PUBLICATIONS

MATHÉMATIQUES

D'ORSAY

Groupe de travail en théorie analytique et élémentaire des nombres

1986 - 1987

88 - 01

Université de PARIS-SUD

Département de Mathématiques
Bâtiment 425
91405 ORSAY France

PUBLICATIONS

MATHÉMATIQUES

D'ORSAY

Groupe de travail en théorie analytique et élémentaire des nombres

1986 - 1987

88 - 01

Université de PARIS-SUD

Département de Mathématiques
Bâtiment 425
91405 ORSAY France

A LA MÉMOIRE DE Paul BELGODÈRE

Le groupe de travail en théorie analytique et élémentaire des nombres à Paris est de création récente (1983). Les conférences ont lieu à l'Institut Henri Poincaré. Grâce à notre regretté Paul Belgodère et à Melle Denise Lardeux, les actes du groupe parurent dans les "Annales de l'Institut Henri Poincaré" de 1983 à 1986. Ils paraissent aujourd'hui dans les "Publications Mathématiques d'Orsay". Le Secrétariat est assuré maintenant par Monique Le Bronnec.

H. DABOUSSI

TABLE DES MATIERES

J.–P. ALLOUCHE.

Tours de Hanoï et automates finis 1

M. BALAZARD

Sur certaines fonctions additives 4

J.–P. BOREL

Suites et mesures ayant des propriétés d'auto–similarité 15

G. GREKOS

Convexité, moyennes, densité, lacunes 39

D.R. HEATH–BROWN

Van der Corput Bounds for the Dedekind Zeta–Function 44

A. IVIC

Les zéros de la fonction zéta de Riemann sur la droite critique 47

M. LANGEVIN

Calculs explicites de constantes de Lehmer 52

M. NAIMI

Les entiers sans facteurs carré $\leq x$ *dont leurs facteurs premiers* $\leq Y$ 69

M. PATHIAUX–DELEFOSSE

Résultat de Cantor et Straus sur la conjecture de Lehmer 77

M. PATHIAUX–DELEFOSSE

Propriété de grands nombres de Pisot 84

A.J. VAN DER POORTEN

Remarks on the continued fractions of algebraic numbers 89

B. ROUSSELET

Inégalités de type Brun–Titchmarsh en moyenne 91

O. SALON
Substitutions, automates et series formelles relatifs aux suites à multi–indices 124
C.J. SMYTH
Generalised Salem numbers 130
C.J. SMYTH
Congruences for Symmetric Functions 136
B. VALLEE
Un problème central en Géométrie Algorithmique des nombres : La réduction des réseaux. Autour de l'algorithme de Lenstra Lenstra Lovasz 144

TOURS DE HANOI ET AUTOMATES FINIS

Jean-Paul Allouche

Cet exposé résume un travail effectué avec F. Dress, sur le problème classique des tours de Hanoi, que nous rappelons brièvement : on dispose de N disques de diamètres respectifs 1,2,...,N, percés en leurs centres, et de trois piquets disposés verticalement. Au départ les disques sont empilés sur le premier piquet, dans l'ordre (le disque N au-dessous), et il faut les transporter sur un autre piquet, les mouvements permis étant de prendre un disque au sommet d'un piquet et de le déposer sur un autre piquet, à condition que ce soit sur un disque plus gros.

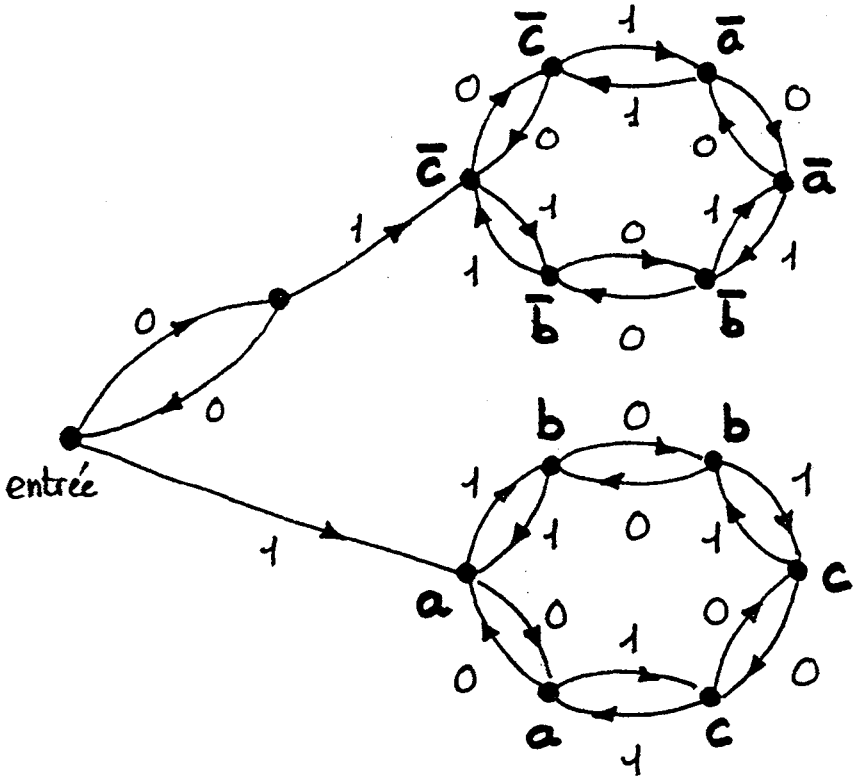
L'algorithme classique, archétype des algorithmes récursifs, permet de transporter N disques du premier au deuxième piquet si N est impair, du premier au troisième si N est pair : supposons par exemple N pair, alors on commence par déplacer les N-1 plus petits disques du premier au second piquet, puis le Nième du premier au troisième piquet, et il ne reste plus qu'à transférer les N-1 plus petits disques du second au troisième piquet.

Le lecteur courageux peut essayer d'appliquer cet algorithme pour N=5, mais nous lui déconseillons formellement cet exercice terrible. S'il essaie malgré tout, il retrouvera vraisemblablement un des résultats de "périodicité" connus dans le folklore et plusieurs fois retrouvés, qui lui simplifiera la tâche.

Nous montrons dans [1] d'une part qu'on peut supposer la pile initiale infinie (et que les (2^N-1) premiers coups permettent de déplacer les N plus petits disques), ce qui était déjà connu, mais aussi d'autre part que la suite infinie de mouvements ainsi construite est engendrée par un automate fini à 14 états que nous reproduisons ci-dessous. Cet automate permet de retrouver les résultats de "périodicité" et de montrer quelles propriétés arithmétiques de l'entier n déterminent le nième coup. De plus le calcul de ce nième coup peut alors se faire uniquement à partir de l'écriture binaire de n et ne nécessite de connaître ni les coups précédents, ni l'état des piquets à cet instant.

Le détail et une bibliographie se trouvent dans [1], on pourra aussi consulter [2] pour d'autres types de résultats inspirés par la démonstration de l'automaticité de la suite infinie des mouvements évoquée ci-dessus.

Un automate à lecture inverse minimal pour jouer aux tours de Hanoi :



On a noté a le déplacement du disque du sommet du premier piquet au second piquet, de même pour b, c, $\bar{a}$, $\bar{b}$, $\bar{c}$:

Pour calculer le 23ème coup, on écrit 23 en base 2 : 10111, on entre ce mot lu de droite à gauche dans l'automate, ce qui donne c, le 23ème coup consiste donc (dans l'algorithme classique) à prendre le disque situé au sommet du 3ème piquet et de le poser sur le premier piquet.

BIBLIOGRAPHIE

[1] J.-P. Allouche et F. Dress.- Tours de Hanoi et automates finis, 1987, preprint.

[2] J.-P. Allouche, J. Betrema et J.O. Shallit.- Sur des points fixes de morphismes d'un monoïde libre, 1987, preprint.

Jean-Paul ALLOUCHE
UA 226
Mathématiques et Informatique
351, Cours de la libération
33405 TALENCE CEDEX

SUR CERTAINES FONCTIONS ADDITIVES

Michel Balazard

I - Introduction

1. Cet exposé présente les résultats obtenus au cours de la préparation d'une thèse de doctorat à l'Université de Limoges. Il s'agit de travaux de théorie probabiliste des nombres portant principalement sur le problème de la répartition des valeurs de fonctions additives comptant, de diverses façons, le nombre de facteurs premiers d'un entier. Le lecteur intéressé trouvera la totalité des démonstrations dans [3], et des aperçus sur différentes étapes de ce travail dans deux exposés précédents : [1] et [2].

II - Etude locale de la fonction Ω_E.

2. Soit E une partie de l'ensemble P des nombres premiers, $p_1<p_2<\dots$ la suite ordonnée des éléments de E. Pour tout entier $n\geq 1$, $\Omega_E(n)$ est le nombre de diviseurs de n qui sont des puissances d'éléments de E. La fonction Ω_E est complètement additive. Le théorème suivant (cf. [8] et [13]) montre que, pour $1\leq n\leq x$, elle suit approximativement une loi de Poisson :

Théorème A (G. Halász, 1972; A. Sárközy, 1977). Pour $x\geq 3$, on pose :

$$E(x)=\sum_{p\leq x,\, p\in E}\frac{1}{p}. \tag{1}$$

Soit $\delta\in]0,1[$. On a :

$$P_x(\Omega_E(n)=k) \ll_\delta e^{-E(x)}\,\frac{E(x)^k}{k!} \tag{2}$$

pour tout E, tout $x\geq 3$ et tout entier naturel k tels que

$$k+1 \leq (2-\delta)\,E(x);$$

$$P_x(\Omega_E(n)=k) \gg_\delta e^{-E(x)}\,\frac{E(x)^{k-1}}{(k-1)!} \tag{3}$$

pour tout E, tout $x\geq 3$ et tout entier naturel k tels que $1\leq k\leq(2-\delta)E(x)$, $E(x)\geq c(\delta)$, où $c(\delta)$ est une constante positive ne dépendant que de δ, et P_x la probabilité uniforme sur $\{n/1\leq n\leq x\}$.

Observons que la valeur moyenne $E(x)$ est toujours inférieure ou égale à $\sum_{p\leq x} \frac{1}{p} = \log\log x + O(1)$, alors que $\Omega_E(n)$ prend, quand n décrit $[1,x]$, toutes les valeurs entières k, $0\leq k\leq \log x/\log p_1$. Il est donc intéressant de se demander ce que deviennent les estimations (2) et (3) lorsque k est supérieur ou égal à $(2-\delta)E(x)$. On dispose principalement du résultat suivant (cf. [11]) :

Théorème B (K.K. Norton 1981). Pour x réel ≥ 1 et k entier ≥ 0 on a :

$$P_x(\Omega_E(n)=k) \ll_{p_1} p_1^{-k} \exp((p_1-1)E(x))\sqrt{1+E(x)}. \tag{4}$$

Les relations (2), (3) et (4) montrent un changement de comportement pour la loi locale $P_x(\Omega_E(n)=k)$: pour les valeurs moyennes de k on a une loi de Poisson et pour les grandes valeurs de k une loi géométrique. Nous avons pu montrer que pour la totalité des valeurs de k, cette loi locale se comporte comme la convolution d'une loi de Poisson et d'une loi géométrique. Soit $S_k(X)= 1+X+\ldots+\frac{X^k}{k!}$ la k-ième somme partielle de la série exponentielle et :

$$E_1(x)= \sum_{\substack{p_1<p\leq x \\ p\in E}} \frac{1}{p}\,. \tag{5}$$

Nous avons le :

Théorème 1. Il existe une constante absolue $K>0$ telle que :

$$P_x(\Omega_E(n)=k)\leq e^{Kp_1\log\log(1+p_1)}\, p_1^{-k}\, e^{-E_1(xp_1^{-k})}\, S_k(p_1E_1(xp_1^{-k})), \tag{6}$$

pour tout E, tout $x\geq 1$ et tout entier naturel k;

$$P_x(\Omega_E(n)=k)\geq e^{-Kp_1\log p_1}\, p_1^{-k}\, e^{-E_1(xp_1^{-k})}\, S_{k-1}(p_1E_1(xp_1^{-k})), \tag{7}$$

pour tout E, tout $x\geq 1$ et tout entier naturel k tels que

$$1\leq k\leq \log x/\log p_1 \quad \text{et} \quad E_1(x\,p_1^{-k}) \geq e^{Kp_1\log\log(1+p_1)}$$

Nous pouvons ici faire quelques remarques :

1°) Lorsque dans (6) on majore le polynôme S_k par la série exponentielle, on obtient :

$$(8)\qquad P_x(\Omega_E(n)=k) \le e^{Kp_1 \mathrm{loglog}(1+p_1)}\, p_1^{-k} \exp((p_1-1)E_1(x\,p_1^{-k}))$$

ce qui améliore (4). On peut d'ailleurs supprimer ici le $\mathrm{loglog}(1+p_1)$, et c'est probablement aussi le cas dans (6).

2°) Quand $k\le(2-\delta)E(x)$, (6) et (7) sont moins précis que (2) et (3) respectivement, car dans ces dernières inégalités les constantes implicites ne dépendent que de δ et non de p_1.

3°) Il serait intéressant de donner une analyse aussi fine que possible du phénomène de décalage d'exposant que l'on observe entre les relations (2) et (3). Si E est un ensemble infini de densité $\alpha<1$ par rapport à l'ensemble de tous les nombres premiers, on a :

$$(9)\qquad P_x(\Omega_E(n)=k) \asymp e^{-E(x)}\,\frac{E(x)^k}{k!}, \qquad k \text{ fixé}, \qquad x \longrightarrow +\infty,$$

alors que

$$(10)\qquad P_x(\Omega_E(n)=k) \asymp e^{-E(x)}\,\frac{E(x)^{k-1}}{(k-1)!}, \qquad k \text{ fixé } \ge 1, \qquad x \longrightarrow +\infty,$$

si $E=P$ (cf. [4]).

3. A la fin du chapitre 21 de [6], P.D.T.A. Elliott mentionne un résultat non publié de G. Halâsz précisant (2) et (3) au moyen d'une formule asymptotique. Dans [3], nous donnons la démonstration complète de ce résultat, sous la forme suivante :

Théorème C (G. Halâsz, 1975). Soit $\delta\in]0,1[$. On a :

$$(10)\qquad P_x(\Omega_E(n)=k)=\frac{E(x)^k}{k!}\,e^{-k}\,x^{-1}\sum_{n\le x} r^{\Omega_E(n)}\,(1+0_{p_1,\delta}(E(x)^{-1/2}))$$

où $r=\dfrac{k}{E(x)}$, pour tout k, tout $x\ge 3$ et tout E vérifiant :

$$(11)\qquad \delta\,E(x) \le k \le (p_1-\delta)E(x).$$

Si E est un ensemble de densité $\alpha>0$ par rapport à l'ensemble de tous les nombres premiers et si $k=rE(x)$, où r est une constante, $r\in]0,2[$, (10) et le Satz 1 de l'article [16] de E. Wirsing donnent :

$$(12)\quad P_x(\Omega_E(n)=k) \sim \frac{e^{-\alpha\gamma(r-1)}}{\Gamma(1+\alpha(r-1))}\prod_{p\in E}\frac{(1-\frac{1}{p})e^{(1-r)/p}}{1-r/p}\,e^{-E(x)}\,\frac{E(x)^k}{k!}, \quad x\longrightarrow+\infty,$$

où γ désigne la constante d'Euler.

4. Dans le cas de E=P, on note $\Omega_E=\Omega$. On dispose de formules asymptotiques précisant les théorèmes A et 1 (cf. [14] et [9]) :

Théorème D (L.G. Sathe, A. Selberg , 1954). Soit $\delta\in]0,1[$. On a :

(13)
$$P_x(\Omega(n)=k)= \frac{1}{\Gamma(r+1)} \prod_{p\geq 2}(1-\frac{1}{p})^r(1-\frac{r}{p})^{-1}(\log x)^{-1}\frac{(\log\log x)^{k-1}}{(k-1)!}(1+O_\delta((\log\log x)^{-1}))$$

uniformément pour $x\geq 3$ et $0<r=\frac{k}{\log\log x}\leq 2-\delta$. Le produit infini porte sur les nombres premiers.

Théorème E (J.-L. Nicolas, 1984). Soit $\delta>0$. On a :

$$(14)\qquad P_x(\Omega(n)=k)= C\,2^{-k}\log\frac{x}{2^k}+O_\delta(2^{-k}(\log\frac{3x}{2^k})^{c(\delta)})$$

uniformément pour $x\geq 3$ et $(2+\delta)\log\log x\leq k\leq \log x/\log 2$, où

$$C=\frac{1}{4}\prod_{p\geq 3}(1+1/p(p-2)),\quad \text{et}\quad 0<c(\delta)<1$$

Comme précédemment, c'est la notion de produit de convolution qui permet de comprendre la transition entre la loi "de type Poisson" de (13) et la loi "de type géométrique" de (14) :

Théorème 2. On a uniformément pour x réel et k entier positif tels que $y=\frac{x}{2^k}\geq 3$:

$$(15)\quad P_x(\Omega(n)=k)= f(r)2^{-k}(\log y)^{-1}S_{k-1}(2\log\log y)(1+O((\log\log y)^{-1/3}))$$

où $r=\min(2,\frac{k-1}{\log\log y})$ et $f(z)=\frac{2^{1-z}}{\Gamma(z+1)}\prod_{p\geq 3}(1-\frac{1}{p})^z(1-\frac{z}{p})^{-1}$ (c'est une fonction holomorphe de z pour $|z|<3$).

Si l'on remplace dans (15) r par la quantité plus compliquée $2\frac{S_{k-2}(2\log\log y)}{S_{k-1}(2\log\log y)}$, on obtient un terme reste nettement meilleur (cf. [3]). Le théorème 2 contient et précise (13) et (14), eu égard aux estimations asymptotiques de la quantité $S_{k-1}(2\log\log y)$. On retrouve aussi un résultat non publié de H. Delange :

Théorème F. On a uniformément pour $x\geq 3$ et k entier ≥ 2 tels que $\frac{x}{2^k}\geq 3$:

(16) $$P_x(\Omega(n)=k)= C\, 2^{-k}\log \frac{x}{2^k}\,(G(t)+O(\frac{1+|t|}{\sqrt{\log\log x}}))$$

où $t= \frac{k-2\ \log\log x}{\sqrt{2\ \log\log x}}$ et $G(t)= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} e^{-\frac{u^2}{2}}\, du.$

5. La présence d'un produit de convolutioon dans les théorèmes 1 et 2 provient du fait suivant : dans l'écriture

$$\Omega_E= v_{p_1} + \Omega_{E_1},$$

où v_{p_1} est la valuation p_1-adique et $E_1= E\setminus\{p_1\}$, les deux variables aléatoires v_{p_1} et Ω_{E_1} sont "presque" indépendantes sur $\{n/1\leq n\leq x\}$. La première suivant une loi "presque" géométrique et la seconde une loi "presque" de Poisson, le résultat devient alors intuitif. Bien entendu, la même analyse permet d'isoler dans la loi de Ω_{E_1} une composante géométrique prépondérante de raison $\frac{1}{p_2}$, mais celle-ci est négligeable par rapport à la loi de v_{p_1}.

Pour terminer ce paragraphe, signalons que G.H. Hardy et S. Ramanujan avaient déjà en 1917 indiqué l'importance des polynômes S_k dans l'étude locale de Ω. Leur mémoire classique "The normal number of prime factors of a number n" contient en effet l'inégalité :

(17) $$P_x(\Omega(n)=k) \leq A_1(\frac{10}{9})^{-k}(\log x)^{-1}\ \delta_{k-1}(\frac{10}{9}\ \log\log x+A_2)$$

valable pour $x\geq 2$, $k\geq 1$, où A_1 et A_2 sont des constantes absolues positives (cf. [12]).

III - Distributions conjointes et grandes déviations

6. L'idée générale de ce paragraphe est que l'obtention d'une formule asymptotique précise pour la probabilité de grande déviation d'une fonction arithmétique donnée est parfois facilitée par l'étude préalable de la distribution conjointe de plusieurs fonctions arithmétiques. Nous illustrerons cette idée par la résolution de deux problèmes.

7. Problème 1 (A. Ivic) : $f(n)$ désignant la moyenne des exposants dans la décomposition en facteurs premiers de l'entier $n\geq 2$, et α étant un réel >1, estimer $P_x(f(n)\geq\alpha)$.

On peut montrer que cette probabilité est $o(1)$ si et seulement si $(\alpha-1)\ \log\log x$ tend vers l'infini. Si α est fixé et x tend vers

l'infini, il s'agit bien d'une grande déviation par rapport à la moyenne 1 de $f(n)$ (pour une démonstration simple de la relation $\sum_{2\le n\le x} f(n)= x(1+O(\frac{1}{\log\log x}))$ on peut consulter l'article [5]).

Nous avons obtenu le résultat suivant :

Théorème 3. Si $\alpha>1$ est fixé et x tend vers l'infini, on a

$$(18)\qquad P_x(f(n)\ge\alpha)= C(\alpha)(\log x)^{2^{1-\alpha}-1}(1+O_\alpha(R(x,\alpha)))$$

où $C(\alpha)= g(\alpha)\,\frac{2^{1-\alpha}}{2^{2^{1-\alpha}}\,\Gamma(2^{1-\alpha})}\prod(1-\frac{1}{\ell})^{2^{1-\alpha}}(1+\frac{2^{1-\alpha}}{\ell-2})$, le produit portant sur les nombres premiers impairs et $g(\alpha)$ désignant $\frac{1}{2q(1-2^{-1/q})}$ si $\alpha=\frac{p}{q}$ est rationnel écrit sous forme irréductible ou $\frac{1}{2\log 2}$ si α est irrationnel; dans le premier cas on a

$R(x,\alpha)= (q-1)(\log x)^{-2^{1-\alpha}}.2\sin^2\pi/q + (\log\log x)^{3/2}(\log x)^{3^{1-\alpha}-2^{1-\alpha}}$ et dans le second $R(x,\alpha)= D^*_\alpha(\sqrt{2^{1-\alpha}\log\log x})$ où $D^*_\alpha(t)= D_\alpha(t) + \frac{1}{t^2}\int_1^t u\,D_\alpha(u)du$,

$D_\alpha(t)= \sup_{0\le a<b\le 1}\left|\frac{1}{t}\sum_{\substack{1\le n\le t\\ a\le\{n\alpha\}<b}} 1-(b-a)\right|$ désignant la discrépance à l'ordre t de la suite $\{n\alpha\}$.

Signalons ici qu'un résultat similaire, et dans certains cas meilleur, a été obtenu indépendamment par G. Tenenbaum (cf. [15]).

L'intervention dans (18) de propriétés diophantiennes du nombre α est a priori surprenante. On peut l'expliquer qualitativement par le fait que f est à valeurs rationnelles et que presque tous les entiers n tels que $f(n)\ge\alpha$ sont tels que $f(n)$ est proche de α.

Indiquons les grandes lignes de la démonstration du théorème 3. Si on note $\omega(n)= \sum_{p\mid n} 1$, la relation $f(n)\ge\alpha$ équivaut à $\Omega(n)=h$, $\omega(n)=k$ et $h\ge\alpha k$: on se ramène à l'étude de la distribution conjointe de Ω et ω. La formule de Cauchy nous donne :

$$(19)\qquad P_x(\Omega(n)\ge\alpha\omega(n))= \frac{1}{(2i\pi)^2}\iint_{|z_1|=r_1,\,|z_2|=r_2} x^{-1}\sum_{n\le x} z_1^{\Omega(n)}z_2^{\omega(n)}M(z_1,z_2)dz_1dz_2$$

où les cercles sont parcourus dans le sens positif, $M(z_1,z_2)= \sum_{h\ge\alpha k} z_1^{-h-1}z_2^{-k-1}$ (c'est une fonctioon holomorphe de (z_1,z_2) pour

$|z_2||z_1|^{\alpha}>1$, $|z_1|>1$; on prend donc $r_1>1$ et $r_2 r_1^{\alpha}>1$).

Un résultat classique de A. Selberg permet d'écrire :

$$\sum_{n\leq x} z_1^{\Omega(n)} z_2^{\omega(n)} = \frac{B(z_1,z_2)}{z_1-2}\, x(\log x)^{z_1 z_2 - 1} + \text{reste}$$

si $|z_1|<2$, où $B(z_1,z_2)= \dfrac{2^{-z_1 z_2}}{\Gamma(z_1 z_2)}\,(z_1-2-z_1 z_2)\prod(1-\frac{1}{\ell})^{z_1 z_2}(1+\frac{z_1 z_2}{\ell-z_1})$ (c'est une fonction holomorphe de (z_1,z_2) pour $|z_1|<3$, $z_2\in\mathbb{C}$).

En prenant $1<r_1<2$, on obtient donc

$$P_x(\Omega(n)\geq\alpha\omega(n)) = \frac{1}{(2i\pi)^2}\iint_{|z_1|=r_1,\,|z_2|=r_2} (\log x)^{z_1 z_2 - 1}\,\frac{B(z_1,z_2)}{z_1-2}\,M(z_1,z_2)\,dz_1 dz_2$$

$$+ \text{reste}$$

$$= -\frac{1}{2i\pi}\int_{|z_2|=r_2} (\log x)^{2z_2-1} B(2,z_2)M(2,z_2)\,dz_2$$

$$+ \text{reste}$$

d'après le théorème des résidus, si $r_2>2^{-\alpha}$.

