Séminaire Dubreil.
Algèbre et théorie
des nombres

G. Poitou

Sur la démonstration de K. F. Roth du théorème de Thue-Siegel

Séminaire Dubreil. Algèbre et théorie des nombres, tome 9 (1955-1956), exp. n° 20, p. 1-7
<http://www.numdam.org/item?id=SD_1955-1956__9__A14_0>

© Séminaire Dubreil. Algèbre et théorie des nombres
(Secrétariat mathématique, Paris), 1955-1956, tous droits réservés.

L'accès aux archives de la collection « Séminaire Dubreil. Algèbre et théorie des nombres » implique l'accord avec les conditions générales d'utilisation (http://www.numdam.org/conditions). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

Numdam
Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques
http://www.numdam.org/

Faculté des Sciences de Paris
-:-:-:-
Séminaire P. DUBREIL et C. PISOT
(ALGÈBRE et THÉORIE DES NOMBRES)
Année 1955/56
-:-:-:-

16 avril 1956

Exposé N° 20

SUR LA DÉMONSTRATION DE K.F. ROTH DU THÉORÈME DE THUE - SIEGEL,

par G. POITOU

-:-:-:-

Soit α un nombre algébrique de degré n, c'est-à-dire la racine d'une équation irréductible,

$$x^n + a_1 x^{n-1} + \dots + a_n = 0 \qquad (1)$$

à coefficients $a_1, \dots, a_n$ rationnels ; supposons qu'il existe une infinité de fractions p/q telles que :

$$\left|\alpha - \frac{p}{q}\right| < \frac{1}{q^k} \qquad (2)$$

où k est une constante fixe.

Il se trouve que cette constante k ne peut être très grande, ce qui signifie que les nombres algébriques s'approchent assez mal par des rationnels. Le premier résultat de cette sorte est dû à Liouville, qui a démontré que $k \leqslant n$. (Sur des classes très étendues de quantités dont la valeur n'est ni algébrique, ni même réductible à des irrationnelles algébriques, C.R. Acad. Sc. Paris, 18 (1844), p.883-885 et p.910-911 ; J. Math. pures et appl., 16 (1851), p.133-142). Ce résultat est intéressant historiquement, car c'est la première démonstration de l'existence des nombres transcendants. Par exemple $\sum \frac{1}{2^{n!}}$ s'approche par des rationnels mieux que ne le permet le résultat de Liouville aux nombres algébriques, donc il est transcendant.

On sait par ailleurs que l'équation (2) a, pour $k = 2$, une infinité de solutions p/q, quel que soit le nombre α réel irrationnel (non nécessairement algébrique). L'inégalité $k \leqslant n$ ne peut donc être améliorée lorsque $n = 2$, c'est-à-dire pour les nombres quadratiques ; pour ceux-ci, elle résulte d'ailleurs de la périodicité du développement en fraction continue, démontrée par Lagrange en 1770 (Additions au mémoire sur la résolution des équations numériques, Mém. Ac. roy. Sc. et Belles-Lettres de Berlin, t. 24, Oeuvres, t.2, p.581-652)

Par contre, pour $n > 2$, on a démontré successivement des majorations de k de plus en plus fortes :

$k \leqslant \frac{n}{2} + 1$ par Axel Thue en 1908 ; cf. notamment :"Über Annäherungswerte algebraischer Zablen", Journal de Crelle, 135 (1909), p. 284-305 ; les autres mémoires sont peu accessibles.

$k \leqslant s + \frac{n}{s+1}$ pour $s = 1 , 2 , \ldots , n-1$ (au mieux, de l'ordre de $2\sqrt{n}$), par C. Siegel dans sa thèse en 1921 : "Approximation algebraischer Zahlen", Math. Zeitschrift, 10 (1921), p. 173-213 ; "Über Näherungswerte algebraischer Zahlen", Math. Annalen, 84(1921) , p. 80-99.

$k \leqslant \sqrt{2n}$ par F.J. Dyson en 1947 : "The approximation to algebraic numbers by rationals", Acta Math., 79 (1947), p.225-240.

La démonstration de Dyson a été simplifiée par Mahler en 1949 : "On Dyson improvement of the Thue - Siegel theorem", Proc. K. Akad. Wet. Amsterdam, 52 (1949), p. 1175-1184.

Le résultat a d'ailleurs été retrouvé par une autre voie par Schneider : "Über eine Dysonsche Verschärfung des Siegel-Thueschen Satzes", Arch. Math., 1 (1948/49), p. 288-295, et aussi, apparemment par Gelfand (cf. son livre : "Nombres transcendants et algébriques", publié à Moscou en 1952, chapître 1).

