Rapport de stage de L3 - ENS de Rennes

Théorèmes classiques de la théorie de la transcendance

Ilan Zysman

Juin 2022

Ce stage de 3-ème année de licence à été réalisé à l'Institut Fourier à Grenoble sous la direction de Tanguy Rivoal, directeur de recherche au CNRS.

Je remercie, Tanguy Rivoal, pour ses précieux conseils et son implication qui me confortent dans l'idée de poursuivre dans la voie de chercheur.

Je tiens également à remercier, Bertrand Rémy, Professeur des Universités à l'ENS de Lyon, qui par sa constante disponibilité, m'a mis en relation avec le laboratoire de Grenoble.

Table des matières

Table des matières **2**
0.1 Introduction et rappels 3

1 Théorème de Lindemann-Weierstrass **8**
1.1 Énoncé du théorème et cas particuliers 8
1.2 Approximants de Padé et d'Hermite-Padé de la fonction exponentielle 9
1.3 Preuve du théorème de Lindemann-Weierstrass 14

2 Lemmes de Siegel **16**
2.1 Les résultats 16
2.2 Applications à la construction de fonctions auxiliaires 18

3 Théorème de Gelfond-Schneider **19**
3.1 Le théorème et son contexte 19
3.2 Construction de la fonction auxiliaire et déterminant de Vandermonde 20
3.3 Preuve du théorème 22

4 Critère de Schneider-Lang **24**
4.1 Enoncé et conséquences du théorème 24
4.2 Principe de la démonstration et outils nécessaires 25
4.3 Preuve du critère 29

5 Applications aux fonctions elliptiques **32**
5.1 Fonction $\wp$ de Weierstrass 32
5.2 Formes modulaires 33
5.3 Les résultats de transcendance 35

Bibliographie **45**

0.1 Introduction et rappels

Position du problème

Notre objectif est d'étudier certains des résultats les plus percutants en théorie de la transcendance, qui ont occupé la fin du XIX-ème siècle jusqu'au milieu du XX-ème siècle. Les deux résultats les plus célèbres obtenus pendant cette période sont le théorème de Lindemann-Weierstrass et le théorème de Gelfond-Schneider qui occupent respectivement les chapitres 1 et 3. Ils sont séparés par un chapitre de transition sur l'important lemme de Siegel, qui, nous le verrons, permet de s'affranchir de formules explicites. Enfin on étudiera le critère de schneider-Lang dans le chapitre 4, qui permet notamment d'obtenir des résultats intéressants sur les fonctions elliptiques que nous aborderons dans un dernier chapitre. En particulier on donnera la preuve d'un théorème récent sur l'invariant modulaire. On commence par quelques rappels et notations.

Rappels analytiques

On rappelle le théorème des résidus de Cauchy concernant les fonctions de la variable complexe :

Proposition 1. *Soit U un ouvert simplement connexe et f une fonction holomorphe sur $U - \{z_1, ..., z_n\}$. Soit γ un lacet inclus dans U ne rencontrant pas les z_i. Alors on a l'égalité :*

$$\int_\gamma f(\zeta)d\zeta = 2i\pi \sum_i Res(f, z_i)Ind(\gamma, z_i).$$

On utilisera ce résultat sous la forme :

$$f(z) = \frac{1}{2i\pi} \int_C \frac{f(\zeta)}{\zeta - z} d\zeta,$$

où C est un cercle, positivement orienté autour de z, inclus dans U.

On dit qu'une fonction f est **méromorphe** sur un ouvert U de $\mathbb{C}$, si f est holomorphe sur $U - D$ où D est discret dans U et si f admet un pôle en chaque point de D. Une fonction méromorphe sur $\mathbb{C}$ s'écrit comme quotient de deux fonctions entières. Une écriture $f = g/h$ où g, h sont entières est dite irréductible si il n'existe pas ϕ entière, non constante, telle que g/ϕ et h/ϕ soient entières. Une telle écriture est alors unique à multiplication par un scalaire près. Dans ce cas on appelle dénominateur de f, l'unique fonction h telle que $f = g/h$ soit irréductible et dont le coefficient de plus petit degré de h est 1.

On notera $\mathscr{D}$ l'opérateur dérivation $f \longmapsto f'$ définit sur l'espace des fonctions méromorphes. Si $S = \sum_n a_n X^n$ est une série entière, on notera $S(\mathscr{D})$, l'opérateur $f \longmapsto \sum_n a_n f^{(n)}$. On n'aura pas à se soucier des questions de convergence puisqu'on appliquera cet opérateur que si S ou f est un polynôme de sorte que la série du dessus est finie.

On définit également l'opérateur $\mathscr{J}$, "inverse" de la dérivation, $\mathscr{J}$: $f \longmapsto \int_0^x f(t)dt$. Il vérifie évidement $\mathscr{D} \circ \mathscr{J}(f) = f$ et $\mathscr{J} \circ \mathscr{D} = f - f(0)$. Sa composée $(n+1)$-ème s'écrit :

$$\mathscr{J}^{n+1}(f)(x) = \int_0^x \frac{(x-t)^n}{n!} f(t)dt.$$

Rappels algébriques

On rappelle qu'un nombre est dit algébrique s'il est racine d'un polynôme non nul à coefficients entiers. On note $\overline{\mathbb{Q}}$ l'ensemble des nombres algébriques, c'est la clôture algébrique de $\mathbb{Q}$ dans $\mathbb{C}$. Si de plus un tel nombre admet un polynôme annulateur unitaire (à coefficients entiers) on dira alors que c'est un entier algébrique (ou parfois juste entier). On appelle conjugués galoisiens d'un nombre algébrique, l'ensemble des racines (prises dans $\overline{\mathbb{Q}}$) de son polynôme annulateur minimal.

Remarque 1. Soit α un nombre algébrique et $P = p_0 + ... + p_n X^n$ un polynôme annulateur à coefficients entiers. Posons $q = \text{ppcm}(p_i)_i$. Alors $q\alpha$ annule le polynôme $p_0 q^n/p_n + p_1 q^{n-1}/p_n X + ... + X^n$ qui est à coefficients entiers par construction. Donc $q\alpha$ est un entier algébrique. Cela justifie la définition suivante :

Définition 1. Soit α un nombre algébrique. On appelle **dénominateur** de α, noté $\text{den}(\alpha)$, le plus petit entier naturel non nul s tel que $s\alpha$ soit entier.

On appelle corps de nombres, toute extension de $\mathbb{Q}$ de dimension finie sur $\mathbb{Q}$ en tant que $\mathbb{Q}$-espace vectoriel. Le théorème de l'élément primitif affirme qu'un tel corps est de la forme $\mathbb{Q}(\alpha)$ où α est un nombre algébrique (qui peut être choisi entier).

Définition 2. Soit $\mathbb{K}$ un corps de nombres, et $G = \text{Gal}(\mathbb{K}/\mathbb{Q})$ le groupe des automorphismes de $\mathbb{Q}$-algèbre de $\mathbb{K}$. Une extension $\mathbb{Q} \hookrightarrow \mathbb{K}$ est dite **galoisienne**, si elle satisfait l'une des propriétés équivalentes :

i) Pour tout α dans $\mathbb{K}$, les conjugués de α sont aussi dans $\mathbb{K}$.

ii) L'ensemble des points fixes par tout élément de G dans $\mathbb{K}$ est égal à $\mathbb{Q}$.

Remarque 2. Soit $\mathbb{K}$ un corps de nombres. On peut toujours trouver une extension finie $\mathbb{L}$ de $\mathbb{K}$ telle que $\mathbb{Q} \hookrightarrow \mathbb{L}$ soit galoisienne. Il suffit d'écrire

$\mathbb{K} = \mathbb{Q}(\alpha)$ et de prendre pour $\mathbb{L}$ le corps de décomposition du polynôme minimal de α. Nous pourrons donc toujours supposer si l'on souhaite, qu'une telle extension $\mathbb{Q} \hookrightarrow \mathbb{K}$ est galoisienne.

Nous aurons souvent besoin de contrôler, non pas un seul nombre algébrique, mais tous ses conjugués simultanément.

Définition 3. Soit α un nombre algébrique. On appelle **maison** de α, noté $\overline{|\alpha|}$ la quantité $\sup_i |\alpha_i|$ où les α_i parcourent les conjugués de α.

Lemme 1. *Soient α et β deux nombres algébriques. On a les inégalités sur les maisons suivantes :*

$$\overline{|\alpha+\beta|} \leq \overline{|\alpha|} + \overline{|\beta|} \text{ et } \overline{|\alpha\beta|} \leq \overline{|\alpha|}\,\overline{|\beta|}.$$

Démonstration. Soit $\mathbb{Q} \hookrightarrow \mathbb{L}$ une extension galoisienne contenant α, β. Alors on a :

$$\begin{aligned}\overline{|\alpha+\beta|} = \sup_{\sigma\in Gal(\mathbb{L}/\mathbb{Q})} |\sigma(\alpha+\beta)| &= \sup_{\sigma\in Gal(\mathbb{L}/\mathbb{Q})} |\sigma(\alpha)+\sigma(\beta)| \\ &\leq \sup_{\sigma\in Gal(\mathbb{L}/\mathbb{Q})} |\sigma(\alpha)| + \sup_{\sigma\in Gal(\mathbb{L}/\mathbb{Q})} |\sigma(\beta)| = \overline{|\alpha|} + \overline{|\beta|}.\end{aligned}$$

De la même manière, on a :

$$\begin{aligned}\overline{|\alpha\beta|} = \sup_{\sigma\in Gal(L/\mathbb{Q})} |\sigma(\alpha\beta)| &= \sup_{\sigma\in Gal(\mathbb{L}/\mathbb{Q})} |\sigma(\alpha)\sigma(\beta)| \\ &\leq \sup_{\sigma\in Gal(\mathbb{L}/\mathbb{Q})} |\sigma(\alpha)| \sup_{\sigma\in Gal(\mathbb{L}/\mathbb{Q})} |\sigma(\beta)| = \overline{|\alpha|}\,\overline{|\beta|}.\end{aligned}$$

□

Notre sujet traite de l'existence ou non de relations algébriques entre les nombres. Il reste donc à définir les notions de transcendance, et plus généralement, d'indépendance algébrique.

Définition 4. On dit qu'un nombre α est **transcendant**, s'il n'est pas algébrique. Autement dit, si pour tout polynôme non nul à coefficients entiers P on a $P(\alpha) \neq 0$. On dit plus généralement qu'une famille $(\alpha_i)_{i\leq N}$ est **algébriquement indépendante** sur un corps K, si pour tout polynôme à plusieurs indéterminés P, non nul à coefficients dans K on a $P(\alpha_1, ..., \alpha_N) \neq 0$. Dans le cas contraire une telle famille est dite algébriquement dépendante sur K.

Remarque 3. La notion d'indépendance algébrique est la même sur $\mathbb{Q}$ ou $\overline{\mathbb{Q}}$ (et donc sur tout corps intermédiaire) bien qu'à priori la première notion

soit plus faible. En effet, supposons l'existence d'un polynôme P, non trivial, à coefficients algébriques tel que $P(\alpha_1, ..., \alpha_N) = 0$. On peut trouver une extension $\mathbb{Q} \subset \mathbb{L}$, galoisienne, contenant tous les coefficients de P. Pour $\sigma \in \mathrm{Gal}(\mathbb{L}/\mathbb{Q})$, on écrit P^σ, le polynôme obtenu à partir de P en appliquant σ à ses coefficients. On forme alors le polynôme : $Q = \prod_{\sigma \in \mathrm{Gal}(\mathbb{L}/\mathbb{Q})} P^\sigma$. Les coefficients de Q sont dans $\mathbb{L}$ et on a pour tout $\sigma \in \mathrm{Gal}(\mathbb{L}/\mathbb{Q})$: $Q^\sigma = \prod_{\sigma' \in \mathrm{Gal}(\mathbb{L}/\mathbb{Q})} P^{\sigma'\sigma} = \prod_{\sigma' \in \mathrm{Gal}(\mathbb{L}/\mathbb{Q})} P^{\sigma'} = Q$. Donc Q est à coefficients rationnels et annule $(\alpha_1, ..., \alpha_N)$. Ainsi être dépendant algébriquement sur $\overline{\mathbb{Q}}$ implique bien être dépendant algébriquement sur $\mathbb{Q}$.

Nous avons également besoin de la notion de trace et de norme d'un nombre algébrique.

Définition 5. Soit α un nombre algébrique. On pose $\mathbb{K} = \mathbb{Q}(\alpha)$ le corps de nombre engendré par α. Alors l'application

$$L_\alpha \ : \ \left\{ \begin{array}{ccc} \mathbb{K} & \longrightarrow & \mathbb{K} \\ x & \longmapsto & \alpha x \end{array} \right.$$

est $\mathbb{Q}$-linéaire. On définit la **trace** et la **norme** de α par : $Tr(\alpha) := Tr_{\mathbb{Q}}(L_\alpha)$ et $\mathscr{N}(\alpha) := Det_{\mathbb{Q}}(L_\alpha)$ qui sont des nombres rationnels. Le polynôme minimal de L_α est le polynôme minimal de α. Ainsi, si α est entier algébrique, son polynôme minimal est à coefficients entiers et donc sa norme et sa trace sont des entiers relatifs.

Remarque 4. Le fait que la norme d'un entier algébrique non nul est un entier relatif non nul est très important. En effet dans toutes nos preuves par contradiction nous montrerons l'existence d'un entier algébrique non nul dont la norme est inférieure à 1 strictement ce qui est impossible et conclut en général la démonstration.

Nous disposons de formules directes pour calculer la norme et la trace. On note S l'ensemble des $\mathbb{Q}$-plongements de $\mathbb{K}$ dans $\overline{\mathbb{Q}}$. Alors,

$$Tr(\alpha) = \sum_{\sigma \in S} \sigma(\alpha) \ \text{ et } \ \mathscr{N}(\alpha) = \prod_{\sigma \in S} \sigma(\alpha).$$

En effet les $\sigma(\alpha)$ parcourent les racines du polynôme minimal de α. Par ailleurs, la forme trace

$$Tr \ : \ \left\{ \begin{array}{ccc} \mathbb{K} \times \mathbb{K} & \longrightarrow & \mathbb{Q} \\ (x, y) & \longmapsto & Tr(xy) \end{array} \right.$$

est bilinéaire et symétrique. Nous aurons besoin pour le chapitre 2 du lemme suivant :

Lemme 2. *La forme bilinéaire trace est non dégénérée.*

Démonstration. Il suffit de montrer que la trace est non identiquement nulle sur $\mathbb{K}$. Or l'expression de la trace ci dessus et l'indépendance des $\mathbb{Q}$-plongements de $\mathbb{K}$ dans $\overline{\mathbb{Q}}$ garantit le résultat. □

Chapitre 1

Théorème de Lindemann-Weierstrass

1.1 Énoncé du théorème et cas particuliers

Nous commençons par un résulat sur la transcendance des valeurs de l'exponentielle de 1882 que l'on peut trouver présenté dans [**4**] chapitre 1, à savoir le théorème de Lindemann-Weierstrass.

