perte de généralité

$$U_t = \lambda_0 U_0 + ... + \lambda_{t-1}U_{t-1} \quad (\lambda_i \in \mathbb{Q}).$$

Alors

$$A(X,Y) = \sum_{k=0}^{t-1} U_k(X)\big(V_k(Y) - \lambda_k V_t(Y)\big).$$

Or ceci contredit la minimalité de t, donc $U_0, ..., U_t$ sont libres sur $\mathbb{Q}$. De manière analogue on démontre que $V_0, ..., V_t$ sont également libres sur $\mathbb{Q}$. □

Maintenant nous avons tous les outils nécessaires pour majorer convenablement l'ordre de $A(X, p_2/q_2)$ en $X = p_1/q_1$.

Lemme 19. *Si* $\log(q_1) > c_2$ *alors il existe* $j \leq \varepsilon\, d\ell + d^2$ *tel que*

$$\partial_1^j A(\frac{p_1}{q_1}, \frac{p_2}{q_2}) \neq 0.$$

Démonstration. Notons j l'ordre de $A(X, p_2/q_2)$ en p_1/q_1.

Tout d'abord remarquons que $A(X, p_2/q_2)$ est non nul car sinon cela impliquerait que

$$\forall k \in \{0, ..., r\}, \quad \sum_{h=0}^{s} a_{k,h}\left(\frac{p_2}{q_2}\right)^h = 0,$$

et comme $A(X,Y)$ est non nul à coefficients entiers par le corollaire 5 on avait que

$$q_2 \leq \mathrm{H}(A).$$

Or ceci donne $\log(q_1) \leq c_2$ et contredit l'hypothèse sur q_1. D'où $A(X, p_2/q_2) \neq 0$.

Ensuite, prenons $t, U_0, ..., U_t, V_0, ..., V_t$ comme dans le lemme 18. La non nullité de $A(X, p_2/q_2)$ impose que l'un des $V_i(p_2/q_2)$ est non nul. Sans perte de généralité disons $V_0(p_2/q_2) \neq 0$. Par ailleurs, posons

$$W(X) := W(U_0, ..., U_t)(X) \in \mathbb{Q}[X].$$

Les U_i étant libres sur $\mathbb{Q}$ qui est de caractéristique nulle, par le lemme 17

$$W(X) \neq 0.$$

Remarquons alors que

$$W(X)V_0(Y) = \begin{vmatrix} A(X,Y) & U_1(X) & ... & U_t(X) \\ . & . & ... & . \\ . & . & ... & . \\ . & . & ... & . \\ \partial_1^t A(X,Y) & U_1^{(t)}(X) & ... & U_t^{(t)}(X) \end{vmatrix}.$$

On pose v l'ordre de $W(X)$ en p_1/q_1. Comme $V_0(p_2/q_2) \neq 0$, en utilisant la formule du multinôme on obtient

$$j \leq t + v.$$

Maintenant il faut donc majorer v en fonction de ℓ. Une difficulté est qu'a priori $W(X)$ est à coefficients rationnels donc on ne va pas pouvoir utiliser le corollaire 5. On sait déjà que $v \leq \deg(W) \leq r(t+1)$, alors pour affiner la majoration nous allons trouver des facteurs irréductibles de W sur $\mathbb{Q}$ qui n'annulent pas p_1/q_1.

Notons P_α le polynôme minimal de α sur $\mathbb{Q}$. Comme $0 \leq \deg(V_0) \leq s < d$, alors on sait que $V_0(\alpha) \neq 0$. Ainsi comme α annule $A(X, \alpha)$ jusqu'au moins l'ordre ℓ alors $W(X)$ annule α jusqu'au moins l'ordre $\ell - t \geq 0$ ($\ell \geq d \geq s \geq t$). Puis par définition de P_α, $(P_\alpha)^{\ell-t}$ divise W dans $\mathbb{Q}[X]$. Enfin, $P_\alpha(p_1/q_1) \neq 0$ d'où

$$v \leq \deg(W) - d(\ell - t) \leq r(s+1) - d(\ell - s) \leq \varepsilon\, d\ell + sd.$$

En conclusion

$$j \leq t + v \leq \varepsilon\, d\ell + s(d+1) \leq \varepsilon\, d\ell + d^2.$$

□

Dans la suite, on supposera (comme on le peut) que $\log(q_1) > c_2$ et donc on peut définir j comme dans le lemme 19 et poser

$$\gamma = \frac{1}{j!}\partial_1^j A(\frac{p_1}{q_1}, \frac{p_2}{q_2}) \neq 0.$$

De plus on supposera $\varepsilon \leq \frac{1}{2d}$ et $\ell > 2d^2$ de sorte que

$$j \leq \varepsilon\, d\ell + d^2 < \ell.$$

Étape 3 : Minoration de γ.