Si nous écrivons $-B(2,z_2)= \sum_{a\geq 0}\beta_a z_2^a$ pour tout $z_2\in\mathbb{C}$, la dernière intégrale est le coefficient de z^{-1} dans le produit $(\log x)^{-1}\sum_{a\geq 0}\beta_a z^a . \sum_{b\geq 0}\frac{(2\log\log x)^b}{b!} z^b . \sum_{h\geq\alpha k} 2^{-h-1}z^{-k-1}$, à savoir :

$$(\log x)^{-1}\sum_{a,b\geq 0}\beta_a\,\frac{(2\log\log x)^b}{b!}\sum_{h\geq\alpha(a+b)} 2^{-h-1} =$$

$$= (\log x)^{-1}\sum_{a\geq 0}\beta_a(2^{-\alpha})^a\sum_{b\geq 0}\frac{(2^{1-\alpha}\log\log x)^b}{b!}\,2^{-\{-\alpha(a+b)\}}$$

où $\{\ \}$ désigne la partie fractionnaire.

Pour estimer la somme sur b, on utilise un raffinement d'un théorème abélien de G. Tenenbaum sur le procédé de sommation de Borel. Cela fournit, pour le résidu cherché, le terme principal :

$$(\log x)^{-1} \sum_{a\geq 0} \beta_a (2^{-\alpha})^a g(\alpha)(\log x)^{2^{1-\alpha}} = C(\alpha)(\log x)^{2^{1-\alpha}-1}.$$

Les propriétés diophantiennes de α interviennent ici pour estimer la différence $\sum_{b\geq 0} \frac{(2^{1-\alpha}\log\log x)^b}{b!} 2^{-\{\alpha(a+b)\}} - g(\alpha)(\log x)^{2^{1-\alpha}}$.

8. Problème 2 (J. Steinig) : $\tau(n)$ désignant le nombre de diviseurs de l'entier positif n, estimer $P_x(\tau(n)\geq \log x)$.

L'intérêt de ce problème est que la moyenne $\frac{1}{x}\sum_{n\leq x} \tau(n) = \log x + O(1)$ est une grande déviation par rapport à l'ordre normal de $\tau(n)$, $1\leq n\leq x$, qui vaut $(\log x)^{\log 2+o(1)}$. Dans [10], K.K. Norton a observé que l'étude des grandes déviations des fonctions additives $\omega(n)$ et $\Omega(n)$ est plus simple que dans le cas de la fonction $\frac{\log \tau(n)}{\log 2}$, qui n'est pas à valeurs entières. Compte tenu de l'encadrement connu : $\omega(n) \leq \frac{\log \tau(n)}{\log 2} \leq \Omega(n)$, il obtenait ainsi :

$$P_x(\tau(n)\geq \log x) \asymp (\log x)^{-\delta}(\log\log x)^{-1/2}$$

où $\delta = 1-(1+\log\log 2)/\log 2 = 0{,}08607\ldots$

P. Erdös et J.-L. Nicolas donnèrent dans leur article [7] une formule asymptotique dûe à H. Delange pour $P_x(\omega(n)\geq \lambda\log\log x)$, $\lambda>1$ fixé. La démonstration de H. Delange utilisait fortement le fait que ω est à valeurs entières, ce qui ne permettait pas de la généraliser immédiatement à $\frac{\log \tau(n)}{\log 2}$. Dans un travail en commun avec J.-L. Nicolas, C. Pomerance et G. Tenenbaum, nous avons obtenu une formule asymptotique pour $P_x(\tau(n)\geq \log x)$:

Théorème 4. Pour $x\geq 3$, on a :

$$P_x(\tau(n)\geq \log x) = C\, K\left(\left\{\frac{\log\log x}{\log 2}\right\}\right) \frac{1+O(1/\log\log x)}{(\log x)^{\delta}\sqrt{\log\log x}},$$

où $C = \frac{\sqrt{\log 2}}{(1-\log 2)\sqrt{2\pi}} \frac{1}{\Gamma(1+\frac{1}{\log 2})} \prod_p (1-\frac{1}{p})^{1/\log 2}(1+\frac{1}{p\log 2}) = 0{,}378\ldots$, et

$$K(\theta) = \frac{1}{(\log 2)^{\theta}} \sum_{d\in S} d^{-1} \prod_{p\mid d} (1+\frac{1}{p\log 2})^{-1} (\log 2)^{-\left[\frac{\log \tau(d)}{\log 2}+1-\theta\right]} \qquad (0\leq\theta<1),$$

S désignant l'ensemble des entiers d tels que $p\mid d \Rightarrow p^2\mid d$.

Ce résultat est en fait un corollaire d'un théorème plus général concernant les fonctions de la forme $f(n)= \omega(n)+\rho(n)$, où ρ est une fonction additive à support dans S (c'est-à-dire vérifiant $\rho(p)=0$ pour tout premier p) soumise à certaines conditions de croissance. L'évaluation d'une probabilité de grande déviation pour f se fait ainsi : on écrit

$$P_x(f(n)\geq z)= \sum_k P_x(\omega(n)= k, \rho(n)\geq z-k) \tag{20}$$

et on est ramené à l'étude de la distribution conjointe de ω et ρ. Cette étude est accomplie par la méthode des séries génératrices, suivant les idées de A. Selberg (cf. [14]); le résultat est suffisamment précis pour permettre l'estimation de la somme du second membre de (20).

BIBLIOGRAPHIE

[1] M. Balazard.- Sur la fonction $\Omega_E(n)$, Groupe d'Etude en Théorie Analytique des Nombres, 2ème année, 1984/85, N°34, 10 pages.

[2] M. Balazard. Sur un théorème de Halász et Sárközy, Groupe d'Etude en Théorie Analytique des Nombres, 3ème année, 1985/86, N° 17, 9 pages.

[3] M. Balazard.- Sur la répartition des valeurs de certaines fonctions arithmétiques additives, Thèse, Université de Limoges, 1987.

[4] H. Delange.- A theorem on integral-valued additive functions, Illinois Journal of Maths., t. 18 N° 3, 1974, p. 357-372.

[5] R.L. Duncan.- On the factorization of integers, Proc. Amer. Math. Soc., t. 25, 1970, p. 191-192.

[6] P.D.T.A. Elliott.- Probabilistic number theory. Vol. I-II. New-York, Heidelberg, Berlin, Springer-Verlag, 1979-1980 (Grundlehren der mathematischen Wissenschaften, 239-240).

[7] P. Erdös et J.-L. Nicolas.- Sur la fonction : nombre de facteurs premiers de n, Enseignement Math., t. 27, 1981, p. 3-27.

[8] G. Halász.- Remarks to my paper "On the distribution of additive and the mean values of multiplicative arithmetic functions", Acta Math. Acad. Scient. Hungar., t. 23, 1972, p. 425-432.

[9] J.-L. Nicolas.- Sur la distribution des nombres entiers ayant une quantité fixée de facteurs premiers, Acta Arithm., t. 44, 1984, p. 191-200.

[10] K.K. Norton.- On the number of restricted prime factors of an integer, I, Illinois Journal of Maths., t. 20, 1976, p. 681-705..

[11] K.K. Norton.- On the number of restricted prime factors of an integer, III, Enseignement Math., t. 28, 1982, p. 31-52.

[12] S. Ramanujan.- Collected Papers, Chelsea Publishing Company, 1962.

[13] A. Sárközy.- Remarks on a paper of G. Halász, Period. Math. Hungar., t. 8, 1977, p. 135-150.

[14] A. Selberg.- Note on a paper by L.G. Sathe, J. Indian Math. Soc., t. 18, 1954, p. 83-87.

[15] G. Tenenbaum.- Sur la distribution conjointe des deux fonctions "nombre de facteurs premiers", à paraître dans Aequationes Mathematicae.

[16] E. Wirsing.- Das asymptotische Verhalten von Summen über multiplikative Funktionen, Math. Annalen, t. 143, 1961, p. 75-102.

Michel Balazard
Département de Mathématiques
Université de Limoges
123, avenue ALbert Thomas
87060 LIMOGES CEDEX

Suites et mesures ayant des propriétés d'auto-similarité

J.-P. Borel

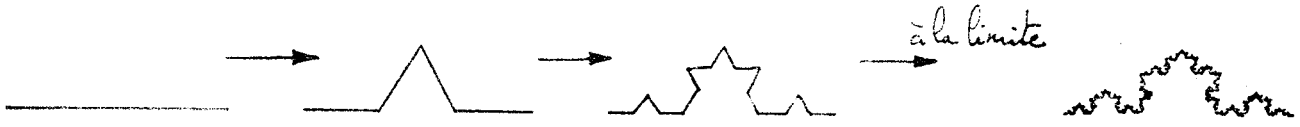
1.- Introduction : l'auto-similarité

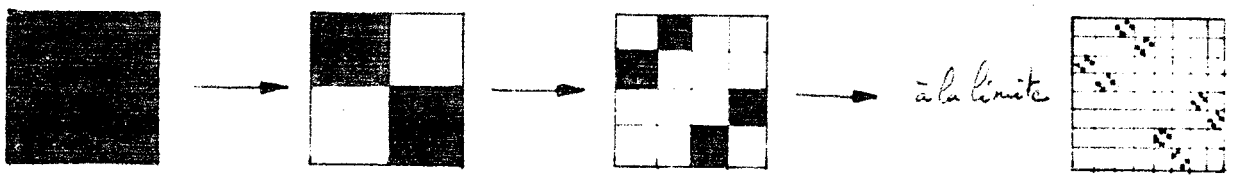
1.1. Essentiellement, un objet mathématique est dit auto-similaire (ou self-similaire) si en regardant un morceau à la loupe, on voit exactement l'objet de départ, du moins en plaçant la loupe à un endroit judicieux.

Quelques exemples de tels objets :

- la spirale $\rho=e^{\theta}$, la loupe étant placée à l'origine
- les objets de type fractal, introduits par Mandelbrot. Ces objets s'obtiennent souvent comme limite, après itération d'une transformation simple.

Deux exemples bien connus :



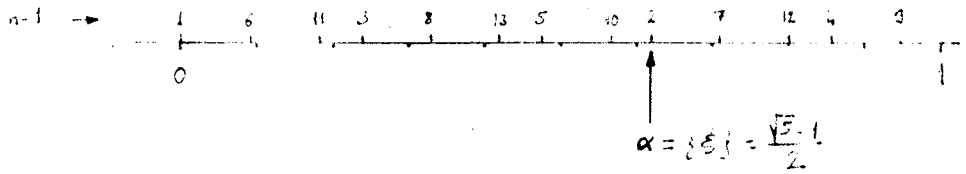


Ces deux types d'objets ont une "dimension" intermédiaire entre 1 et 2 (plusieurs types de dimension peuvent être considérés).

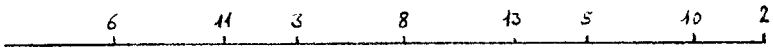
Ici, nous ne nous intéressons pas à ce type de question.

Contrairement à la spirale, la loupe peut, pour les deux objets précédents, être placée en plusieurs endroits (une infinité en fait) mais quand même pas n'importe où... .

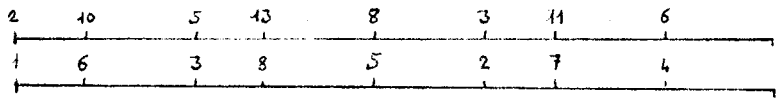
1.2. En théorie des nombres, on rencontre aussi de tels objets : considérons la suite des parties fractionnaires $\{n\xi\}$, où ξ est le nombre d'or $\frac{1+\sqrt{5}}{2}$. En voici les 13 premiers termes :



Si on place la loupe sur l'intervalle $]0,\alpha]$, on obtient :

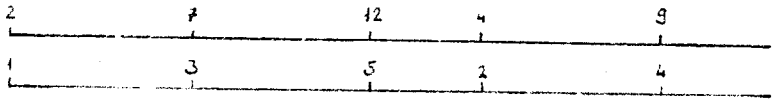


ce qui donne en reversant l'intervalle :



qui est la reproduction exacte des 8 premiers termes de la suite totale, l'ordre étant conservé.

De même, si on place la loupe sur l'intervalle $[\alpha,1]$, on trouve exactement les 5 premiers termes de la suite, dans le bon ordre :



En fait, cette propriété est toujours vraie si l'on considère les n premiers termes de cette suite, quel que soit l'entier n. Cela se montre facilement en utilisant le développement en fraction continue de ξ, qui est $\xi = [1,1,1,\ldots,1,\ldots]$.

1.3. Regardons maintenant la répartition de cette suite dans l'intervalle $[0,1[$, c'est-à-dire les propriétés asymptotiques de la suite des probabilités sur $[0,1[$:

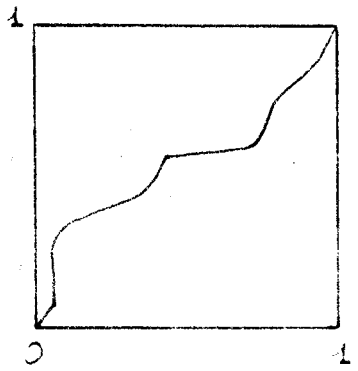
$$\mu_N = \frac{1}{N} \sum_{n=1}^{N} \delta_{\{(n-1)\xi\}}.$$

Il est bien connu que cette suite est équirépartie modulo 1, c'est-à-dire que μ_N converge étroitement vers la probabilité uniforme λ sur $[0,1]$. Cela entraîne donc que la proportion des $\{n\xi\}$ qui sont dans l'intervalle $[0,\alpha]$ tend vers α, quand on prend $0 \leq n \leq N-1$ et N tendant vers $+\infty$.

Supposons maintenant connue cette seule propriété (qui s'obtient d'ailleurs directement, sans passer par l'équirépartition). Si μ est une mesure adhérente à la suite (μ_N), on a donc $\mu(\alpha)=\alpha$.

Représentons μ à l'aide de son graphe $gr(\mu)$, défini par :

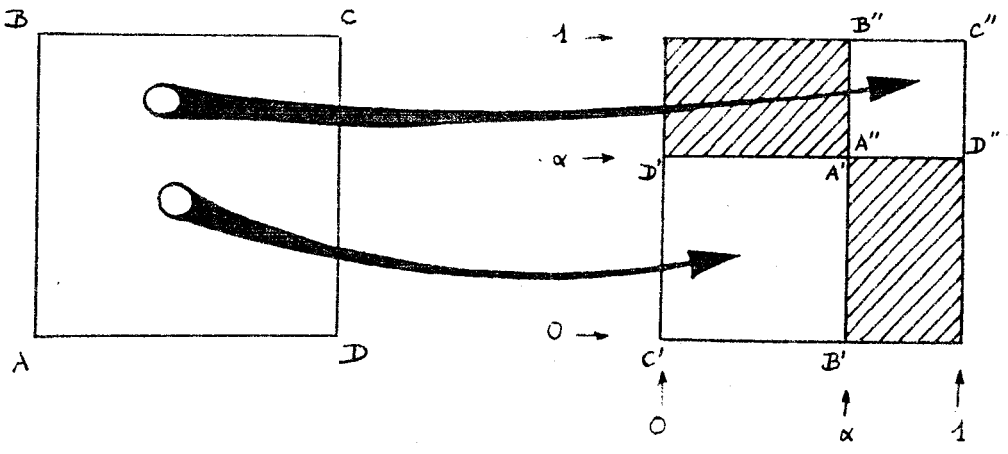
$$(x,y) \in gr(\mu) \iff F_\mu(x) \leq y \leq F_\mu(x+0)$$



où F_μ est la fonction de répartition de μ. Ce graphe est donc une courbe continue qui joint les points $(0,0)$ et $(1,1)$, telle que x et y croissent simultanément, et qui s'obtient très simplement en ajoutant à la représentation graphique de F_μ les segments verticaux correspondant aux sauts.

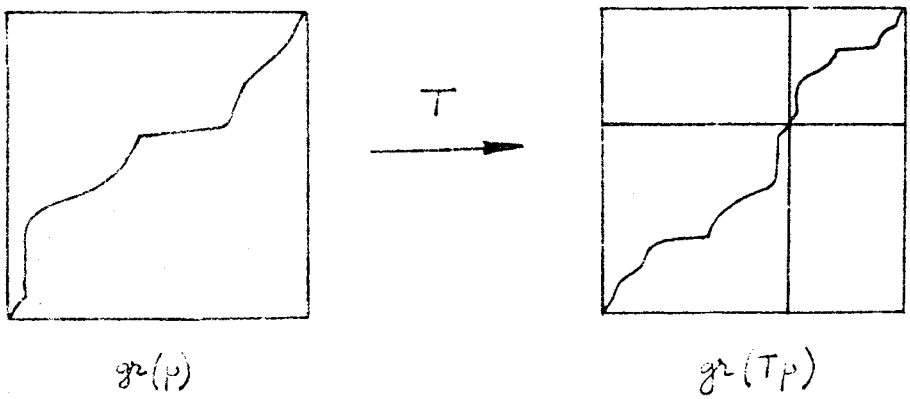
La propriété vue plus haut sur la suite des $\{n\xi\}$ se traduit alors de la façon suivante :

la mesure obtenue à partir de μ par la transformation définie sur le carré $[0,1] \times [0,1]$ par :



(les deux images du carré ABCD sont obtenues par des transformations affines de chaques coordonnées)
cette transformation étant prise au niveau des graphes, est aussi adhérente à la suite (μ_N), par morceaux.

En particulier, si on note T la transformation ainsi définie dans l'ensemble des probabilités sur [0,1]



et si la suite $\{n\xi\}$ admet une mesure de répartition asymptotique (i.e. μ_N admet une limite faible) μ, on a nécessairement $\mu = T\mu$.

Une telle mesure a des propriétés d'autosimilarité au niveau de son graphe, puisque l'on retrouve dans A"B"C"D" la copie (homothétique) du graphe total.

Il est facile de voir que, dans ce cas, T a un seul point fixe, qui est la probabilité uniforme λ, de graphe la diagonale AC.

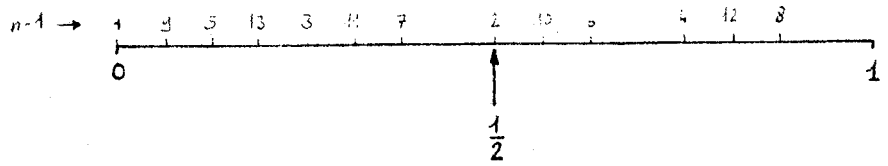
1.4. On observe des propriétés tout à fait analogue si l'on regarde la suite de van der Corput, définie par :

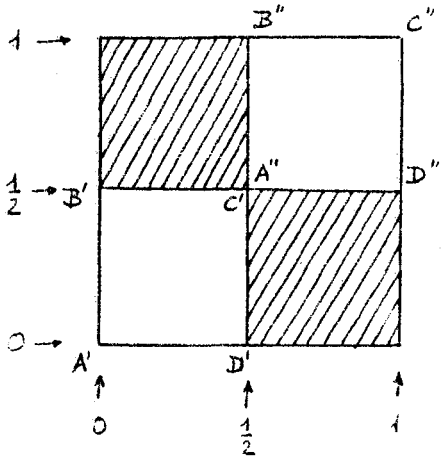
$$u_n = \sum_{i=0}^{\infty} \epsilon_i \, 2^{-i-1} \quad \text{si} \quad n \quad \text{s'écrit en base 2} \quad n = \sum_{i=0}^{\infty} \epsilon_i 2^i$$

puisque l'on a les relations

$$u_{2n} = \frac{1}{2} u_n$$
$$u_{2n+1} = \frac{1}{2} + \frac{1}{2} u_n$$

ce qui traduit le fait que si l'on place la loupe sur l'intervalle $[0,\frac{1}{2}[$ ou sur l'intervalle $[\frac{1}{2},1[$, on retrouve la suite globale :





Un raisonnement analogue conduit à la transformation T' définie par l'image du carré ci-contre. En effet, il est ici immédiat que la proportion des u_n, $0\leq n\leq N-1$, qui sont dans $[0,\frac{1}{2}[$ tend vers $\frac{1}{2}$, puisque $u_n<\frac{1}{2}$ équivaut à n pair.

Ici aussi, la probabilité uniforme est le seul point fixe de T'. Il est en fait classique que la suite de van der Corput est équirépartie modulo 1.

2.- Suites auto-similaires

2.1. On dira qu'une suite U à valeurs dans [0,1] est auto-similaire, si elle vérifie des propriétés :

$$(*) \qquad U|_I = \varphi(U)$$

où l'on a :

I sous-intervalle de [0,1];

φ bijection monotone (i.e. continue) de [0,1] dans l'adhérence $\overline{I}$ de I;

si $U=(u_n)_{n\geq 1}$, $\varphi(U)$ est la suite de terme général $\varphi(u_n)$ pour $n\geq 1$

$U|_I$ est la sous-suite de U formée par les u_n appartenant à I

propriétés vérifiées pour un ensemble fini de sous-intervalles I, formant une quasi-partition de l'intervalle [0,1] :

$\mathrm{card}(I \cap I') < \infty$ (donc égal à 0 ou 1)
$\mathrm{card}([0,1] - \bigcup I) < \infty$

Les propriétés (*) sont des propriétés d'auto-similarité (on place la loupe sur I), mais φ traduit le fait que l'on peut avoir une loupe déformante. Il n'y a pas déformation lorsque φ est affine, et φ décroissante signifie que l'on renverse l'image (comme cela est le cas en 1.2, pour l'intervalle $]0,\alpha]$).

2.2. Comment construire des suites auto-similaires, en général ? Le schéma de construction consiste à découper $[0,1]$ en morceaux, à définir une fonction φ pour chacun des morceaux, et à préciser l'ensemble des indices n tels que u_n est dans ce morceau.

Cela donne le modèle suivant :

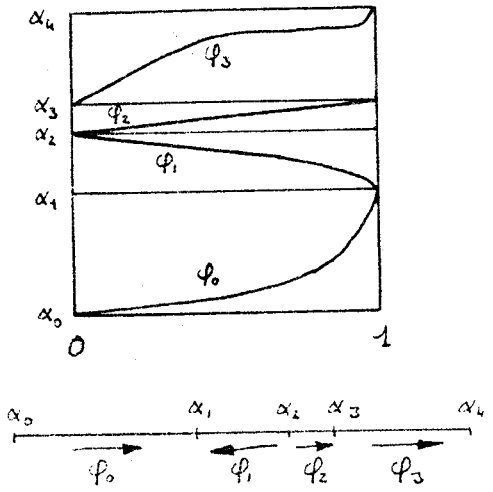
• $\mathcal{A}$ alphabet fini, en pratique $\mathcal{A} = \{0,1,\ldots,k-1\}$ avec $k \geq 2$

• $\mathcal{F}$ fractionnement de $[0,1]$, indicé sur $\mathcal{A}$: $\mathcal{F} = (I_a)_{a \in \mathcal{A}}$ avec des I_a sous-intervalles de $[0,1]$, formant une quasi-partition. En pratique, on choisit donc :

$$0 = \alpha_0 < \alpha_1 < \ldots < \alpha_k = 1$$

$I_a = I\alpha_a, \alpha_{a+1}I$ où I signifie] ou [indifféremment

• Φ déformation de support $\mathcal{F}$, c'est-à-dire $\Phi = (\varphi_a)_{a \in \mathcal{A}}$ où φ_a est bijective continue de $[0,1]$ dans $\overline{I}_a$.



deux représentations possibles de Φ

Ces déformations peuvent avoir certaines propriétés. Nous serons amenés à considérer les suivantes :

Φ <u>contractante</u> si toutes les φ_a sont contractantes, c'est-à-dire qu'il existe une constante $c<1$ telle que :

$$\forall a \in \mathcal{A} \quad \forall (x,y)\in[0,1]^2 \quad |\varphi_a(x)-\varphi_a(y)| \leq c.|x-y|$$

Φ <u>contractante large</u> si le même résultat est vrai avec $c=1$ seulement.

Φ <u>monotone</u> si toutes les applications φ_a ont même sens de monotonic. On dira alors que Φ est <u>croissante</u>, ou <u>décroissante</u>.

Φ <u>homogène</u> si on a :

$$\forall x\in[0,1], \quad \sum_{a\in\mathcal{A}} |\varphi_a(x)-\varphi_a(0)| = x$$

ce qui est vérifié dans le cas particulier suivant :

Φ <u>linéaire</u> si toutes les applications φ_a sont affines. Elles sont donc uniquement déterminées par les I_a et leur sens de variation. Ce cas est celui observé dans les deux exemples du § 1.

• E <u>règle de partage</u>, c'est-à-dire $E=(\epsilon_n)_{n\geq 1}$ suite à valeurs dans $\mathcal{A}$, et telle que l'équation $\epsilon_n=a$ a une infinité de solutions n pour chaque $a\in\mathcal{A}$.