Siegel avait conjecturé l'inégalité $k \leqslant 2$; elle a été établie l'an dernier par K.F. Roth, dans le mémoire dont je vais parler : " Rational approximations to algebraic numbers", Mathematika, 2, 1, p. 1-20.

On voit par là que le titre de cet exposé devrait être plutôt :

"Sur le théorème de Liouville - Thue - Siegel - Dyson - Roth".

Le résultat de Roth est évidemment définitif sous cette forme. Sous cette forme, car on peut encore s'interroger par exemple sur les développements en fraction continue des nombres algébriques (cf. H. Davenport - K.F. Roth : Rational approximation to algebraic numbers, Mathematika, 2, 2, p. 160-167) sans parler des généralisations à l'approximation des nombres algébriques par les nombres d'un corps algébrique, généralisations inaugurées par Siegel.

Avant de parler des démonstrations, je veux souligner l'importance de ces résultats pour l'étude de certaines équations diophantiennes. Ainsi le résultat de Thue a pour conséquence le théorème suivant, qui explique l'intérêt qui s'est attaché ensuite à ces questions :

"Si $f(x , y)$ est un polynome homogène irréductible à coefficients entiers de degré $\geqslant 3$, l'équation $f(x , y) = a$ a au plus un nombre fini de solutions en nombres entiers x , y ".

Comme $f(x , y) = a_0(x - \alpha_1 y)(x - \alpha_2 y) \ldots (x - \alpha_n y)$, s'il y avait une infinité de solutions, il faudrait qu'une des formes linéaires $x - \alpha_i y$ prenne des valeurs arbitrairement petites ; supposons que ce soit $x - \alpha_1 y$; les autres formes sont alors de l'ordre de y , de sorte que $(x - \alpha_1 y)y^{n-1}$ resterait borné, et $\left| \frac{x}{y} - \alpha_1 \right| < \frac{1}{y^k}$ aurait une infinité de solutions pour tout $k < n$, ce qui contredit le théorème de Thue. Le même raisonnement prouve qu'il n'y a qu'un nombre fini de solutions pour une équation $f(x , y) = g(x , y)$ si g est un polynome quelconque de degré assez petit ; en fait, de degré $n - 3$ au plus, d'après le théorème de Roth. On peut naturellement chercher une estimation de ce nombre fini de solutions (cf. l'article cité de Davenport - Roth).

Remarquons que, pour la démonstration d'une majoration de k , on peut se limiter à considérer les nombres α algébriques entiers ; car si α est algébrique, il y a un entier M tel que $M\alpha$ est entier algébrique et

$$\left| M\alpha - \frac{Mp}{q} \right| < \frac{M}{q^k} < \frac{1}{q^{k'}}, \quad \text{pour tout } k' < k .$$

Examinons la démonstration de Liouville. Posons $f(x) = x^n + a_1 x^{n-1} + \ldots + an$; soit B une borne supérieure de la valeur absolue de la dérivée de f , de sorte que $f(\beta) = f(\beta) - f(\alpha) = (\beta - \alpha) f'(\gamma)$ d'où $f(\beta) < B|\beta - \alpha|$. Si, en particulier, $\beta = \frac{p}{q}$, on a par contre $\left| f(\frac{p}{q}) \right| \geqslant \frac{1}{q^n}$ ($\frac{p}{q}$ n'est pas racine de f) d'où $\frac{1}{q^n} \leqslant B\left| \alpha - \frac{p}{q} \right| \leqslant \frac{B}{q^k}$ et l'on trouve bien $k \leqslant n$.

Les autres démonstrations sont incomparablement plus difficiles, mais on y retrouve des idées voisines. Il y intervient des polynomes à coefficients entiers, qui, d'une part, ont des coefficients point trop grands, de sorte qu'ils ne sont pas trop grands pour des nombres voisins de α , et, d'autre part, ne sont pas de degré trop élevé, de sorte que, s'ils ne sont pas nuls pour des valeurs rationnelles des variables, ils ne sont pas trop petits ; et la comparaison de ces estimations supérieure et inférieure sont la clef de la démonstration. Thue introduit des polynomes de la forme $\alpha Q(x) - P(x)$, dont α est une racine d'ordre assez élevé, les autres auteurs cités utilisent en

général des polynomes à deux variables $R(x , y)$, nuls ainsi que certaines de leurs dérivées pour $x = y = \alpha$, mais non nuls pour $x = \frac{p_1}{q_1}$, $y = \frac{p_2}{q_2}$, ces fractions étant choisies de façon à vérifier(2) , avec des ordres de grandeur savamment dosés. Roth utilise, au contraire, des polynomes à un nombre suffisamment grand de variables, ce qui ne va pas sans quelques complications. Devant les difficultés ainsi soulevées, Dyson notamment avait échoué ; il semble que Roth ait obtenu là le succès décisif.