En écrivant explicitement γ

$$\gamma = \sum_{k=j}^{r}\sum_{h=0}^{s}\binom{k}{j}a_{k,h}\left(\frac{p_1}{q_1}\right)^{k-j}\left(\frac{p_2}{q_2}\right)^{h},$$

on remarque avec le même raisonnement que celui fait dans le chapitre 5 que

$$|\gamma| \geq q_1^{j-r}q_2^{-s}.$$

Étape 4 : Majoration de γ.

Posons pour simplifier les notations

$$R(X,Y) := \frac{1}{j!}\partial_1^j A(X,Y).$$

Tout d'abord par l'égalité des accroissements finis, il existe un réel x tel que $|x| \leq |\alpha| + 1$ et

$$\left|R\left(\frac{p_1}{q_1},\frac{p_2}{q_2}\right) - R\left(\frac{p_1}{q_1},\alpha\right)\right| = \left|\frac{p_2}{q_2} - \alpha\right| |\partial_2 R(\frac{p_1}{q_1}, x)|.$$

Puis

$$\begin{aligned}
|\partial_2 R(p_1/q_1, x)| &\leq \sum_{k=j}^{r}\sum_{h=1}^{s}\binom{k}{j}|a_{k,h}(p_1/q_1)^{k-j}hx^{h-1}| \\
&\leq (r+1)(s+1)2^r\,\mathrm{H}(A)(|\alpha|+1)^{r+s} \\
&\leq 2^{2r(s+1)+r}e^{c_2\ell}(|\alpha|+1)^{r+s} \\
&\leq e^{c_4\ell},
\end{aligned}$$

avec $c_4 = 3(1+\varepsilon)d\log(2) + c_2 + \big((1+\varepsilon)d+1\big)\log(|\alpha|+1)$. D'où

$$\left|R\left(\frac{p_1}{q_1},\frac{p_2}{q_2}\right) - R\left(\frac{p_1}{q_1},\alpha\right)\right| \leq \frac{1}{q_2^{\rho}}e^{c_4\ell}.$$

Par ailleurs en utilisant le même raisonnement que dans le chapitre 5

$$\left|R\left(\frac{p_1}{q_1},\alpha\right) - R(\alpha,\alpha)\right| \leq q_1^{-\rho(\ell-j)}\sum_{k=j}^{r}2^k\left|\frac{1}{k!}\partial_1^k A(\alpha,\alpha)\right|$$

et

$$\left|\frac{1}{k!}\partial_1^k A(\alpha,\alpha)\right| \leq (2^3\alpha^*)^r e^{c_1\ell}$$

où on rappelle que $\alpha^* = \max(|\alpha|, 1)$. D'où

$$\left|R\left(\frac{p_1}{q_1},\alpha\right) - R(\alpha,\alpha)\right| \leq q_1^{-\rho(\ell-j)}e^{c_5\ell},$$

où $c_5 = (1+\varepsilon)d\log(2^6\alpha^*) + c_1$. Finalement, en regroupant tous les termes on obtient la majoration suivante de $|\gamma|$:

$$\begin{aligned}
|\gamma| &= |R(p_1/q_1, p_2/q_2) - R(\alpha,\alpha)| && (j<\ell)\\
&\le |R(p_1/q_1, p_2/q_2) - R(p_1/q_1,\alpha)| + |R(p_1/q_1,\alpha) - R(\alpha,\alpha)|\\
&\le e^{c_6\ell}(q_1^{-\rho(\ell-j)} + q_2^{-\rho}) && (c_6 = c_4 + c_5).
\end{aligned}$$

Étape 5 : Conclusion.