<u>Définition</u>. <u>On dit que la suite</u> $U=(u_n)_{n\geq 1}$ <u>est auto-similaire assoicée à</u> (Φ,E) <u>si elle vérifie les deux conditions suivantes</u> :

$$\forall a\in\mathcal{A} \quad U|_{I_a} = \varphi_a^o(U) \quad \underline{\text{où}} \quad \varphi_a^o \ \underline{\text{est la restriction de}} \ \varphi_a \ \underline{\text{à}} \ \varphi_a^{-1}(I_a);$$

$$\forall n\geq 1 \quad u_n \in I_{\epsilon_n} \quad \underline{\text{ou}} \quad u_n \notin F = \bigcup_{a\in\mathcal{A}} I_a.$$

<u>2.3</u>. L'existence, et alors le nombre, de suites auto-similaires associées à (Φ,E) donné n'est pas immédiate. En s'imposant (et pour toute la suite) la condition que les u_n sont deux à deux distincts, quatre problèmes apparaissent :

1 $\varphi_a(x)=x$ a beaucoup de solutions x, pour un a donné.
Si $\epsilon_1=a$, il y a donc beaucoup de choix possibles pour u_1 dans I_a.
$\longrightarrow$ non unicité

2 $\varphi_a(x)=x$ a peu de solutions x, pour un certain a, soit un nombre fini m. Si $\epsilon_1=\epsilon_2=\ldots=\epsilon_{m+1}=a$, on ne peut pas placer à la fois $u_1, u_2 \ldots u_{m+1}$ dans l'intervalle I_a $\longrightarrow$ non existence

3 $[0,1]-F$ est non vide, on peut donc y placer certains u_n.
$\longrightarrow$ non unicité

4 $I_a \cap I_{a'}$ non vide, pour certains couples (a,a') avec $a \neq a'$. Un u_n commun pose alors des problèmes. $\longrightarrow$ non existence

Les problèmes 3 et 4 sont des phénomènes de bord. Ils apparaissent clairement si l'on cherche à déterminer E associée à la suite $\{n\xi\}$ introduite au § 1 :

$$I_0 =]0,\alpha] \qquad I_1 = [\alpha,1[$$

et donc : ϵ_1 indéterminé car $0 \notin I_0 \cup I_1$ (problème 3)
ϵ_2 indéterminé car $\alpha \in I_0 \cap I_1$ (problème 4)
ϵ_n déterminé pour $n \geq 2$, et l'on voit facilement que :

$$\epsilon_n = 1 - [(n-1)(\xi-1)] + [(n-2)(\xi-1)]$$

ce qui donne la suite :

$$E = ??01001010010\ldots$$

et si on remplace les ?? par 01 pour que la formule donnant ϵ_n reste vraie, on obtient :

$$E = 0101001010010\ldots$$

suite qui est le point fixe commençant par 0 engendré par la substitution :

$$0 \longrightarrow 010$$
$$1 \longrightarrow 10$$

(cf. [3] pour une démonstration).

2.4. En faisant un certain nombre d'hypothèses (simples mais très techniques) sur Φ et E, on peut éviter les problèmes 1 et 2, et faire en sorte que les phénomènes de bord 3 et 4 se compensent. On obtient alors :

Proposition (sous certaines hypothèses non précisées ici). Il existe une et une seule suite auto-similaire $U=(u_n)_{n\geq 1}$ associée à (Φ,E). Elle est définie par, en posant $\kappa=\mathrm{card}([0,1[\cap F)$:

si $\kappa=0$ $\quad u_1=u[\epsilon_1]$ ($u[a]$ désignant le point fixe unique de φ_a)
  si $n\geq 2$ $\quad u_n=\varphi_{\epsilon_n}(u_m)$
si $\kappa\geq 1$ $\quad u_1,u_2,\ldots,u_\kappa$ choisis dans $[0,1[\cap F$
  si $n\geq\kappa+1$ $\quad u_n=\varphi_{\epsilon_n}(u_m)$

où m s'obtient à partir de n par la relation :

$$m=1+\mathrm{card}\,\{1\leq i\leq n-1/u_i\in I_{\epsilon_n}\}.$$

2.5. Considérons les quantités :

$$\eta(n,a)=\mathrm{card}\,\{1\leq i\leq n/\epsilon_i=a\}\qquad a\in\mathcal{A},\quad n\geq 1$$

qui décrivent la suite E.

En supposant vérifiées quelques autres hypothèses techniques, on obtient alors la relation :

$$m=1+\mathrm{card}\,\{1\leq i\leq n-1/u_i\in I_{\epsilon_n}\}=\eta(n,\epsilon_n)$$

qui sera supposée vérifiée dans toute la suite.

Cela revient à dire que les phénomènes de bord n'existent pas (par exemple, si les intervalles I_a sont deux à deux disjoints, le résultat ci-dessus est immédiat) ou se compensent : si l'on revient à l'exemple $\{n\xi\}$ du § 1, remplacer les ?? par 01 permet le résultat pour $n\geq 2$.

2.6. Nous allons nous intéresser aux problèmes suivants :
dans les bons cas du 2.5, est-ce que la suite (Φ,E) a une répartition asymptotique dans l'intervalle $[0,1]$? Il se trouve que la réponse ne dépend que de la suite E, et plus précisément de l'existence des limites :

$$\lim_{n\to\infty}\frac{1}{n}\eta(n,a)=\gamma_a\qquad(a\in\mathcal{A})$$

c'est-à-dire de l'existence d'une fréquence asymptotique de chacune des lettres dans le mot infini E.

La mesure de répartition est, lorsqu'elle existe, pure. Elle dépend de propriétés relatives de Φ et de E, et son graphe vérifie des propriétés d'auto-similarité (en fait, elle est le point fixe unique d'une certaine transformation T du type de celles introduites au § 1).

Enfin, une formule explicite est donnée, permettant de mesurer l'écart entre la mesure limite μ et les mesures $\mu_N = \frac{1}{N} \sum_{n=1}^{N} \delta_{u_n}$. Cet écart permet d'évaluer la discrépance, dans le cas où U est équirépartie.

3.- Mesures auto-similaires

3.1. Soient I et J deux sons-intervalles fermés de [0,1] :

$$I = [\alpha,\alpha'] \qquad J = [\beta,\beta']$$

obtenus comme image de [0,1] par deux applications φ et ψ :

φ : $[0,1] \longrightarrow U$ bijective continue , constante si $\alpha=\alpha'$;
ψ : $[0,1] \longrightarrow J$ bijective croissante, constante si $\beta=\beta'$.

On définit alors une transformation $T_{\varphi\psi}=T$ sur l'ensemble $\mathcal{P}$ des probabilités sur [0,1] :

ν probabilité sur [0,1] : $\mu \in \mathcal{P}$
ν mesure image de μ par φ : on a donc $\nu \in \mathcal{P}$, ν de support contenu dans I
T_μ défini par sa fonction de répartition $F_{T_\mu} = \psi \circ F_\nu - \beta$.

T_μ est donc une mesure positive, à support contenu dans I, et de masse totale $\beta'-\beta=|J|$.

Il est alors facile de voir que l'on a :

$$F_{T_\mu}(x) = \begin{cases} \tilde{\psi}(F_\mu(\varphi^{-1}(x))) & \text{si } \varphi \text{ croissante} \\ \tilde{\psi}(1-F_\mu(\varphi^{-1}(x)+0)) & \text{si } \varphi \text{ décroissante} \end{cases}$$

si l'on note $\tilde{\psi} = \psi - \beta$.

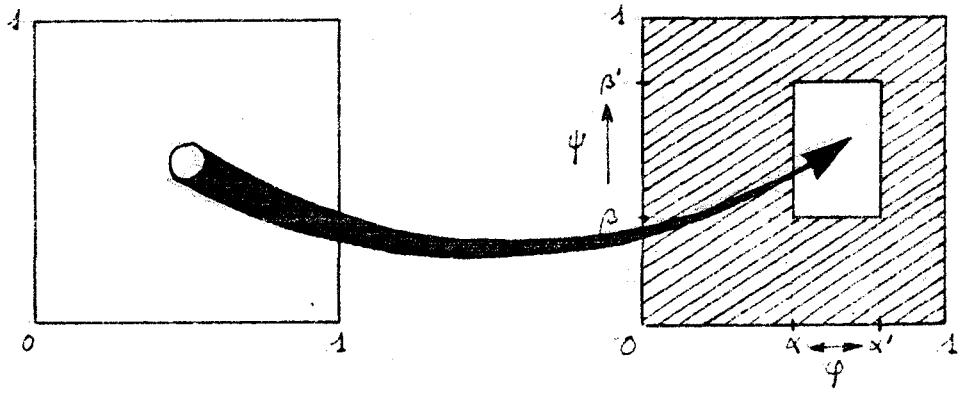
Si on suppose $x \in I$ et $y \in J$, on obtient

$$(x, y-\beta) \in gr(T\mu) \iff (\varphi^{-1}(x), \hat{\psi}^{-1}(y)) \in gr(\mu)$$

(φ et ψ non constantes), et dans les cas dégénérés :

si $|J|=0$, $\quad T_\mu = 0$
si $|I|=0$, $\quad T_\mu = |J| . \delta_\alpha$.

On peut représenter T de la façon suivante :



3.2. Soient Φ et Ψ deux déformations, de supports respectifs $\mathcal{I}$ et $\mathcal{J}$ indicés sur le même alphabet fini $\mathcal{A}$. On suppose que :

Ψ est croissante;
Φ et Ψ peuvent être dégénérées, c'est-à-dire que certains des intervalles associés I_a et J_a peuvent être réduits à un point. Les applications associées sont alors constantes.

Définition. On dit que la mesure $\mu \in \mathcal{P}$ est auto-similaire associée à (Φ,Ψ) si elle est point fixe de la transformation de $\mathcal{P}$ dans lui-même définie par :

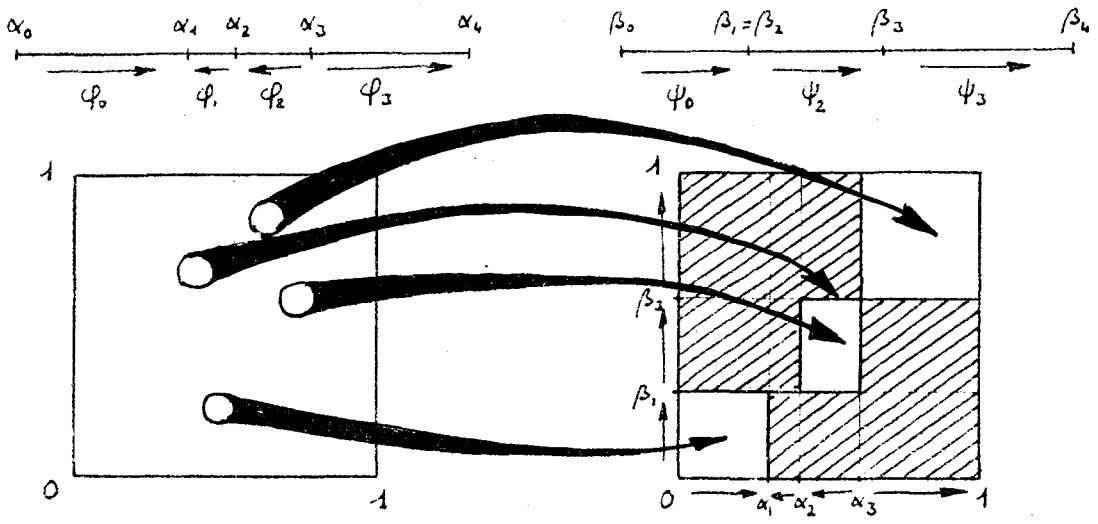
$$T = T(\Phi,\Psi) = \sum_{a\in\mathcal{A}} T_{\varphi_a \psi_a} .$$

Si $\mu \in \mathcal{P}$, $T\mu$ est donc une mesure positive, de masse totale :

$$\|T\mu\| = \sum_{a\in\mathcal{A}} \|T_{\varphi_a \psi_a} \mu\| = \sum_{a\in\mathcal{A}} |J_a| = 1$$

et donc $T\mu \in \mathcal{P}$.

La transformation T peut se représenter de la façon suivante, si par exemple :



3.3. Ni l'existence, ni l'unicité, ne sont assurées dans le cas général lorsque l'on cherche les mesures auto-similaires associées à (Φ,Ψ) donné.

Φ étant choisi, on notera $\hat{\Psi}$ la déformation de support $\mathcal{G}$ définie par :

$\hat{\psi}_a = \psi_a$ si φ_a est croissante ou constante;

$\hat{\psi}_a = \psi_a \circ s$ si φ_a est décroissante, s étant la bijection sur $[0,1]$: $x \longmapsto 1-x$. On obtient alors :

Lemme. Soit $a \in \mathcal{A}$, et $(x,y) \in gr(\mu)$. Alors $(\varphi_a(x), \hat{\psi}_a(x)) \in gr(T\mu)$.

Comme les points $(0,0)$ et $(1,1)$ sont dans le graphe de μ, on en déduit $k+1$ points du graphe de $T\mu$, et par itération k^n+1 points du graphe de $T^n\mu$ (si l'on suppose que $|I_a| = |J_a| = 0$ est exclu, ce que l'on peut faire : sinon, on enlève la lettre a de $\mathcal{A}$...).

Si μ est point fixe de T, on peut espérer que tous les points ainsi obtenus caractérisent la mesure μ : c'est ce qui se passe dans les bons cas. Pour préciser cela, il faut introduire une nouvelle notion : celle de numération liée à une déformation.

3.4. Une autre approche possible est de considérer T comme une application de $\mathcal{P}([0,1]^2)$ dans lui-même, donnée par :

$$TA = \bigcup_{a\in\mathcal{A}} \{(\varphi_a(x), \tilde{\psi}_a(y)), (x,y)\in A\}$$

La suite des itérés $T^n[0,1]^2$ décroît vers un ensemble limite L, qui est donc stable par T. Si l'on contient un rectangle vrai (non aplati ou réduit à un point), il est difficile de conclure. Sinon, L est le graphe d'une certaine mesure $\mu\in\mathcal{P}$, qui est alors l'unique point fixe de T. Cela est le cas, lorsque Φ et Ψ sont contractants, et on peut dire mieux :

Proposition. Si Φ et Ψ sont contractantes, T est contractante au sens de la distance de Paul Levy sur $\mathcal{P}$.

La distance de Paul Levy correspond à la convergence étroite des mesures de probabilité. On peut trouver sa définition dans [12].

T a donc bien un point fixe unique μ^0, et pour toute probabilité $\mu\in\mathcal{P}$, la suite $T^n\mu$ converge étroitement vers μ^0. On peut donc espérer obtenir, d'après un résultat de Méla [11], un résultat de pureté sur μ^0.

C'est ce que nous préciserons par la suite. Pour cela, nous aurons encore besoin de la notion de numération associée à une déformation.

4.- Système de numération associé à une déformation

4.1. On considère une déformation Φ, éventuellement dégénérée :

$$\mathcal{A} = \{0,1,\ldots,k-1\}$$
$$\varphi_a : [0,1] \longrightarrow [\alpha_a, \alpha_{a+1}]$$

et on note :

$\mathcal{A}^*$ ensemble des mots finis $= a_1a_2\ldots a_n$ sur l'alphabet $\mathcal{A}$ (donc $\mathcal{A}^* = \bigcup_{n=0}^{\infty} \mathcal{A}^n$)

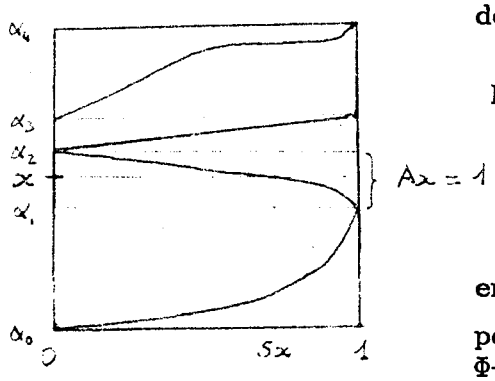
$\mathcal{A}^\infty$ ensemble des mots infinis $\_ = a_1a_2\ldots a_n\ldots$ sur l'alphabet $\mathcal{A}$ ($\mathcal{A}^\infty = \mathcal{A}^{\mathbb{N}^*}$)

ainsi que

$D = \{\alpha_1, \alpha_2, \ldots, \alpha_{k-1}\}$ ensemble des points communs à deux intervalles.

4.2. Les points de $[0,1]-D$ sont dans un seul intervalle $[\alpha_a, \alpha_{a+1}]$, ou I_a. Cela permet de définir deux applications S et A :

$$A : [0,1]-D \longrightarrow \mathcal{A} \quad \text{par} \quad x\in I_{Ax}$$
$$S : [0,1]-D \longrightarrow [0,1] \quad \text{par} \quad Sx = \varphi_{Ax}^{-1}(x)$$



On ne peut itérer S directement, mais on peut itérer ce procédé, en définissant S^n sur $[0,1]-D_n$, où D_n est défini par :

$$D_n = \bigcup_{i<n} S^{-i}(D)$$

Si on pose $D^* = \bigcup_{n=0}^{\infty} S^{-n}(D) = \lim\uparrow D_n$, ensemble dénombrable, donc petit, cela permet de définir sur $[0,1]-D^*$ le Φ-développement de x, de la façon suivante :

$$\begin{array}{ccc} [0,1]-D^* & \longrightarrow & \mathcal{A}^\infty \\ & \mathbb{F} & \\ x & \longrightarrow & (AS^{n-1}x)_{n\geq 1} \end{array}$$

4.3. Soit $a = a_1 a_2 \ldots a_n$ un mot fini. On pose alors :

$$\varphi_a(x) = \varphi_{a_1} \circ \varphi_{a_2} \circ \ldots \circ \varphi_{a_n}(x) \qquad (0\leq x\leq 1)$$

Soit $\underline{a}$ un mot infini, et a_n son préfixe de longueur n. Alors $K_n = \varphi_{a_n}([0,1])$ est une suite décroissante d'intervalles fermés, et donc a une limite K, intervalle fermé non vide. On notera :

$$K = I_{\underline{a}} = \varphi_{\underline{a}}([0,1]) = [x_{\underline{a}}, y_{\underline{a}}] \qquad \text{(et aussi } I_a = \varphi_a([0,1]) \text{ pour } a\in\mathcal{A}^*)$$

$$\mathcal{A}_2^\infty = \{\underline{a}\in\mathcal{A}^\infty / |I_{\underline{a}}| = 0\}.$$

Cela permet, en identifiant l'intervalle $[x,x]$ au point x, de définir une application :

$$\mathcal{A}_r^\infty \longrightarrow [0,1]$$

$$\underline{a} \xmapsto{G} x = \lim_{n\to\infty} \varphi_{a_1} \circ \varphi_{a_2} \circ \ldots \circ \varphi_{a_n}(\xi) \qquad 0\leq\xi\leq 1$$

Dans le cas particulier où Φ est contractante, alors $\mathcal{A}_r^\infty = \mathcal{A}^\infty$: en effet, si c est associée à Φ, on a (pour tout $\underline{a}$ et $n\geq 1$) $|K_n| \leq c^n$, et donc $|I_a| = 0$ pour tout mot infini $\underline{a}$.

4.4. On peut maintenant préciser la notion de Φ-développement de $x\in[0,1]$, généralisant celle introduite en 4.2 :

est dit Φ-développement infini de x lorsque $x\in I$;
est dit Φ-développement fini de x lorsque $x=\varphi(0)$.

Proposition. Tout $x\in[0,1]$ a au moins un Φ-développement infini
Si Φ est contractante :

tout $x\notin D^*$ a exactement un Φ-développement infini, égal à $\mathbb{F}(x)$, et ce développement caractérise x.

Si Φ est non dégénérée :

tout $x\in D^*$ a exactement deux Φ-développements infinis

Si Φ est monotone croissante :

tout $x\in D^*$ a exactement un Φ-développement fini (qui le caractérise...).

Cela signifie essentiellement que, dans les bons cas, on a des bijections réciproques l'une de l'autre :

$$[0,1]-D^* \underset{G}{\overset{\mathbb{F}}{\rightleftarrows}} X([0,1]-D^*).$$

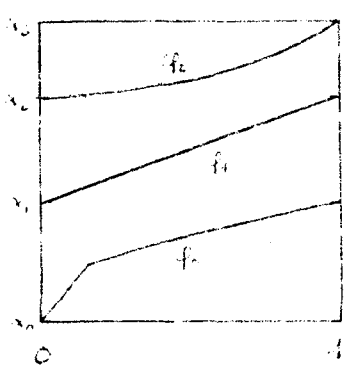
4.5. Lorsque Φ est monotone, on retrouve une façon de représenter les nombres réels, étudiée par Bissinger [2], Everett [5] et Renyi [13]. Il s'agit d'écrire le nombre $x\in[0,1]$ sous la forme :

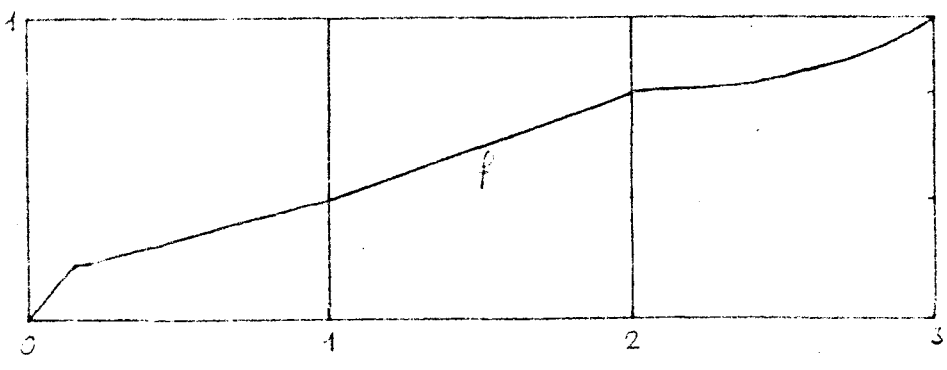
$$x= f(a_1+f(a_2+f(a_3+\ldots)))$$

où f vérifie certaines conditions.

Le lien entre f et Φ est donné par :

$$f : [0,K] \longrightarrow [0,1]$$
$$f(a+x)= \varphi_a(x) \qquad a\in\mathcal{A},\ x\in[0,1]$$





4.6. On retrouve donc les développements classiques. Si Φ est croissante linéaire, avec $\alpha_a = \frac{a}{k}$ $(0 \leq a \leq k)$, on obtient le développement en base k habituel, qui correspond chez Renyi à $f(x) = \frac{x}{k}$.

5.- Propriétés ergodiques

5.1. Soit μ une probabilité sur $[0,1]$, telle que $\mu(D^*)=1$ (c'est par exemple le cas si μ est diffuse), D^* étant associé à une déformation Φ donnée.

Si Φ est contractante, $\mathbb{F}$ définit (presque sûrement) une bijection entre $[0,1]$ et $\mathcal{A}^*$, qui permet de transférer μ. Les propriétés relatives de μ, S, A peuvent alors s'étudier dans l'espace d'arrivée $\mathcal{A}^\infty$.

Dans cet espace :

A devient la projection 1ère coordonnée $\quad \mathcal{A}^\infty \longrightarrow \mathcal{A}$, $\ \underline{a} \longrightarrow a_1$

S devient le shift classique $\quad \mathcal{A}^\infty \longrightarrow \mathcal{A}^\infty$, $\ \underline{a} \longrightarrow S = a_2 a_3 \cdots a_{n+1} \cdots$

φ_a devient l'application $\quad \mathcal{A}^\infty \longrightarrow \mathcal{A}^\infty$, $\ \underline{a} \longrightarrow a\underline{a} = a a_1 a_2 \cdots a_{n-1} \cdots$

5.2. Soient Φ et Ψ deux déformations, comme en 3.2, et $T = T(\Phi,\Psi)$.

Proposition. T a un point fixe unique $\mu^0 = \mu^0(\Phi,\Psi)$ dès que l'on a :

$$\mathcal{A}^\infty = \mathcal{A}_r^\infty(\Phi) \cup \mathcal{A}_r^\infty(\Psi)$$

et le graphe de μ^0 est donné par :

$$gr(\mu^0) = \bigcup_{\underline{a} \in \mathcal{A}^\infty} (\varphi_{\underline{a}}([0,1]) \times \hat{\Psi}_{\underline{a}}([0,1]))$$

La condition est vraiment utile : s'il existe $a \in \mathcal{A}$ tel que $\varphi_a(x) = \psi_a(x) \equiv x$ sur un intervalle non réduit à un point, T a une infinité de points fixes.

Elle est réalisée si $\mathcal{A}_r^\infty(\Phi) = \mathcal{A}^\infty$ (ou la même égalité avec Ψ), ce qui est vrai dès que $D^*(\Phi)$ est dense dans $[0,1]$, ce qui est une conséquence de Φ contractante.

Dans les bons cas (par exemple Ψ contractante et Φ non-dégénérée), on peut expliciter la fonction de répartition de μ^o :

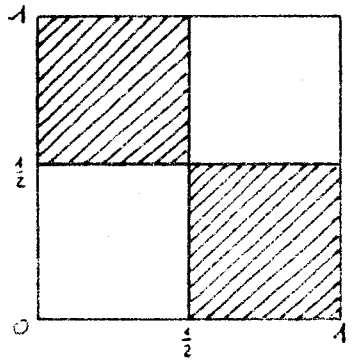
$$F_{\mu_o} = G_{\hat{\Psi}} \circ F_{\Phi} \quad \text{sur} \quad D^*(\Phi).$$

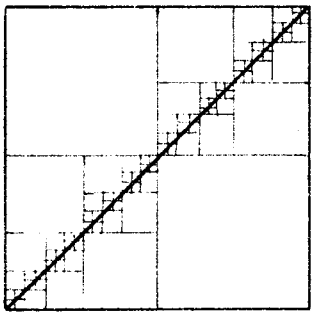
La démonstration de ce résultat provient de l'utilisation du lemme 3.3, et de son itération : on passe des lettres $a \in \mathcal{A}$ aux mots finis **a**, puis aux mots infinis $\underline{a}$. Cette démonstration conduite aussi à :

Proposition. Si $\mathcal{A}^\infty = \mathcal{A}_r^\infty(\Phi) \cup \mathcal{A}_r(\Psi)$, et si $\mu \in \mathcal{P}$, la suite de probabilités $T^n\mu$ converge en loi vers μ^o.