Théorème 1. *Soient $a_1, ..., a_n$ des nombres algébriques, linéairement indépendants sur $\mathbb{Q}$. Alors les nombres $e^{a_1}, ..., e^{a_n}$ sont algébriquements indépendants sur $\overline{\mathbb{Q}}$.*

On a un énoncé équivalent, même s'il peut sembler plus faible à priori :

Théorème 2. *Soient $a_1, ..., a_n$ des nombres algébriques, distincts. Alors les nombres $e^{a_1}, ..., e^{a_n}$ sont linéairements indépendants sur $\overline{\mathbb{Q}}$.*

Étudions le cas $n = 1$ dans le théorème (1) qui est un résultat que l'on appelle le théorème d'Hermite-Lindemann. On obtient que si a est un nombre algébrique non nul, alors e^a est transcendant. On retrouve par exemple la transcendance de e. On peut également retrouver la transcendance de π. En effet si π était algébrique, alors il en serait de même pour $i\pi$. D'après le théorème, $e^{i\pi} = -1$ serait transcendant ce qui est absurde.

On va montrer l'équivalence entre les deux théorèmes :

Théorème (1) $\Longrightarrow$ *Théorème* (2). Prenons $a_1, ..., a_k$ des nombres algébriques distincts. On suppose sans perte de généralité que $a_1, ..., a_l$ sont linéairements indépendants sur $\mathbb{Q}$ et que $a_{l+1}, ..., a_k$ sont dans $\mathrm{Vect}_{\mathbb{Q}}(a_1, ..., a_l)$. Supposons donnée une relation $\sum_{i=1}^{k} c_i e^{a_i} = 0$ avec les c_i dans $\mathbb{Z}$. Or $a_{l+1}, ..., a_k$ sont linairement dépendants des $a_1, .., a_l$ donc on a en fait une relation algébrique entre $e^{a_1}, ..., e^{a_l}$. D'après les hypothèses du théorème (1), cette relation est

triviale ce qui montre que les $e^{a_1}, ..., e^{a_k}$ sont linéairement indépendants sur $\mathbb{Q}$ d'où le résultat. □

Théorème (2) $\Longrightarrow$ *Théorème* (1). Soient $a_1, ..., a_k$ des nombres algébriques, linéairements indépendants sur $\mathbb{Q}$. Soit P un polynôme de $\overline{\mathbb{Q}}[X_1, ..., X_n]$ tel que $P(e^{a_1}, ..., e^{a_k}) = 0$. Alors en développant on obtient une relation de la forme

$$\sum c_{\lambda_1,...,\lambda_n} e^{\lambda_1 a_1 + ... + \lambda_n a_n} = 0.$$

Or par hypothèse sur les (a_i), les nombres $\lambda_1 a_1 + ... + \lambda_n a_n$ sont distincts, donc d'après l'énoncé (2), les nombres $e^{\lambda_1 a_1 + ... + \lambda_n a_n}$ sont linéairements indépendants. Ainsi P est nul et on obtient le résultat (1). □

1.2 Approximants de Padé et d'Hermite-Padé de la fonction exponentielle

Etant donné une fonction entière f, se pose la question de savoir si l'on peut correctement approcher f, au voisinage de 0 par une fraction rationnelle. Plus précisément, si l'on se donne deux entiers n et m, peut-on trouver deux polynômes $P_{n,m}$ et $Q_{n,m}$ de degrés respectivements au plus n et m, tels que $f - \frac{P_{n,m}}{Q_{n,m}}$ admette un zéro d'ordre au moins $n + m + 1$ à l'origine ?

La bonne façon de formuler le problème est de le rendre linéaire en les coefficients des polynômes, à savoir rechercher un tel couple de polynômes, non nul, tel que $Q_{n,m}f - P_{n,m}$ ait un zéro d'ordre au moins $n + m + 1$. Un tel couple de polynômes est appelé **approximant de Padé** d'indice $[n/m]$ de f.

Quel que soit l'indice $[n/m]$, on est assuré de l'existence d'un tel approximant de Padé. En effet, on doit résoudre un système linéaire homogène en les coefficients de $P_{n,m}$ et $Q_{n,m}$ qui comporte $n+m+1$ équations de la forme

$$(Q_{n,m}f - P_{n,m})^{(k)} = 0, \quad k = 0, ..., n + m$$

et $n + m + 2$ inconnues que sont les coefficients de $P_{n,m}$ et $Q_{n,m}$. Puisque la matrice correspondant à ce système admet plus de colonnes que de lignes, elle possède un noyau non trivial, donc le système admet une solution non triviale. Un tel couple n'est pas en général unique, mais la fraction $P_{n,m}/Q_{n,m}$ l'est. En effet, si deux couples $(P, Q), (\tilde{P}, \tilde{Q})$ de polynômes satisfont aux hypothèses, on a les égalités suivantes :

$$R = Qf - P,$$
$$\tilde{R} = \tilde{Q}f - \tilde{P},$$

où R et $\tilde{R}$ admettent un zéro d'ordre au moins $n+m+1$. En multipliant par $\tilde{Q}$ la première et par Q la seconde puis en faisant la différence on trouve :

$$\tilde{Q}R - Q\tilde{R} = \tilde{P}Q - P\tilde{Q}.$$

Le terme de gauche admet un zéro d'ordre au moins $n+m+1$ tandis que celui de droite est un polynôme de degré au plus $n+m$ par hypothèse, il est donc nul.

On s'intéresse maintenant au cas $f = \exp$ et $n = m$. Nous allons voir que dans le cas de l'exponentielle, ou comme on le verra dans la partie qui suit, de plusieurs exponentielles, on peut expliciter les polynômes obtenus. On cherche donc P_n, Q_n de degrés au plus n, tels que

$$Q_n(z)\exp(z) - P_n(z) = cz^{2n+1} + ... \quad . \tag{1.1}$$

En dérivant $n+1$ fois l'égalité du dessus on obtient :

$$\mathscr{D}^{n+1}(Q_n(z)e^z) = \underbrace{c(2n+1)!/n!}_{c_0} z^n + ... \quad .$$

Donc

$$e^z(1+\mathscr{D}^{n+1})(Q_n) = c_0 z^n + ... \quad .$$

D'où finalement :

$$(1+\mathscr{D}^{n+1})(Q_n) = e^{-z}(c_0 z^n + ...) = c_0 x^n + ... = c_0 x^n,$$

car le polynôme de gauche est de degré au plus n. En choisissant $c_0 = 1$, les polynômes Q_n et P_n sont de degré exactement n et la valeur c plus haut est non nulle de sorte que la fonction $Q_n \exp - P_n$ s'annule exactement à l'ordre $2n+1$ en 0. Finalement, on trouve que $Q_n(z) = c_0(1+\mathscr{D})^{-n-1}(z^n)$. En partant de la même égalité puis en multipliant par e^{-z} on peut également trouver une formule pour P_n, à savoir, $P_n(z) = c_0(\mathscr{D}-1)^{-n-1}(z^n)$.

On note $R_n(z) = Q - n(z)e^z - P_n(z)$ le n-ième reste que l'on appelle généralement **fonction auxiliaire**. On a $\mathscr{D}^{n+1}(R_n) = \mathscr{D}^{n+1}(Q_n(z)e^z) = e^z(1+\mathscr{D}^{n+1})(Q_n) = z^n e^z$. Puisque R_n s'annule à l'ordre $2n+1$ en zéro, on obtient une formule explicite de la fonction auxiliaire :

$$R_n(z) = \mathscr{J}^{n+1}(e^z z^n) = \frac{z^{2n+1}}{n!}\int_0^1 (1-t)^n t^n e^{tz} dt. \tag{1.2}$$

On peut facilement obtenir les estimations dont on a besoin grâce à la formule du dessus. En effet on a immédiatemment

$$|R_n(z)| \leq |z|^{n+1} e^{\mathfrak{Re}(z)}/n!,$$

ainsi que $R_n(z) > 0$ pour tout z réel strictement positif.

Remarque 5. Ces estimations de la fonction R_n permettent déjà facilement de montrer l'irrationnalité de e^r, pour r rationnel non nul. En effet, supposons que ce n'est pas le cas, on a l'existence de $(p,q) \in \mathbb{Z} \times \mathbb{Z}$ non tous les deux nuls tels que $qe^r = p$. Quitte à prendre une racine n-ième, il suffit de montrer le résultat pour r entier. On remarque que P_n, Q_n étant à coefficients entiers, $P_n(r), Q_n(r)$ sont également entiers ainsi que $sR_n(r)$. Or les estimations ci-dessus donnent $0 < sR_n(r) < se^r r^{2n+1}/n!$. En prenant n assez grand on obtient une contradiction.

Si l'on souhaite aller plus loin, notamment en vu de démontrer notre théorème, il faut généraliser notre construction d'approximants de Padé. Plus précisément, si l'on se donne un système $(f_1, ..., f_n)$ de fonctions entières, on cherche des polynômes P_i, tels que $\sum P_i f_i$ s'annule à un grand ordre en zéro. On appellera **approximants d'Hermite-Padé** de $(f_1, ..., f_n)$ et d'indice $(n_1, ..., n_k)$ tout k-uplet de polynômes P_i vérifiant :

$$\begin{cases} \deg(P_i) \leq n_i - 1 \\ \mathrm{ord}_{z=0}(\sum_i P_i f_i) \geq N - 1 \end{cases} \qquad \text{où} \quad N = n_1 + ... + n_k.$$

Notre problème nous incite à considérer les f_i de la forme $f_i(z) = e^{\rho_i z}$, où les ρ_i sont des complexes deux à deux distincts. On notera

$$R_N := \sum_i P_i(z) e^{\rho_i z} = cx^N + ..., \tag{1.3}$$

et on dira que R_N est une fonction auxiliaire. On prend $c = 1$. Dans notre cas très particulier où les f_i sont des exponentielles, on dispose d'expressions explicites des P_i et de R_N que l'on obtient par récurrence sur k. L'idée est la même qu'avec les approximants de Padé.

Si $k = 1$, P_1 vaut clairement $z^N/N!$. On suppose donc $k \geq 2$. On multiplie par $e^{-\rho_k z}$ puis on dérive n_k fois l'expression définissant R_N dans l'égalité au dessus. On obtient donc

$$\begin{aligned} S(z) := \mathscr{D}^{n_k}(R_N(z)e^{-\rho_k z}) &= \sum_j \mathscr{D}^{n_k}(P_j(z)e^{(\rho_j - \rho_k)z} \\ &= \sum_j e^{(\rho_j - \rho_k)z}(\rho_j - \rho_k + \mathscr{D})^{n_k} P_j(z) = \sum_j e^{(\rho_j - \rho_k)z} Q_j(z). \end{aligned}$$

D'un autre coté on a

$$\mathscr{D}^{n_k}(R_N(z)e^{-\rho_k z}) = \mathscr{D}^{n_k}\left(\frac{z^{N-1}}{(N-1)!}e^{-\rho_k z} + ...\right) = \frac{z^{N-n_k-1}}{(N-n_k-1)!} + ... \ .$$

Les polynômes $(Q_j)_{j \leq k-1}$ sont de même degré que les $(P_j)_{j \leq k}$, et on remarque que ce sont des approximants d'Hermite Padé pour les fonctions

$g_i(z) = e^{(\rho_i-\rho_k)z}$. Enfin, la constante devant le terme $z^{N-n_k-1}/(N-n_k-1)!$ est toujours égale à 1. On a donc $P_j = (\rho_j - \rho_k + \mathscr{D})^{-n_k-1}(Q_j)$ et par récurrence avec le cas $k = 1$ on obtient :

$$P_j(z) = \prod_{i\neq j, i=1}^{k} (\rho_j - \rho_i + \mathscr{D})^{-n_i-1}(z^{n_j}/n_j!). \tag{1.4}$$

De plus, puisque R_N s'annule à un ordre suffisemment grand en zéro, on a pour $x > 0$,

$$R_N(x) = \mathscr{J}^{n_k}\mathscr{D}^{n_k}(R_N(x)) = e^{\rho_k x}\mathscr{J}^{n_k}S(x) = \frac{e^{\rho_k x}}{n_k!}\int_0^x (x-t)^{n_k}S(t)dt. \tag{1.5}$$

où S est la fonction auxiliaire associée aux $e^{(\rho_j-\rho_k)z}$. Cette remarque nous permet de trouver une expression explicite de R_N par récurrence sur k.

Lemme 3. *On a l'expression intégrale valable pour $x > 0$:*

$$R_N(x) = \underset{t_1+\ldots+t_k=x, t_i\geq 0}{\int\cdots\int} \prod_{j=1}^{k} \frac{t_j^{n_j}}{n_j!} e^{\rho_j t_j} dt_1 \cdots dt_{k-1}. \tag{1.6}$$

Démonstration. Pour $k = 2$, l'expression de droite de (3) est

$$\int_0^x \frac{t^{n_1}(x-t)^{n_2}}{n_1!n_2!} e^{\rho_1 t+\rho_2(x-t)} dt = \frac{e^{\rho_2 x}}{n_2!}\int_0^x (x-t)^{n_2}\frac{t^{n_1}}{n_1!}e^{t(\rho_1-\rho_2)}dt.$$

Ce terme vaut $e^{\rho_2 x}\mathscr{J}^{n_2}(S(x))$, avec $S(x) = e^{x(\rho_1-\rho_2)}x^{n_1}/n_1!$. La fonction S est bien la fonction auxiliaire pour la fonction $e^{x(\rho_1-\rho_2)}$, donc d'après l'égalité (1.5), on obtient bien le résultat.

Supposons le résultat vrai au rang $k-1$. On écrit l'égalité précédent le lemme en appliquant l'hypothèse de récurrence à S :

$$\begin{aligned}
R_N(x) &= \frac{e^{\rho_k x}}{n_k!}\int_0^x (x-t)^{n_k}\left[\underset{t_1+\ldots+t_k=t, t_i\geq 0}{\int\cdots\int} \prod_{j=1}^{k}\frac{t_j^{n_j}}{n_j!}e^{(\rho_j-\rho_k)t_j}dt_1\cdots dt_{k-2}\right]dt \\
&= \frac{e^{\rho_k x}}{n_k!}\int_0^x t^{n_k} \underset{t_1+\ldots+t_k=x-t, t_i\geq 0}{\int\cdots\int} \prod_{j=1}^{k}\frac{t_j^{n_j}}{n_j!}e^{(\rho_j-\rho_k)t_j}dt_1\cdots dt_{k-2}dt \\
&= \underset{t_1+\ldots+t_{k-1}+t=x, t_i\geq 0}{\int\cdots\int} e^{(t_1+t_2+\ldots)\rho_k}\prod_{j=1}^{k}\frac{t_j^{n_j}}{n_j!}e^{(\rho_j-\rho_k)t_j}dt_1\cdots dt_{k-1}dt \\
&= \underset{t_1+\ldots+t_k=x, t_i\geq 0}{\int\cdots\int} \prod_{j=1}^{k}\frac{t_j^{n_j}}{n_j!}e^{\rho_j t_j}dt_1\cdots dt_{k-1}.
\end{aligned}$$

□

Déterminant d'une famille de fonctions auxiliaires

Les formes explicites des P_j et de R_N nous permettent de les estimer convenablement. Notons $M = \sup_{i\neq j}|\frac{1}{\rho_i-\rho_j}|$. Si ζ est non-nul on peut écrire le développement classique :

$$(\zeta+\mathscr{D})^{-n-1} = \zeta^{-n-1}\sum_{k=n+1}^{\infty}\binom{k}{n+1}\zeta^{n+1-k}\mathscr{D}^{k-n-1}.$$

L'expression (1.4) de P_j obtenue couplée à ce développement donne :

$$|P_l(x)| \leq \prod_{i\neq l}(M^{-1}+\mathscr{D})^{-n_i-1} \leq (M^{-1}+\mathscr{D})^{n_l-N}(x^{n_l}/n_l!).$$

On peut donc trouver une majoration de $P_l(x)$ à savoir :

$$\begin{aligned} P_l(x) &\leq M^{N-n_l}\sum_{k=N-n_l}^{N}\binom{k}{N-n_l}\left|(M\mathscr{D})^{k-N+n_l}\left(\frac{x^{n_l}}{n_l!}\right)\right| \\ &\leq M^{N-n_l}\sum_{k=0}^{n_l}\binom{k+N-n_l}{N-n_l}M^k\binom{n_l}{k}|x^{n_l-k}| \\ &\leq \sum_{k=0}^{n_l}\binom{N}{k}M^{N-n_l+k}|x^{n_l-k}| \leq 2^N(M+|x|)^N. \quad (1.7) \end{aligned}$$

Le nombre $P_l(1)$ est un nombre algébrique si les ρ_i sont algébriques, au vu de l'expression obtenue de ce polynôme. On cherche maintenant à estimer son dénominateur. Si l'on prend pour h un dénominateur commun des $1/(\rho_i-\rho_j)$, alors $\prod(h(\rho_j-\rho_l+\mathscr{D}))^{-n_j-1}(z^{n_j})$ est à coefficients entiers. Il suit de l'expression de P_j que le nombre $n_l!h^NP_l(1)$ est entier.