Voici comment on peut énoncer la partie finale de la démonstration :

Soit δ un nombre positif très petit, et $\frac{p_1}{q_1} , \frac{p_2}{q_2} , \ldots , \frac{p_m}{q_m}$ des solutions de (2) telles que :

$$\log q_1 > \frac{Cm}{\delta^2} , \quad \frac{\log q_{i+1}}{\log q_i} > \frac{2}{\delta} .$$

C étant une certaine constante ne dépendant que de α .

Soient $r_1 , r_2 , \ldots , r_m$ des entiers tels que

$$r_1 > \frac{10 \log q_m}{\delta \log q_1}$$ et que les quantités $q_1^{r_1} , q_2^{r_2} , \ldots , q_m^{r_m}$ soient à peu près égales.

Supposons qu'il existe un polynome $Q(x_1 , x_2 , \ldots , x_m)$ possédant les propriétés suivantes :

(A) $\frac{j_1}{r_1} + \frac{j_2}{r_2} + \ldots + \frac{j_m}{r_m} \geqslant \Theta(n,m) + 0(\delta)$ pour les indices $j_1 , j_2 , \ldots , j_m$ tels que $D_1^{j_1} D_2^{j_2} \ldots D_m^{j_m} Q(\alpha, \ldots , \alpha) \neq 0$

(B) les coefficients de Q sont majorés par $q_1^{\delta r_1 (1+0(\delta))}$

(C) Le degré de Q en x_i est au plus r_i et $Q(\frac{p_1}{q_1} , \ldots , \frac{p_m}{q_m}) \neq 0$

Dans (A) $\Theta(n,m)$ désigne une fonction donnée, et D_i la dérivation par rapport à la i-ème variable.

Dans ces conditions, on a l'inégalité $k \leqslant \frac{m}{\Theta}$.

L'énoncé précédent, et surtout les démonstrations, sont grandement

facilités par la définition suivante : on appelle indice du polynome $Q(x_1 , \dots , x_m)$ en $(\alpha_1 , \dots , \alpha_m)$ par rapport à $(r_1 , \dots , r_m)$ le minimum de $\frac{j_1}{r_1} + \dots + \frac{j_m}{r_m}$, pour les $(j_1 , \dots , j_m)$ tels que $D_1^{j_1} , \dots , D_m^{j_m} Q(\alpha_1 , \dots , \alpha_m) \neq 0$. Alors (A) ci-dessus s'énonce : l'indice de Q en $(\alpha, \dots , \alpha)$ par rapport à $(r_1 , \dots , r_m)$ est au moins $\Theta + 0(\delta)$.

Démonstration : D'après (C) , on a :

$$Q(\frac{p_1}{q_1} , \dots , \frac{p_m}{q_m}) \geqslant q_1^{-r_1} \dots q_m^{-r_m} \quad \text{à peu près égal à} \quad q_1^{-mr_1}$$

D'autre part, d'après (A) ,

si $Q(\alpha + \xi_1 , \dots , \alpha + \xi_n) = \sum \dots \sum a_{i_1 \dots i_m} \xi_1^{i_1} \dots \xi_m^{i_m}$,

on a $a_{i_1 \dots i_m} = 0$ pour les i tels que $\frac{i_1}{r_1} + \dots + \frac{i_m}{r_m} < \Theta + 0(\delta)$;

en particulier $a_{0,0,\dots,0} = 0$. Donc

$$\left| Q(\frac{p_1}{q_1} , \dots , \frac{p_m}{q_m}) \right| \leqslant \left| \sum \dots \sum a_{i_1 \dots i_m} (\frac{p_1}{q_1} - \alpha)^{i_1} \dots (\frac{p_m}{q_m} - \alpha)^{i_m} \right| \leqslant$$

d'après (2)

$$\sum \dots \sum \left| a_{i_1 \dots i_m} \right| (q^{-i_1} \dots q^{-i_m})^k <$$

$$< (r_1 + 1) \dots (r_m + 1) q_1^{+\delta r_1 (1+0(\delta))} q_1^{-kr_1(\Theta + 0(\delta))} < q_1^{-kr_1(\Theta + 0(\delta))}$$

car $(r_1 + 1) \dots (r_m + 1) < e^{r_1} \dots e^{r_m} < e^{mr_1} < q_1^{\frac{r_1 \delta^2}{C}}$