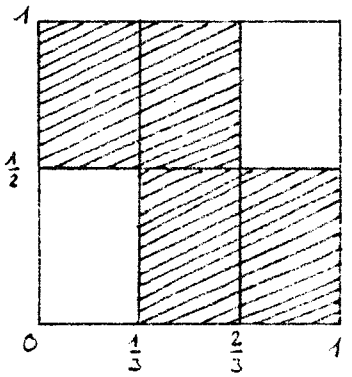
5.3. Il est possible de construire des exemples (avec des fonctions φ_a et ψ_a compliquées...) où μ^o existe, et a une composante absolument continue et une composante purement discontinue.

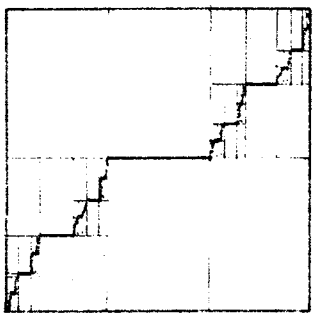
Voici deux exemples simples, où Φ et Ψ sont linéaires croissantes :





$p^\circ = \lambda$





mesure de Cantor

Nous allons voir que ces exemples reflètent bien ce qui se passe.

5.4. Commençons par déterminer les cas où μ^o a des masses ponctuelles, et donc où l'on ne peut pas appliquer le principe énoncé en 5.1.

Proposition. μ^0 n'est pas diffuse si et seulement si $\mathcal{A}_r^\infty(\Psi)$ est strictement contenu dans $\mathcal{A}^\infty$. Si Ψ est contractante, cela équivaut à :

$$\exists a \in \mathcal{A} \qquad |I_a| = 0 \quad \text{et} \quad |J_a| > 0.$$

Dans ce cas, μ^0 est purement discontinue, de support $S = \{\varphi_{aa}(0),\ a \in \mathcal{A}^*\}$.

Le cas Cantor (μ^0 ayant un support Lebesgue-négligeable) se déduit du précédent par symétrie entre Φ et Ψ.

5.5. Regardons maintenant les liens entre la mesure μ^0 et la transformation S. On montre facilement (la définition de homogène est faite pour cela) que si μ^0 existe :

Proposition. μ^0 est S-invariante si et seulement si Ψ est homogène.

Il n'est pas surprenant que seul Ψ joue un rôle : en effet, dès que l'on peut passer dans l'espace image $\mathcal{A}^\infty$, la déformation Φ n'apparaît plus dans la mesure image υ^0 de μ^0 : elle est en effet caractérisée par

$$\forall \in \mathcal{A}^* \qquad \upsilon^0(a\mathcal{A}^\infty) = |\hat{\psi}_a(1) - \hat{\psi}_a(0)| = |J_a|.$$

La question qui se pose alors est de savoir si μ^0 est ergodique, ce qui revient à dire que tout évènement S-invariant B vérifie $\mu^0(B)=0$ ou $\mu^0(B)=1$.

Théorème. Si Ψ est homogène, et s'il existe une constante absolue C telle que :

$$\forall a \in \mathcal{A}^* \qquad \sup|\psi'_a| \leq C \inf|\psi_a|$$

alors μ^0 est ergodique pour S.

L'hypothèse (les ψ'_a existent sauf sur un ensemble au plus dénombrable) est faite pour pouvoir utiliser un théorème de Knopp [9]. Elle est vérifiée avec C=1 si Ψ est linéaire.

Deux mesures ergodiques pour une même transformation sont étrangères : c'est une conséquence facile du théorème ergodique. Si Ψ et Ψ^* vérifient les hypothèses du théorème, on a donc :

$$\Psi \neq \Psi' \Longrightarrow \mu^0(\Phi,\Psi) \perp \mu^0(\Phi,\Psi^*).$$

Nous allons retrouver ce résultat, en précisant les supports respectifs.

5.5. Supposons Φ et Ψ contractantes non dégénérées. μ^o est donc diffuse, et $X_n = AS_{n-1}$ définit (pour $n \geq 1$) une suite de variables aléatoires à valeurs dans $\mathcal{A}$.

Théorème. Si Ψ est linéaire (et Φ contractante non dégénérée), les variables aléatoires X_n sont indépendantes, de même loi donnée par :

$$\forall a \in \mathcal{A} \qquad \mu^o(X_n = a) = |J_a|$$

ce qui peut s'écrire sous la forme équivalente :

$$\nu^o = \bigotimes_{n=1}^{\infty} \sum_{a \in \mathcal{A}} |J_a| \delta_a .$$

Alors μ^o est concentré sur l'ensemble $\tilde{N} = \tilde{N}(\Phi,\Psi)$ des nombres réels x dont le Φ-développement infini vérifie

$$\forall a \in \mathcal{A}, \quad \lim_{n \to \infty} \frac{1}{n} \operatorname{card} \{1 \leq m \leq n \ AS^{m-1}x = a\} = |J_a| .$$

Plus précisément, on peut montrer dans ce cas que μ^o est concentré sur l'ensemble $N = N(\Phi,\Psi)$ des nombres Ψ-normaux en base Φ, défini de la façon suivante :

si $\mathbb{F}_\Phi(x) = a_1 a_2 \ldots a_n \ldots$ défini μ^o-presque sûrement

$x \in N \iff \forall a \in \mathcal{A}^*, \ \lim_{n \to \infty} \frac{1}{n} \operatorname{card} \{1 \leq m \leq n, \ a_m a_{m+1} \ldots a_{m+\ell-1} = a\} = |J_a| \quad (\ell = |a|)$.

Les ensembles $\tilde{N}(\Phi,\Psi)$ étant deux à deux disjoints, on obtient le résultat de pureté suivant :

Corollaire. Soient Φ et Ψ deux déformations linéaires.

si $\mathcal{G} = \mathcal{F}$, $\mu^o = \lambda$;

si $\mathcal{G} = \mathcal{F}$, μ^o est purement singulière.

6.- Répartition des suites auto-similaires

6.1. On reprend ici les notations du §2, dans le cadre simplifié défini en 2.5 : U est auto-similaire associée à (Φ,E). On supposera ici que Φ est contractante (non dégénérée).

Si I est un sous-intervalle de $[0,1]$, on posera :

$$\mathcal{N}(N,I) = \mathcal{N}(N,I,U) = \operatorname{card} \{1 \leq n \leq N / u_n \in I\}.$$

Les hypothèses faites entraînent que $N(N,I_a)=\eta(N,a)$ pour tout $a\in\mathcal{A}$ et N assez grand (la différence ne pouvant jamais dépasser 1, liée aux phénomènes de bord).

On peut montrer qu'une mesure de répartition asymptotique de U ne peut avoir de masse sur les bords des intervalles I_a. Une condition nécessaire pour qu'il existe une telle mesure de répartition est donc que :

$$(*) \qquad \forall a\in\mathcal{A},\ \lim_{N\to\infty}\frac{1}{N}\eta(N,a)=\gamma_a \quad \text{existe}$$

(cette propriété se lit sur la suite E). Alors $\sum_{a\in\mathcal{A}}\gamma=1$, et il existe un unique fractionnement linéaire croissant ψ tel que :

$$\forall a\in\mathcal{A},\quad |J_a|=\gamma_a.$$

ψ étant linéaire, les résultats du § 5 s'appliquent : $T(\Phi,\Psi)$ a un unique point fixe $\mu^o=\mu^o(\Phi,\Psi)$, mesure diffuse vérifiant le théorème 5.5.

6.2. Supposons toujours (*) vérifiée, et considérons les écarts :

$$\mathcal{E}(N,I)=\mathcal{E}(N,I,U)=N(N,I,U)-\mu^o(I).N$$
$$\rho(N,a)=\eta(N,a)-\gamma_a N.$$

Les hypothèses techniques faites entraînent alors :

Formule descendante sur les écarts. Soit $x\in\overline{I}_a$, et $N\in\mathbb{N}^*$. Posons :

$$y=\varphi_a^{-1}(x) \qquad Na=\eta(N,a)$$

Alors :

$$\begin{cases} \sum_{a'<a}\rho(N,a')+\rho(N,a)F_{\mu^o}(y)+\mathcal{E}(Na,[0,y[) & \text{si } \varphi_a\uparrow \\ \sum_{a'\leq a}\rho(N,a')-\rho(N,a)F_{\mu^o}(y)-\mathcal{E}(Na,]0,y]) & \text{si } \varphi_a\downarrow \end{cases}$$

Cette formule est dite descendante car on a, avec les hypothèses faites sur E :

soit $Na<N$
soit $Na=N=1$ et $a=\epsilon_1$.

Elle provient de formules analogues pour $N(N,[0,x[)$ et $\mu^o([0,x[)$.

Si on pose :

$$[a]=\begin{cases}0\\ \\ 1\end{cases} \qquad \boxed{a}=\begin{cases}\{0,1,\dots,a-1\} & \text{si } \varphi_a \text{ est croissante}\\ \\ \{0,1,\dots,a-1,a\} & \text{si } \varphi_a \text{ est décroissante}\end{cases}$$

on obtient alors :

Lemme. Si $a\in\mathcal{A}$, $N\in\mathbb{N}^*$ et $x=u_n$ avec $n\leq N$, alors $y=\varphi_a^{-1}(x)$ vérifie $y=u_m$, avec $m=na\leq Na$, et :

$$\mathcal{E}(N,[0,x])=\sum_{a'\in\boxed{a}}\rho(N,a')+(-1)^{[a]}\rho(N,a).F_{\mu^0}(y)+\mathcal{E}(Na,[0,y[)$$

L'intérêt de ce résultat est que l'on peut l'itérer.

6.3. Cela permet tout d'abord de déterminer la répartition de U dans [0,1], avec les hypothèses faites précédemment :

Théorème. La suite auto-similaire U a une répartition asymptotique dans [0,1] si et seulement si E vérifie la condition (*) :

$$\forall a\in\mathcal{A},\quad \lim_{N\to\infty}\frac{1}{N}\eta(N,a)=\gamma_a \text{ existe}$$

la mesure de répartition étant alors la mesure auto-similaire $\mu^0=\mu^0(\Phi,\Psi)$ associée.

Si Φ est linéaire, μ^0 est pure et l'on a :

• si $\gamma_a=1$ pour un $a\in\mathcal{A}$, μ^0 est la mesure de Dirac δ_x, où $x=G_\Phi(a)$ est le point fixe unique de φ_a;

• si $\gamma_a=|I_a|$ pour tout $a\in\mathcal{A}$, μ^0 est la probabilité uniforme λ sur [0,1];

• si aucune des deux conditions précédentes n'est remplie, μ^0 est purement singulière.

Corollaire. U est équirépartie dans [0,1] si et seulement si

$$\begin{cases}\Phi \text{ est linéaire}\\ \text{et}\\ E \text{ vérifie } \forall a\in\mathcal{A}\ \lim_{N\to\infty}\frac{1}{N}\eta(N,a)=|I_a|.\end{cases}$$

En effet, il est possible de préciser exactement les couples (Φ,Ψ) tels que $\mu^{o}(\Phi,\Psi)=\lambda$.

6.4. On peut faire opérer le semi-groupe $\mathcal{A}^{*}$ sur $\mathbb{N}$, en posant :

$0a=0$ pour tout $a\in\mathcal{A}$;
$Na=(Na')a$ si a,a' sont deux mots finis tels que $a=a'.a$.

Si $a=a_1a_2\ldots a_n$ est un mot fini de longueur n, on pose pour $1\leq i\leq n$:

$a_i=a_1a_2\ldots a_i=[i]$ (préfixe de a de longueur i)

$[a_i]=\sum_{j=1}^{i}[a_j]$ (nombre d'indices j tels que φ_{a_j} est décroissante, $1\leq j\leq i$).

Théorème. Si $u_1=0$, et si $x=u_n$ de Φ-développement fini minimal, on a pour tout $N\geq n$:

$$\mathcal{E}(N,[0,u_n[)=\sum_{i=1}^{|a|}\Big((-1)^{[a_{i-1}]}\sum_{a\in\boxed{a_i}}\rho(N_{i},a)+(-1)^{[a_i]}\rho(Na_i,a_i)F_{\mu^{o}}(u_{na_i})\Big).$$

La condition $N\geq n$ n'est pas pas gênante ici, puisque seuls les points $u_1,u_2,\ldots,u_n$ interviennent dans l'écart $\mathcal{E}(N,I)$. De cette formule "explicite" pour les écarts, on déduit une estimation de la discrépance lorsque U est équirépartie.

On rappelle que la discrépance de la suite U est définie par (voir par exemple [10]) :

$$D_N^{*}(U)=\sup_{0<x\leq 1}|N(N,[0,x[)-xN|$$

L'équirépartition équivaut à $D_N^{*}=o(1)$, et on sait qu'il existe une constante absolue C_o telle que l'on ait pour toute suite à valeurs dans $[0,1]$:

$$\overline{\lim_{N\to\infty}}\frac{N\,D_N^{*}}{\mathrm{Log}N}\geq C_o$$

(on pourra trouver des estimations de C_o dans Bejian [1] et Faure [6]).

Corollaire. Si $u_1=0$, si Φ est linéaire, et si f est une fonction arithmétique croissante telle que $f(N)=o(N)$, et :

$$\forall N\in\mathbb{N}^*,\ \forall a\in\mathcal{A},\quad |\eta(N,a)-|I_a|.N| \le f(N)$$

Soit $\gamma = \max_{a\in\mathcal{A}} |I_a|$. Alors la discrépance de U vérifie :

$$ND_N^*(U) \le f(N)\frac{k \operatorname{Log} N}{2 \operatorname{Log}\frac{1}{\gamma} - \epsilon} + C_\epsilon) \qquad \text{pour tout } \epsilon>0.$$

Si en particulier, les restes $\eta(N,a)-|I_a|.N$ sont bornés, la discrépance $D_N^*(U)$ a le plus petit ordre de grandeur possible :

$$\overline{\lim_{N\to\infty}} \frac{N\, D_N^*(U)}{\operatorname{Log} N} \le \frac{k}{2 \operatorname{Log}\frac{1}{\gamma}} M$$

où M majore les restes.

Deux exemples :

	k	γ	M	$\frac{k\,M}{2\operatorname{Log}\gamma^{-1}}$	$\overline{\lim}$ (connue)
suites $\{n\xi\}$ $\xi=\frac{1+\sqrt{5}}{2}$	2	$\frac{\sqrt{5}-1}{2}$	$\frac{\sqrt{5}-1}{2}$	$\frac{\sqrt{5}-1}{2\operatorname{Log}\xi}$	$\frac{3}{20\operatorname{Log}\xi}$ (Dupain, [4])
suite de van der Corput	2	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2\operatorname{Log}2}$	$\frac{1}{3\operatorname{Log}2}$ (Haber, [8])

Les résultats obtenus ici ne sont donc pas optimaux ! il y a deux façons de les améliorer :

si U est auto-similaire associée à (Φ,E), elle est aussi auto-similaire associée aux itérés (non définis ici) de (Φ,E);
si les quantités $\eta(N,a)-|I_a|.N$ sont connues avec précision, on peut espérer obtenir exactement la valeur de la $\overline{\lim}$. Si E est périodique, ce travail a été fait par Faure [7].

BIBLIOGRAPHIE

[1] R. Bejian.- Sur certaines suites présentant une faible discrépance à l'origine, C.R.A.S. Paris t. 285 s.A (1977), 313-316.

[2] B.H. Bissinger.- A generalization of continued fractions, Bull. Amer. Math. Soc. 50 (1944), 868-876.

[3] J.-P. Borel.- Suites ayant de bonnes discrépances, Journées Mathématiques de Valenciennes (1982).

[4] Y. Dupain.- Discrépance à l'origine de la suite $(n\frac{1+\sqrt{5}}{2})$, Annales de l'Institut Fourier, t. 29 (1979), 81-106.

[5] C.I. Everett.- Representations for real numbers, Bull. Amer. Math. Soc. 52 (1946), 861-876.

[6] H. Faure.- Discrépances de suites associées à un système de numération (en dimension m), Bull. SMF 109 (1981), 143-182.

[7] H. Faure.- Thèse d'Etat, Université d'Aix Marseille I.

[8] S. Haber.- On a sequence of points of interest for numerical quadrature, J. Rest. Nat. Bur. Standards Sect. B 70 (1966), 127--296.

[9] K. Knopp.- Mengentheorische Behandlung einiger Probleme der diophantischen Approximationen und der transfiniten Wahrscheinlichkeiten, Math. Annalen 95 (1926), 409-426.

[10] L. Kuipers, H. Niederreiter.- Uniform distribution of sequences, J. Wiley & Sons, New York (1974).

[11] J.-F. Mela.- Exposé au colloque "des fractales en mathématiques et en physique", CIRM, janvier 1986.

[12] A. Renyi.- Calcul des probabilités, Dunod (1966).

[13] A. Renyi.- Representations for real numbers and their ergodic properties, Acta Math. Acad. Sci. Hongaricae (1957), t. 8, 477-493.

Convexité, moyennes, densité, lacunes

Georges GREKOS

1.- Contenu de l'exposé.

L'étude de la répartition des couples ordonnés constitués des deux densités asymptotiques supérieure et inférieure de chaque sous-ensemble d'une partie donnée de $\mathbb{N}$, a conduit à montrer [1] que ces couples-là forment un convexe de $\mathbb{R}^2$. Ici, nous présentons des développements récents de ce travail dans deux directions différentes.

En commun avec Bodo Volkmann [3], nous poursuivons l'étude de l'ensemble des couples des densités. Nous établissons des liens entre la forme du convexe obtenu et la répartition des lacunes dans la partie considérée de $\mathbb{N}$.

Ensuite, on démontre [2] un théorème analogue pour les moyennes limite des fonctions bornées non négatives définies sur $\mathbb{R}_+$.

2.- Densités des sous-ensembles.

On appelle convexe de type K toute partie C de $\mathbb{R}^2$ vérifiant les trois propriétés ci-dessous :

(i) Il existe deux nombres réels α et β, $0 \leq \beta \leq \alpha$, tels que les points $(0,0)$, $(\alpha,0)$ et (α,β) appartiennent à C, et que tout point (x,y) de C vérifie $0 \leq y \leq \beta$ et $y \leq x \leq \alpha$.

(ii) C est un convexe.

(iii) C est un fermé.

Soit A une partie de $\mathbb{N}$. Pour n appartenant à $\mathbb{N}^*$, on pose :

$$A(n) = \mathrm{Card}(A \cap [1,n]). \tag{1}$$

On définit les densités asymptotiques supérieure et inférieure de A :

$$\overline{d}A = \limsup_{n \to +\infty} n^{-1} A(n), \qquad \underline{d}A = \liminf_{n \to +\infty} n^{-1} A(n). \tag{2}$$

Si A est une partie donnée de $\mathbb{N}$, on pose :

$$S_d(A) = \{(\overline{d}B, \underline{d}B) \in \mathbb{R}^2 ;\ B \subset A\}. \tag{3}$$

Théorème 1 ([1]).

(I) Pour toute partie A de $\mathbb{N}$, $S_d(A)$ est un convexe de type K, avec $\alpha = \overline{d}A \leq 1$ et $\beta = \underline{d}A$.

(II) Etant donné un convexe C de type K avec $\alpha \leq 1$, il existe une partie A de $\mathbb{N}$ telle que $S_d(A) = C$.

3.- Densité et lacunes.

On définit la densité lacunaire d'une partie infinie $A = \{a_1 < a_2 < \ldots\}$ de $\mathbb{N}$ par la formule :

$$\lambda(A) = \limsup_{j \to +\infty} (a_{j+1}/a_j). \tag{4}$$

Alors $\lambda(A)$ appartient à $[1, +\infty]$.

D'autre part, la somme de deux parties A, B de $\mathbb{N}$ est

$$A+B = \{a+b ; a \in A, b \in B\}, \tag{5}$$

et on pose

$$hA = A+(h-1)A, \quad (h \text{ entier} \geq 2). \tag{6}$$

Voici les propriétés additives élémentaires de la densité lacunaire.

Proposition ([3]).

a) Soient A,B deux parties infinies de $\mathbb{N}$. Si $1 < \lambda(A) < +\infty$ et $1 < \lambda(B) < +\infty$, alors

$$\lambda(A+B) \leq \frac{\lambda(A)\,\lambda(B)-1}{\lambda(A)+\lambda(B)-2}.$$

b) Soit A une partie infinie de $\mathbb{N}$.

b.1) On a

$$\max\left(1, \frac{\lambda(A)}{2}\right) \leq \lambda(2A) \leq \frac{\lambda(A)+1}{2}.$$

b.2) La limite de $\lambda(hA)$, quand h tend vers $+\infty$, vaut soit 1, soit $+\infty$.

Relativement à une partie A de $\mathbb{N}$, on note f la fonction ayant pour graphe la partie supérieure de la frontière de $S_d(A)$

$$f : [0, \overline{d}A] \longrightarrow [0, \underline{d}A] : w \longmapsto z = f(w) = \max\{y ; (w,y) \in S_d(A)\}. \tag{7}$$

Cette fonction est concave et par conséquent elle possède en chaque point $w\in[0,\overline{d}A[$ une dérivée à droite, notée $f'_D(w)$.

Dans les énoncés qui suivent on convient de ce que $(+\infty)^{-1}=0$.

Théorème 2 ([3]).

(I) Soient A une partie infinie de $\mathbb{N}$ et f la fonction correspondante à $S_d(A)$. On a $f'_D(0) \leq (\lambda(A))^{-1}$.

(II) Soient C un convexe de type K de $\mathbb{R}^2$ avec $\alpha\leq 1$, f la fonction associée à C dans le sens de la formule (7), et λ_o un élément de $[1,+\infty]$ vérifiant $f'_D(0)\leq\lambda_o^{-1}$. Alors, il existe une partie infinie A de $\mathbb{N}$ telle que $S_d(A)=C$ et $\lambda(A)=\lambda_o$.

Soient A une partie de $\mathbb{N}$ avec $\overline{d}A>0$, et ϵ un réel appartenant à $[0,\overline{d}A[$. Pour n entier positif, on pose

$$\ell(A,\epsilon,n)= \max \{k\in\mathbb{N};\ A(n+q)-A(n)\leq\epsilon q,\ (q=0,1,\ldots,k)\},$$

et on définit

$$\lambda(A,\epsilon)= \limsup_{n\to+\infty} (1+n^{-1}\ell(A,\epsilon,n)). \tag{8}$$

On constate facilement que $\lambda(A,0)$ est égal à la densité lacunaire $\lambda(A)$. Pour un ensemble A donné, $\lambda(A,\epsilon)$ est une fonction croissante de ϵ. Ceci permet de considérer la $\lim_{\substack{\epsilon\to 0\\ \epsilon>0}} \lambda(A,\epsilon)$, notée $\lambda(A,0+0)$, qui est au moins égale à $\lambda(A)$.

Théorème 3 ([3]).- Soient A une partie de $\mathbb{N}$ avec $\overline{d}A>0$, et f la fonction associée à $S_d(A)$. Alors $f'_D(0)= (\lambda(A,0+0))^{-1}$.

4.- Moyennes limite et convexité.

On désigne par F l'ensemble des fonctions définies sur $\mathbb{R}_+$, à valeurs réelles non négatives, bornées et dont l'intégrale - au sens de Lebesgue - existe sur tout intervalle $[0,t]$. On munit F de la relation d'ordre partiel suivante : soient a, b deux fonctions de F; on pose $b\leq a$ si et seulement si, quel que soit $t\geq 0$, on a $b(t)\leq a(t)$.

A une fonction a de F, on associe le couple $L(a)= (\overline{L}(a),\underline{L}(a))\in\mathbb{R}^2$, où

(9) $$\overline{L}(a)= \limsup_{t\to+\infty} t^{-1}\int_0^t a(s)ds,\quad \underline{L}(a)= \liminf_{t\to+\infty} t^{-1}\int_0^t a(s)ds,$$

et on pose

$$S_L(a)= \{L(b)\in\mathbb{R}^2;\ b\in F,\ b\leq a\}$$

Théorème 4 ([2]).

(I) Si $a\in F$, alors $S_L(a)$ est un convexe de type K avec $\alpha= \overline{L}(a)$ et $\beta= \underline{L}(a)$.

(II) Etant donné un convexe C de type K, il existe une fonction a de F telle que $S_L(a)= C$.

Remarques.

1) L'obtention d'un convexe $S_L(a)$ non trivial de type K est bien due au fait que l'on considère les moyennes. Car si l'on pose

$$L_1(a)= (\limsup_{t\to+\infty} a(t),\ \liminf_{t\to+\infty} a(t)),$$

alors, pour toute fonction a de F, l'ensemble $S_1(a)= \{L_1(b);b\in F,b\leq a\}$ est un trapèze.

2) Si la fonction $a\in F$ est la fonction caractéristique d'une partie A de $\mathbb{N}$, à savoir $a(t)=1$, si $[t]\in A$, $a(t)=0$, sinon, alors les limites $\overline{L}(a)$ et $\underline{L}(a)$, relation (9), deviennent les densités asymptotiques $\overline{d}A$ et $\underline{d}A$, relation (2), respectivement. Ainsi le théorème 1 est un cas particulier du théorème 4.