La non-nullité de $R_N(1)$ est essentielle pour obtenir une contradiction dans la preuve du théorème de Lindemann-Weierstrass. Pour pallier à cette difficulté on va considérer une famille de fonctions auxiliaires, dont on sera assurés que l'une d'entre elles ne s'annulera pas en 1.

On pose $R_j(x) = P_{j1}(x)e^{\rho_1 x} + ... + P_{j1}(x)e^{\rho_k x}$ où les $P_{j1}, ..., P_{jk}$ forment un système d'approximants d'Hermite-Padé d'indice $(n-1, ..., n-1, n, ..., n)$. On considère maintenant le déterminant :

$$\Delta(x) := \det((P_{ij}(x))_{i,j\leq k}) \quad \text{qui est un polynôme en } x.$$

En développant ce déterminant, on a

$$\Delta(x) = \sum_{\sigma\in S_k}\varepsilon(\sigma)\prod_{i=1}^{k}P_{i\sigma(i)}(x),$$

et l'on remarque que seul le terme correspondant à $\sigma = id$ est de degré kn, là où les autres sont de degré strictement inférieur. Donc $\Delta(x) = \gamma x^{kn} + r(x)$, avec $r \in \mathbb{C}[X]$ de degré strictement inférieur à kn et γ non nul.

On note maintenant $\Delta_1, \cdots, \Delta_k$ les mineurs des éléments de la première colonne. On peut multiplier la matrice par des transvections de la forme $C_1 \leftarrow C_1 + e^{(\rho_j - \rho_1)x} C_j$ puis développer le déterminant par rapport à la première colonne ce qui donne :

$$\Delta(x) = (\Delta_1(x)R_1(x) + \cdots + \Delta_k(x)R_k(x))e^{-\rho_k x}. \tag{1.8}$$

Les R_k s'annulent à l'ordre nk en 0, par définition des approximants de Hermite-Padé. Il s'ensuit alors que $\Delta(x) = \gamma x^{kn}$ et $\Delta(1) = \gamma \neq 0$. L'égalité (1.8) implique alors que l'un des $R_j(1)$ est non nul.

1.3 Preuve du théorème de Lindemann-Weierstrass

On suppose par l'absurde qu'il existe un polynôme $P(X_1, ..., X_p)$ à coefficients entiers non tous nuls tel que $P(e^{a_1}, ..., e^{a_p}) = 0$. On note $\mathbb{K}$ le corps de nombre engendré par $a_1, ..., a_p$ et h son degré. On note également d le degré total du polynôme P. On note b un entier relatif, supérieur à d, à définir, qui ne dépendra pas de n.

Il y a exactement $m = \binom{b+p}{p}$ monômes de degré au plus b et $r := \binom{b+p-d}{p}$ monômes de degré au plus $b-d$ dans $\mathbb{K}[X_1, ..., X_p]$. On note ces monômes respectivement $Y_1, ..., Y_m$ et $Z_{m-r+1}, ..., Z_m$. Les polynômes $Z_k P$ ont un degré au plus b. On peut donc écrire, pour tout k, $Z_k P = a_{k1}Y_1 + ... + a_{km}Y_m$, où les a_{kj} sont des coefficients (éventuellement nuls) de P.

On remarque que la matrice $m \times r$ (a_{kj}) est de rang r. En effet, dans le cas contraire, on aurait l'existence d'une relation non triviale entre les r colonnes de cette matrices, c'est à dire une relation non triviale entre les $Z_k P$ et donc entre les Z_k ce qui est absurde. On écrit $Y_l = X_1^{h_1} ... X_p^{h_p}$ et $\rho_l = h_1 a_1 + ... + h_p a_p$. Par hypothèse du théorème, les ρ_i sont distincts et par hypothèse sur P on a :

$$\sum_i a_{ki} e^{\rho_i} = 0, \quad \text{pour tout } k = m - r + 1, ..., m.$$

On peut considérer la famille de fonctions auxiliaires construites à la partie précédente, $R_k(z) = P_{k1}(z)e^{\rho_1 z} + ... + P_{km}(z)e^{\rho_m z}$. On a vu que le déterminant $\Delta(1)$ de $(P_{kj}(1))_{k,j}$ était non nul. Par le théorème de la base incomplète, on peut trouver $m-r$ lignes de $(P_{kj}(1))$, telle que si l'on concatène ces lignes et la matrice (a_{kj}) on obtient une matrice $m \times m$ inversible. Si $k_1, ..., k_{m-r}$ sont ces lignes, on pose $a_{tl} = P_{k_t l}(1)$ et $\beta_t = R_{k_t}(1)$ pour $t = 1, ..., m-r$. On

obtient ainsi les relations :

$$a_{k1}e^{\rho_1} + ... + a_{km}e^{\rho_1} = \beta_k \quad \text{pour } k = 1, ..., m,$$

où β_k vaut zéro si k est supérieur à $m - r + 1$.

En appliquant les transvections $C_1 \leftarrow C_1 + (e^{\rho_1} - e^{\rho_k})C_k$ sur la première colonne on obtient en développant par rapport à cette première colonne, que le déterminant de (a_{kj}) vaut $A := (\Delta_1\beta_1 + ... + \Delta_m\beta_m)e^{-\rho_1}$ où Δ_i est le i-ème mineur de la première colonne.

On peut appliquer la majoration du dénominateur des $P_l(1)$ de la partie précédente et trouver un dénominateur commun D aux $\{a_{kl}\}_{k\leq m-r, l\leq m}$ tel que $D \leq c^n n^n$. Les autres a_{kl} sont entiers par hypothèse donc D^{m-r} est un dénominateur de A. On trouve une minoration de la norme de l'entier algébrique $D^{m-r}A$:

$$\mathscr{N}(D^{m-r}A) \geq 1 \quad \text{d'où} \quad \mathscr{N}(A) \geq D^{h(r-m)} \geq c_1^n n^{nh(r-m)}.$$

D'autre part l'estimations (1.7) de la partie précédente nous donne $\overline{|a_{kl}|} \leq c_2^n$ donc $\overline{|A|} \leq c_3^n$ et par ailleurs, $|R_{k_t}(1)| = |\beta_t| \leq c_3 n^m n^{-nm}$. Ainsi on trouve une majoration de la norme :

$$\mathscr{N}(D^{m-r}A) \leq D^{h(m-r)}\overline{|A|}^{h-1}|A| \leq c_4^n n^{-mn}.$$

En faisant tendre n vers l'infini et en comparant la minoration et la majoration obtenue on trouve que $m \leq h(r-m)$ c'est à dire $r \leq (1 - 1/h)m$. On remarque que r et m sont equivalents en tant que fonction de b. En effet ce sont des polynômes de même degré et coefficient dominant. Ainsi en choisissant b suffisament grand on aboutit à une contradiction avec l'inégalité $r \leq (1 - 1/h)m$. Le résultat est démontré.

Chapitre 2

Lemmes de Siegel

2.1 Les résultats

Jusqu'à présent, lorsque nous avons eu affaire à un système d'équations linéaires homogènes avec plus d'inconnues que d'équations, nous avons simplement utilisé l'existence d'une solution non triviale. Cela était possible car on avait accès de fait aux formules explicites des solutions. Par la suite, on n'aura pas accès à des formules explicites, mais on va voir que l'on peut de manière générale contrôler de manière adéquate la taille de nos solutions grâce au lemme de Siegel. On pourra se référer aux dernières pages de **[3]** pour une présentation de ces résultats.

Lemme 4. *Supposons données M équations linéaires homogènes de la forme :*

$$\sum_{j=1}^{N} a_{ij}x_j = 0, \quad i = 1, ..., M,$$

*avec $N > M$ et les a_{ij} entiers **relatifs**. On note $A = \sup|a_{ij}|$. Alors il existe une solution $(x_1, ..., x_N)$ en entiers relatifs non tous nuls, qui vérifie*

$$|x_j| \leq (NA)^{\frac{N}{N-M}} + 2.$$

Démonstration. On désigne par $[\![a, b]\!]$, l'intervalle d'entiers $[a, b] \cap \mathbb{Z}$. Soit X un entier naturel pair strictement supérieur à $(NA)^{\frac{N}{N-M}}$. Notons ϕ_i les formes linéaires $\sum_j a_{ij}x_j$. Lorsque $(x_1, ..., x_N)$ parcourt le cube $[\![-\frac{X}{2}, \frac{X}{2}]\!]^N$, les formes ϕ_i prennent des valeurs entières, majorées en valeurs absolues par $NAX/2$.

Donc les $(x_1, ..., x_N)$ sont à valeurs dans un ensemble à $(X + 1)^N$ éléments, tandis que les $(\phi_1(x_1, ..., x_N), ..., \phi_M(x_1, ..., x_N))$ sont à valeurs dans un ensemble à $(NAX + 1)^M$ éléments.

Puisque $N > M$, $(X+1)^{\frac{N}{M}} > X^{\frac{N}{M}} + 1$. De plus, par hypothèse, $X^{\frac{N}{M}-1} = X^{\frac{N-M}{M}} > NA$. Ainsi on trouve finalement

$$(X+1)^N \geq (X^{\frac{N}{M}} + 1)^M \geq (NAX + 1)^M.$$

Il existe donc, par principe des tiroirs, deux N-uplets distincts x et $\tilde{x}$ de $[\![-\frac{X}{2}, \frac{X}{2}]\!]^N$ qui ont mêmes images par les ϕ_i. Au final, le N-uplet $x - \tilde{x} \neq 0$ annule les ϕ_i et est dans $[\![-X, X]\!]^N$. C'est bien le résultat voulu. □

Nous pouvons généraliser le résultat précédent dans un corps de nombres quelquonque. On sera essentiellement amené à décomposer les coefficients dans une base de l'anneau d'entiers et à appliquer le lemme 4.

Lemme 5. *Supposons données M équations linéaires homogènes de la forme :*

$$\sum_{j=1}^{N} a_{ij} x_j = 0, \quad i = 1, ..., M$$

*avec $N > M$ et les a_{ij} entiers **algébriques** situés dans un corps de nombres $\mathbb{K}$. On note $A = \sup |\overline{a_{ij}}|$. Alors il existe une solution $(x_1, ..., x_N)$ en entiers algébriques non tous nuls, qui vérifie*

$$|\overline{x_j}| \leq c(cNA)^{\frac{N}{N-M}} + 2,$$

où c ne dépend que de K.

Démonstration. On peut se donner une $\mathbb{Z}$-base de $\mathscr{O}_K$ de la forme $(\omega_1, ..., \omega_d)$. Cette base admet une structure multiplicative que l'on peut d'écrire par $\omega_i \omega_j = \sum_k b_{ij}^{(k)} \omega_k$ avec $b_{ij}^{(k)}$ dans $\mathbb{Z}$. Les entiers $a_{ij} \in \mathscr{O}_K$ se décomposent dans cette base :

$$a_{ij} = \sum_{k=1}^{d} a_{ij}^{(k)} \omega_k, \qquad a_{ij}^{(k)} \in \mathbb{Z}. \tag{2.1}$$

Si maintenant $x_1, ..., n_N$ sont des éléments de $\mathscr{O}_K$, on a quel que soit $i \in \{1, ..., M\}$,

$$\sum_{j=1}^{N} a_{ij} x_j = \sum_{j=1}^{N} \sum_{k,l=1}^{d} a_{ij}^{(k)} \omega_k x_j^{(l)} \omega_l = \sum_{j=1}^{N} \sum_{k,l=1}^{d} \sum_{p=1}^{d} a_{ij}^{(k)} x_j^{(l)} b_{kl}^{(p)} \omega_p,$$

où les $x_j^{(l)}$ sont les coordonnées de x_j.

On est donc amené à résoudre le système

$$\sum_{j=1}^{N} \sum_{l=1}^{d} \left(\sum_{k=1}^{d} a_{ij}^{(k)} b_{kl}^{(p)} \right) x_j^{(l)} = 0 \qquad i = 1, ..., M, p = 1, ..., d$$

d'inconnues $x_j^{(l)}$. On peut appliquer le lemme 4 car on a un système de dM équations à dN inconnues. Il existe donc des $x_j^{(l)}$ solutions du système en entiers relatifs non tous nuls tels que

$$|x_j^{(l)}| \leq \left(dN \sup_{i,p} \left| \sum_{k=1}^{d} a_{ij}^{(k)} b_{kl}^{(p)} \right| \right)^{\frac{dN}{dN-dM}} \leq \left(Cd^2 N \sup_{i,k,p} |a_{ij}^{(k)}| \right)^{\frac{N}{N-M}}.$$

Ainsi on obtient la majoration :

$$\overline{|x_j|} \leq \tilde{C} \left(Cd^2 N \sup_{i,k,p} |a_{ij}^{(k)}| \right)^{\frac{N}{N-M}}.$$

Il reste à contrôler les $|a_{ij}^{(k)}|$ en fonction de $\overline{|a_{ij}|}$. La matrice $(\sigma(\omega_k))_{\sigma,k}$, où les σ parcourent le groupe $\mathrm{Gal}(K/\mathbb{Q})$, est inversible. En effet, le produit de cette matrice et de sa transposée donne

$$\begin{pmatrix} \sum_\sigma \sigma(\omega_1\omega_1) & \cdots & \sum_\sigma \sigma(\omega_1\omega_d) \\ \vdots & \vdots & \vdots \\ \sum_\sigma \sigma(\omega_d\omega_1) & \cdots & \sum_\sigma \sigma(\omega_d\omega_d) \end{pmatrix} = \begin{pmatrix} Tr_K(\omega_1\omega_1) & \cdots & Tr_K(\omega_1\omega_d) \\ \vdots & \vdots & \vdots \\ Tr_K(\omega_d\omega_1) & \cdots & Tr_K(\omega_d\omega_d) \end{pmatrix},$$

qui est la matrice de la forme bilinéaire trace. Or celle-ci est non dégénérée, d'après le lemme 2, donc sa matrice est inversible. On applique les différents σ à l'égalité (2.1) et on trouve donc une constante C_1 telle que $|a_{ij}^{(k)}| \leq C_1 \overline{|a_{ij}|}$. On a bien la majoration

$$\overline{|x_j|} \leq \tilde{C} \left(CC_1 d^2 N \sup_{i,j} \overline{|a_{ij}|} \right)^{\frac{N}{N-M}}.$$

ce qui donne le résultat.