A l'aide des étapes 3 et 4, on obtient l'encadrement de $|\gamma|$ suivant :

$$q_1^{j-r}q_2^{-s} \le |\gamma| \le e^{c_6\ell}(q_1^{-\rho(\ell-j)} + q_2^{-\rho}).$$

Ainsi

$$e^{c_6\ell} \ge q_1^{\rho(\ell-j)+j-r}q_2^{\rho-s} \ge q_1^{-j(\rho-1)-r}q_2^{\rho-s}.$$

En passant au logarithme, on obtient

$$\begin{aligned}
\frac{c_4}{\log(q_1)}\ell &\ge (\rho-s)\frac{\log(q_2)}{\log(q_1)} - \big(j(\rho-1)+r\big)\\
&\ge (\rho-s)\ell - \left((\varepsilon\, d\ell + d^2)(\rho-1) + \frac{1+\varepsilon}{s+1}d\ell\right) \qquad (j \le \varepsilon\, d\ell + d^2).
\end{aligned}$$

Ensuite, par définition de ρ on sait que $\frac{\rho-\left(s+d/(s+1)\right)}{3d\left(\rho+1/(s+1)\right)} > 0$ et donc comme on peut prendre ε aussi petit qu'on veut, prenons

$$0 < \varepsilon < \frac{\rho-\big(s+d/(s+1)\big)}{3d\big(\rho+1/(s+1)\big)}.$$

Par ailleurs, on peut également prendre $\log(q_1)$ aussi grand qu'on veut, donc prenons

$$\log(q_1) > \frac{c_4}{\varepsilon}.$$

Ainsi on a

$$\begin{aligned}
d^2(\rho-1) &\ge \left(\rho - (s + \frac{d}{s+1} + \frac{\varepsilon}{s+1}d + \varepsilon\, d\rho + \varepsilon)\right)\ell\\
&\ge \left(\rho - s - \frac{d}{s+1} - 2d\,\varepsilon\,\big(\frac{1}{s+1} + \rho\big)\right)\ell\\
&\ge d\,\varepsilon\left(\frac{1}{s+1} + \rho\right)\ell\\
&\ge \varepsilon(1+d\rho)\ell.
\end{aligned}$$

En conclusion, nous avons la majoration de ℓ suivante :

$$\ell \leq \frac{d^2(\rho-1)}{\varepsilon(1+d\rho)}.$$

Or ceci est absurde car on peut prendre ℓ aussi grand qu'on veut. Ainsi nous avons démontré l'amélioration de Siegel du théorème de Thue.

Pour mmieux quatifier l'important gain obtenu par Siegel, montrons cette propriété.

Propriété 11. *Soit $d \geq 4$ un entier. Alors*

$$\min_{1\leq s\leq d-1}\left(s+\frac{d}{s+1}\right) \leq 2\sqrt{d}.$$

Démonstration. Posons la fonction $f : x \in [1,+\infty[\longmapsto x + \frac{d}{x+1} \in \mathbb{R}$. Une étude de celle-ci nous donne qu'elle atteint son minimum en $y := \sqrt{d}-1$. Pour montrer la propriété il suffit alors de démontrer

$$f(\lfloor y \rfloor) \leq 2\sqrt{d} \quad \text{et} \quad f(\lfloor y \rfloor + 1) \leq 2\sqrt{d}.$$

Tout d'abord

$$\begin{aligned} f(\lfloor y \rfloor + 1) - 2\sqrt{d} &\leq y + 1 + \frac{d}{y+1} - 2\sqrt{d} \\ &\leq 0. \end{aligned}$$

Ensuite posons $\delta = \{y\}$, alors

$$\begin{aligned} f(\lfloor y \rfloor) - 2\sqrt{d} &= y - \delta + \frac{d}{y-\delta+1} - 2\sqrt{d} \\ &= g(\delta) - (1+\sqrt{d}) \end{aligned}$$

où $g : x \in [0,1] \longmapsto \frac{d}{\sqrt{d}-x} - x \in \mathbb{R}$. Or toujours par une étude de fonction, on montre que g est croissante sur $[0,1]$. D'où

$$\begin{aligned} f(\lfloor y \rfloor) - 2\sqrt{d} &\leq \frac{d}{\sqrt{d}-1} - 1 - (\sqrt{d}+1) \\ &\leq \frac{1}{\sqrt{d}-1} - 1 \\ &\leq 0 \qquad\qquad (\sqrt{d} \geq 2). \end{aligned}$$

□

Ainsi nous avons un exposant de l'ordre de $\sqrt{d}$ au lieu de d ce qui est une nette amélioration par rapport au théorème de Thue.