3) L'affirmation (II) du théorème 4 s'obtient comme corollaire de la deuxième partie du théorème 1, compte tenu du fait que si $\mu\in\mathbb{R}_+$, alors $S_L(\mu a)= \mu\, S_L(a)$.

4) Dans la démonstration [2] de la partie (I) du théorème 4, la preuve de la convexité de $S_L(a)$ suit les idées de [1]. Pour établir la fermeture de $S_L(a)$, nous utilisons un procédé où l'on s'approche du point sur la frontière par des points de l'intérieur, et on construit par blocs une fonction qui correspond au point limite.

BIBLIOGRAPHIE

[1] G. Grekos.- Répartition des densités des sous-suites d'une suite d'entiers, J. Number Theory 10 (1978), 177-191.

[2] G. Grekos.- Moyennes limite et convexité, Manuscrit.

[3] G. Grekos et B. Volkmann.- On densities and gaps, J. Number Theory, 26 (1987), 129-148.

Georges GREKOS
Université de Grenoble I
Institut Fourier (Laboratoire associé au C.N.R.S.)
B.P. 74
38402 SAINT-MARTIN-D'HERES CEDEX
FRANCE

Van der Corput Bounds for the Dedekind Zeta-Function

D.R. Heath-Brown

There is a long standing programme to generalize results already known for the Riemann Zeta-function to the Dedekind Zeta-function. This has been carried out successfully for the analytic continuation and functional equation, and for the zero-free region. On the other hand, for example, we do not know in general whether a finite proportion of the zeros of $\zeta_K(s)$ lie on the critical line $\mathrm{Re}(s)=\frac{1}{2}$.

Here we consider the function

$$\mu_K(\sigma)= \inf \{\xi : \zeta_K(\sigma+it) \ll t^{\xi}, \ t\geq 2\}.$$

It is easily shown that $\mu_K(\sigma)$ is non-negative, continuous and downwards convex. Moreover, we have $\mu_K(\sigma)=0$ for $\sigma\geq 1$, and $\mu_K(\sigma)= (\frac{1}{2} - \sigma)n$ for $\sigma\leq 0$, where n is the degree $[K:\mathbb{Q}]$. It follows that $\mu_K(\frac{1}{2}) \leq \frac{n}{4}$. For the Riemann Zeta-function (for which $K=\mathbb{Q}$, $n=1$) We have the better result $\mu_{\mathbb{Q}}(\frac{1}{2}) \leq \frac{1}{6}$, which may be proved by van der Corput's method. It is this estimate which we wish to generalize. We have :

Theorem (Heath-Brown [1]). *If K is an algebraic number field of degree n over $\mathbb{Q}$, then $\mu_K(\frac{1}{2}) \leq \frac{n}{6}$.*

This is easily proved if K is an abelian extension of $\mathbb{Q}$, since $\zeta_K(s)$ then factorizes into a product of Dirichlet L-functions. In general, however, an n-dimensional version of van der Corput's method is required. Those forms of the method currently available (Srinivasan [2], for example) are extremely complicated. A considerable simplification is made possible by introducing smooth weights.

As an example we consider the well-known bound

$$(*) \qquad \int_N^{2N} e^{2\pi i g(x)} dx \ll \lambda^{-\frac{1}{2}}, \quad \lambda= \inf |g''(x)|.$$

To formulate an n-dimensional version of this we suppose that g is defined on $[N,2N]^n$. Let λ be defined in terms of the Hessian of g by

$$\lambda = \inf \left| \det\left(\frac{\partial^2 g}{\partial x_i \partial x_j}\right)_{1 \le i,j \le n} \right|.$$

Suppose moreover that the partial derivatives of g satisfy

$$\frac{\partial^{r_1+\ldots+r_n} g}{\partial x_1^{r_1} \ldots \partial x_n^{r_n}} \ll_{r_1,\ldots,r_n} \lambda^{1/n} N^{2-r_1-\ldots-r_n}$$

for all non-negative integers r_i for which

$$\sum_1^n r_i \ge 2.$$

Now let $w(x_1,\ldots,x_n)$ be a smooth weight function with support in $[N,2N]^n$, and satisfying

$$\frac{\partial^{r_1+\ldots+r_n} w}{\partial x_1^{r_1} \ldots \partial x_n^{r_n}} \ll_{r_1,\ldots,r_n} N^{-r_1-\ldots-r_n}$$

for all non-negative integers r_i. The analogue of (*) is then the bound

$$\int_N^{2N} \ldots \int_N^{2N} w(x_1,\ldots,x_n) e^{2\pi i g(x_1,\ldots,x_n)} dx_1 \ldots dx_n \ll \lambda^{\frac{1}{2}+\epsilon},$$

for any $\epsilon > 0$.

BIBLIOGRAPHY

[1] D.R. Heath-Brown.- The growth rate of the Dedekind Zeta-function on the critical line, Acta Arith., 49, to appear.

[2] B.R. Srinivasan.- The lattice point problem of many dimensional hyperboloids. III, Math. Ann., 160, (1965), 280-311.

D.R. Heath-Brown
Magdalen College
Oxford OX1 4AU
ENGLAND

LES ZEROS DE LA FONCTION ZETA DE RIEMANN SUR LA DROITE CRITIQUE

Aleksandar IVIĆ

Soient $0<\gamma_1\leq\gamma_2\leq\ldots$ les ordonnées des zéros de la fonction $\zeta(s)$ sur la "droite critique" $\sigma=\mathrm{Res}=\frac{1}{2}$ avec $\mathrm{Im}s>0$. On s'intéresse ici à deux problèmes concernant la distribution des différences $\gamma_{n+1}-\gamma_n$.

Problème 1 : Estimer

$$\theta=\inf\{c\geq 0 : \gamma_{n+1}-\gamma_n \ll \gamma_n^c\}.$$

Problème 2 : Estimer le nombre de "grandes" différences $\gamma_{n+1}-\gamma_n$ entre les zéros consécutifs de $\zeta(\frac{1}{2}+it)$ pour $\gamma_n\leq T$.

Dans le second problème, bien sûr, il faut définir précisément les grandes différences $\gamma_{n+1}-\gamma_n$. Si on suppose que l'hypothèse de Riemann est vraie, alors il en découle (voir E.C. Titchmarsh [11], th. 14.13) que

$$\gamma_{n+1}-\gamma_n \ll \frac{1}{\log\log\gamma_n},$$

donc nous cherchons les réponses aux Problèmes 1 et 2 sans aucune hypothèse.

Dans le Problème 1, les majorations $\theta\leq\frac{1}{4}$, $\theta\leq\frac{1}{6}$, $\theta\leq\frac{5}{32}=0,15625$ et $\theta\leq 0,15594583\ldots$ sont dûes à Hardy-Littlewood [2], R. Balasubramanian [1] et J. Moser [9] (indépendemment), A.A. Karacuba [8], et A. Ivić [4], respectivement. Pour obtenir la dernière majoration de θ, on étudie le comportement dans les intervalles courts de la fonction

$$Z(t)=\chi^{-\frac{1}{2}}(\tfrac{1}{2}+it)\zeta(\tfrac{1}{2}+it),\ \chi(s)=\frac{\zeta(s)}{\zeta(1-s)}=2^s\pi^{s-1}\sin(\tfrac{\pi s}{2})\Gamma(1-s),$$

de sorte que $|Z(t)|=|\zeta(\frac{1}{2}+it)|$ et que les valeurs de $Z(t)$ sont réelles pour t réel. Les zéros de $\zeta(\frac{1}{2}+it)$ sont alors précisément les

zéros de $Z(t)$, et on note que, si $Z(u)\neq 0$ pour $t - \frac{1}{4}V \leq u \leq t + \frac{1}{4}V$, alors $I_1(t) = |I_2(t)|$, où

$$I_1(t) = \int_{t-\frac{1}{4}V}^{t+\frac{1}{4}V} |Z(u)|\, e^{-(t-u)^2 G^{-2}}\, du,$$

$$I_2(t) = \int_{t-\frac{1}{4}V}^{t+\frac{1}{4}V} Z(u)\, e^{-t-u)^2 G^{-2}}\, du.$$

Tenant compte que $|Z(t)| = |\zeta(\frac{1}{2} + it)|$ on peut minorer $I_1(t)$, et en utilisant l'intégrale classique

$$\int_{-\infty}^{\infty} \exp\,(Ax - Bx^2)\, dx = (\pi/B)^{1/2} \exp\,(\frac{A^2}{4B}) \quad (\mathrm{Re}B > 0)$$

et l'équation fonctionnelle approchée pour $\zeta(s)$, on peut évaluer et majorer $I_2(t)$ (voir [4], [5] Ch. 10, [6], [7]). Les résultats sont contenus dans le lemme suivant :

Lemme. Supposons que $T \geq T_0$, $\frac{1}{2}T \leq t \leq T$, $500 \log T \leq V \leq T^{1/6}$, $G = VL^{-1}$, $L = 10\sqrt{\log T}$, $X = T^{1/2} V^{-1} L^2$. Alors $I_1(t) \gg G$ et

$$(1) \qquad I_2(t) \ll \int_{-\frac{1}{4}V}^{\frac{1}{4}V} \Big| \sum_{-X \leq n \leq \sqrt{\frac{t+u}{2\pi}}} n^{-1/2-it-iu} \Big|\, e^{-u^2 G^{-2}}\, du + 1.$$

On peut majorer la somme dans (1) et, pour V suffisament petit, en déduire une contradiction avec $I_1(t) = |I_2(t)|$, ce qui nous fournit une valeur pour θ. La majoration triviale donne immédiatement le résultat classique $\theta \leq \frac{1}{4}$ de Hardy-Littlewood, mais une estimation beaucoup plus précise a été faite ([4], [5], Ch. 10 et [6]) avec la théorie des paires d'exposants de van der Corput ([5], Ch. 2). Le résultat est

Théorème 1. Si (κ,λ) est une paire d'exposants, alors

$$(2) \qquad \theta \leq \frac{\kappa+\lambda}{4(\kappa+\lambda)+2}.$$

Le résultat classique de R.A. Rankin [10] que

$$(3) \qquad \inf(\kappa+\lambda) \leq 0{,}8290213565\ldots$$

donne alors que $\theta \leq 0{,}15594583\ldots$. Très récemment M.N. Huxley et N. Watt [3] ont établi que

$$(\kappa_0,\lambda_0)= (\frac{9}{56} + \epsilon, \frac{37}{56} + \epsilon)$$

est une paire d'exposants avec $\kappa_0+\lambda_0= \frac{23}{28} + \epsilon= 0{,}821428571\ldots+\epsilon$, ce qui est plus petit que la constante dans (1). La preuve du Théorème 1 montre en fait qu'on doit avoir $(\kappa,\lambda)= BA(p,q)$, où (p,q) est une paire d'exposants arbitraires et

$$A(p,q)= (\frac{p}{2p+2} , \frac{1}{2} + \frac{q}{2p+2}), \quad B(p,q)= (q - \frac{1}{2}, p + \frac{1}{2})$$

sont les procédés A et B bien connus de la théorie des paires d'exposants. Donc on ne peut pas pour le moment utiliser la nouvelle paire (κ_0,λ_0) directement dans (2). Mais si on forme la paire

$$(\kappa_1,\lambda_1)= B\ A\ B\ A^3\ B\ A^2\ B\ A(\kappa_0,\lambda_0)= (\frac{612}{3227} + \epsilon, \frac{1179}{1844} + \epsilon),$$

alors (κ_1,λ_1) est de la forme $B\ A(p,q)$ est

$$\kappa_1 + \lambda_1= \frac{10701}{12908} + \epsilon= 0{,}82902076231\ldots + \epsilon, \tag{4}$$

ce qui améliore la valeur de Rankin dans (3). En rapportant (4) dans (2) il s'ensuit que

$$\theta \leq \frac{\kappa_1+\lambda_1}{4(\kappa_1+\lambda_1)+2} = 0{,}15594578839988\ldots + \epsilon.$$

Je tiens à remercier J.-P. Massias (Limoges), qui a calculé (κ_1,λ_1).

En ce qui concerne le Problème 2, on peut utiliser aussi le lemme 1 pour l'étude des "grandes" différences de $\gamma_{n+1}- \gamma_n$. Les résultats, dont une démonstration détaillée vient d'être donnée par A. Ivić et M. Jutila [7], sont contenus dans le

Théorème 2. Soit R le nombre de γ_n pour lesquels $0<\gamma_n\leq T$ et $\gamma_{n+1}-\gamma_n\geq V$ pour $V>0$ arbitraire. Alors on a uniformément

$$R \ll T\ V^{-2}\ \log T \tag{5}$$

et

$$R \ll T\ V^{-3}\ \log^5 T. \tag{6}$$

Une estimation triviale pour R est $R \ll T\, V^{-1}$. On voit que (6) améliore (5) pour $V \gg \log^4 T$. Un corollaire simple de (6) est

$$(7) \qquad \sum_{0<\gamma_n\leq T} (\gamma_{n+1}-\gamma_n)^3 \ll T \log^6 T.$$

En effet, la somme dans (7) est $O(T)$ plus

$$\sum_{V=T^{1/6}2^{-j},\, j=1}^{[\log T]} \sum_{\gamma_n\leq T,\, V\leq\gamma_{n+1}-\gamma_n<2V} (\gamma_{n+1}-\gamma_n)^3$$

$$\leq 8V^3 \sum_{V=T^{1/6}2^{-j},\, j=1}^{[\log T]} \sum_{\gamma_n\leq T,\, \gamma_{n+1}-\gamma_n\geq V} 1$$

$$\ll V^3 \sum_{V=T^{1/6}2^{-j},\, j=1}^{[\log T]} T\, V^{-3} \log^5 T \ll T \log^6 T,$$

où nous avons utilisé (6) et le fait qu'on a $O(\log T)$ valeurs pour V.

En concluant, on peut remarquer qu'on peut généraliser les Problèmes 1 et 2 et étudier les zéros d'autres fonctions similaires à la fonction $\zeta(s)$. Signalons seulement qu'une possibilité est l'étude des zéros consécutifs de $Z^{(k)}(t)$ pour k un nombre naturel fixé (voir [6]), et qu'une autre possibilité est l'étude des zéros des fonctions zéta associées aux coefficients des formes modulaires. Cette dernière étude vient d'être effectuée en détail par A. Ivić et M. Jutila [7], et un résultat analogue au Th. 2 y est démontré.

BIBLIOGRAPHIE

[1] R. Balasubramanian.- An improvement of a theorem of Titchmarsh on the mean square of $|\zeta(\frac{1}{2} + it)|$, Proc. London Math. Soc. 35 (1978), 540-576.

[2] G.H. Hardy and J.E. Littlewood.- Contributions to the theory of the Riemann zeta-function and the distribution of primes, Acta Math. 41 (1918), 119-196.

[3] M.N. Huxley and N. Watt.- Exponential sums and the Riemann zeta-function, subm. J. London Math. Soc.

[4] A. Ivić.- Topics in recent zeta-function theory, Publ. Math. d'Orsay, Université Paris-Sud, Orsay, 1983.

[5] A. Ivić.- The Riemann zeta-function, John Wiley & Sons, New York, 1985.

[6] A. Ivić.- On a problem connected with zeros of $\zeta(s)$ on the critical line, Monatsh. Math. (in press).

[7] A. Ivić and M. Jutila.- Gaps between consecutive zeros of the Riemann zeta-function on the critical line, Monatsh. Math. (in press).

[8] A.A. Karacuba.- On the distance between consecutive zeros of the Riemann zeta-function on the critical line (Russian), Trudy Mat. Inst. Steklova AN SSSR 157 (1981), 49-63.

[9] J. Moser.- On a theorem of Hardy-Littlewood in the theory of the Riemann zeta-function (Russian), Acta Arith. 31 (1976), 45-51 and ibid. 35 (1979), 403-404.

[10] R.A. Rankin.- Van der Corput's method and the theory of exponent pairs, Quart. J. Math. (Oxford) 6 (2) (1955), 147-153.

[11] E.C. Titchmarsh.- The theory of the Riemann zeta-function, Oxford University Press, Oxford, 1951.

A. IVIĆ
Katedra Matematike RGF-a
Universiteta u Beogradu
11000 Beograd, Đušina 7
JUGOSLAVIJA

Calculs explicites de constantes de Lehmer

Michel Langevin

Résumé : Soit V un voisinage d'un point du cercle-unité. On étudie comment expliciter une constante C(V) strictement supérieure à 1 telle que tout entier algébrique $x_1 \neq 0$, non racine de l'unité, dont aucun des conjugués $x_1, x_2, \ldots, x_d$ ne se trouve dans V, vérifie

$$(\sup(1, |x_1|)\sup(1, |x_2|)\ldots\sup(1, |x_d|))^{1/d} \geq C(V)$$

Mots-clés : Problème de Lehmer, mesure de Mahler, diamètre transfini.

Sommaire :

§ 1.- Introduction
§ 2.- L'idée du calcul
§ 3.- La méthode de Fekete et Szegö
§ 4.- Existence de C(V)
§ 5.- Calculs explicites de constantes de Lehmer C(V)
§ 6.- Bibliographie
§ 7.- Dernière minute (juin-juillet 1987)

§1.- Introduction.

Soit

$$P(z) = a_o(z-x_1)(z-x_2)\ldots(z-x_d) = a_o z^d + a_1 z^{d-1} + \ldots + a_d$$

un polynôme à coefficients complexes; la mesure de Mahler M(P) est, par définition :

$$M(P) = |a_o| \prod_i \sup(1, |x_i|) \tag{1}$$

(lorsque P est à coefficients réels, en notant k le nombre de racines de module strictement supérieur à 1 et en rangeant ces racines par modules décroissant, l'écriture ci-dessus devient, quitte à changer a_o en $-a_o$, $M(P) = a_o x_1 x_2 \ldots x_k$).

Cette mesure de Mahler est une "hauteur" au sens habituel qu'on donne à ce terme. On le voit aisément en écrivant $M(P)$ sous la forme

$$(2) \qquad M(P)= \exp\left(\int_0^1 \log|P(e^{2i\pi t})|dt\right)$$

(lorsque P ne s'annule pas sur le cercle-unité, le calcul ci-dessous donne une preuve facile de l'identité de (1) et de (2) (plus généralement, on peut appliquer la formule de Jensen) :

$$\exp\left(\int_0^1 \log|P(e^{2i\pi t})|dt\right)= \lim_{n\to+\infty}|\mathrm{Res}(P,z^n-1)|^{1/n}= |a_0| \prod_i \sup(1,|x_i|)\ ).$$

En effet, la forme (2) montre que $M(P)$ est une moyenne géométrique, i.e. pour la norme "L_0", qu'on peut donc majorer par la moyenne quadratique; globalement, il vient :

$$M(P)\leq M_2(P)=$$

$$= \left(\int_0^1 |P(e^{2i\pi t})|^2 dt\right)^{\frac{1}{2}}= \left(\sum_i |a_i|^2\right)^{\frac{1}{2}}\leq \sup_{|z|=1}|P(z)| \leq \sum_i |a_i| \leq \prod_i (1+|x_i|) \leq 2^d M(P).$$

Soit $x\neq 0$ un nombre algébrique. Par définition, la mesure de Mahler de x est celle de son polynôme minimal. Par exemple, la mesure d'une racine de 1 est 1. La réciproque est vraie ("théorème de Kronecker") : soit x tel que $M(x)=1$; c'est donc un entier algébrique dont tous les conjugués $x_1= x, x_2, \ldots, x_d$ sont de module 1; si x n'est pas une racine de l'unité, on a, pour tout entier $n>0$,

$$|x^n-1| \geq 1/\left(\prod_{i\geq 2}|x_i^n-1|\right) \geq 2^{1-d}$$

ce qui prouve que le groupe multiplicatif engendré par x est discret et ce qui contredit donc l'hypothèse.

Le théorème de Kronecker suggère la question naturelle (où $R(1)$ désigne l'ensemble des racines de 1) :

(Problème de Lehmer) A-t-on bien $\inf_{x\notin R(1)} M(x) > 1$?

Ce problème est maintenant ouvert depuis plus d'un demi-siècle. Nous ne reviendrons pas sur les progrès réalisés en 1971 par P. Blanksby et H.L. Montgomery d'une part et par C. Smyth d'autre part, pas plus que sur ceux réalisés plus récemment, suite aux idées nouvelles introduites par Dobrowolski, par Cantor-Straus puis R. Louboutin; nous nous bornerons à rappeler l'inégalité prouvée par ce dernier auteur : si $x\notin R(1)$ est un nombre algébrique de degré d assez grand, on a

$$M(x) > 1 + (9/4 - \epsilon)(\mathrm{loglog}d/\log d)^3 \qquad \text{(cf. (Lo))}.$$

Pour une vision plus générale des problèmes de théorie des nombres - et d'autres spécialités - liés à la mesure de Mahler, on pourra consulter l'exposé de D. Boyd (mai 85, Sydney) (cf. (Bo)).

Soit V un voisinage donné d'un point du cercle-unité. Le but de ce travail est d'étudier les mesures de Mahler des entiers algébriques x ($x\neq 0$, $x\notin R(1)$) dont aucun (ou du moins presqu'aucun...) ne se trouve dans V. Pour de tels x, on établira l'inégalité :

Théorème 1. $M(x) > C(V)^{d^\circ x}$ (C(V) : constante >1 explicitable en fonction de V)

et on montrera comment calculer explicitement de telles cons-tantes, dénommées dans le titre "constantes de Lehmer".

Un exemple où l'hypothèse ci-dessus sur x est vérifiée est celui des entiers algébriques totalement réels où l'on a démontré :

$$M(x) \geq ((1+5^{1/2})/2)^{d/2}$$

(ce résultat est contenu dans un travail de A. Schinzel de 1973 (cf. (Sc)) et est le premier pas d'une description des valeurs prises par les $M(x)^{1/d}$ quand x parcourt l'ensemble des entiers totalement réels (cf. (Sm)).

Les résultats qui suivent sont ceux de deux Notes de l'auteur; la première (cf. (L1)) où est établie l'existence de C(V), la seconde où l'on explique comment calculer C(V) et retrouver ainsi très simplement le facteur $(1+5^{1/2})/2$ obtenu par Schinzel (cf. (L2)). On décrira de plus ici le rôle essentiel - en filigrane dans ces Notes - joué par la notion de "diamètre transfini". A ce développement près, ces travaux ainsi que ceux de R. Louboutin (lesquels relèvent aussi des "méthodes transcendantes" par la forme des démonstrations) ont déjà fait l'objet d'un exposé (ICM 86, Berkeley).

§2.- L'idée du calcul.

Soient :

Q un polynôme réciproque de degré q (i.e. $Q(z)= z^q Q(1/z)$),

$R(z)= |Q(z)|^{1/q}$,

$r= \sup_{|z|\leq 1} R(z)$,

t un paramètre positif ou nul,

s un paramètre vérifiant $0<s<1$,

X l'ensemble des points z du disque-unité vérifiant $|z|^t R(z)\leq s$,

P un polynôme $(P(z)= a_o z^d+\ldots+a_d= a_o(z-x_1)\ldots(z-x_d))$,

d' le nombre des zéros x_i de P vérifiant $x_i \in X$ ou $1/x_i \in X$.

Avec ces notations, on va montrer comment majorer la mesure de P à partir d'une minoration du résultant généralisé :

$$\prod_{1\leq i\leq d} |x_i|^t R(x_i)$$

Pour cela, on partage l'ensemble des indices en deux parties :

- ceux des i pour lesquels $|x_i|\leq 1$ et on majore alors $|x_i|^t R(x_i)$ par s si $x_i \in X$ et par r si $x_i \notin X$,
- ceux des i pour lesquels $|x_i|>1$, on écrit alors

$$|x_i|^t R(x_i)= |x_i|^{1+2t}.(1/|x_i|)^t R(1/x_i)$$

ce qui permet de se ramener à l'intérieur du disque-unité et donc de majorer $(1/|x_i|)^t R(1/x_i)$ comme ci-dessus suivant que $1/x_i$ se trouve dans X ou non. Globalement, il vient donc :

$$(3) \qquad \prod_{1\leq i\leq d} |x_i^t R(x_i)| \leq (M(P)/|a_o|)^{1+2t} s^{d'} r^{d-d'}$$

Exemple : On suppose les polynômes P et Q unitaires à coefficients entiers et sans zéro commun. Le membre de gauche de (3) est alors minoré par 1 d'où

$$M(P) \geq s^{-d'} r^{d'-d}.$$

Pour que ce résultat soit non trivial, il faut que X ne soit pas trop petit. Soit V un voisinage d'un point du cercle-unité. On verra aux paragraphes suivants qu'il existe des fonctions de la forme

$$z \longrightarrow |z|^t R(z) \qquad \text{(pour t assez grand)}$$

majorées strictement par 1 sur le complémentaire X de V dans le disque-unité de sorte qu'il existe des valeurs de s où le calcul ci-dessus s'applique.

On décrit maintenant brièvement le rôle du paramètre t; supposons pour simplifier qu'aucun zéro de P ne se trouve dans V; quitte à réduire V, on peut supposer que, pour $i= 1,2,\ldots,d$, $x_i \notin V$ et $1/x_i \notin V$; le calcul précédent montre alors que :

$$M(P) \geq ((\sup_{z\in X} |z|^t R(z))^{-1/(1+2t)})^d \qquad \text{(pour } t \text{ assez grand)}$$

et le problème est alors celui d'un choix optimal du paramètre t.