□

2.2 Applications à la construction de fonctions auxiliaires

Notre problème consiste en le suivant : si une famille de fonctions méromorphes $f_1, ..., f_k$ est donnée, on souhaite construire une **fonction auxiliaire** $F = \sum_i P_i f_i$ qui s'annule en différents points à certains ordres (selon la situation). Par exemple, on a déjà vu dans le chapitre précèdent dans la définition (1.3), la construction d'une fonction auxilaire R_N, qui s'annulait à l'ordre N en 0. Ce problème se ramène toujours à un système linéaire homogène où les équations sont données par l'annulation de F aux différents points et les inconnues sont les coefficients des P_i. Le lemme de Siegel nous permettra de majorer la taille des coefficients.

Chapitre 3

Théorème de Gelfond-Schneider

3.1 Le théorème et son contexte

Le résultat qui suit que l'on peut trouver dans [**3**] ou[**4**], constitue le septième problème de la liste des 23 problèmes que David Hilbert avait exposés au congrès International des Mathématiciens en 1900. Il a même ajouté à son intention que la résolution de ce problème lui semblait appartenir à un avenir encore plus lointain que celle de l'hypothèse de Riemann ou du dernier théorème de Fermat. Pourtant, une preuve de ce théorème est publiée en 1934, simultanément par Gelfond et Schneider. Nous présenterons dans ce chapitre seulement la preuve de Schneider, puisque la méthode de Gelfond se généralise grandement en un résultat que nous présenterons au chapitre suivant.

Théorème 3. *Soient α, β deux nombres algébriques avec β irrationnel et α différent de* 0 *et* 1. *Alors α^β est transcendant.*

Ainsi, les nombres $2^{\sqrt{2}}$ et $e^\pi = i^{-2i}$ par exemple sont transcendants.

Remarque 6. Nous pouvons reformuler le théorème de cette façon : on note $\mathscr{L}$ l'ensemble $\{\log(\alpha),\ \alpha \in \overline{\mathbb{Q}}^*\}$, où la branche du logarithme choisi n'importe pas. Si α, β sont algébriques, $\log(\alpha^\beta) = \beta \log(\alpha)$ et $\log(\alpha)$ sont indépendants sur $\mathbb{Q}$ dès que β est irrationnel. Dans ce cas, le théorème 3 énonce que $\log(\alpha^\beta)$ n'est pas dans $\mathscr{L}$. Réciproquement, supposons que $\log(\alpha), \log(\beta)$ dans $\mathscr{L}$, sont indépendants linéairement sur $\overline{\mathbb{Q}}$ dès qu'ils le sont sur $\mathbb{Q}$. Si β est irrationnel, $\log(\alpha)$ et $\log(\alpha^\beta)$ sont indépendants sur $\mathbb{Q}$ mais pas sur $\overline{\mathbb{Q}}$. L'hypothèse montre que α^β n'est pas algébrique. Le théorème 3est donc équivalent au théorème suivant :

Théorème 4. *Soient α, β, deux éléments de $\mathscr{L}$ qui sont $\mathbb{Q}$-linéairement indépendants. Alors α et β sont $\overline{\mathbb{Q}}$-linéairement indépendants.*

Cette re-écriture permet de citer la généralisation que Baker a apporté en 1966 que nous ne démontrerons pas, mais que l'on peut trouver dans [**7**] en première partie.

Théorème 5. *Soient $\alpha_1, ..., \alpha_n$, deux éléments de $\mathscr{L}$ qui sont $\mathbb{Q}$-linéairement indépendants. Alors α_1,... , α_n sont $\overline{\mathbb{Q}}$-linéairement indépendants.*

3.2 Construction de la fonction auxiliaire et déterminant de Vandermonde

Notons $\mathbb{K}$ le corps engendré par $\alpha, \beta, \alpha^\beta$ et supposons que $\mathbb{K}$ est un corps de nombre de degré d. Si α est une racine de l'unité, alors α^β n'en est pas une car β est irrationnel. Dans ce cas, on peut remplacer α par $\alpha' = \alpha^\beta$, β par $\beta' = \beta^{-1}$ et α^β par $(\alpha^\beta)^{\beta^{-1}} = \alpha = \alpha'^{\beta'}$ sans changer le corps $\mathbb{K}$. On peut donc supposer que α n'est pas une racine de l'unité.

On se donne un entier m à définir plus tard qui dépend uniquement de ces données. On pose $n = q^2/m$, où q est un entier, indépendant de m, que l'on va faire tendre vers $+\infty$. Nous désignerons par $c_1, c_2, ...$ des constantes dont seule l'indépendance avec le paramètre n (et donc q) sera importante pour nous. Certaines de ces valeurs peuvent être quantifiées mais nous ne en préoccuperons pas ici.

On va s'intéresser aux fonctions $f_i(z) = \alpha^{(m-i)z}$ et on pose

$$R(z) = P_1(z)\alpha^{mz} + ... + P_m(z)\alpha^z = \sum_i P_i(z)\alpha^{(m-i+1)z}$$

la fonction auxiliaire. On va chercher des polynômes P_i tels que R s'annule en de nombreux points.

Lemme 6. *Il existe des polynômes $P_1, ..., P_m$ de degrés au plus $2n-1$, à coefficients entiers algébriques qui vérifient :*

(i) Chaque coefficient de ces polynôme a une maison majorée par $c^n n^n$.

(ii) La fonction R s'annule en les $q^2 = mn$ points $\lambda + \mu\beta$ où $1 \leq \lambda, \mu \leq q$.

Démonstration. La seconde condition nous donne mn équations linéaires homogènes de la forme $R(\lambda + \mu\beta) = P_1(\lambda + \mu\beta)\alpha^{m\mu\beta}\alpha^{m\lambda} + ... + P_m(\lambda + \mu\beta)\alpha^{\mu\beta}\alpha^\lambda = 0$, à $2mn$ inconnues que sont les coefficients des polynômes P_i. Les coefficients de l'équation sont de la forme $(\lambda + \mu\beta)^k\alpha^\lambda\alpha^{\mu\beta}$, dont un dénominateur commun est $\mathrm{den}(\beta)^{2n-1}\mathrm{den}(\alpha)^{mq}\mathrm{den}(\alpha^\beta)^{mq} \leq c_1^n$. On peut multiplier ce système par ce dénominateur pour le rendre à coefficients entiers algébriques. Il ne reste plus qu'à majorer les maisons des $(\lambda+\mu\beta)^k\alpha^\lambda\alpha^{\mu\beta}$

pour pouvoir appliquer le lemme 5 de Siegel. Ceci se fait simplement en utilisant la sous additivité et sous multiplicativité de la maison :

$$\begin{aligned}\left|\overline{(\lambda+\mu\beta)^k\alpha^{\lambda}\alpha^{\mu\beta}}\right| &\leq (\overline{|\lambda|}+\overline{|\mu|}\overline{|\beta|})^k\overline{|\alpha|}^{mq}\overline{|\alpha^{\beta}|}^{mq}\\ &\leq (q+q\overline{|\beta|})^{2n}(\overline{|\alpha|}\overline{|\alpha^{\beta}|})^{mq} \leq c_2^n n^n.\end{aligned}$$

D'après le lemme de Siegel, il existe donc une solution non triviale $(P_1,...,P_m)$, dont les coefficients sont entiers et de maison majorée par $c_0 \times (c_0c_1^nc_2^nn^n)^{nm/(2nm-nm)} \leq c^nn^n$. □

Notre but est de construire un entier algébrique dont la norme est comprise strictement entre 0 et 1 pour obtenir une contradiction. Pour cela il faut s'assurer que notre nombre est différent de 0. Nous allons utiliser une matrice de Vandermonde, ce qui sera pratique pour appliquer l'équation fonctionnelle $\alpha^{x+y}=\alpha^x\alpha^y$ et montrera la non-annulation de l'un des $R(\lambda+\mu\beta)$ où $q+1 \leq \lambda \leq 2q, 1 \leq \mu \leq q$.

On définit pour $k=1,...,m$ le k-ième translaté du polynôme P_ℓ comme : $P_{k\ell}(z) := \alpha^{(m-\ell+1)k}P_\ell(z+k)$. De ce fait on a :

$$R(z+k)=\sum_{\ell} P_{k\ell}(z)\alpha^{(m-\ell)z}.$$

On note $P_{t_1},...,P_{t_g}, t_1<...<t_g$ les polynômes qui sont non nuls parmi les $P_1,...,P_m$ qui définissent R. On écrit enfin $P_{t_\ell}(z)=a_{r_\ell,\ell}z^{r_\ell}+a_{r_\ell-1,\ell}z^{r_\ell-1}+...+a_{0,\ell}$ et $\Delta(z)$ le déterminant de la matrice :

$$\begin{pmatrix} P_{1,t_1}(z) & P_{1,t_2}(z) & \cdots & P_{1,t_g}(z)\\ P_{2,t_1}(z) & P_{2,t_2}(z) & \cdots & P_{2,t_g}(z)\\ \vdots & \vdots & \ddots & \vdots\\ P_{g,t_1}(z) & P_{g,t_2}(z) & \cdots & P_{g,t_g}(z)\end{pmatrix}.$$

Lemme 7. *Il existe un nombre ξ parmi les $2mn=2q^2$ points $\lambda+\mu\beta$, $1\leq\lambda\leq 2q, 1\leq\mu\leq q$, tel que $\Delta(\xi)$ est non nul.*

Démonstration. On a $\Delta(z)=\det\left((\alpha^{(m-\ell)k}(a_{r_\ell,\ell}(z+k)^{r_\ell}+...+a_{0,\ell}))_{\ell,k}\right)$ donc par multilinéarité du déterminant on obtient :

$$\begin{aligned}\Delta(z) &= \det\left((\alpha^{(m-\ell)k}a_{r_\ell,\ell}(z+k)^{r_\ell})_{\ell,k}\right)+Q(z)\\ &= a_{r_1,1}\cdots a_{r_g,g}\det\left((\alpha^{(m-\ell)k})_{k,\ell}\right)z^{r_1+...+r_g}+Q'(z),\end{aligned}$$

où Q et Q' sont des polynômes de degré au plus $r_1+...+r_g-1$.

Or $\det\left((\alpha^{(m-\ell)k})_{k,\ell}\right)$ est un déterminant de Vandermonde égal à $\prod_{k<l}(\alpha^{m-t_\ell}-\alpha^{m-t_k})$ qui est non-nul car α n'est pas une racine de l'unité.

Donc Δ est un polynôme de degré exactement $r_1+...+r_g$ inférieur à $2mn$ strictement car les P_i sont de degré au plus $2n-1$. Ainsi parmi les $2mn$ points $\lambda+\mu\beta$ où $1\leq\lambda\leq 2q, 1\leq\mu\leq q$, il existe un nombre ξ, tel que $\Delta(\xi)$ est non nul. □

3.3 Preuve du théorème

D'après le lemme 7, toute combinaison non triviale de lignes de (P_{k,t_ℓ}) est non nulle donc en particulier, pour un certain k, on a :

$$\gamma := R(\xi+k)=\sum_\ell P_{kl}(\xi)\alpha^{\xi(k-\ell)}\neq 0.$$

Il nous reste à majorer successivement dénominateur, maison et valeur absolue de ce nombre γ pour conclure.

On a $\mathrm{den}(\gamma)=\mathrm{den}(R(\xi+k))$ qui divise :

$$(\mathrm{den}(\alpha)\mathrm{den}(\alpha^\beta))^{3mq}\times \mathrm{ppcm}(\mathrm{den}(P_{k1}(\xi+k)),...,\mathrm{den}(P_{km}(\xi+k))).$$

Donc

$$\mathrm{den}(\gamma)\leq c_3^{3mq}(\mathrm{den}(\beta)^{2n-1})^m\leq c_4^n. \tag{3.1}$$

D'autre part, en utilisant le lemme 6 et les propriétés sur la maison on obtient :

$$\begin{aligned}\overline{|\gamma|}&\leq\sum_\ell \overline{|P_\ell(\xi+k)|}\overline{|\alpha|}^{mq}\overline{|\alpha^\beta|}^{2mq}\\ &\leq 2mnc_2^n n^n(2q+q\overline{|\beta|})^{2n}\overline{|\alpha|}^{mq}\overline{|\alpha^\beta|}^{mq}\leq c_5^n n^{2n}.\end{aligned}$$

Il n'y a plus qu'à établir une majoration de $|\gamma|$. On profite de l'existence de nombreux zéros de R pour définir la fonction **entière** suivante :

$$S(z)=R(z)\prod_{1\leq\lambda,\mu\leq q}\frac{\xi+k-\lambda-\mu\beta}{z-\lambda-\mu\beta}.$$

On note que $S(\xi+k)=R(\xi+k)$. D'après le théorème de Cauchy on a donc :

$$\gamma=S(\xi+k)=\int_{\mathscr{C}}\frac{S(z)}{z-\xi-k}dz,$$

où $\mathscr{C}$ est un cercle centré en $\xi + k$ et de rayon n assez grand. Si z est sur ce cercle on a les majorations :

$$|R(z)| \leq 2mn^{2n-1}(c_2^n n^n)|\overline{\alpha}|^{mn} \leq c_6^n n^{3n}$$

et

$$\left| \prod_{1 \leq \lambda, \mu \leq q} \frac{\xi + k - \lambda - \mu\beta}{z - \lambda - \mu\beta} \right| \leq (n - 2|\beta|q)^{-q^2}(c_7 q)^{q^2} \leq c_8^n n^{-mn/2},$$

car $n \geq c_9 q^2$. On obtient finalement un majoration de $|\gamma|$:

$$|\gamma| \leq \frac{1}{|2i\pi|} 2\pi n c_6^n n^{3n} c_8^n n^{-mn/2} \leq c_{10}^n n^{n(3-m/2)}. \tag{3.2}$$

On peut regrouper les 3 inégalités obtenues en remarquant que $\mathrm{den}(\gamma)\gamma$ est un entier algébrique non nul :

$$1 \leq \mathscr{N}(\mathrm{den}(\gamma)\gamma) \leq \mathrm{den}(\gamma)^d |\overline{\gamma}|^{d-1}|\gamma| \leq (c_4^n)^d (c_5^n n^{2n})^{d-1} c_{10}^n n^{n(3-m/2)}. \tag{3.3}$$

L'exposant de n^n qui intervient dans l'inégalité precédente est $2(d-1) + 3 - m/2$. En choisissant m suffisamment grand au début on aboutit bien à une contradiction. Ainsi, si β est irrationnel et $\alpha \neq 0, 1$, l'un des nombres $\alpha, \beta, \alpha^\beta$ est transcendant. Cela implique en particulier l'énoncé.

Chapitre 4

Critère de Schneider-Lang

4.1 Enoncé et conséquences du théorème

Dans ce chapitre, nous aborderons un important résultat dû à Schneider [**3**] et Lang [**1**], qui généralise la démonstration de Gelfond du théorème de Gelfond-Schneider à une large classe de fonctions méromorphes qui satisfont à une équation différentielle algébrique. L'un de ses intérêts est son application aux fonctions elliptiques que nous présenterons dans une dernière partie, résultats inconnus avant. Avant d'énoncer le critère, rappelons ce que l'on entend par ordre de croissance d'une fonction méromorphe.

Définition 6. Soit f une fontion entière. On dit que f est d'ordre au plus $\rho \geq 0$ si elle vérifie :

$$\exists R > 0,\ \exists M > 0,\ \forall r \geq R,\ \sup_{|z| \leq r} |f(z)| \leq M^{r^\rho}.$$

On dit qu'une fonction méromorphe est d'ordre au plus ρ si ses numérateur et dénominateur sont d'ordre au plus ρ.