En comparant les inégalités précédentes, on trouve :

$$kr_1(\Theta + 0(\delta)) < mr_1 \quad \text{d'où} \quad k \leqslant \frac{m}{\Theta} .$$

Pour achever la démonstration, il suffit donc de prouver l'existence d'un polynome Q avec les propriétés (A) , (B) , (C) pour un Θ .

convenable, par exemple $\theta = \frac{m}{2} - 3n\sqrt{m}$. Naturellement c'est la partie difficile de la démonstration.

Voici une esquisse de cette partie :
Considérons l'ensemble des polynomes $W(x_1 , \ldots , x_m)$ de la forme

$$\sum_{s_1=0}^{r_1} \ldots \sum_{s_m=0}^{r_m} C(s_1 \ldots s_m) x_1^{s_1} \ldots x_m^{s_m}$$ où les C sont tous les entiers compris entre 0 et B . Il y en a $(B+1)^r$ avec $r = (r_1 + 1) \ldots (r_m + 1)$

Pour l'un d'eux, considérons les dérivées $W_{j_1 \ldots j_m}(x_1 , \ldots , x_m) =$

$$= \frac{1}{j_1! \ldots j_m!} D_1^{j_1} \ldots D_m^{j_m} W \quad \text{pour} \quad 0 \leqslant j_i \leqslant r_i , \quad \frac{j_1}{r_1} + \ldots + \frac{j_m}{r_m} \leqslant \theta ;$$

leur nombre D est majoré par $\frac{r}{3n}$, ce qu'on établit par récurrence.

Soit alors $T_{j_1 \ldots j_m}(W ; x)$ le reste de la division de $W_{j_1 \ldots j_m}(x , x , \ldots x)$ par $f(x)$. A chaque W est ainsi associé un ensemble de D restes ; comme les coefficients de ces restes sont majorés par $B^{1+3\delta}$, il y a au plus $(1 + 2B^{1+3\delta})^{nD}$ ensembles de tels restes, nombre qui est inférieur à $(B+1)^2$ pour δ assez petit ; donc il existe deux polynomes W , soient W' et W'' , ayant exactement les mêmes restes ; leur différence W^* est telle que les dérivées $W^*_{j_1 \ldots j_m}(x , x , \ldots , x)$ sont divisibles par $f(x)$, donc $W^*_{j_1 \ldots j_m}(\alpha , \ldots , \alpha) = 0$
Donc l'indice de W^* en $(\alpha , \ldots , \alpha)$ par rapport à $(r_1 , \ldots , r_m)$ est au moins θ . De plus, les coefficients de W^* sont majorés par B , de sorte que W satisfait à (A) et (B) avec $B = q_1^{\delta r_1}$; malheureusement, il n'est pas assuré qu'il vérifie (C) et c'est même la principale difficulté .

Du fait que W^* est un polynome à coefficients entiers, non identiquement nul, à coefficients bornés par B et de degré au plus r_i en x_i , on déduit que son indice en $(\frac{p_1}{q_1} , \ldots , \frac{p_m}{q_m})$ par rapport à $(r_1 , \ldots , r_m)$

est au plus $10^m \delta (\frac{1}{2})^m$; cette démonstration, par récurrence sur m , peut sans doute être considérée comme la véritable nouveauté de cet article.

Il en résulte que W^* a un dérivé $W^*_{k_1 \ldots k_m} = Q$ tel que

$$Q(\frac{p_1}{q_1} , \ldots , \frac{p_m}{q_m}) \neq 0 \quad \text{avec} \quad \frac{k_1}{r_1} + \ldots + \frac{k_m}{r_m} < 10^m \delta (\tfrac{1}{2})^m .$$

Par suite, le polynome Q , qui a la propriété (C) , a encore les propriétés (A) et (B) .

En effet, son indice en $(\alpha , \ldots , \alpha)$ par rapport à $(r_1 , \ldots , r_m)$ est au moins

$$\Theta - (\frac{k_1}{r_1} + \ldots + \frac{k_m}{r_m}) \geqslant \Theta - 10^m \delta (\tfrac{1}{2})^m = \Theta + 0(\delta)$$

-:-:-:-:-:-:-