Ce qui précède explique le plan des paragraphes ultérieurs :

Au §3, on établit une condition nécessaire et suffisante sur une partie compacte E du plan complexe pour qu'il existe un polynôme unitaire à coefficients entiers Q vérifiant

$$\sup_{z\in E} |Q(z)| < 1;$$

Au §4, on verra comment déduire du §3 l'existence de fonctions

$$z \longmapsto |z|^t R(z)$$

ce qui permettra d'établir l'existence de la constante $C(V)$.

Au §5, on décrira sur un exemple comment choisir t.

§3.- La méthode de Fekete et Szegö.

Soit X une partie compacte du plan complexe. On appelle "diamètre transfini" $t(X)$ (ou seulement t quand il n'y a pas d'ambiguïté) la borne inférieure des $(\sup_{z\in X} |U(z)|)^{1/d^oU}$ lorsque U décrit l'ensemble des polynômes unitaires à coefficients complexes dont les zéros appartiennent à X (on pourrait en fait supprimer cette contrainte ou bien l'aggraver en imposant aux zéros de U de se trouver sur le "bord extérieur" de X sans changer le résultat, mais c'est la définition ci-dessus qui sera la plus commode dans le contexte de ce travail; toutefois, les autres définitions classiques du diamètre transfini - comme limite des bornes supérieures des moyennes géométriques des distances séparant les points de X ou comme coefficient directeur dans la représentation conforme de l'extérieur du disque-unité sur l'extérieur de X ... - ainsi que les remarques précédentes ne doivent pas être perdues de vue même si elles n'interviennent que ponctuellement dans la suite).

Comme l'ont observé Fekete et Szegö, une idée de Kakeya suffit pour établir le lemme suivant :

Lemme 1. On suppose $t(X)<1$ et X symétrique par rapport à l'axe réel (i.e. $X=\overline{X}$). Il existe alors un polynôme U unitaire à coefficients entiers vérifiant

$$\|U\|_X = \sup_{z\ X} |U(z)| < 1$$

(réciproquement, l'existence d'un tel polynôme pour une partie X du plan montre que $t(X\cup\overline{X})<1$).

On va à la fois établir ce lemme et prouver un résultat plus général (sur l'approximation par des fonctions polynômes à coefficients entiers) qui prendra des formes bien différentes suivant la position de $t(X)$ par rapport à 1. Un deuxième lemme nous sera utile :

Lemme 2. On suppose $X=\overline{X}$ et soit t' un réel vérifiant : $t'>t(X)$. Il existe alors une constante $k>0$ et une suite (F_d) de polynômes unitaires à coefficients réels de degré d telle que :

$$\|F_d\|_X < k.t'^d.$$

Démonstration du lemme 2 : La définition du diamètre transfini $t(X)$ et l'hypothèse $X=\overline{X}$ montrent l'existence d'un polynôme F unitaire à coefficients réels vérifiant $\|F\|_X<t'^f$ où f désigne le degré de F. Les polynômes F_d définis par : $F_d(z)= z^r F(z)^q$ avec $d= fq+r$, $0\leq r<f$ répondent à la question.

On passe des polynômes F_d aux polynômes à coefficients entiers en itérant le procédé suivant :

- Il existe clairement un réel a_1 qu'on peut supposer, ou bien positif (ou nul) et strictement majoré par 1 (c'est le choix fait dans la suite) ou bien de valeur absolue au plus 1/2, tel que les coefficients de z^d et de z^{d-1} de $F_d+a_1F_{d-1}$ soient entiers;
- de même, on voit qu'il existe a_2 $(0\leq a_2<1)$ tels que les coefficients de z^d, z^{d-1}, z^{d-2} du polynôme $F_d+ a_1F_{d-1}+ a_2F_{d-2}$ soient entiers ...

On voit donc qu'on peut ainsi obtenir par récurrence une combinaison linéaire

$$F_d+ a_1F_{d-1}+\ldots+a_d$$

telle que le polynôme obtenu ci-dessus soit à coefficients entiers.

Supposons alors que X soit de diamètre transfini au moins 1, le paramètre t' est alors strictement supérieur à 1 et on a donc :

$$\|F_d+a_1F_{d-1}+\ldots+a_d\|_X^{1/d} < (k(t'^{d+1}-1)/(t'-1))^{1/d}$$

ce qui établit l'égalité suivante :

Si $X=\overline{X}$ et $t(X)\geq 1$, on a : $t(X)= \inf_{U\in\mathbb{Z}[z]} \|U\|_X^{1/d\circ U}$.

On suppose maintenant X de diamètre transfini strictement inférieur à 1. Soit t' vérifiant $t(X)<t'<1$. Soit d_o un entier vérifiant

$$k\, t'^{d_o}/(1-t') < 1/3$$

où k désigne la constante associée par le lemme 2 à X et t'.

Pour tout entier $d \geq d_o$, soit G_d la combinaison linéaire

$$G_d = F_d + a_1 F_{d-1} + \ldots + a_{d_o} F_{d_o} \qquad (\text{avec } 0 \leq a_1, a_2, \ldots, a_{d_o} < 1)$$

construite par le procédé déjà employé ci-dessus à partir des polynômes F_d fournis par le lemme 2. Chaque polynôme G_d peut donc être écrit

$$G_d = H_d + H'_d$$

où H_d est unitaire, à coefficients entiers, de degré d,
où H'_d est à coefficients réels positifs majorés par 1 et de degré au plus d_o.

D'autre part, l'hypothèse faite sur d_o montre que les normes sur X des G_d sont toutes strictement majorées par $1/3$ et le principe des tiroirs montre que, les polynômes H'_d étant de hauteur et de degré bornés, l'on a pour un couple (d_1, d_2) d'entiers distincts convenables :

$$\|H'_{d_1} - H'_{d_2}\|_X < 1/3.$$

Globalement, on obtient donc, pour le polynôme à coefficients entiers $H_{d_1} - H_{d_2}$,

$$\|H_{d_1} - H_{d_2}\|_X \leq \|G_{d_1}\|_X + \|G_{d_2}\|_X + \|H'_{d_1} - H'_{d_2}\|_X < 1/3 + 1/3 + 1/3 = 1,$$

ce qui établit le lemme 1.

§4.- Existence de C(V).

Soient u un nombre complexe de module 1 et V un voisinage de u. Pour $\epsilon>0$ assez petit, le disque admettant pour diamètre le segment joignant les points $u/(1+\epsilon)$ et $u(1+\epsilon)$ est inclus dans V et est stable par l'inversion de pôle 0 et de puissance 1. Par conséquent, si l'on désigne désormais par V la réunion de ce disque et de son symétrique par rapport à l'axe réel, on voit que les conditions :

$$v\in V \quad \text{et} \quad 1/v\in V$$

sont équivalentes et que le choix d'un tel V n'amène aucune perte de généralité pour le théorème à démontrer, c'est-à-dire l'existence de C(V) (cf. §1).

Soit X le disque-unité privé des points de V. Pour pouvoir appliquer le lemme 1 du §3, on va établir l'inégalité :

$$t(X) < 1.$$

Deux procédés sont à notre disposition :

1°) Appliquer à l'enveloppe convexe Y de X l'énoncé suivant qui précise la classique inégalité isopérimétrique (cf. (L.3)) :

Lemme 3. Soit Y une partie convexe compacte du plan d'aire A, de périmètre L, de diamètre transfini t(Y), alors :

$$(A/\pi)^{1/2} \leq t(Y) \leq L/2\pi.$$

L'inégalité cherchée se déduit clairement du lemme. L'avantage de ce procédé est de fournir un encadrement, ce qui permettra de situer les limites de la méthode pour le calcul de C(V).

2°) Considérer l'ensemble X' des points de X de module 1; établir le lemme 4.

Lemme 4. Le diamètre transfini d'un arc du cercle-unité de longueur ℓ est égal à $\sin(\ell/4)$ (démonstration à la fin de ce paragraphe);

- construire grâce au lemme 1 un polynôme unitaire Q à coefficients entiers vérifiant $\|Q\|_{X'} < 1$;
- observer qu'on peut supposer le polynôme Q réciproque en formant au besoin

$$Q(z).z^q.Q(1/z) \qquad \text{(où q désigne le degré de Q)}$$

puisque, lorsque $|z|=1$,

$$|z^q.Q(1/z)| = |Q(\bar{z})| = |Q(z)|;$$

- introduire un exposant e convenable permettant de passer de

$$\sup_{z\in X'} |Q(z)| < 1$$

à

$$\sup_{z\in X} |z^e.Q(z)| < 1,$$

ce qui établira l'inégalité cherchée $t(X) < 1$.

Ce deuxième procédé fournit donc directement la fonction auxiliaire R introduite au §2; en fait, on n'a besoin d'appliquer le lemme 1 qu'à l'ensemble X' mais ce qu'on vient de voir prouve que les inégalités $t(X)<1$ et $t(X')<1$ sont équivalentes... . En revenant aux notations du §2, on voit qu'en choisissant convenablement les paramètres t et s, les ensembles X définis ci-dessus et au §2 coïncident. Par conséquent, si P est le polynôme minimal d'un entier algébrique dont aucun conjugué ne se trouve dans V, le calcul du §2 établit l'existence de C(V), c'est-à-dire le théorème 1. On voit même qu'on peut affaiblir l'hypothèse : "aucun conjugué ne se trouve dans V" en "presqu'aucun (dans un sens à préciser en fonc-tion de V ...) ne se trouve dans V". La démonstration montre de plus qu'on ne peut espérer obtenir pour C(V) une valeur meilleure que $1/t(X)$ (on rappelle que X est le complémentaire de V dans le disque-unité).

Il reste à établir le lemme 4. Un nouveau lemme sera nécessaire.

Lemme 5. Soit X une partie compacte du plan.

(i) Soient U un polynôme unitaire à coefficients complexes, u le degré de U, $U^{-1}(X)$ l'ensemble des points z vérifiant $U(z)\in X$, alors : $t(U^{-1}(X)) = t(X)^{1/u}$.

(ii) Soient U/V une fraction rationnelle dont le numérateur est unitaire et de degré u strictement supérieur au degré du dénominateur, X_r l'image réciproque de X par la fonction U/V (i.e., comme dans (i), l'ensemble des z vérifiant $U(z)/V(z)\in X$), on a alors :

$$(\inf_{z\in X_r} |V(z)|)\ t(X) \le t(X_r)^u \le (\sup_{z\in X_r} |V(z)|)\ t(X)$$

Il est clair que (ii) est une généralisation de (i) mais qu'il suffit de prouver (i) pour obtenir sans réelle difficulté une preuve de (ii).

Pour tout polynôme unitaire U_1, on a :

$$\sup_{z\in U^{-1}(X)} |U_1(U(z))| = \sup_{z\in X} |U_1(z)| \qquad \text{d'où} \quad t(U^{-1}(X)) \le t(X)^{1/u}.$$

Réciproquement, pour tout polynôme unitaire U_2 à zéros dans $U^{-1}(X)$, le polynôme transformé $Rés_y(z-U(y), U_2(y))$ s'écrit comme le produit des valeurs prises par U_2 en les divers points de $U^{-1}(z)$; en particulier, lorsque z décrit X, on a :

$$\sup_{z\in X} |Rés_y(z-U(y),U_2(y))| \leq \sup_{z\in U^{-1}(X)} |U_2(z)|^u \quad \text{d'où} \quad t(X) \leq t(U^{-1}(X))^u$$

On va maintenant déduire les valeurs du diamètre transfini d'un segment et d'un arc de cercle de celle d'un disque (ou d'un cercle, ce qui revient au même). Le diamètre transfini d'un cercle ou d'un disque est clairement égal à son rayon (il suffit de l'établir dans le cas du disque-unité, ce diamètre transfini est majoré par 1 puisque la norme du polynôme z est égale à 1 sur le disque-unité, il est minoré par 1 puisque la norme (sur le disque-unité) d'une fonction polynôme est toujours au moins égale à la valeur absolue de chacun de ses coefficients, en particulier 1 dans le cas d'un polynôme unitaire...). On applique le lemme 5 (ii) au cas où X est le segment réel $[-2,2]$, où $U(z)= z^2+1$, $V(z)=z$. L'image réciproque X_r est alors le cercle-unité et le lemme 5 montre donc que le diamètre transfini d'un segment est égal au quart de sa longueur. On choisit maintenant pour X le segment réel $[2\cos(\ell/2),2]$ qui est donc de diamètre transfini $1/2(1-\cos(\ell/2))= (\sin(\ell/4))^2$ et on conserve les mêmes valeurs que précé-demment pour U et V; l'image réciproque X_r est alors un arc du cercle-unité de longueur ℓ et le lemme 4 se déduit maintenant du lemme 5.

§5.- Calculs explicites de constantes de Lehmer C(V).

On revient aux notations du §2. On voit que se posent deux problèmes :
- Construire des fonctions auxiliaires R permettant d'obtenir, pour des valeurs convenables des paramètres s, t, des parties X du disque-unité aussi grandes que possible; on a vu que cela était possible sur le plan théorique par un procédé effectif; en pratique, on utilise des produits de polynômes cyclotomiques (et donc réciproques) élevés à des puissances réelles bien choisies de sorte que la fonction obtenue soit strictement majorée par 1 en valeur absolue sur une grande portion du cercle-unité.

- Choisir l'exposant t de manière optimale.

En fait, on va voir qu'une fonction R étant choisie, ce choix optima conduit à un calcul très simple de C(V), au moins lorsque R satis-fait à certaines conditions techniques. On va exposer ce calcul sur un exemple.

Soit $Q(z)= (z+1)^2(z^2+1)(z^2+z+1)^2$, c'est un produit de polynômes cyclotomiques et donc un polynôme réciproque. Soit $Z= z+(1/z)$, lorsque $|z|=1$, on a :

$$|Q(z)| = |(Z+2)Z(Z+1)^2|.$$

Si, de plus Re $z \leq 0$, Z reste compris entre -2 et 0 et donc :

$$\sup_{\mathrm{Re}z\leq 0,\ |z|=1} |Q(z)| = \sup_{-2\leq Z\leq 0} |Z(Z+1)^2(Z+2)| = \sup_{0\leq Z'\leq 1} Z'(1-Z') = 1/4.$$

Soit B défini par : $|B|=1$, $\mathrm{Re}B = 1/20$, $\mathrm{Im}B>0$. $\overline{B}$ désigne le conjugué de B. Un calcul facile sur l'axe réel (cf. ci-dessus) montre que, lorsque z décrit l'arc de cercle passant par -1 joignant B à $\overline{B}$ (qu'on note $A(B\overline{B})$), on a :

$$\sup_{z\in A(B\overline{B})} |Q(z)| = |Q(B)| = |Q(\overline{B})| = (1/10)(1+(1/10))^2(2+(1/10)) = 0{,}2541 > 1/4.$$

Soit X la partie du disque-unité bordée d'une part par l'arc $A(B\overline{B})$, d'autre part par le segment $S(B\overline{B})$ joignant B à $\overline{B}$. Soit $R(z) = |Q(z)|^{1/8}$.

Le principe du maximum montre que, pour tout exposant $t \geq 0$, la borne supérieure, lorsque z décrit X, de la fonction $z \longmapsto |z|^t R(z)$ est atteinte pour $z \in S(B\overline{B})$. Le calcul du §2 montre alors qu'il convient d'effectuer un choix de t pour que

$$\left(\sup_{z\in S(B\overline{B})} |z|^t R(z)\right)^{-1/(1+2t)}$$

soit le plus grand possible.

Soit z_o le point de $S(B\overline{B})$ de module le plus élevé vérifiant

$$|z_o| = R(z_o)^2 \qquad (\text{i.e. } |z_o|^4 = |Q(z_o)|);$$

alors, pour toute valeur du paramètre t, on a

$$(|z_o|^t R(z_o))^{-1/(1+2t)} = |z_o|^{1/2} \leq \left(\sup_{z\in S(B\overline{B})} |z|^t R(z)\right)^{-1/(1+2t)}.$$

Par conséquent, s'il existe une valeur du paramètre t pour laquelle le maximum sur le segment $S(B\overline{B})$ de la fonction $|z|^t R(z)$ est atteint en z_o, alors $|z_o|^{-1/2}$ est égal à la constante C(V) cherchée et ce calcul est nécessairement optimal.

E. Reyssat a calculé z_o et vérifié que la condition technique ci-dessus est bien valide dans le cas particulier envisagé, ce qui se ramène à une

étude locale d'une famille à un paramètre réel de fonctions d'une variable réelle. On a ainsi prouvé le théorème suivant :

Théorème 2. Soient $B,\overline{B}$ les points du cercle-unité de partie réelle 1/20 et D(B) le disque circonscrit au triangle $OB\overline{B}$. Tout polynôme P irréductible à coefficients entiers différent des polynômes cyclotomiques

$$\phi_2(z)= z+1,\quad \phi_3(z)= z^2+z+1,\quad \phi_4(z)= z^2+1,$$

dont les zéros sont :
ou bien de partie réelle au plus égale à 1/20
ou bien à l'extérieur de D(B),
vérifie :

$$M(P) > (1{,}08)^d \qquad (d : \text{degré de } P).$$

Remarques :

1°) A l'aide de la plus simple des fonctions auxiliaires, c'est-à-dire $Q(z)= 1-z$, on prouve, par le même procédé :

Théorème 3 (cf. (L.2)). Soient u un nombre complexe de module 1, P un polynôme à coefficients complexes :

$$P(z)= a_o z^d+ a_1 z^{d-1}+\ldots+ a_d= a_o(z-x_1)(z-x_2)\ldots(z-x_d),$$

r le nombre des indices i vérifiant $\mathrm{Arg}(x_i)= \mathrm{Arg}(u)$, alors :

$$M(P) \geq 2^{-d}\left(|P(u)|^{1/r}+ \left(|P(u)|^{2/r}+ 4^{d/r}|a_o a_d|^{1/r}\right)^{1/2}\right)^r.$$

L'inégalité ci-dessus, avec $|P(u)|= |a_o|= |a_d|= 1$ et d=r, nous fournit le corollaire suivant, redonnant le résultat déjà cité de Schinzel (cf. §1) :

Corollaire. Soit $x\neq 1,2$ un entier algébrique totalement positif (i.e. dont tous les conjugués sont des réels positifs) de degré d, alors :

$$M(x) \geq \left((1+5^{1/2})/2\right)^d \quad (\text{résultat optimal avec } x= (3+5^{1/2})/2= \left((1+5^{1/2})/2\right)^2).$$

On notera sur ce théorème et son corollaire combien intervient peu l'hypothèse, a priori essentielle, du problème de Lehmer : "P à coefficients entiers".

On voit de plus que l'introduction du nombre d'or $(1+5^{1/2})/2$ dans le corollaire est naturelle; elle correspond à $z_o^{1/2}$ où z_o est la racine inférieure à 1 de l'équation en z_o vue au début de ce paragraphe et qui s'écrit dans ce cas :

$$z_o = (1-z_o)^2.$$

2°) Le procédé décrit ci-dessus permet donc un calcul explicite de constantes de la forme $C(V)$ à partir de fonctions auxiliaires convenables et est assez facile à mettre en oeuvre; cependant, le fait que, par construction, le maximum de

$$|z|^t R(z) \quad \text{lorsque} \quad z \quad \text{décrit} \quad X$$

ne soit atteint que sur $S(B\overline{B})$ suggère bien qu'on ne tire pas le meilleur parti du calcul du §2. En fait, le problème général est celui du choix des exposants réels (e_i) à affecter aux polynômes réciproques Q_i (cyclotomiques ou autres) permettant de construire de <u>bonnes</u> fonctions

$$R(z) = \left(\prod_i |Q_i(z)|^{e_i}\right)^{1/q} \qquad \text{avec} \quad q = \sum_i e_i d^\circ Q_i$$

(c'est-à-dire petites sur de larges portions du cercle-unité). Ce problème est assez vaste pour qu'on y intègre le choix du paramètre t sans en augmenter sensiblement la difficulté. Toutes ces perspectives sont certes plus techniques mais devraient fournir de meilleurs résultats; des travaux conduits dans ce sens par G. Rhin le suggèrent.

§6.- <u>Bibliographie</u>.

(Bo) D. Boyd.- Inverse problems for Mahler's measure (titre manquant dans le manuscrit), Diophantine Analysis (Proc. ed. by Loxton and Van der Poorten) London Math. Soc. (L.N. Ser. 109), Camb. Univ. Press, 1986.

(L.1) M. Langevin.- Méthode de Fekete-Szegö et problème de Lehmer, C.R. Acad. Sc. Paris, 301, 1985 p. 463-466.

(L.2) M. Langevin.- Minorations de la maison et de la mesure de Mahler de certains entiers algébriques, C.R. Acad. Sc. Paris, 303, 1986, p. 523-526.

(L.3) M. Langevin.- Solution des problèmes de Favard (à paraître aux Ann. Inst. Fourier, Grenoble).

(Lo) R. Louboutin.- Sur la mesure de Mahler d'un nombre algébrique, C.R. Acad. Sc. Paris, 296, 1983, p. 707-708.

(Sc) A. Schinzel.- On the product of the conjugates outside the unit circle of an algebraic number, Acta Arithmetica, XXIV, 1973, p. 385-399.

(Sm) C.J. Smyth.- On the measure of totally real algebraic integers (I and II) J. Australian Math. Soc. (Ser. A), 30, 1980, p. 137-149 and Math. of Computation, 37, 1981, p. 205-208.

§7.- Dernière minute (juin-juillet 1987)...

On signale maintenant quelques développements de dernière minute non mentionnés dans l'exposé oral du 19 janvier 1987.

Dans (L.1), est cité sous la référence (6) un résultat de M. Mignotte, déduit d'un travail de Erdös et Turán, établissant que les arguments des zéros des polynômes (irréductibles à coefficients entiers ...) de petite mesure et de grand degré sont bien distribués. En fait, comme son auteur l'obervait à la lecture de (L.1), les arguments donnés dans (6) (cf. (L.1) ou ci-dessous) permettent d'établir un énoncé de la forme suivante :

Théorème (M. Mignotte, à paraître). Soient u un nombre complexe de module 1 et ϵ un réel strictement positif. Alors, tout polynôme P irréductible, à coefficients entiers, dont les zéros sont hors du secteur angulaire de sommet O, d'axe Ou, d'angle ϵ, et de degré $d>D(\epsilon)$, vérifie :

$$M(P) > \exp(c\,\epsilon^3\,d) \qquad (c \text{ constante absolue}).$$

Cet énoncé demeure exact, comme ceux des paragraphes antérieurs, en supposant seulement qu'une proposition suffisante des racines de P soient hors du secteur angulaire. De cette propriété, on peut déduire avec l'auteur précité une évaluation explicite de C(V) (de forme analogue à celle donnée ci-dessus) par une mise en forme convenable de l'argument suivant :

Soit V le voisinage de u obtenu par intersection du secteur angulaire et de la couronne de centre O et de rayons $1/(1+\epsilon'),(1+\epsilon')$ (ϵ' fonction convenable de ϵ) et supposons seulement que P soit sans zéro dans V; alors :

- ou bien P n'a que peu de zéros dans le secteur angulaire ce qui permet d'appliquer le théorème ci-dessus,

-ou bien P a une certaine proportion de zéros dans le secteur angulaire mais de modules éloignés de 1 d'au moins ϵ' ce qui permet à nouveau de conclure.

On revient maintenant sur l'outil principal de la démonstration de théorème, i.e. le résultat d'Erdös et Turán, pour en décrire brièvement les liens avec la notion de diamètre transfini.

Théorème (Erdös-Turán). Soit Q un polynôme unitaire à coefficients complexes dont tous les d zéros sont de module 1. Pour tout arc A du cercle-unité, soient d(A) le nombre de zéros de Q appartenant à A et l(A) la longueur de A. Alors, la discrépance associée aux zéros de Q vérifie :

$$\sup_A |d(A)/d-1(A)/2\pi| < 10(\log\|Q\|^{1/d})^{1/2} \quad \text{où} \quad \|Q\|= \sup_{|z|=1} |Q(z)|.$$

La notion de diamètre transfini n'est pas évoquée dans la démonstration donnée par ces auteurs, ni dans le résumé rappelé par Mignotte dans (6). Elle est néanmoins clairement sous-jacente (d'ailleurs, M. Mignotte auquel je signalais ce point m'a dit avoir trouvé la même remarque dans la littérature sous la forme d'une communication orale de Fekete relative à ce théorème). Nous allons en esquisser la preuve. La démonstration d'Erdös et Turán repose sur une propriété des polynômes de Čebičev (à l'intérieur d'une famille de polynômes unitaires, est "de Čebičev de degré d pour une partie compacte $X\subset\mathbb{C}$" tout élément U de la famille, de degré d, tel que $\|U\|_X$ soit minimal) qui prennent leur maximum en valeur absolue en un nombre de points égal au degré ... Or, X désignant toujours une partie compacte de $\mathbb{C}$, on rappelle que le diamètre transfini de X est égal à

$$t(X)= \inf(\|U\|_X)^{1/d\circ U}$$

où U décrit l'ensemble des polynômes unitaires et on a vu que cette propriété restait valable même en limitant le domaine de variation à ceux de ces polynômes dont les zéros appartiennent au bord (extérieur) de X et donc, si Q désigne un tel polynôme, on a $t(X) \leq (\|Q\|_X)^{1/d\circ Q}$

Ce qu'établit quantitativement le théorème d'Erdös et Turán, dans le cas où X est le disque-unité, c'est que, si $(\|Q\|_X)^{1/d\circ Q}$ est assez proche de $t(X)=1$, alors Q est voisin du polynôme de Čebičev de même degré pour le cercle-unité (i.e. (z^d-v) avec $d= d\circ Q$ et $|v|=1$), ce qui prouve en particulier une "unicité" (à isométrie près) de ces polynômes de Čebičev.