Théorème 6. *Soit $\mathbb{K}$ un corps de nombres de degré d. Soient $f_1, ..., f_N$ des fonctions méromorphes d'ordre au plus ρ. On suppose que l'anneau $\mathbb{K}[f_1, ..., f_N]$ est stable par la dérivation d/dz et que le corps $\mathbb{K}(f_1, ..., f_N)$ est de degré de transcendance au moins 2* ***sur*** *$\mathbb{K}$. Alors si $\zeta_1, ..., \zeta_m$ vérifient $f_i(\zeta_j) \in \mathbb{K}$, pour tous i, j, on a l'inégalité : $m \leq 6\rho d$. En particulier, m est fini.*

Ce théorème permet déjà de récupérer une partie des résultats obtenus dans les chapitres précédents notamment le théorème de Gelfond-Schneider et le théorème d'Hermite-Lindemann.

- On se donne α, β algébriques, avec β irrationnel et $\alpha \neq 0, 1$ et on suppose α^β algébrique. On pose alors $f_1(z) = e^z$ et $f_2(z) = e^{\beta z}$, toutes les deux d'ordre fini égal à 1. Alors f_1 et f_2 sont algébriquement indépendante sur le corps de nombres $\mathbb{K} = \mathbb{Q}(\alpha, \beta, \alpha^\beta)$. Dans le cas contraire on aurait

l'existence d'un polynôme non trivial P tel que $P(e^z, e^{\beta z}) = \sum c_{k,j} e^{(k+j\beta)z}$ qui serait identiquement nul. Puisque les $k+j\beta$ sont non nuls par irrationalité de β, les $e^{(k+j\beta)z}$ sont $2i\pi/(k+j\beta)$ périodiques. On a donc une combinaison linéaire non triviale de fonctions périodiques, de périodes distinctes ce qui est impossible. Or, par définition de $\mathbb{K}$, l'anneau $\mathbb{K}[e^z, e^{\beta z}]$ est stable par dérivation. Si on pose $\zeta_j = j\log(\alpha)$ les valeurs $f_i(\zeta_j) = \alpha^j$ où $(\alpha^\beta)^j$ sont dans $\mathbb{K}$ quel que soit j. Cela contredit en particulier la finitude de m dans le théorème (6), aboutissant bien à une contradiction.

- Si maintenant on souhaite retrouver la transcendance de e^α, α algébrique non nul, on suppose le contraire, c'est à dire que $\mathbb{K} = \mathbb{Q}(\alpha, e^\alpha)$ est un corps de nombres. On pose $f_1(z) = z$, $f_2(z) = e^z$ et $\zeta_j = j\alpha$. Ces fonctions sont d'ordre fini. Puisque l'exponentielle n'est pas une fonction algébrique, f_1 et f_2 sont algébriquement indépendantes sur $\mathbb{K}$ (vu comme corps des fonctions constantes à valeurs dans $\mathbb{K}$). Enfin, l'anneau $\mathbb{K}[z, e^z]$ est évidemment stable par dérivation. Supposons e^α algébrique. Alors les nombres $\zeta_j = j\alpha$ vérifient $f_i(\zeta_j) \in \mathbb{K}$ pour tout j entier. Cela contredit encore une fois la finitude de m et on aboutit à une contradiction.

4.2 Principe de la démonstration et outils nécessaires

La preuve de ce résultat passe par la contruction d'une fonction auxiliaire que l'on va chercher à annuler à un grand ordre en différents points donnés, en contrôlant grâce au lemme de Siegel, la taille des paramètres. Les hypothèses du théorème sur l'anneau $\mathbb{K}[f_1, ..., f_N]$ nous encouragent à établir des résultats au préalable sur la dérivation de polynômes de $\mathbb{K}[X_1, ..., X_N]$.

Dérivations de polynômes à plusieurs variables

On se donne un polynôme P de $\mathbb{K}[X_1, ..., X_N]$ que l'on écrit comme une somme finie :

$$P = \sum_\alpha p_\alpha X^\alpha \qquad \text{où } \alpha = (\alpha_1, ..., \alpha_N) \text{ et } X^\alpha = X_1^{\alpha_1} ... X_N^{\alpha_N}.$$

On rappelle que le degré total de P, noté $\deg(P)$, désigne l'entier $\max\{\alpha_1 + ... + \alpha_N,\ p_\alpha \neq 0\}$ et on notera $\mathrm{den}(P)$, le plus petit dénominateur strictement positif commun à ses coefficients.

Enfin, on rappelle que la dérivation partielle par rapport à la i-ème variable, est définie par linéarité sur les monômes :

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

Définition 7. Une dérivation D sur $\mathbb{K}[X_1, ..., X_N]$ est une application linéaire dans lui-même qui vérifie l'identité du produit : $D(PQ) = D(P)Q +$

$D(Q)P$ pour tous polynômes P, Q. Elle est donc uniquement déterminée par ses valeurs sur les X_i.

Remarque 7. Soit D une dérivation sur $\mathbb{K}[X_1, ..., X_N]$. On vérifie par linéarité que l'on a la formule :

$$D(P) = \sum_i D(X_i)\frac{\partial P}{\partial X_i}. \tag{4.1}$$

Nous allons maintenant pouvoir définir une relation d'ordre sur les polynômes qui sera compatible avec les opérations.

Définition 8. Soient P et M deux polynômes de $\mathbb{K}[X_1, ..., X_N]$ où M est à coefficients positifs.
On dit que M domine P et on écrit $P \ll M$ si pour tout indice α, on a l'inégalité sur les coefficients correspondants : $\overline{|p_\alpha|} \leq m_\alpha$.

Remarque 8. La relation de domination est compatible avec la somme, le produit et la dérivation partielle. Pour la somme il suffit d'écrire, avec les notations évidentes, que si $P \ll M$ et $Q \ll N$ alors $\overline{|(p+q)_\alpha|} = \overline{|p_\alpha + q_\alpha|} \leq \overline{|p_\alpha|} + \overline{|q_\alpha|} \leq m_\alpha + n_\alpha$ donc $P + Q \ll M + N$. On vérifie le résultat concernant le produit et la dérivation partielle par additivité sur les cX^α.

Première estimation

Remarque 9. Si M domine P et si $(x_1, ..., x_N)$ est un N-uplet du corps K, alors on a l'inégalité $\overline{|P(x_1, ..., x_N)|} \leq M(\overline{|x_1|}, ..., \overline{|x_N|})$. Il suffit en effet de fixer un $\mathbb{Q}$-plongement σ de $\mathbb{K}$ dans $\mathbb{C}$ et d'utiliser sa multiplicativité et additivité pour obtenir le résultat :

$$\begin{aligned}|\sigma(P(x_1, ..., x_N))| &= |P(\sigma(x_1), ..., \sigma(x_N))| \\ &\leq \overline{|P|}(|\sigma(x_1)|, ..., |\sigma(x_N)|) \leq M(\overline{|x_1|}, ..., \overline{|x_N|}).\end{aligned}$$

On peut maintenant énoncer le résultat principal de cette partie.

Lemme 8. *Soient $f_1, ..., f_N$ des fonctions holomorphes au voisinage d'un point ζ. On suppose que les dérivées f_j' sont des polynômes en les $(f_i)_i$, c'est à dire que l'anneau de fonctions $\mathbb{K}[f_1, ..., f_N]$ est stable par l'opérateur de dérivation. Alors il existe un entier $C > 0$ tel que pour tout polynôme P et tout entier naturel k, on ait les majorations :*

$$\overline{|f^{(k)}(\zeta)|} \leq \overline{|P|} r^k k! C^{k+r}, \tag{4.2}$$

$$den(f^{(k)}(\zeta)) \mid den(P) C^{k+r}, \tag{4.3}$$

où $f = P(f_1, ..., f_N)$ et $r = deg(P)$.

Démonstration. Par hypothèse on a quel que soit i l'existence d'un polynôme P_i vérifiant $f_i' = P_i(f_1, ..., f_N)$. On note D l'unique dérivation qui vérifie $D(X_i) = P_i$. On vérifie facilement que $f^{(k)} = D^k(P)(f_1, ..., f_N)$. On note $T_n = (1 + X_1 + ... + X_N)^n$ et h le maximum des degrés de P_i.

Pour commencer on va établir par récurrence la majoration $D^k(P) \leq \overline{|P|} r^k C_1^k k! T_{r+k(h-1)}$, où C_1 ne dépend ni de k ni de P.

Si $k = 0$, le résultat est évident car le polynôme T_r contient tous les monômes de degré au plus r. On suppose donc $k \geq 1$ et que le résultat est vrai pour $k - 1$.

D'après la formule (4.1) on a :

$$\begin{aligned} D^{k+1}(P) = \sum_i P_i \frac{\partial D^k(P)}{\partial X_i} &\leq \sum_i \overline{|P_i|} T_h \frac{\partial \left(\overline{|P|} r^k C_1^k k! T_{r+k(h-1)} \right)}{\partial X_i} \\ &\leq \overline{|P|} r^k C_1^k k! \big(r + (h-1)k \big) \big(\sum_i \overline{|P_i|} \big) T_{r+(k+1)(h-1)-1}, \end{aligned}$$

grâce à la compatibilité de la domination aux opérations. D'où l'inégalité

$$D^{k+1}(P) \leq \overline{|P|} r^{k+1} C_1^{k+1} (k+1)! T_{r+(k+1)(h-1)}.$$

Ainsi on obtient

$$\begin{aligned} \overline{\left| f^{(k)}(\zeta) \right|} &= \overline{\left| D^k(P)(f_1(\zeta), ..., f_N(\zeta)) \right|} \\ &\leq \overline{|P|} (rC_1)^k k! T_{r+k(h-1)} (\overline{|f_1(\zeta)|}, ..., \overline{|f_N(\zeta)|}) \leq \overline{|P|} r^k C_2^{r+k} k!, \end{aligned}$$

ce qui donne la majoration (4.2).

Passons à la majoration (4.3) et notons $C_3 = \text{ppcm}(\text{den}(P_i))$. On va vérifier par récurrence sur k que $\text{den}(D^k(P))$ divise $\text{den}(P) C_3^k$. Le cas $k = 0$ est immédiat, on suppose $k \geq 1$ et le résultat au rang $k - 1$. D'après l'identité (4.1) on a :

$$\begin{aligned} \text{den}(P) C_3^{k+1} D^{k+1}(P) = \text{den}(P) C_3^{k+1} \sum_i P_i \frac{\partial D^k(P)}{\partial X_i} \\ = \sum_i C_3 P_i \frac{\partial (\text{den}(P) C_3^k D^k(P))}{\partial X_i}. \end{aligned}$$

Or les polynômes $C_3 P_i$ et $\text{den}(P) C_3^k D^k(P))$ sont dans $\mathscr{O}_K[X_1, ..., X_N]$ qui est stable par somme, produit et dérivation partielle. On en déduit que $\text{den}(P) C_3^{k+1} D^{k+1}(P)$ est à coefficients dans $\mathscr{O}_K$ ce qui achève la récurrence. D'autre part, le polynôme $D^k(P)$ est de degré au plus $r + kh$. Donc en spécialisant pour les $f_i(\zeta)$ on trouve que

$$\begin{aligned} &\text{ppcm}(\text{den}(f_i(\zeta)))^{r+kh} \text{den}(P) C_3^k D^k(P)(f_1(\zeta), ..., f_N(\zeta)) \\ &\qquad = \text{ppcm}(\text{den}(f_i(\zeta)))^{r+kh} \text{den}(P) C_3^k f^{(k)}(\zeta) \in \mathscr{O}_K. \end{aligned}$$

Finalement, grâce a l'inégalité :

$$|\mathrm{ppcm}(\mathrm{den}(f_i(\zeta)))^{r+kh}\mathrm{den}(P)C_3^k| \leq \mathrm{den}(P)C_4^{k+r}$$

pour un $C_4 > 0$ indépendant de k et P, l'assertion (4.3) sur le dénominateur de $f^{(k)}(\zeta)$ est démontrée.

□

Construction d'une fonction auxiliaire

L'hypothèse sur le corps $K(f_1, ..., f_N)$ nous permet de considérer deux fonctions f et g parmi les f_i qui sont algébriquement indépendantes sur K. On se donne également un entier r tel que $n := \frac{r^2}{2m}$ soit entier. On définit F par $F(z) := \sum_{0\leq i,j\leq r} c_{i,j} f^i g^j$ où les coefficients $c_{i,j}$ sont à ajuster.

Lemme 9. *Il existe une famille $(c_{i,j})_{i,j}$ dans $\mathscr{O}_K$, tel que chaque $c_{i,j}$ a sa maison majorée par $C_1^n n^{3/2n}$, et qui vérifie les conditions :*

$$\forall k < n, \forall p \leq m, \quad F^{(k)}(\zeta_p) = 0. \tag{4.4}$$

Démonstration. Les équations (4.4) se traduisent en un système linéaire qui fait intervenir $r^2 = 2mn$ inconnues que sont les $c_{i,j}$, pour mn équations. Pour appliquer le lemme 5 de Siegel il nous reste à trouver un dénominateur commun des $(f^i g^j)^{(k)}(\zeta_p)$ à k, p fixés et à majorer leurs maisons.

On applique le lemme 8 de la partie précédente à $P = X_1^i X_2^j$ et $r = i + j$:
(*i*) $\overline{|(f^i g^j)^{(k)}(\zeta_p)|} \leq \overline{|P|}(i+j)^k k! C^{k+i+j} \leq (2r)^n n! C^{3n}$,
(*ii*) $\mathrm{den}((f^i g^j)^{(k)}(\zeta_p)) \mid \mathrm{den}(P) C^{k+i+j} \mid C^{3n}$.

Les éléments $C^{3n}(f^i g^j)^{(k)}(\zeta_p)$ sont des entiers algébriques. D'après le lemme 5 de Siegel, il existe donc des $c_{i,j}$ entiers algébriques, non tous nuls, qui satisfont à ((4.4)) et qui vérifient la majoration :

$$\overline{|c_{i,j}|} \leq (1/C)^{3n}\, b(bmn(2r)^n n! C^{3n})^{\frac{mn}{2mn-mn}} \leq C_1^n n^{3/2n}.$$

□

Par l'hypothèse d'indépendance algébrique sur f et g, la fonction F est non identiquement nulle. Il existe donc un entier $s \geq n$, qui vérifie :
(*i*) Pour $k < s$, $F^{(k)}(\zeta_p) = 0$ quel que soit p.
(*ii*) Il existe un entier p, que l'on peut supposer égal à 1 sans perte de généralité, qui vérifie $\gamma := F^{(s)}(\zeta_p) \neq 0$.

Le nombre γ est algébrique par les hypothèses sur les f_i. On note q son dénominateur.

4.3 Preuve du critère

La preuve du théorème repose principalement sur des estimations concernant le nombre γ défini précédemment.

Estimations de γ

On commence par donner une borne supérieure sur le dénominateur et la maison de γ.