6.2 Améliorations ultérieures

L'exposant trouvé par Siegel n'est pas optimal et a été amélioré par la suite. Nous avons notamment le résultat suivant dû à Dyson et Gelfond.

Théorème 13. (Dyson–Gelfond) *Soit α un réel algébrique de degré $d \geq 2$. Si $\rho > \sqrt{2d}$ alors il existe un nombre fini de solutions rationnelles p/q de l'inéquation*

$$\left|\alpha - \frac{p}{q}\right| \leq \frac{1}{q^\rho}.$$

La démonstration de ce résultat que l'on peut retrouver dans [6] s'effectue toujours avec la même méthode. Seulement, on va imposer d'autres conditions d'annulation de la fonction auxiliaire $A(X, Y)$ en considérant aussi des dérivées partielles selon la variable Y. De plus ce résultat est meilleur que celui de Siegel par la propriété suivante.

Propriété 12. *Si $d \geq 3$ est un entier alors*

$$\sqrt{2d} < \min_{1 \leq s \leq d-1} \left(s + \frac{d}{s+1}\right).$$

Démonstration. En reprenant le raisonnement et les notations de la propriété 11, si on démontre que $f(\sqrt{d} - 1) > \sqrt{2d}$ alors on a le résultat. Comme

$$f(\sqrt{d} - 1) - \sqrt{2d} = 2\sqrt{d} - 1 - \sqrt{2d} \geq \sqrt{3}(2 - \sqrt{2}) - 1 > 0,$$

on conclut. □

Finalement c'est Roth qui démontra en 1955 un résultat d'approximation avec une minoration de ρ optimale.

Théorème 14. (Roth)
Soit α un réel algébrique irrationnel. Si $\rho > 2$ alors il n'existe qu'un nombre fini de solutions rationnelles p/q de l'inéquation

$$\left|\alpha - \frac{p}{q}\right| \leq \frac{1}{q^\rho}.$$

Le corollaire 3 nous montre que le théorème n'est pas vrai si $\rho = 2$ ce qui montre l'optimalité de cette minoration. On peut retrouver la démonstration dans [6] et [7]. Elle repose sur la construction d'une fonction auxiliaire qui est un polynôme $A(X_1, ..., X_m)$ à un grand nombre m d'indéterminées auquel on impose des conditions d'annulation au point $(\alpha, ..., \alpha)$. On va donc considérer m approximations rationnelles au lieu de 2. Cependant, cela rend plus difficile de trouver un rationnel non nul γ. Ce théorème valu à Roth la médaille fields en 1958.

Références

[1] F. J. Dyson, *The approximation to algebraic numbers by rationals*, Acta Mathematica **79** (1947), 225–240.

[2] A. O. Gelfond, *Sur le septième problème de D. Hilbert*, Izv. Akad. Nauk SSSR, ser. VII (1934), 623–630.

[3] A. O. Gelfond, *Approximation to algebraic numbers and their logarithms*, Vestnik Moskov. Univ. **9** (1948), 3–25, en russe.

[4] L. Guth, *Introduction to Diophantine Equations*, *Proof of Thue's Theorem – Part I*, *Proof of Thue's Theorem – Part II*, *Proof of Thue's Theorem – Part III* [notes de cours], (2025).

[5] S. Lang (1965) *Algebra*, Graduate Texts in Mathematics, Springer New-York, edition 3, 1965.

[6] M. Mignotte, *Approximations des nombres algébriques* [notes de cours], Publications mathématiques d'Orsay, (1975), N°77-74.

[7] K. F. Roth, *Rational approximations to algebraic numbers*, Mathematika **2** (1), (1955), 1–20.

[8] T. Schneider, *Transzendenzuntersuchungen periodischer Funktionen. I. Transzendenz von Potenzen* J. reine angew. Math. **172** (1935), 65–69.

[9] T. Schneider, *Introduction aux nombres transcendants*, Gauthier–Villars, (1959), 149 pages.

[10] C. L. Siegel, *Approximation algebraischer Zahlen*, Mathematische Zeitschrift **10** (3) (1921), 173–213.

[11] C. L. Siegel, *Transcendental Numbers*, Annals of Math Studies, Princeton University Press, 102 pages, 1949.