On abandonne maintenant le théorème d'Erdös-Turán et ses possibles extensions à des domaines autres que le disque-unité (grâce aux techniques déjà évoquées au §4) pour décrire son application au présent problème, toujours en suivant ces auteurs et (6). On commence par un intéressant lemme dont on va voir qu'il se déduit du célèbre théorème de Ptolémée.

Lemme. Soient

$$P(z)= a_o z^d+a_1 z^{d-1}+\ldots+a_d= a_o(z-x_1)\ldots(z-x_d) \quad \text{et}$$

$$Q(P)(z)= (z-x_1/|x_1|)\ldots(z-x_d/|x_d|), \quad \text{alors :}$$

$$\|Q(P)\| \leq \|P\|/(|a_o a_d|)^{1/2}$$

($\|\ \|$: norme de la convergence uniforme sur le disque-unité).

Rappelons l'énoncé du théorème de Ptolémée (démonstration au chapitre "Inversion" de tout bon ouvrage de géométrie élémentaire ou, en deux lignes, grâce au birapport et à l'inégalité triangulaire) :

Théorème (Ptolémée). Le produit des longueurs des diagonales d'un quadrilatère est inférieur ou égal à la somme des produits des longueurs des côtés opposés.

Corollaire. Soient z, x deux nombres complexes, alors :

$$|z-x|^2 = |z||x||z/|z|-x/|x||^2+(|z|-|x|)^2;$$

en particulier, si $|z|=1$,

$$|z-x| \leq (|x|)^{1/2}(|z-x/|x||) \qquad \text{d'où le lemme.}$$

Reporté dans le théorème d'Erdös et Turán, en remplaçant $\|Q\|$ par $\|P\|/(|a_o a_d|)^{1/2}$, le lemme ci-dessus montre qu'on peut s'affranchir de toute hypothèse sur le polynôme. Il reste maintenant à introduire $M(P)$ à la place de $\|P\|$ dans l'énoncé obtenu pour l'adapter au résultat cherché. L'application brutale de l'inégalité : $\|P\| \leq 2^{d^oP} M(P)$ (optimale pour des polynômes à coefficients complexes) ne conduit à aucun résultat. C'est ici qu'intervient l'hypothèse "P irréductible à coefficients entiers" qui permet à Mignotte, grâce au lemme de Siegel, de construire un multiple de P de norme majorée en fonction de $M(P)$. De façon précise, soit d_1 un entier strictement supérieur à d, alors, il existe un multiple P_1 de P de degré au plus d_1 et de hauteur au plus

$$(2(d_1+1)^d M(P)^{d_1})^{1/(d_1+1-d)}$$

On obtient alors le théorème énoncé au début de ce paragraphe 7 en appliquant l'inégalité générale (i.e. sans hypothèse sur le polynôme) d'Erdös et Turán au polynôme P_1 dont la norme sur le disque-unité est majorée par $(1+d_1)$ fois la hauteur, cette dernière étant majorée en fonction de la mesure de P comme ci-dessus.

Références :

(6) de (L.1) : M. Mignotte.- Sur la répartition des racines des polynômes, Journées de théorie élémentaire et analytique des nombres, Caen, septembre 1980, dactylographié.

Lemme de Siegel: M. Mignotte.- Estimations élémentaires effectives sur les nombres algébriques, Publ. IRMA Strasbourg, 1979 ou (même titre) Proc. "Journées Arithmétiques d'Exeter", 1980, London Math. Soc. (L.N. Ser 56) Cambr. Univ. Press.

Théorème d'Erdös-Turán : P. Erdös - P. Turán.- On the distribution of roots of polynomials, Ann. Math. 51, 1950, p. 105-119.

Enfin, une démonstration complète des résultats qu'on vient d'esquisser sera publiée par M. Mignotte.

Michel LANGEVIN
U.A. 763 "Problèmes Diophantiens"
Institut Henri Poincaré
11, rue Pierre et Marie Curie
75231 PARIS Cedex 05

Les entiers sans facteurs carré $\leq x$ dont leurs facteurs premiers $\leq Y$

Mongi NAIMI

1. Introduction.

On considère la fonction $\mu^2(n)$ fonction caractéristique des entiers sans facteur carré, et on s'intéresse à la quantité

$$\Psi_2(x,y)= \sum_{\substack{n\leq x \\ P(n)\leq y}} \mu^2(n)$$

et son lien avec

$$\Psi(x,y)= \sum_{\substack{n\leq x \\ P(n)\leq y}} 1.$$

$\Psi(x,y)$ a été étudié par de nombreux auteurs dont A. Hildebrand (1985) [1] qui a montré que :

$$(1) \qquad \Psi(x,y)= x\ \rho(u)\left(1 + O\left(\frac{u\ \mathrm{Log}(u+1)}{\log x}\right)\right)$$

uniformément pour $y\geq \exp(\mathrm{Log}\ \log x)^{5/3+\epsilon}$, où $u= \frac{\mathrm{Log}\ x}{\mathrm{Log}\ y}$ et $\rho(u)$ est la fonction de Dickman définie par le système suivant :

$$\begin{cases} \rho(u)=1 & 0\leq u\leq 1 \\ -u\rho'(u)= \rho(u-1) & u>1 \end{cases}$$

A. Hildebrand et G. Tenenbaum (1986) [3] ont montré par une méthode analytique le résultat suivant :

$$(2) \qquad \Psi(x,y)= \frac{x^{\alpha}\ \zeta(\alpha,y)}{\alpha\sqrt{2\Pi}\ \Phi_2(\alpha,y)}\left(1 + O(1/u) + O\left(\frac{\mathrm{Log}\ y}{y}\right)\right)$$

pour $x,y\geq 2$ uniformément où

$$\zeta(s,y)= \prod_{p\le y} (1-p^{-s})^{-1}$$

$$\Phi_2(s,y)= \frac{\partial^2}{\partial s^2} \operatorname{Log} \zeta(s,y)$$

et α la racine de l'équation $\operatorname{Log} x= \sum_{p\le y} \frac{\operatorname{Log} p}{p^{\alpha}-1}$.

A. Ivic (1985) [4] en écrivant $\mu^2(n)= \sum_{d^2/n} \mu(d)$ et en utilisant (1) a montré que

$$\Psi_2(x,y) \sim \frac{6}{\Pi^2} \Psi(x,y) \tag{3}$$

pour $y \ge \exp(\operatorname{loglog}^{2+\epsilon} x)$

dans ce même papier il conjecturait que l'équivalence (3) est vraie pour $y \ge \operatorname{Log}^{1+\epsilon} x$.

Dans ce travail en reprenant la méthode utilisée dans [4] nous démontrons le résultat suivant.

Théorème 1. Pour $y \ge \operatorname{Log}^{1+\epsilon} x$ et $u \ge (\operatorname{Log} \operatorname{Log} 2y)^2$ on a :

$$\Psi_2(x,y)= \frac{x^{\beta}\, \zeta_2(\beta,y)}{\beta\sqrt{2\Pi\, \varphi_2(\beta,y)}} (1+0(1/u)) \tag{4}$$

où

$$\zeta_2(s,y)= \prod_{p\le y} (1+p^{-s})$$

$$\varphi_2(s,y)= \frac{\partial^2}{\partial^2 s} \operatorname{Log} \zeta_2(s,y)$$

et β la racine de l'équation $\operatorname{Log} x= \sum_{p\le y} \frac{\operatorname{Log} p}{p^{\beta}+1}$.

En comparant les quantités qui interviennent dans (2) et (4) nous déduisons le résultat suivant.

Théorème 2. Pour $y \geq \mathrm{Log}^{1+\epsilon} x$

$$(5) \qquad \frac{1}{\zeta(2\beta,y)}\Psi(x,y)(1+0(1/u)) \leq \Psi_2(x,y) \leq \frac{1}{\zeta(2\alpha,y)}\Psi(x,y)(1+0(1/u))$$

en particulier pour $y \geq \mathrm{Log}^{2+\epsilon} x$

$$(6) \qquad \Psi_2(x,y) \sim \frac{1}{\zeta(2\beta_o,y)}\Psi(x,y)$$

où $\beta_o = 1 - \frac{\mathrm{LogLog}\ x}{\mathrm{Log}\ x}$ (partie principale commune de α et β).

On remarque d'après (6) que la relation (3) n'est vraie que si $\frac{\mathrm{LogLog}\ x}{\mathrm{Log}\ y} \longrightarrow 0$ quand $x,y \longrightarrow \infty$.

Par une autre méthode et dans un cadre un peu plus général A. Ivic et G. Tenenbaum [5] ont récemment démontré que

$$(7) \quad \text{pour } y \geq \mathrm{Log}^{2+\epsilon} x \qquad \Psi_2(x,y) \sim \frac{1}{\zeta(2\beta_o)}\Psi(x,y) \quad \text{et que}$$

$$(8) \quad \text{pour } y \leq \mathrm{Log}^{2-\epsilon} x \qquad \Psi_2(x,y) \ll \Psi(x,y)\exp - \mathrm{Log}^{\epsilon/3} y.$$

Notre méthode justifiant ainsi l'introduction du point β permet d'obtenir dans (5) une minoration de $\frac{\Psi_2(x,y)}{\Psi(x,y)}$ et une majoration plus fine que (8), pour $\mathrm{Log}^{1+\epsilon} x \leq y \leq \mathrm{Log}^{2-\epsilon} x$; dans la zone $y \geq \mathrm{Log}^{2+\epsilon} x$, notre résultat (6) est identique à leur résultat (7).

2. Lemmes.

La méthode utilisée est basée essentiellement sur la formule de Perron.

$$(9) \qquad \sum_{\substack{n \leq x \\ P(n) \leq y}} \mu^2(n) = \frac{1}{2i\pi}\int_{\beta-iT}^{\beta+iT} \zeta_2(s,y)\frac{x^s}{s}\,ds + 0\Big(x^{\beta} \sum_{\substack{n \geq 1 \\ P(n) \leq y}} \frac{\mu^2(n)}{n^{\beta}} \min\Big(1, \frac{1}{T(\mathrm{Log}\frac{x}{n})}\Big)\Big)$$

le point β est choisi de façon que $\frac{x^s \zeta_2(s,y)}{s}$ soit maximale dans un voisinage de β et relativement petite ailleurs, pour cela on prendra β comme zéro de la dérivée logarithmique de $x^s \zeta_s(s,y)$ ou encore β solution de l'équation :

$$\mathrm{Log}\ x = \sum_{p \leq y} \frac{\mathrm{Log}\ p}{p^{\beta}+1}$$

une étude du point $\beta = \beta(x,y)$ permet d'obtenir les résultats suivants :

Lemme 1. Pour $y \geq \text{Log}^{1+\epsilon} x$

$$\beta = 1 - \frac{\xi(u)}{\text{Log } y} + O\left(\frac{1}{u \text{ Log } y}\right) + O\left(\frac{\log x}{y}\right) + O(\exp - \sqrt{\text{Log } y}) \tag{10}$$

où $\xi(u)$ est la solution de l'équation $e^{\xi(u)} = 1+u\,\zeta(u)$.

Remarque :

$$\xi(u) = \text{Log}(u \text{ Log } u) + O\left(\frac{\text{Log Log } u}{\text{Log } y}\right) \quad \text{pour} \quad u \geq 0 \tag{11}$$

voir [2].

Lemme 2. Pour $y \geq \text{Log}^{1+\epsilon} x$ et $u \geq u_o$

$$\varphi_k(\beta,y) \asymp u \text{ Log}^k y \qquad 1 \leq k \leq 3 \tag{12}$$

$$\varphi_4(\beta+it,y) \ll u \text{ Log}^4 y \qquad t \leq \frac{1}{\text{Log } y} \tag{13}$$

$$\varphi_2(\beta,y) = \Phi_2(\alpha,y)(1+O(1/\text{Log } y)) \tag{14}$$

où on note par $\varphi_k(s,y) = \dfrac{\partial^k}{(\partial s)^k} \text{Log } \zeta_2(s,y)$.

Lemme 3. Pour $y \geq \text{Log}^{1+\epsilon} x$ et $u \geq u_o$

$$\left|\frac{\zeta_2(\beta+it,y)}{\zeta_2(\beta,y)}\right| \ll \begin{cases} \exp(-c_1 t^2 \varphi_2(\beta,y)) & |t| \leq 1/\text{Log } y \\ \exp\left(-c_2 \dfrac{u\,t^2}{(1-\beta)^2+t^2}\right) & 1/\text{Log } y \leq |t| \exp \text{Log}^{3/2-\epsilon} y \end{cases} \tag{15}$$

de telles majorations permettent de montrer que la contribution de l'intégrale dans (9) est dans un voisinage de β et qu'ailleurs la quantité $\dfrac{x^s \zeta_2(x,y)}{sx^\beta \zeta_2(\beta,y)}$ est relativement petite.

3. Démonstration des théorèmes

Preuve du théorème 1 : La preuve du théorème 1 se résume dans les deux lemmes suivants :

Lemme 4. Pour $y \geq \mathrm{Log}^{1+\epsilon}$, $y \geq (\mathrm{loglog}\ 2y)^2$

$$\Psi_2(x,y) = \frac{1}{2i\Pi} \int_{\beta-iT_o}^{\beta+iT_o} \frac{\zeta_2(s,y)x^s}{s}\, ds + O\left(\frac{x^B \zeta_2(\beta,y)}{\sqrt{\varphi_2(\beta,y)}} \frac{1}{u}\right) \tag{16}$$

où

$$T_o = \frac{1}{u^{1/3} \log y}$$

Lemme 5. Pour $y \geq \mathrm{Log}^{1+\epsilon}$

$$\frac{1}{2i\Pi} \int_{\beta-iT_o}^{\beta+it_o} \zeta_2(s,y)\, \frac{x^s}{s}\, ds = \frac{x^\beta \zeta_2(\beta,y)}{\beta\sqrt{2\Pi\varphi_2(\beta,y)}} (1+o(1/u)) \tag{17}$$

le théorème se déduit alors des formules (16) et (17).

Indication pour la preuve du lemme 4.

On part de la formule (9) par le choix de $T = (\exp -\log y^{3/2-\epsilon} + \exp - \frac{cu}{\log^2})^{-2}$ et en utilisant (15) on montre que le terme du reste de la formule (9) est :

$$O\left(\frac{x^\beta \zeta_2(\beta,y)}{\sqrt{\varphi_2(\beta,y)}}\, 1/u\right) \quad \text{pour} \quad u \geq (\mathrm{loglog}\ 2y)^2$$

de nouveau par la majoration (15) on montre que la contribution de l'intégrale dans la formule (9) dans le domaine défini par

$$\{\beta+it,\ T_o \leq |T| \leq T\} \quad \text{est aussi}$$

$$O\left(\frac{x^\beta \zeta_2(\beta,y)}{\sqrt{\varphi_2(\beta,y)}} \frac{1}{u}\right)$$

Preuve du lemme 5 : Pour un développement de Taylor de $\log(\zeta_2(s,y)x^s)$ au voisinage de β et en tenant compte du fait que $\log x + \varphi_1(\beta,y) = 0$ on a :

$$\zeta_2(s,y)x^s = x^\beta\,\zeta_2(\beta,y)e^{-\frac{t^2}{2}\varphi_2(\beta,y)}\exp(-\frac{t^3}{3!}\varphi_3(\beta,y)+O(t^4\varphi_4^*))$$

où $$\varphi_4^* = \sup\ |\varphi_4(\beta+it,y)| \qquad |t|\le T_o$$

par un développement de

$$\exp(-\frac{t^3\varphi_3(\beta,y)}{3!} + O(t^4\varphi_4^*))$$

d'une part et de $\frac{1}{s} = \frac{1}{\beta}(\frac{1}{1-it/\beta})$ d'autre part on a :

$$\frac{\zeta_2(s,y)x^s}{s} = \frac{x^\beta\zeta_2(\beta,y)}{\beta}e^{-\frac{t}{2}\varphi_2(\beta,y)}\{1-\frac{it}{\beta}-\frac{it^3}{3!}\varphi_3(\beta,y)+O(R(t))\}$$

$$R(t) = O(\frac{t^2}{\beta^2} + t^6\varphi_3^2 + t^4\varphi_4^*)$$

si on prend l'intégrale sur le segment $[\beta+iT_o,\ \beta-iT_o]$ on a

$$\frac{1}{2i\pi}\int_{\beta-iT_o}^{\beta+iT_o}\zeta_2(s,y)\frac{x^s}{s}\,ds =$$

$$= \frac{x^\beta\zeta_2(\beta,y)}{2\pi\,\beta}\left(\int_{-T_o}^{T_o}e^{-\frac{t^2}{2}\varphi_2(\beta,y)}dt + \int_{-T_o}^{T_o}e^{-\frac{t^2}{2}\varphi_2(\beta,y)}R(t)dt\right)$$

si on calcule les intégrales du membre de droite après le changement de variable $\tau = t\sqrt{\varphi_2(\beta,y)}$ on trouve le lemme.

<u>Preuve du théorème 2</u> : Le théorème 2 est un comparaison de $\Psi_2(x,y)$ et $\Psi_2(x,y)$. Nous le démontrerons dans la zone $y\ge\log x^{1+\epsilon}$ et $u\ge(\text{loglog } 2y)^2$ puisque pour $u\le(\text{loglog } 2y)^2$ la formule (3) est vraie. D'après (2), (4) et (14) les seules quantités qui restent à comparer sont : $x^\beta\zeta_2(\beta,y)$ et $x^\alpha\zeta(\alpha,y)$. C'est le but du lemme suivant :

Lemme 6. Pour $y \geq \log x^{1+\epsilon}$

$$(18) \qquad \frac{x^{\alpha-\beta}}{\zeta(2\beta,y)} \leq \frac{\zeta_2(\beta,y)}{\zeta(\alpha,y)} \leq \frac{x^{\alpha-\beta}}{\zeta(2\alpha,y)}$$

en particulier pour $y \geq \log x^{2+\epsilon}$

$$(19) \qquad \frac{\zeta_2(\beta,y)}{\zeta(\alpha,y)} \sim \frac{x^{\alpha-\beta}}{\zeta(2\beta_o y)} \quad \text{avec} \quad \beta_o = 1 - \frac{\log\log x}{\log y} .$$

Preuve du lemme 6 :

On a :
$$\zeta_2(\beta,y) = \frac{\zeta(\beta,y)}{\zeta(2\beta,y)}$$

$$\frac{\zeta_2(\beta,y)}{\zeta(\alpha,y)} = \frac{1}{\zeta(1\beta,y)} \frac{\zeta(\beta,y)}{\zeta(\alpha,y)}$$

$$\frac{\zeta(\beta,y)}{\zeta(\alpha,y)} = \exp \int_\alpha^\beta \Phi_1(\sigma,y)\, d\sigma$$

$$\Phi_1(s,y) = \frac{\partial}{\partial s} \log \zeta(s,y)$$

d'où

$$\frac{\zeta(\beta,y)}{\zeta(\alpha,y)} = \exp(\alpha-\beta) \sum_{p \leq y} \frac{\log p}{p^c - 1} \quad \text{où} \quad \beta < c < \alpha$$

$$\geq \exp(\alpha-\beta) \sum_{p \leq y} \frac{\log p}{p^\alpha - 1} = x^{\alpha-\beta}, \quad \text{puisque} \quad \log x = \sum_{p \leq y} \frac{\log p}{p^\alpha - 1})$$

ce qui donne la minoration de (18). La majoration se fait de la même manière en écrivant

$$\frac{\zeta_2(\beta,y)}{\zeta(\alpha,y)} = \frac{1}{\zeta(2\alpha,y)} \frac{\zeta_2(\beta,y)}{\zeta(\alpha,y)}$$

$$= \frac{1}{\zeta(\alpha,y)} \exp \int_\alpha^\beta \varphi_1(\sigma,y)\, d\sigma$$

en tenant compte du fait que $\varphi_1(\beta,y) = \log x.$

BIBLIOGRAPHIE

[1] A. Hildebrand.- On the number of positif integers $\leq$ and free of Prime factors $>y$, J. Number Theory, 22 (1986) 289-307.

[2] A. Hildebrand.- On the local behaviour of $\Psi(x,y)$, Trans. Amer. Math. Soc. (à paraître).

[3] A. Hildebrand and G. Tenenbaum.- On integers free of large prime factors, Trans. Amer. Math. Soc., vol. 296 N° 1 (1986), 265-289.

[4] A. Ivic.- On square free number with restricted primes factors, Studia. Scient. Math. Hungarica, Vol. 20 (1985), 189-192.

[5] A. Ivic and G. Tenenbaum.- Local density over integers of large factors, Quart J. Math. Oxford (2) 37 (1986), 401-417.

Mongi NAIMI
Département de Mathématiques
Faculté des Sciences
Tunis
TUNISIE

Résultat de Cantor et Straus sur la conjecture de Lehmer

M. Pathiaux-Delefosse

I - Introduction.

Etant donné un entier algébrique, α, non nul, de degré d, de conjugués $\alpha_1,\alpha_2,\ldots,\alpha_d$, la mesure de Mahler est définie par

$$M(\alpha)= \prod_{i=1}^{d} \max(1, |\alpha_i|).$$

Kronecker [4] a montré que si $M(\alpha)=1$, alors α est une racine de l'unité. En 1933 D.H. Lehmer a émis la conjecture suivante : il existe une constante $c>1$ telle que $M(\alpha)\geq c$ pour tout α entier algébrique non racine de l'unité.

Notons $C(d)= \inf M(\alpha)$ pour α entier algébrique, non racine de l'unité, de degré $\leq d$. Etant donné qu'il n'existe qu'un nombre fini d'entiers algébriques de degré $\leq d$ vérifiant $M(\alpha)\leq 2$, on a donc $c(d)>1$, $\forall d\in\mathbb{N}^*$; la fonction $d\in\mathbb{N}^* \longrightarrow C(d)$ est décroissante; la conjecture de Lehmer se réduit à montrer que $\lim_{d\to+\infty} C(d)>1$.

Quelques réponses à cette conjecture :

A l'heure actuelle, on n'a pas trouvé d'entier algébrique, non racine de l'unité, tel que $M(\alpha)<M(\alpha_o)= 1,1762808\ldots$, où α_o est le nombre de Salem de degré 10, de polynôme minimal

$$z^{10}+ z^9- z^7- z^6- z^5- z^4- z^3+ z + 1.$$

En 1977, E. Drobrowolski [3], puis M. Mignotte [8] ont prouvé que :

$$C(d) \geq 1 + 10^{-6} (\text{loglog}d/\log d)^3.$$

En 1982, D.C. Cantor et E.G. Straus [2] avec une preuve beaucoup plus élémentaire on montré que :

$$C(d) \geq 1 + 2(1+o(1))(\text{loglog}d/\log d)^3.$$

Une faute résidait dans cette preuve; elle a été corrigée [2] par les mêmes auteurs. La constante 2 a été améliorée en $\frac{9}{4}$ par R. Louboutin [6].

Nous nous proposons d'exposer la démonstration de Cantor et Strauss, en y apportant quelques modifications.

II - Lemmes algébriques :

Lemme 2.1 (Mahler). Soit α un entier algébrique non nul, α' un conjugué de α; s'il existe des entiers naturels r et s, distincts, tels que $\alpha'^{r}=\alpha^{s}$ alors α est une racine de l'unité.

Preuve : Soit σ un $\mathbb{Q}$-automorphisme de l'extension galoisienne $\mathbb{Q}[\alpha_1,\alpha_2,\ldots,\alpha_d]$ tel que $\sigma(\alpha)=\alpha'$; alors $(\sigma(\alpha))^r=\alpha^s$; d'où $\sigma^2(\alpha^{r^2})=\sigma(\sigma(\alpha^{r^2}))=\sigma(\sigma(\alpha^r))^r=(\sigma(\alpha^s))^r=\alpha^{s^2}$; par récurrence on obtient $\sigma^k(\alpha^{r^k})=\alpha^{s^k}$ $\forall k\in\mathbb{N}^*$; soit k un entier tel que $\sigma^k=1$, on obtient alors $\alpha^{r^k-s^k}=1$ et α est une racine de l'unité.

Lemme 2.2. Soit α un nombre algébrique, non racine de l'unité et p un nombre premier. Alors si le nombre α^p a un degré inférieur à celui de α, il existe un entier algébrique β, non racine de l'unité tel que :

i) $d^o\beta < d^o\alpha$
ii) $M(\beta) \leq M(\alpha)$.

Preuve : Si α est de degré p sur $\mathbb{Q}[\alpha^p]$, on prend $\beta=\alpha^p$, on a alors $d^o\beta= d^o\alpha/_p$ et $M(\beta)= M(\alpha)$.