Lemme 10. *Il existe deux constantes $C_3, C_5 > 0$, indépendantes de n, s, telles que $q \leq C_3^s$ et $\overline{|\gamma|} \leq C_5^s s^{3s}$.*

Démonstration. Nous allons encore appliquer le lemme 8 de la partie précédente. On pose $P = \sum_{i,j} c_{i,j} X_1^i X_2^j$ et $h := P(f,g)$. On a $\deg(P) \leq 2r$. Puisque les coefficients $c_{i,j}$ obtenus précédemment sont entiers et d'après l'inégalité (4.3), on a :

$$q = \mathrm{den}(F^{(s)}(\zeta_1)) = \mathrm{den}(P(f,g)^{(s)}(\zeta_1)) \leq \mathrm{den}(P) C_2^{s+2r} = C_3^s.$$

On passe à la majoration de $\overline{|\gamma|}$ en utilisant l'inégalité (4.2) :

$$\overline{|\gamma|} = \overline{\left|P(f,g)^{(s)}(\zeta_1))\right|} \leq \overline{|P|}(2r)^s s! C_4^{s+2r}.$$

De plus les maisons des $c_{i,j}$ sont majorées par $C_1^n n^{3/2n}$. On obtient donc les inégalités :

$$\overline{|\gamma|} \leq C_1^n n^{3/2n} (2r)^s s! C_4^{s+2r} \leq C_1^s s^{3/2s} \sqrt{2m}^s \sqrt{s}^s s^s C_4^{s+2r} \leq C_5^s s^{3s}.$$

□

Revenons à notre objectif qui est de majorer m. Notre fonction auxiliaire F s'annule en les m points $(\zeta_i)_{i \leq m}$ à un ordre au moins s ce qui va permettre de la diviser par une grande fonction sans rajouter de pôles. Nous pourrons donc grâce à la formule de Cauchy obtenir une très bonne majoration de $|\gamma|$. Il reste au préalable à s'affranchir des défauts d'holomorphie de F, donc de f et g. On rappelle que f, g sont méromorphes d'ordre au plus ρ.

Lemme 11. *Il existe une constante $C_9 > 0$, indépendante de s, telle que $|\gamma| < C_9^s s^{(5/2 - \frac{m}{2\rho})s}$.*

Démonstration. Considérons la fonction entière θ, définie comme le produit des dénominateurs de f et g. Par hypothèse c'est un produit de fonctions d'ordre au plus ρ, donc θ est entière d'ordre au plus ρ. De plus par définition du dénominateur d'une fonction méromorphe $\theta f, \theta g$ sont également entières. Enfin, puisque ζ_1 n'est pôle ni de f ni de g, θ ne s'annule pas en ζ_1.

Par définition de F, $\theta^{2r}F$ est entière. Nous allons appliquer la formule de Cauchy à la fonction holomorphe H, définie par :

$$H(z) = \frac{\theta^{2r}(z)F(z)}{\prod_{p=1}^{m}(z-\zeta_p)^s}.$$

On écrit la formule intégrale pour un cercle C, de centre ζ_1 et de rayon R dépendant de s, à expliciter plus tard :

$$H(\zeta_1) = \frac{1}{2i\pi}\int_C \frac{H(\zeta)}{\zeta-\zeta_1}d\zeta = \frac{1}{2i\pi}\int_C \frac{\theta^{2r}(\zeta)F(\zeta)}{\prod_{p=1}^{m}(\zeta-\zeta_p)^s(\zeta-\zeta_1)}d\zeta. \tag{4.5}$$

L'hypothèse sur l'ordre de $\theta, \theta f$ et θg assure que pour R assez grand il existe une constante C_6, indépendante de s, qui vérifie pour tout z de module R :

$$|\theta(z)| \leq C_6^{R^\rho},\ |\theta(z)f(z)| \leq C_6^{R^\rho},\ |\theta(z)g(z)| \leq C_6^{R^\rho}.$$

On peut donc majorer $H(\zeta_1)$ en utilisant (4.5) :

$$\begin{aligned}|H(\zeta_1)| &\leq \frac{1}{2\pi}\int \sum_{i,j} |\overline{c_{i,j}}| \frac{|\theta f(\zeta)|^i |\theta g(\zeta)|^j |\theta(\zeta)|^{2r-i-j}}{\prod|\zeta-\zeta_p|^s}d\zeta \\ &\leq \frac{r^2 C_1^s s^{3/2s} C_6^{2rR^\rho}}{2\pi(R-M)^{ms}} \leq \frac{C_7^s s^{3/2s} C_6^{2rR^\rho}}{R^{ms}}.\end{aligned} \tag{4.6}$$

où M est le maximum des modules des ζ_p.

Il nous faut encore trouver le lien entre γ et $H(\zeta_1)$ et on procède alors comme suit. Au facteur $s!$ près, γ n'est autre que le coefficient a_s du développement analytique, au point ζ_1, de F. On écrit, au voisinage du point ζ_1, $F(z) = (z-\zeta_1)^s(a_s + ...) = H(z)\prod_{p=1}^{m}(\zeta-\zeta_p)^s/\theta^{2r}(z)$. Donc par non-annulation de θ au voisinnage de ζ_1 on a :

$$a_s = H(\zeta_1)\prod_{p=2}^{m}(\zeta_1-\zeta_p)^s/\theta^{2r}(\zeta_1) = \gamma/s!$$

Ainsi on a finalement $|\gamma| \leq s! C_7^s s^{3/2s} C_6^{2rR^\rho}/R^{ms}$. Nous pouvons choisir $R = s^{1/2\rho}$ qui tend bien vers l'infini quand s tend vers l'infini. Avec ce choix on a $|\gamma| \leq s^s C_7^s s^{3/2s}(C_6^{\sqrt{m}})^s/s^{\frac{ms}{2\rho}} \leq C_9^s s^{(5/2-\frac{m}{2\rho})s}$ □

Conclusion

La preuve du théorème est presque terminée, il reste juste à rassembler les divers résultats obtenus. Le nombre $q\gamma$ est entier non nul donc on a les inégalités :

$$1 \leq \mathscr{N}(q\gamma) \leq q|\gamma|\overline{|\gamma|}^{d-1} \leq C_3^s (C_5^s s^{3s})^{d-1} C_9^s s^{(5/2-\frac{m}{2\rho})s}. \tag{4.7}$$

En identifiant la puissance de s intervenant dans l'inégalité précédente on trouve $3(d-1)+5/2-m/2\rho \geq 0$ donc $m \leq 6\rho d$.

Remarque 10. La majoration $m \leq 6\rho d$ n'est pas optimale, mais sa finitude permet d'obtenir les résultats de transcendance indiqués.

Chapitre 5

Applications aux fonctions elliptiques

Dans cette dernière partie on va s'intéresser aux fonctions elliptiques et modulaires. On appliquera notamment le critère de Schneider-Lang à la fonction $\wp$ de Weierstrass. Au vu du nombre important de nouveaux concepts on ne présentera pas toutes les démonstrations. On peut trouver plus de détails dans [**6**] ou dans [**8**] chapitre 7.

5.1 Fonction $\wp$ de Weierstrass

Définition 9. Une fonction ϕ de la variable complexe est dite elliptique si elle est méromorphe sur $\mathbb{C}$ et doublement périodique : il existe deux nombres w_1, w_2 non $\mathbb{R}$-colinéaires qui verifient $\phi(z + w_2) = \phi(z + w_1) = \phi(z)$ pour tout z.

On peut remarquer que la donnée des périodes w_1, w_2 est la donnée d'un réseau Λ de $\mathbb{C}$, c'est à dire d'un sous groupe de rang 2 de $\mathbb{C}$ de la forme $\Lambda = w_1\mathbb{Z} \oplus w_2\mathbb{Z}$. Une fonction elliptique sur un tel réseau est donc une fonction méromorphe qui prend les mêmes valeurs sur chaque domaine dit "fondamental" de la forme $\{(n+\lambda)w_1 + (m+\nu)w_2,\ \lambda, \mu \in [0, 1[\}$ où n et m sont des entiers fixés.

Cette remarque nous conduit à considérer des fonction de la forme $g(z) = \sum_{\lambda\in\Lambda}(z-\lambda)^{-k}$. Malheuresement une telle série ne converge pas pour toutes les valeurs de k mais seulement si $k \geq 3$. Il serait cependant intéressant de pouvoir considérer $k = 2$ dans la série ci-dessus, les autres s'obtenant par dérivés successives. On introduit donc la série modifiée :

$$\wp(z) = \frac{1}{z^2} + \sum_{\lambda\in\Lambda\setminus\{0\}} \left(\frac{1}{z+\lambda}\right)^2 - \left(\frac{1}{\lambda}\right)^2. \tag{5.1}$$

On peut montrer que la série ci-dessus converge normalement sur tout compact en dehors de Λ où elle admet des pôles doubles. Il est facile de vérifier alors qu'elle est bien doublement périodique selon Λ. Nous avons construit notre première fonction elliptique, dite de Weierstrass. Sa dérivée, également elliptique, s'écrit :

$$\wp'(z) = -2\left(\frac{1}{z^3} + \sum_{\lambda\in\Lambda-\{0\}}\left(\frac{1}{z+\lambda}\right)^3\right) = -2\sum_{\lambda\in\Lambda}\left(\frac{1}{z+\lambda}\right)^3.$$

La fonction $\wp$ satisfait l'équation différentielle elliptique suivante :

$$\wp'(z)^2 = 4\wp(z)^3 - G_2\wp(z) - G_3 \quad \text{où } G_k := \sum_{\lambda\in\Lambda\setminus\{0\}}\left(\frac{1}{\lambda}\right)^{2k}. \tag{5.2}$$

On dit que G_2 et G_3 sont les invariants de la fonction elliptique $\wp$ associée au réseau Λ.

Pour le voir on considère la différence des deux membres de (5.2) qui est méromorphe et Λ-périodique. Le seul pôle éventuel (modulo périodicité) est en zéro mais un développement limité montre que la fonction obtenue est nulle à l'origine. Elle est holomorphe et donc bornée par périodicité. Le théorème de Liouville permet de conclure qu'elle est nulle.

En dérivant l'équation (5.2) on remarque que $\wp''$ s'écrit comme un polynôme en $\wp$. On peut se demander plus généralement, quelles fonctions elliptiques sur Λ s'écrivent comme fraction rationnelle en $\wp$. La parité de $\wp$ force une telle fonction à être paire. Cette condition est en faite suffisante. De même, toute fonction elliptique impaire s'écrit comme fraction rationnelle en $\wp'$. On conclut cette partie par la formule dite d'addition concernant la fonction $\wp$. Si z et z' sont tels que z, z' et $z+z'$ n'appartiennent pas aux pôles de $\wp$, on a l'identité suivante :

$$\wp(z+z') + \wp(z) + \wp(z') = \frac{1}{4}\left(\frac{\wp'(z)-\wp'(z')}{\wp(z)-\wp(z')}\right)^2. \tag{5.3}$$

5.2 Formes modulaires

On désigne par $\mathfrak{H}$ le demi-plan de Poincaré, c'est à dire l'ensemble $\{z \in \mathbb{C}, \mathfrak{Im}(z) > 0\}$. Le groupe $SL_2(\mathbb{Z})$ agit sur $\mathfrak{H}$ de la façon suivante :

$$\begin{array}{ccc} SL_2(\mathbb{Z})\times\mathfrak{H} & \to & \mathfrak{H} \\ (A,z) & \mapsto & \frac{az+b}{cz+d} \end{array}$$

On note cette action $A\cdot z$.

Définition 10. Soit f une fonction définie, holomorphe sur $\mathfrak{H}$. On dit que f est une forme modulaire de poids k si elle satisfait les deux conditions suivantes :

- $f(A \cdot z) = (cz + d)^{-k} f(z)$
- f est bornée en ∞. On dit alors que f est holomorphe aux pointes.

Remarque 11. Pour k fixé, l'ensemble M_k des formes modulaires de poids k forme un $\mathbb{C}$-espace vectoriel. De plus, le produit d'une forme modulaire de poids k et d'une forme modulaire de poids k' est une forme modulaire de poids $k+k'$. Ainsi l'ensemble des formes modulaires $\mathscr{M}$ admet une structure de $\mathbb{C}$-algèbre graduée par le poids : $\mathscr{M} = \bigoplus_k M_k$.

On note que le groupe $SL_2(\mathbb{Z})$ est engendré par les deux matrices :

$$T = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \quad S = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}.$$

La matrice T agit sur $\mathfrak{H}$ par translation de 1. Au vu de ces remarques, la première condition peut s'écrire de manière plus simple. En effet il suffit que f soit 1-périodique et vérifie $f(-\frac{1}{z}) = z^k f(z)$. En appliquant la première condition à $A = -Id$ on obtient la relation $f(z) = (-1)^k f(z)$. Ainsi pour k impair il n'y a pas de forme modulaire non triviale. La 1-périodicité de f a pour conséquence qu'elle se factorise par la fonction $e : z \longrightarrow e^{2i\pi z}$. Plus précisément on a la décomposition de Fourier suivante de f :

$$f(z) = \sum_{n\in\mathbb{Z}} a_n e^{2i\pi nz} = \sum_{n\in\mathbb{Z}} a_n q^n =: \tilde{f}(q), \tag{5.4}$$

où $q = e^{2i\pi z}$ est dans le disque unité ouvert. La condition d'holomorphie aux pointes implique précisément que les a_n sont nuls pour n strictement négatif.

Le lien avec les courbes elliptiques est le suivant. On se donne un réseau $\Lambda = w_1\mathbb{Z} \oplus w_2\mathbb{Z}$ et k un enier supérieur ou égal à 2. On va définir un type important de forme modulaire appelées séries d'Eisenstein, qui seront associées aux fonction elliptiques sur le réseau Λ :

$$G_k(\tau) := \sum_{(n,m)\in\mathbb{Z}^2\setminus\{(0,0)\}} \left(\frac{1}{n+m\tau}\right)^{2k} = \sum_{\lambda\in\Lambda\setminus\{(0,0)\}} \frac{1}{\lambda^{2k}}\,. \tag{5.5}$$

Un calcul simple montre que cette série converge bien sur $\mathfrak{H}$ et que G_k est modulaire de poids $2k$. Or au facteur w_1^{2k} près, $G_2(\tau)$ et $G_3(\tau)$ sont les invariants de la courbe elliptique associée aux périodes w_1 et $w_2 = \tau w_1$. Ainsi nos premières formes modulaires non triviales sont données par G_2 et G_3.

Proposition 2. *Pour tout entier k, L'espace des formes modulaires de poids k est engendré par les $G_2^n G_3^m$ où $4n + 6m = k$.*

Les formes modulaires sont de fait très rigides ce qui va amener à considérer une classe plus large de fonctions comprenant les quotient de formes modulaires :

Définition 11. On appelle fonction modulaire toute fonction méromorphe sur $\mathfrak{H}$ qui vérifie :
- $f(A \cdot z) = (cz + d)^{-k} f(z)$
- f admet un pôle en ∞, *i.e* on dit que f est méromorphe aux pointes.

Remarque 12. Ces fonctions admettent un développement de Fourier de la forme :

$$f(z) = \sum_{n \geq M}^{\infty} a_n q^n = \tilde{f}(q) \text{ où } q = e^{2i\pi z}.$$

On s'intéresse particulièrement pour les théorèmes de transcendance à deux fonctions modulaires, le discriminant modulaire Δ et l'invariant modulaire J, vu comme fonction sur le disque unité ouvert :

$$\Delta(z) = \frac{(G_2^3 - G_3^2)(z)}{1728} = q \prod_{n \geq 1} (1 - q^n)^{24} = \tilde{\Delta}(q), \quad j(z) = \frac{G_2^3(z)}{\Delta(z)} = J(q). \tag{5.6}$$

Δ est une forme modulaire de poids 12 et j est une fonction modulaire de poids 0.

5.3 Les résultats de transcendance

Théorème 7. *Soit τ dans $\mathfrak{H}$. Si τ et $j(\tau)$ sont tous les deux algébriques, alors τ est un irrationnel quadratique sur $\mathbb{Q}$.*

Démonstration. On va utiliser la fonction de Weierstrass $\wp$ associée à un réseau bien choisi. Supposons que τ et $j(\tau)$ soient algébriques. On définit nos deux périodes par $w_1 = 2\pi\tilde{\Delta}(q)^{1/12}$ et $w_2 = \tau w_1$, où $q = e^{2i\pi\tau}$ et $\Delta(q)^{1/12}$ est une racine 12-ème quelconque de $\tilde{\Delta}(q)$.

On rappelle que $\wp$, associée au réseau $\Lambda = w_1\mathbb{Z} \oplus w_2\mathbb{Z}$, vérifie

$$\wp'(z)^2 = 4\wp(z)^3 - 60G_2\wp(z) - 140G_3$$
$$\text{où } G_k := \sum_{\lambda \in \Lambda - \{0\}} \lambda^{-2k} = w_1^{-2k} G_k(\tau).$$

Or G_2, G_3 sont algébriques car ils vérifient $G_2^3 - 27G_3^2 = 1$ et $1728G_2^3 = j(\tau)$. On va appliquer le critère de Schneider-Lang à $f_1 = \wp(z), f_2 = \wp(\tau z)$. On pose $K = \mathbb{Q}(\tau, G_2, G_3, \wp(w_1/2), \wp(w_2/2))$ qui est un corps de nombres. En effet, les nombres $\wp(w_1/2)$ et $\wp(w_2/2)$ vérifient $4\wp(w_i/2)^3 - G_2\wp(w_i/2) - G_3 = \wp'(w_i/2)^2$. Or $\wp'$ est w_i-périodique et impaire donc cette dernière quantitée est nulle ce qui assure que les $\wp(w_i/2)$ sont algébriques.

L'anneau $\mathbb{K}[f_1, f_2, f_1', f_2']$ est stable par dérivation d'après l'équation différentielle (5.2) satisfaite par $\wp$ et puisque τ est dans $\mathbb{K}$. La fonction $\wp$ est de plus d'ordre fini ce qui est assuré par sa double périodicité. Enfin puisque $\wp(w_1/2), \wp(w_2/2)$ sont algébriques, les nombres $\wp(w_1(n+1/2)), \wp(w_2(n+1/2))$ le sont aussi par périodicité et les fonctions f_1, f_2, f_1', f_2' prennent simultanément une infinité de valeurs dans le corps $\mathbb{K}$. Si le degré de transcendance de $\mathbb{K}[f_1, f_2, f_1', f_2']$ était au moins 2, on aurait une contradiction en appliquant le critère de Schneider-Lang. Ainsi, $\wp(z)$ et $\wp(\tau z)$ sont algébriquement dépendantes sur $\mathbb{K}$. Il nous reste à montrer que cela force τ à être quadratique sur $\mathbb{Q}$.

Soit $P(\wp(z), \wp(\tau z)) = 0$ une relation non triviale. Nous allons montrer que les périodes de $\wp(\tau z)$ (w_1/τ et w_1 par exemple) sont nécessairement des multiples rationnels de périodes de $\wp$ à savoir w_1 et τw_1. Cela implique évidemment que τ est quadratique puisque dans ce cas on a $w_1/\tau \in w_1\mathbb{Q} \oplus \tau w_1 \mathbb{Q}$ donc $1 \in \tau\mathbb{Q} \oplus \tau^2\mathbb{Q}$. Soit effectivement T une période de $\wp$. On prend un point a_0, tel que $P(a_0, \zeta) = 0$ implique $\partial P/\partial X_2(a_0, \zeta) \neq 0$. D'après le théorème des fonctions implicites, il existe $\phi_1, ..., \phi_d$ holomorphes au voisinage de a_0, tels que les solutions de l'équation $P(z, \zeta)$, dans ce voisinage, s'écrivent $\zeta = \phi_j(z)$. Soit z_0 tel que $\wp(z_0) = a_0$. On considère la suite $z_0, z_0 + T, ...$ Au voisinage de ces points $z_0 + nT$, la fonction $\wp(\tau z)$ s'écrit donc $\phi_{j_n}(\wp(z))$ pour un certain $j_n \leq d$ car $\wp(z_0 + nT) = a_0$. Par finitude du nombre de ϕ_j, il existe donc deux entiers m et n tels que le valeurs prises par $\wp(\tau z)$ au voisinage de $z_0 + nT$ et $z_0 + mT$ soient les mêmes. Il s'ensuit par principe du prolongement analytique que $\wp(\tau z)$ est $(m-n)T$-périodique, ce qui montre le résultat. □

On définit

$$\sigma(z) := z \prod_{w \in \Lambda} \left(1 - \frac{z}{w}\right) e^{z/w + z^2/(2w^2)}$$

appellé fonction σ de Weierstrass, où Λ est un réseau de $\mathbb{C}$. Sa dérivée logarithmique, $\zeta = \sigma'/\sigma$ vérifie $\zeta' = -\wp$ où $\wp$ est associée au réseau Λ. La fonction ζ est alors quasi-périodique au sens suivant : si $\Lambda = w_1\mathbb{Z} \oplus w_2\mathbb{Z}$, alors il existe deux nombres η_1 et η_2 tels que :

$$\zeta(z + w_i) = \zeta(z) + \eta_i \ , \qquad i = 1, 2.$$

On appelle η_1 et η_2 les quasi-périodes de la fonction ζ de Weierstrass associée au réseau Λ.

Théorème 8. *Soit $\alpha \in \mathbb{C}$ qui n'est pas un pôle de $\wp$. Supposons que les invariants G_2 et G_3 de $\wp$ soient algébriques. On suppose en plus qu'il existe deux nombres algébriques a et b, non tous les deux nuls, tels que $a\zeta(\alpha) + b\alpha$ soit algébrique. Alors $\wp(\alpha)$ est transcendant.*

Démonstration. On suppose que les hypothèses du théorème sont vérifiées. On choisit $f_1(z) := az + b\zeta(z)$ et $f_2(z) = \wp(z)$. La fonction f_1 satisfait à l'équation différentielle

$$f_1^{(3)}(z) + \frac{6}{b} f_1'(z)^2 - \frac{12a}{b} f_1'(z) + \frac{12a^2 + G_2 b^2}{2b} = 0,$$

et f_2 satisfait également à une équation différentielle algébrique. Supposons $\wp(\alpha)$ algébrique. Du théorème d'addition (5.3) et de l'équation différentielle (5.2) satisfaite par $\wp$, il suit que les nombres $\wp(k\alpha), \wp'(k\alpha)$, tant que ces quantitées sont définies, sont des nombres algébriques et appartiennent à $\mathbb{K} = \mathbb{Q}(a, b, G_2, G_3, \wp(\alpha), \wp'(\alpha), f_1(\alpha))$. La fonction f_1 vérifie également une formule d'addition, à savoir

$$f_1(z + z') = f_1(z) + f_1(z') + \frac{b}{2} \frac{\wp'(z) - \wp'(z')}{\wp(z) - \wp(z')}. \tag{5.7}$$

Puisque $f_1(\alpha)$ est supposé algébrique, la formule (5.7) implique que les valeurs $f_1(k\alpha)$ sont dans $\mathbb{K}$. On peut appliquer le critère de Schneider-Lang pour aboutir à une contradiction. La fonction σ est d'ordre fini donc sa dérivée σ' l'est aussi (voir **[5]**) et finalement ζ et $\wp$ sont d'ordre fini. Enfin les fonction f_1 et f_2 sont algébriquement indépendantes sur $\mathbb{K}$. En effet, si un polynôme P non nul vérifiait $P(az + b\zeta(z), \wp(z)) = 0$, alors on aurait par périodicité $P(az + b\zeta(z) + k(aw_i + b\eta_i), \wp(z)) = 0$ quelque soit k. C'est une équation polynomiale en k de la forme $Q(\wp(z))(aw_i + b\eta_i)^d k^d + p_1 k^{d-1} + ... = 0$ avec Q non nul, donc il existe z tel que $Q(\wp(z)) \neq 0$. En divisant par k^d et en faisant tendre k vers l'infini, il s'ensuit alors que $aw_i + b\eta_i = 0$. Puisque a et b ne sont pas tous les deux nuls, on obtient $w_1\eta_2 - w_2\eta_1 = 0$ ce qui est contraire à un résultat de Legendre qui montre que cette quantitée vaut $2i\pi$. On pourra se reporter à **[12]** pour une démonstration de ce résultat. Les valeurs $f_i(k\alpha)$ sont dans $\mathbb{K}$, ce qui n'est pas possible d'après le critère de Schneier-Lang. Il suit de cette contradiction que α ne peut être algébrique. □

On énonce finalement un résultat dit théorème Stéphanois, datant de 1996, connu autrefois comme la conjecture de Mahler-Manin. On note $\mathbb{C}_p$ le corps des nombres complexes p-adiques, qui est définit comme la complétion de la clôture algébrique de $\mathbb{Q}_p$, le corps des nombres p-adiques. On rappelle que J est défini en (5.6).

Théorème 9. - *Soit q un élément non nul dans le disque unité ouvert de $\mathbb{C}$ ou de $\mathbb{C}_p$. Alors si q est algébrique, $J(q)$ est transcendant.*

On ne traitera pas le cas p-adique, mais ce dernier se ramène facilement au cas complexe. Le point crucial de la démonstration est l'existence d'une relation algébrique $\Phi_n(J(q), J(q^n)) = 0$ et la bonne compréhension du polynôme modulaire $\Phi_n \in \mathbb{Z}[X, Y]$. Ce dernier est symétrique en les variables

X et Y, son coefficient dominant en chaque indéterminée est 1, et enfin son degré partiel sur chaque indéterminée est

$$\psi(n) := n \prod_{p|n} (1 + 1/p).$$

Lemme 12. *On définit la longueur d'un polynôme P dans $\mathbb{Z}[X, Y]$, notée $L(P)$, comme étant la somme des valeurs absolues de ses coefficients. Quel que soit $\epsilon > 0$, il existe une constante C telle que l'on ait les majorations*
(i) $\psi(n) \leq Cn^{1+\epsilon}$ *et*
(ii) $\log(L(\Phi_n)) \leq Cn^{3/2}$.

On peut trouver une preuve de ce lemme dans **[9]**. La majoration de (i), n'est pas optimale mais se montre facilement comme on va le voir. En effet, si on se donne $\epsilon > 0$ alors quel que soit n on a :

$$\prod_{p|n} (1 + 1/p) \leq \underbrace{\prod_{p|n, p<2/\epsilon} (1 + 1/p)}_{C_\epsilon} \prod_{p|n, p\geq 2/\epsilon} (1 + 1/p) \leq C_\epsilon (1 + \epsilon/2)^G,$$

où G désigne le cardinal de l'ensemble des premiers divisant n et C_ϵ ne dépendant pas de n. Bien sûr, la quantitée G est majorée par $\log_2(n)$ donc on obtient :

$$\psi(n) \leq nC_\epsilon (1 + \epsilon/2)^{\log_2(n)} \leq nC_\epsilon 2^{\epsilon \log_2(n)} \leq C_\epsilon n^{1+\epsilon},$$

car $(1 + \epsilon/2) \leq 2^\epsilon$ pour ϵ suffisamment petit, ce qui montre le résultat.

La fonction J définie sur le disque unité admet un unique pôle simple en 0. Il suit que la fonction $\tilde{J}(z) = zJ(z)$ est holomorphe sur le disque. On peut majorer les coefficients de son développement de Taylor en 0 : $\tilde{J}(z) = \sum_{n\geq 0} c(n) z^n$, par

$$0 \leq c(n) \leq e^{c_0 \sqrt{n}}.$$

Plus généralement, si on note $\sum_{n\geq 0} c_k(n) z^n$ le développement de Taylor en 0 de $\tilde{J}^k(z)$, on a les inégalités :

$$0 \leq c_k(n) \leq e^{c_0 \sqrt{kn}}.$$

On peut en trouver une démonstration dans **[2]**.

Lemme 13. *La série formelle définissant J est transcendante sur $Q(X)$. Autrement dit, tout polynôme P de $Q[X, Y]$ tel que $P(z, J(z))$ soit identiquement nul est nul.*

Démonstration. Soit P un tel polynôme. Par définition de l'invariant J on a donc $P(e^{2i\pi z}, j(z)) = 0$, $z \in \mathfrak{H}$. On fixe t un réel transcendant supérieur à 1.

Par modularité, on a $P(e^{2i\pi A\cdot it}, j(it)) = 0$ quelque soit A dans $SL_2(\mathbb{Z})$. On prend en particulier A de la forme

$$A_k = \begin{bmatrix} 1 & 0 \\ k & 1 \end{bmatrix}$$

où k parcourt $\mathbb{Z}$. Puisque $A_k \cdot (it) = \frac{it}{kit+1}$, les nombres $e^{2i\pi A_k\cdot it} = e^{2i\pi \frac{it}{kit+1}}$ sont tous distincts. En effet, les nombres $\frac{it}{kit+1} - \frac{it}{k'it+1} = \frac{-(k-k')t^2}{(kit+1)(k'it+1)}$ ne peuvent être dans $\mathbb{Z}$ par transcendance de t. Ainsi le polynôme $P(X, j(it))$ admet une infinité de solutions, il est donc nul. On peut écrire

$$P(X,Y) = \sum_k P_k(Y)X^k$$

avec les $P_k(j(it))$ nuls. Or lorsque t parcourt les nombres transcendants supérieurs à 1, $j(it)$ prend une infinité de valeurs. Sinon il existerait une valeur c que j prendrait un nombre indénombrable de fois et puisque l'on peut écrire $\mathfrak{H}$ comme réunion dénombrable de compacts, il existerait un zéro non isolé de $j-c$ et j serait constante. Il s'ensuit que les P_k sont nuls et P également. □

Il nous reste encore deux majorations générales à obtenir. Commençons par une remarque qui sert pour la preuve du lemme qui suit. Si $\alpha = \alpha_1$ est un nombre algébrique, α_2, ...,α_d ses conjugués et a_d le coefficient dominant du polynôme minimal de α_1, les quantitées

$$a_d \cdot \prod_{k\in I} \alpha_k, \qquad I \subset \{1, ..., d\}$$

sont des entiers algébriques. Pour le voir on définit d'abord $\mathrm{den}_{\mathbb{Q}}(\alpha)$ comme le plus petit nombre rationnel strictement positif r tel que $r\alpha$ soit entier. La fonction $\mathrm{den}_{\mathbb{Q}}$ est multiplicative : $\mathrm{den}_{\mathbb{Q}}(\alpha\beta) = \mathrm{den}_{\mathbb{Q}}(\alpha)\mathrm{den}_{\mathbb{Q}}(\beta)$ quels que soient α et β algébriques. Puisque le conjugué d'un entier est un entier, il s'ensuit immédiatemment que $\mathrm{den}_{\mathbb{Q}}(\alpha_i) = \mathrm{den}_{\mathbb{Q}}(\alpha_1)$ pour tout $i = 1, ..., d$. Puisque $a_d \cdot \prod_{k=1,...,d} \alpha_k$ est entier (c'est le coefficient constant du polynôme minimal de α), a_d est un multiple entier de $\mathrm{den}_{\mathbb{Q}}(\prod_{k=1,...,d} \alpha_k) = \mathrm{den}_{\mathbb{Q}}(\alpha)^d$. Si $\mathrm{den}_{\mathbb{Q}}(\alpha) = p/q$ avec $p^q = 1$, il s'ensuit que l'entier a_d est multiple de p^d donc de p^k pour $k \leq d$. Au final le nombre $\prod_{k\in I} p\alpha_k$, et à fortiori $a_d \cdot \prod_{k\in I} \alpha_k$, est entier.

Définissons la mesure de Mahler d'un nombre algébrique $\alpha = \alpha_1$. Si $\alpha_2, ..., \alpha_d$ sont ses conjugués, et a_0 le coefficient dominant de son polynôme minimal, on pose

$$M(\alpha) := |a_0| \prod_{j=1}^{d} \max(1, |\alpha_j|).$$

On admettra que si π_α est le polynôme minimal de α sur $\mathbb{Q}$, cette quantitée est aussi égale à

$$\exp\left(\frac{1}{2\pi}\int_0^{2\pi}\log|\pi_\alpha(e^{it})|dt\right),$$

ce qui découle de la formule de Jensen que l'on peut trouver dans **[11]** aux pages 207 et suivantes.

Lemme 14. *Soit $P\in\mathbb{Z}[X,Y]$ et $(\alpha,\beta)\in\mathbb{Q}^2$ une racine de P. On suppose que $P(\alpha,Y)$ n'est pas un polynôme constant. Alors on a l'inégalité*

$$\log(M(\beta))\leq \deg(\alpha)\big(\log(L(P))+D_X h(\alpha)\big). \tag{5.8}$$

où D_X est le degré partiel de P en X.

Démonstration. On utilise les notations de la remarque qui précède. Soit

$$Q(Y):=a_0^{D_X}\prod_{j=1}^{d}P(\alpha_j,Y).$$

Puisque $P(\alpha,Y)\neq 0$, il en est de même des polynômes $P(\alpha_j,Y)$ et donc $Q(Y)$ est également non nul. Si l'on écrit $P=\sum_{i\leq D_X,j}c_{ij}X^iY^j$, on remarque que les coefficients de Q sont des sommes de termes de la forme $a_1^D c_{D_1 j}\alpha_1^{D_1}\cdots c_{D_d j}\alpha_d^{D_d}$ avec les $D_k\leq D_X$. D'après la remarque précédente, ce sont des entiers algébriques. Il est clair que les coefficients de Q vivent dans $\mathbb{Q}$, car ils sont symétriques en les α_i. Ainsi $Q\in\mathbb{Z}[Y]$ et Q annule β par hypothèse, c'est donc un multiple de son polynôme minimal. Puisque le mesure de Mahler est multiplicative et supérieure à 1 (c'est évident), on a

$$M(\beta)=M(\pi_\beta)\leq M(Q)=a_0^{D_X}\prod_{j=1}^{d}M(P(\alpha_j,Y)).$$

De la définition de la mesure de Mahler on a évidemment $M(R)\leq\sup_{|z|=1}|R(z)|\leq L(R)$, pour tout polynôme R. Ainsi on obtient la majoration

$$M(\beta)\leq a_0^{D_X}\prod_{j=1}^{d}L(P)\max(1,|\alpha_j|^{D_X})\leq L(P)^d M(\alpha)^{D_X}=(L(P)e^{h(\alpha)D_X})^d$$

ce qui est le résultat voulu. □

On peut trouver la démonstration du lemme qui suit dans **[6]**, chapitre 3, lemme 6. On définit d'abord la hauteur de Weil d'un nombre algébrique α par

$$h(\alpha):=\log(M(\alpha)/\deg(\alpha)).$$

On écrit également l'inégalité de Liouville pour un tel α :

$$\log|\alpha|\geq -\deg(\alpha)h(\alpha).$$

Lemme 15. *Soit $P \in \mathbb{Z}[X_1, ..., X_m]$ non nul dont le degré partiel en X_i est D_i. Si $\alpha_1, ..., \alpha_m$ sont des nombres algébriques, on a la majoration :*

$$h(P(\alpha_1, ..., \alpha_m)) \leq \log(L(P)) + \sum_{i=1}^{m} D_i h(\alpha_i).$$

Nous pouvons passer à la démonstration du théorème à proprement parler. On se donne des paramètres N, L_1, L_2, indépendants des constantes qui apparaîtront dans la suite et que l'on pourra choisir arbitrairement grands. On suppose qu'il existe un nombre algébrique $0 < |q| < 1$ tel que $J(q)$ soit également algébrique. Nous allons contruire une fonction auxiliaire et un nombre algébrique non nul et à l'aide de ces hypothèses nous établirons des estimations contradictoires. Précisons que nous désignons par $c_1, c_2, ...$ des constantes qui dépendent de la donnée q mais pas des paramètres N, L_1, L_2.

Preuve du théorème 9. Nous allons construire un polynôme de $\mathbb{Z}[X, Y]$ tel que $P(z, \tilde{J}(z))$ ait un zéro d'ordre N à l'origine et P ait un degré partiel en X (resp. Y) majoré par L_1 (resp L_2). Si l'on écrit $P = \sum_\lambda a_\lambda X_1^{\lambda_1} X_2^{\lambda_2}$, alors

$$\begin{aligned} P(z, \tilde{J}(z)) &= \sum_\lambda a_\lambda z^{\lambda_1} \tilde{J}(z)^{\lambda_2} \\ &= \sum_\lambda a_\lambda z^{\lambda_1} \left(\sum_{n \geq 0} c_{\lambda_2}(n) z^n \right)^{\lambda_2} = \sum_{n \geq 0} \left(\underbrace{\sum_{\lambda, \lambda_1 \leq n} a_\lambda c_{\lambda_2}(n - \lambda_1)}_{b_n} \right) z^n. \end{aligned}$$

On cherche donc à résoudre le système d'équations $b_n = 0, n < N$ et d'inconnues les a_λ. Pour être assuré de l'existence d'une solution il faut plus d'inconnues que d'équations donc on suppose d'ores et déjà $N \leq 2L_1L_2$. Les nombres $c_k(n)$ sont entiers relatifs, donc on peut directement appliquer le lemme de Siegel et trouver une solution $a_\lambda \in \mathbb{Z}$ qui satisfait la majoration :

$$|a_\lambda| \leq (Ne^{c\sqrt{L_2 N}})^{\frac{L_1 L_2}{N - L_1 L_2}} \leq e^{c_1 \sqrt{L_2 N}},$$

et l'on a $b_n \in \mathbb{Z}$ pour tout $n \geq 0$.

Cette construction est terminée, on note $F(z) = P(z, \tilde{J}(z)) = \sum_n b_n z^n \in \mathbb{Z}[[z]]$ la fonction auxiliaire obtenue. Cette fonction n'est pas nulle car $\tilde{J}(z)$ est transcendante sur $\mathbb{Q}(z)$, on note $M \geq N$ son ordre en 0. On remarque que la formule donnant b_n et la construction précédente implique que $|b_n| < e^{c_2 \sqrt{L_2 n}}$. Ainsi pour $k \geq 0$, on a

$$|b_{M+k}| \leq e^{c_2 \sqrt{L_2 (M+k)}} \leq e^{c_2 (k + \sqrt{L_2 M})},$$

dès que $L_2 k \leq 2k\sqrt{L_2 M}$ i.e dès que $L_2 \leq N$. Il s'ensuit finalement une majoration de $F(z)$ pour $|z| \leq 1/2e^{-c_2}$:

$$|F(z)| \leq \sum_{k\geq 0} |b_{M+k}||z|^{M+k} \leq |z|^M \sum_{k\geq 0} \left(\frac{1}{2}\right)^k e^{c_2\sqrt{L_2 M}} \leq 2|z|^M e^{c_2\sqrt{L_2 M}}.$$

Au vu de la relation algébrique existant entre J et les J^n, nous pouvons supposer, quitte à remplacer q par une puissance n-ième adéquate, que $|q| \leq 1/2e^{-c_2}$. La fonction F étant non nulle, soit S le plus petit entier ≥ 1 qui vérifie $F(q^S) \neq 0$ (on a donc $F(q), ..., F(q^{S-1}) = 0$). L'existence de zéros de F nous permet de considérer la fonction holomorphe sur le disque unité

$$G(z) := z^{-M} F(z) \prod_{1\leq s\leq S-1} \frac{|q|^2 - z\overline{q}^s}{|q|(z - q^s)}.$$

D'une part, nous avons

$$|G(0)| = \left| b_M \prod_{1\leq s\leq S-1} q^{2-1-s} \right| \geq \left| q^{-\frac{(S-1)(S-2)}{2}} \right|, \tag{5.9}$$

car b_M est un entier non nul. D'autre part, la majoration de F sur le disque de rayon $1/2e^{-c_2}$ donne

$$\sup_{|z|=|q|} |G(z)| \leq |q|^{-M} \sup_{|z|=|q|} |F(z)| \leq 2e^{c_2\sqrt{L_2 M}}. \tag{5.10}$$

Si l'on combine les inégalités (5.9) et (5.10) avec le principe du maximum $G(0) \leq \sup_{|z|\leq|q|}|G(z)|$, on obtient :

$$\frac{(S-1)(S-2)}{2} \log\left(\frac{1}{|q|}\right) \leq \log(2) + c_2\sqrt{L_2 M},$$

ce qui donne

$$S^2 \leq \max(3, c_3\sqrt{L_2 M}). \tag{5.11}$$

Il nous reste à minorer le nombre algébrique non nul $F(q^S) = P(q^S, \tilde{J}(q^S))$. La relation $\Phi_S(J(q), J(q^S))$ montre que le degré de $J(q^S)$ est de degré au plus $\psi(S)$ sur $\mathbb{Q}(J(q))$. Puisque $J(q)$ est supposé algébrique, et si l'on note c_4 son degré, le degré de $J(q^S)$ sur Q est au plus $c_4\psi(S)$. D'après le lemme (14) appliqué à $P = \Phi_S$ et $(\alpha, \beta) = (J(q), J(q^S))$, dont les hypothèses sont évidemment vérifiées, on a

$$\log\Big(M(J(q^S))\Big) \leq c_4\Big(\log(L(\Phi_S)) + \psi(S)h(J(q))\Big).$$

Puisque le nombre $F(q^S)$ est dans $\mathbb{Q}(q, J(q^s))$, on a

$$\deg(F(q^s)) \leq \deg(q)\deg(J(q^S)) \leq c_5\psi(S).$$

D'après le lemme (15), la hauteur de $F(q^S)$ vérifie

$$h(F(q^S)) \leq \log(L(P)) + \deg_X(P)h(q) + \deg_Y(P)h(q^S J(q^S)).$$

La construction du polynôme P plus haut donne $\log(L(P)) \leq \log(L_1L_2) + c_1\sqrt{L_2N}$. La multiplicativité de la hauteur et l'inégalité de Liouville donnent alors

$$\begin{aligned}\log|F(q^S)| &\geq -\deg(F(q^S))h(F(q^S))\\ &\geq -c_5c_4\psi(S)\Big(\log(L_1L_2) + c_1\sqrt{L_2N} + L_1h(q^S) + L_2\big(h(q^S J(q^S))\big)\Big)\\ &\geq -c_6\psi(S)\Big(\sqrt{L_2N} + S(L_1+L_2) + L_2\big(\psi(S) + \log(L(\Phi_S))\big)\Big).\end{aligned}$$

D'après le lemme 12, on obtient

$$\log|F(q^S)| \geq -c_7S^{1+\epsilon}\Big(\sqrt{L_2N} + S(L_1+L_2)\Big). \tag{5.12}$$

D'autre part on a la majoration de $\log|F(q^S)|$ donné par (5.3) :

$$\log|F(q^S)| \leq \log(2) - MS|q| + c_2\sqrt{L_2M} \tag{5.13}$$

ce qui donne l'encadrement :

$$\begin{aligned}\exp(-c_7S^{1+\epsilon}(\sqrt{L_2N} + S(L_1+L_2)))&\\ \leq |F(q^S)| \leq \exp(\log(2) - MS|q| + c_2\sqrt{L_2M}).&\end{aligned}$$

On rappelle que l'on peut choisir L_1, L_2 et N à loisir dès qu'ils vérifient $L_1L_2 \geq 2N$ et $L_2 \leq N$ et on souhaite obtenir une contradiction à partir des inégalités (5.12) et (5.13), c'est à dire trouver L_1, L_2 et N tels que

$$\begin{aligned}MS > |q|^{-1}\Big(\log(2) + c_2\sqrt{L_2M} + c_7S^{1+\epsilon}(\sqrt{L_2N} + S(L_1+L_2))\Big)&\\ \geq c_8S^{1+\epsilon}\big(\sqrt{L_2N} + S(L_1+L_2)\big).&\end{aligned} \tag{5.14}$$

On pose alors $L_1 = L_2$ la partie entière de $2\sqrt{N}$. Puisque $M \geq N$, l'inégalité du dessus se traduit par $\sqrt{M} > c_9S^{\epsilon}(N^{1/4} + S)$ et d'après l'inégalité (5.11), $S^2 \leq c_3\sqrt{L_2M}$ donc l'inégalité (5.14) est satisfaite pour N suffisamment grand quelque soit le choix de $\epsilon < 1$ ce qui termine la preuve. □

Nous concluons ce rapport en mentionnant le théorème de Mahler sur la transcendance des nombres de la forme $\sum_{k\geq 0} \alpha^{d^k}$ pour α algébrique et $d \geq 2$. La preuve de ce théorème en particulier est remarquable en cela qu'elle n'utilise pas le lemme de Siegel, mais simplement l'existence d'une solution pour les systèmes linéaires. On peut se reporter à [**10**] pour plus de détails.

Bibliographie

[1] Serge Lang (1965) *Algebra*, Graduate Texts in Mathematics, Springer New-York, edition 3, Appendice, 1965

[2] Kurt Mahler, *On the coefficients of transformation polynomial for the modular function*, Bull, Austral, 10, pages 197-218, 1974

[3] Theodor Schneider, *Introduction aux nombres transcendants*, Gauthier-Villars, 149 pages, 1959

[4] Carl Ludwig Siegel, *Transcendental Numbers*, Marson Morse et Emil Artin, Princeton University Press, 102 pages, 1949

[5] Georges Valiron, *Fonctions entières d'ordre fini et fonctions méromorphes*, ETH-Bibliothek, l'enseignement mathématique tome 4, 1958

[6] Michel Waldschmidt, *Fonctions modulaires et transcendance* : Leçons de mathématiques d'aujourd'hui, 1996

[7] Michel Waldschmidt, *Linear independance of logarithm of algebraic numbers*, Institut des sciences mathématiques, Madras, 168 pages, 1992

[8] Jean-Pierre-Serre, *Cours d'arithmétique*, Presse universitaire de France, collection sup, le Mathématicien, 1970

[9] Paul Cohen, *On the coefficients of the transformation polynomials for the elliptic modular function*, Cambridge, 1984, 95, pages 389-402

[10] Kumiko Nishioka, *Mahler Functions and Transcendence*, Springer, 2009, 200 pages

[11] L. V. Ahlfors, *Complex Analysis*, McGraw-Hill, International series in pure and applied mathematics, 1979, page 207

[12] Peter Duren, *The Legendre relation for elliptic integrals*, Paul Halmos, Celebrating 50 years of mathematics, New York, Springer-Verlag, 1991, pages 305-315

[13] Barré-Sirieix, Katia ; Diaz, Guy ; Gramain, François ; Philibert, Georges *Une preuve de la conjecture de Mahler-Manin* , Invent. Math. 124, No. 1-3, 1-9 (1996).