Sinon le polynôme $X^p-\alpha^p$ est réductible sur $\mathbb{Q}[\alpha^p]$ et peut s'écrire sous la forme $X^p-\alpha^p= \prod_{i=1}^{p} (X-\rho^i\alpha)= A.B$ où ρ est une racine primitive p-ième de l'unité et A et B sont éléments de $\mathbb{Q}[\alpha^p][X]$, avec $d^oA\geq 1$, $d^oB\geq 1$. En considérant A(0), on trouve qu'il existe $r\in\{1,\ldots,p-1\}$ et ρ' racine p-ième de l'unité tels que $\rho'\alpha^r\in\mathbb{Q}[\alpha^p]$; or r et p sont premiers entiers; il existe donc 2 entiers u et v tels que ru+pv=1; alors $\rho'^{u}\alpha^{ru}\in\mathbb{Q}[\alpha^p]$ et $\rho'^{u}\alpha\in\mathbb{Q}[\alpha^p]$; on prend alors $\beta=\rho'^{u}$. On a alors $d^o\beta\leq d^o\alpha^p<d^o\alpha$. En outre $\beta^p=\alpha^p$. Donc $\mathbb{Q}[\beta]\subset\mathbb{Q}[\beta^p]=\mathbb{Q}[\alpha^p]$ et degré β^p= degré β; soit $\beta_1,\beta_2,\ldots,\beta_s$ les conjugués de β sur $\mathbb{Q}$; le polynôme $Q=\prod_{i=1}^{s}(X-\beta_i^p)$ est donc le polynôme minimal de $\beta^p=\alpha^p$ sur $\mathbb{Q}$; donc $Q=P^r$ où P est le polynôme minimal de α sur $\mathbb{Q}$ et $(M(\beta))^r=M(\alpha)$ avec $r\geq 2$; d'où $M(\beta)<M(\alpha)$.

Remarque : L'énoncé de ce corollaire figure dans l'article de Mignotte mais avec une démonstration fausse.

Corollaire 1. $C(d)= \inf_{\alpha\in\Omega_d} M(\alpha)$ où Ω_d est l'ensemble des entiers algébriques, de degré $d^o \leq d$, non racine de l'unité, tels que $\alpha_i^p \neq \alpha_j^p$ pour tout nombre premier p et pour tout i et j distincts appartenant à l'ensemble $\{1,2,\ldots,d'\}$.

Lemme 2.3. Soit P un polynôme à coefficients entiers alors

$$P(X^p)= (P(X))^p + p\,G(X) \qquad \text{où} \qquad G\in\mathbb{Z}[X].$$

Preuve : On utilise le petit théorème de Fermat.

III - Introduction de déterminants.

Soit m nombres complexes $\alpha_1,\alpha_2,\ldots,\alpha_m$ et $r_1,r_2,\ldots,r_m$ m nombres entiers, on définit le déterminant de Vandermonde généralisé d'ordre $n= r_1+r_2+\ldots+r_m$, noté $D(\alpha_1,\alpha_2,\ldots,\alpha_m,r_1,r_2,\ldots,r_m)$ par
$D(v_o(\alpha_1),v_1(\alpha_1),\ldots,v_{r_1-1}(\alpha_1),v_o(\alpha_2),\ldots,v_{r_2-1}(\alpha_2),\ldots,v_o(\alpha_m),\ldots,v_{r_m-1}(\alpha_m))$
où

$$v_o(\alpha)=\begin{bmatrix} 1 \\ \alpha \\ \vdots \\ \alpha^{n-1} \end{bmatrix}, \quad v_1(\alpha)=\begin{bmatrix} 0 \\ 1 \\ 2\alpha \\ \vdots \\ (n-1)\alpha^{n-2} \end{bmatrix}, \ldots, v_i(\alpha)=\begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \\ C_i^{i+1}\alpha \\ \vdots \\ C_i^{n-1}\alpha^{n-1-i} \end{bmatrix}=\frac{1}{i!}\,\frac{d^i v_o(\alpha)}{d^i\alpha}.$$

Lemme 3.1. $\qquad \|v_i(\alpha)\|^2 \leq (\mathrm{Max}(1,|\alpha|))^{2n} n^{2i+1}.$

Preuve : $\qquad C_i^k= \dfrac{k(k-1)\ldots(k-(i+1))}{i!} \leq \dfrac{n(n-1)\ldots(n-(i+1))}{i!} \leq n^i$

D'où $\qquad \|v_i(\alpha)\|^2 \leq n\times n^{2i}\times(\mathrm{Max}(1,|\alpha|))^{2n}.$

Lemme 3.2. $\qquad D(\alpha_1,\alpha_2,\ldots,\alpha_m,r_1,r_2,\ldots,r_m)= \pm \prod_{1\leq i<j\leq m} (\alpha_i-\alpha_j)^{r_i r_j}.$

Preuve : [7]

IV - Démonstration de $C(d) \geq 1 + (2+o(1))(\log\log d/\log d)^3$.

Soit α un élément de Ω_d, de degré $d' \leq d$, de conjugués $\alpha=\alpha_1, \alpha_2, \ldots, \alpha_{d'}$, $p_1=2, \ldots, p_s$ les s premiers nombres premiers, et k un élément de $\mathbb{N}^*$; on considère alors les déterminants notés $V(\alpha,d',s,k)$ définis par

$$V(\alpha,d',s,k)= V(\alpha_1,\alpha_2,\ldots,\alpha_{d'},\alpha_1^{p_1},\ldots,\alpha_{d'}^{p_1},\ldots,\alpha_1^{p_s},\ldots,\alpha_{d'}^{p_s},\underbrace{k\ldots k}_{d' \text{ éléments}},\underbrace{1,\ldots,1}_{sd \text{ éléments}})$$

Ce sont des déterminants d'ordre $kd'+sd'= d'(k+s)=n$. Le principe de la démonstration consiste à majorer et minorer ces déterminants.

a) majoration :

Les colonnes de ce déterminant contenant α_1, par exemple sont $v_o(\alpha_1),\ldots,v_{k-1}(\alpha_1)$, $v_o(\alpha_1^{p_1}),\ldots,v_o(\alpha_1^{p_s})$. Le produit A_1 de la norme au carré de ces colonnes, d'après le lemme 3.1 vérifie donc $A_1 \leq n^k\, n^{2(1+\ldots+k-1)}\, n^s (\mathrm{Max}(1,|\alpha_1|))^{2n(k+p_1+\ldots+p_s)}$. De la majoration d'Hadamard on obtient :

$$(1)\quad |V(\alpha,d',s,k)|^2 \leq n^{(k^2+s)d'} (M(\alpha))^{2n(k+p_1+\ldots+p_s)} \quad \text{avec } n=d'(k+s).$$

b) minoration :

i) On remarque que $V^2(\alpha,d',s,k)\in\mathbb{Z}$ car V^2 est une fonction symétrique de $\alpha_1,\alpha_2,\ldots\alpha_{d'}$; en outre $V^2\neq 0$ car le lemme 3.2 montre que les facteurs de V sont soit de la forme $\alpha_i^{p_i}-\alpha_j^{p_j}$ avec $i\neq j$ et $\alpha_i^{p_i}-\alpha_j^{p_j}\neq 0$ d'après le corollaire 1, soit de la forme $\alpha_i^{p_i}-\alpha_j^{p_j}$ avec $p_i\neq p_j$ donc différent de zéro d'après le lemme 2.1. Donc $V^2(\alpha,d',s,k)\in\mathbb{Z}^*$.

ii) Soit p fixé appartenant à l'ensemble $\{p_1,p_2,\ldots,p_s\}$; d'après le lemme 3.2, $V^2(\alpha,d,s,k,)$ contient le facteur

$$\prod_{i=1}^{d'}\prod_{j=1}^{d'}(\alpha_i^p-\alpha_j)^{2k} = \prod_{i=1}^{d'} P(\alpha_i^p)^{2k} \quad \text{où } P(X)= \prod_{j=1}^{d'}(X-\alpha_j).$$

D'après le lemme 2.3 $P(\alpha_i^p)= (P(\alpha))^{p+p\gamma_i}$ où γ_i est entier algébrique.

Donc $P(\alpha_i^p) = p\,\gamma_i$. Or $\prod_{i=1}^{d'} P(\alpha_i^p)$ est un entier; il est donc divisible par $p^{d'}$ et V^2 est divisible par $p^{2kd'}$. De i) et ii) on obtient donc

$$|V(\alpha,d',s,k)|^2 \geq \Big(\prod_{i=1}^{s} p_i\Big)^{2d'k}. \tag{2}$$

c) <u>conclusion</u> :

En combinant (1) et (2) on obtient

$$\Big(\prod_{i=1}^{s} p_i\Big)^{2k} \leq n^{(k^2+s)}\, M(\alpha)^{2(k+s)(k+p_1+\ldots+p_s)}.$$

D'où $$\log M(\alpha) \geq \frac{2k\sum_{i=1}^{s} \log p_i - (s+k^2)\log(d'(k+s))}{2(k+s)(k+p_1+\ldots+p_s)} \qquad \forall \alpha \in \Omega_d.$$

Puisque $d' \leq d$ $\forall \alpha \in \Omega_d$, on obtient

$$\log C(d) \geq \frac{2k\sum_{i=1}^{s} \log p_i - (s+k^2)\log(d(k+s))}{2(k+s)(k+\sum_{i=1}^{s} p_i)} \qquad \forall k \in \mathbb{N}^*. \tag{3}$$

Pour d assez grand on pose $r = \log d$, on choisit alors $k = [r/\log r]$ et $s = [(r/\log r)^2/2]$; du théorème des nombres premiers on a

$$\sum_{i=1}^{s} \log p_i = s \log s(1+o(1))$$

$$\sum_{i=1}^{s} \log p_i = \frac{1}{2} s^2 \log s(1+o(1)).$$

L'inégalité (3) devient alors :

$$\log C(d) \geq \frac{2ks\log s - (k^2+s)\log(d(k+s))}{2(k+s)(k+s^2\log s/2)} \qquad (1+\sigma(1));$$

soit

$$\log C(d) \geq \frac{2r^3/(\log r)^2 - (3/2)r^3/(\log r)^2}{2.\frac{1}{2}(r/\log r)^2.\frac{1}{4}(r^4/(\log r)^3)} \qquad (1+\sigma(1));$$

d'où

$$\log C(d) \geq 2(\log r/r)^3(1+o(1)) = 2(\log\log d/\log d)^3(1+o(1))$$

et

$$C(d) \geq 1 + 2(1+o(1))(\log\log d/\log d)^3.$$

BIBLIOGRAPHIE

[1] D.C. Cantor et E.G. Straus.- On a conjecture of D.H. Lehmer, Acta Arith., Warszawa, t. 42, 1982, 97-100.

[2] D.C. Cantor et E.G. Straus.- Correction to the paper "On a conjecture of D.H. Lehmer", 1983.

[3] E. Dobrowolski.- On a question of Lehmer and the number of irreducible factors of a polynomial, Acta Arith., Warszawa, t. 34, 1979, 391-401.

[4] L. Kronecker.- Zwei Sätze über Gleichungen mit ganzzahligen Coefficienten, J. für reine und angew. Math., t. 53, 1857, 173-175.

[5] D.H. Lehmer.- Factorization of certain cyclotomic functions, Annals of Math. Series 2, t. 34, 1933, 461-479.

[6] R. Louboutin.- Sur la mesure de Malher d'un nombre algébrique, Comptes rendus, 296, serie I, 1983, 707-708.

[7] C. Meray.- Sur un déterminant dont celui de Vandermonde n'est qu'un cas particulier, Revue de Math. Spé., 9 (1899), 217-219.

[8] M. Mignotte.- Entiers algébriques dont les conjugués sont proches du cercle unité, Séminaire D.P.P., Théorie des Nombres, 19ème année, 1977-78, n°39.

[9] C.J. Smyth.- Algebraic integers whose conjugates the unit circle, Bull. London Math. Soc., t. 3, 1971, 169-175.

M. PATHIAUX-DELEFOSSE
Département de Mathématiques
Université de Paris 6
4, place Jussieu
75005 PARIS

Propriété de grands nombres de Pisot

M. Pathiaux-Delefosse

I - Introduction.

Soit σ un nombre de Salem; Boyd [2] a montré qu'il existe des nombres θ de Pisot, arbitrairement grands, tels que si P est le polynôme minimal de θ, alors σ est zéro du polynôme $zP+\tilde{P}$ où $\tilde{P}$ est le polynôme réciproque de P.

Réciproquement Salem en 1945 [4] avait prouvé que si P est le polynôme minimal d'un nombre de Pisot alors le polynôme $zP+\tilde{P}$ admet un unique zéro σ dans $\{z\in\mathbb{C}, |z|>1\}$ et σ est un nombre de Salem. Le but de cet exposé est de déterminer des inégalités liant θ et σ. En notant S_{u_o} l'ensemble des nombres de Pisot dont le polynôme minimal a un terme constant égal à u_o, on montre en particulier le théorème suivant :

Théorème 1. Il existe une constante C, calculable, telle pour tout $u_o\in\mathbb{N}^*$ et pour tout $\theta\in S_{u_o}$, de degré s on ait :

$$\theta-u_o \geq C\,(\text{loglogs}/\text{logs})^6.$$

On notera $P= u_o+\ldots+z^s$ le polynôme minimal de θ, nombre de Pisot

$\frac{P}{\tilde{P}}= u_o+u_1z+\ldots \in \mathbb{Z}[[z]]$ la série de Taylor de $\frac{P}{\tilde{P}}$ en zéro,

$R= zP+\tilde{P}$, σ le nombre de Salem zéro de R, $\theta= \theta_1,\theta_2,\ldots,\theta_s$ les conjugués de θ.

II - Inégalités liant θ, u_o, u_1, σ.

Lemme. Si $u_o\leq -1$ alors $\theta<\sigma$, et si $u_o\geq 1$ alors $\sigma<\theta$.

Preuve : On a $R(o)= \tilde{P}(o)=1$, $R(\frac{1}{\theta})= \frac{1}{\theta}P(\frac{1}{\theta})$; or $|u_o|= \theta|\theta_2|\ldots|\theta_s|$. Donc $|u_o| \leq \theta|\theta_i|$ et $|\theta_i| \geq \frac{|u_o|}{\theta} \geq \frac{1}{\theta}$ pour $2\leq i\leq s$. Donc P n'a pas de zéro dans $D(o,\frac{1}{\theta})$ et $P(\frac{1}{\theta})$ est du signe de $P(o)=u_o$.

D'où $u_o<o$ entraîne $R(\frac{1}{\theta})<o$ et $\frac{1}{\sigma}<\frac{1}{\theta}$ et $u_o>o$ entraîne $\frac{1}{\theta}<\frac{1}{\sigma}$.

Lemme 1.1. Soit $\theta\in S_{u_o}$ où $u_o\geq 2$, on a les inégalités suivantes :

i) $u_o \leq \theta$

ii) $$\frac{u_o(u_1-(u_o^2-1))}{(u_o+1)^2} \leq \theta-u_o \leq \frac{u_o(u_1-(u_o^2-1))}{(u_o-1)^2}.$$

Preuve : Soit g la fonction définie par $g(z)=\frac{P(z)}{\tilde{P}(z)}\frac{1-\theta z}{\theta-z}$; g est analytique sur $D(o,1)$; en outre $|g(z)|=1$ si $|z|=1$; donc $|g(o)|\leq 1$ et $|g'(o)|\leq 1-|g(o)|^2$; en développant ces 2 inégalités on obtient i) et ii).

Lemme 1.1. Soit $D_2^*= z^2-\frac{u_1}{u_o-1}z+u_o$, alors le polynôme $zD_2^*+\tilde{D}_2^*$ est un polynôme réciproque ayant un unique zéro noté $\frac{1}{\sigma_2^*}$ dans $D(o,1)$ et $\sigma<\sigma_2^*$.

Preuve : La démonstration est due à Boyd [1], p. 1253.

Théorème 2. Soit $\theta\in S_{u_o}$ où $u_o\geq 2$ alors

$$\sigma+\frac{1}{\sigma}-2 \leq \frac{u_1-(u_o^2-1)}{u_o-1} \leq (\theta-u_o)\frac{(u_o+1)^2}{u_o(u_o-1)}.$$

Preuve : On calcule aisément que $zD_2^*+\tilde{D}_2^*=(z+1)(z^2+z\frac{(u_o-1)^2-u_1}{u_o-1}+1)$. σ_2^* défini dans le lemme précédent est le zéro >1 du polynôme $Q= z^2+z\frac{(u_o-1)^2-u_1}{u_o-1}+1$; or $\sigma<\sigma_2^*$. Donc $Q(\sigma)<o$, d'où $\sigma+\frac{1}{\sigma}-2 \leq \frac{u_1-(u_o^2-1)}{u_o-1}$ et en utilisant le lemme 1.1 on obtient le résultat énoncé.

III - Démonstration du théorème 1 et conjecture.

On utilise le résultat de Dobrowolski [2], affirmant qu'il existe une constante $A= 10^{-6}$, telle que pour tout nombre de Salem σ de degré d on ait :

$$\sigma > 1 + A\,(\text{loglog}d/\log d)^3.$$

En remarquant que $d^o\sigma \leq d^o\theta + 1$ et en utilisant le théorème 2, on obtient le théorème 1.

Remarque : Si la conjecture de Lehmer est vérifiée, du théorème 2 on peut alors déduire qu'il existe une constante $B>0$ telle que $\theta-u_o \geq B$ $\forall u_o \in \mathbb{N}^*$, $\forall \theta \in S_{u_o}$; on peut donc émettre la conjecture suivante, plus faible que celle de Lehmer.

Conjecture : Il existe une constante $B>0$ telle que $\theta-u_o \geq B$, $\forall u_o \in \mathbb{N}^*$, $\forall \theta \in S_{u_o}$.

IV - Propriété des éléments de S_{u_o}.

Soit $\theta \in S_{u_o}$ où $u_o \geq 2$. Posons $a_o = u_o+1$ $b_o = u_o-1$ et de façon générale $a_n = \det(I+U_n)$ $b_n = -\det(I-U_n)$ où U_n est la matrice

$$U_n = \begin{bmatrix} u_n & u_{n-1} & \cdots & u_0 \\ u_{n-1} & u_{n-2} & \cdots & 0 \\ \vdots & & & \\ u_0 & 0 & \cdots\cdots & 0 \end{bmatrix}$$

On a alors $a_1 = u_1-(u_o^2-1)$ $b_1 = u_1+u_o^2-1$.
On peut alors démontrer le théorème suivant :

Théorème 3. Soit $\theta \in S_{u_o}$, où $u_o \geq 2$ $d^o\theta = s$, alors

$$a_n > o \quad \forall n \geq s, \qquad a_n = o \quad \forall n \geq s+1$$
$$b_n > 0 \quad \forall n \geq s-1 \qquad b_n = o \quad \forall n \geq s$$

et vérifient

$$\begin{cases} a_n b_{n-2}+a_{n-2}b_n = 2\, a_{n-1}b_{n-1} & (1) \quad 2 \leq n \leq s-2 \\ |P(1)|\, b_{s-3}+|P(-1)|\, a_{s-3} = 2a_{s-2} & (2) \end{cases}$$

On obtient alors de (1)

$$\frac{a_n}{b_{n-1}} = \frac{a_{n-1}}{b_{n-2}}\left(1+\frac{a_n b_{n-2}-a_{n-2}b_n}{a_n b_{n-2}+a_{n-2}b_n}\right) \leq 2\,\frac{a_{n-1}}{b_{n-2}}.$$

D'où

$$\frac{a_{s-2}}{b_{s-3}} \leq 2^{s-3}\,\frac{a_1}{b_o}.$$

Or (2) donne $\frac{1}{2} \leq \frac{a_{s-2}}{b_{s-3}}$ car $P(1) \in \mathbb{Z}$;

donc $$\frac{1}{2^{s-2}} \leq \frac{a_1}{b_o} \quad \text{et} \quad a_1 \geq \frac{u_o - 1}{2^{s-2}}$$

et du théorème (2) on déduit que

$$\theta - u_o \geq \frac{1}{2^{s-2}} \frac{(u_o+1)^2}{u_o(u_o-1)} \underset{u_o \to +\infty}{\simeq} \frac{1}{2^{s-2}}.$$

Cette inégalité est trés mauvaise par rapport à celle du théorème (1), mais dans cette démonstration on n'utilise pas le fait que a_n et b_n sont des entiers et en particulier la congruence $a_n b_{n-2} + a_{n-2} b_n \equiv o \quad (a_{n-1} b_{n-1})$.

On peut penser que l'étude du système diophantien permettrait d'avoir un meilleur résultat.

BIBLIOGRAPHIE

[1] D. Boyd.- Pisot and Salem Numbers in intervals of the real line., Math. of computation, volume 32, n° 144, 1978, 1244-1260.

[2] D. Boyd.- Small Salem Numbers, Duke Math. J. t. 44, n°2, 1977, 315-327.

[3] E. Drobowolski.- On a question of Lehmer and the number of irreducible factors of a polynomial, Acta Arith., t. 34, 1979, 391-401.

[4] R. Salem.- Algebraic numbers and Fourier analysis (Heath Mathematical Monographs).

M. Pathiaux-Delefosse
Departement de Mathématiques
Université de Paris 6
4, place Jussieu
75005 PARIS

REMARKS ON THE CONTINUED FRACTIONS OF ALGEBRAIC NUMBERS

A.J. van der Poorten*

Summary: It is well-known that very little is known about the behaviour of the sequence (x_n) of the notorious $3x+1$–problem

$$\begin{aligned} x_{n+1} &= \tfrac{1}{2}x_n && \text{if } x_n \text{ is even} \\ &= \tfrac{1}{2}(3x_n+1) && \text{if } x_n \text{ is odd.} \end{aligned}$$

In general we do not know whether the sequence cycles, whether it is unbounded... . Of course it is conjectured (and this is supported by experimental evidence) that all positive integers x eventually yield the cycle $1,2,1,2,\ldots$.

Suppose that the real irrational algebraic number α has continued fraction expansion $\alpha = [c_0\,;\,c_1,\,c_2,\ldots\,]$ and define the sequence (α_n) by $\alpha = [c_0\,;\,c_1,\,c_2,\ldots\,,\,\alpha_{n+1}]$. Let $f \in \mathbf{Z}[X]$ be a polynomial without multiple zeros (so $\gcd(f,f')=1$) and with $f(\alpha)=0$. Let $\deg f = r$, and write $f(X) = a_0X^r + \ldots + a_r$.

One can show, quite readily, that to an accuracy of $\approx q_n^{-4}$ one has

$$\alpha_{n+1} \approx \frac{(-1)^{n+1}}{q_n^2}\frac{f'(\frac{p_n}{q_n})}{f(\frac{p_n}{q_n})} - \frac{q_{n-1}}{q_n} + \frac{(-1)^n}{q_n^2}\frac{f''(\frac{p_n}{q_n})}{2f'(\frac{p_n}{q_n})} \tag{1}$$

with, as usual, $[c_0\,;\,c_1,\,c_2,\ldots\,,\,c_n] = p_n/q_n$.

* Work supported by grants from the ARGS (Australian Research Grant Scheme)

Roth's theorem implies that for n sufficiently large (ineffectively) one has

$$c_{n+1} = \left\lfloor \frac{(-1)^{n+1}}{q_n^2} \frac{f'(\frac{p_n}{q_n})}{f(\frac{p_n}{q_n})} - \frac{q_{n-1}}{q_n} \right\rfloor \tag{2}$$

recursively. It is little wonder that we know nothing about the expansion, as a regular continued fraction, of algebraic numbers of degree ≥ 3. We do not know whether the partial quotients are unbounded, or even whether infinitely many are greater than 2. Of course, we conjecture that an α of degree ≥ 3 behaves as does almost every real irrational; but even the experimental information is quite sparse. The formula (1) is accurate effectively if $r < 6$ and (2) is correct effectively (for $n > N(f, \alpha)$) if $r < 4$. In particular, if $\alpha = \sqrt[3]{5}$, say, and $f(X) = X^3 - 5$ one has exactly , for $n \geq 0$

$$\sqrt[3]{5} = [c_0\,;\, c_1,\, c_2,\, \ldots] \text{ and } \; c_{n+1} = \left\lfloor \frac{(-1)^{n+1}}{q_n} \frac{3p_n^2}{p_n^3 - 5q_n^3} - \frac{q_{n-1}}{q_n} \right\rfloor .$$

The approximation (1) yields a computationally convenient method for obtaining the continued fraction of α. My remarks arise, almost entirely, from suggestions of Enrico Bombieri.

Alf van der Poorten
School of MPCE
Macquarie University
NSW 2109
AUSTRALIE

INEGALITES DE TYPE BRUN-TITCHMARSH EN MOYENNE

Bruno ROUSSELET

I - Introduction.

La répartition des nombres premiers dans les progressions arithmétiques est un sujet passionnant et difficile. Ici peut-être plus que partout ail-leurs, il y a un fossé très important entre les hypothèses les plus naturel-les et les résultats démontrés. Notant comme d'habitude $\pi(x;q,a)$ le nombre de nombres premiers inférieurs à x et congrus à a modulo q, la forme même du théorème de Dirichlet laisse espérer l'énoncé suivant :

Pour tout $\epsilon > 0$, on a

$$\pi(x;q,a) \sim_{\epsilon} \frac{x}{\varphi(q)\log x} \quad \text{pour} \quad q \leq x^{1-\epsilon}.$$

Mais ce résultat, conséquence de l'hypothèse de Montgomery, est loin d'être démontré. Il faut beaucoup réduire le domaine de validité, et supposer seulement $q \leq (\log x)^A$, pour pouvoir affirmer l'équivalence, par le théorème de Siegel-Walfisz.

La répartition en moyenne est mieux connue, puisque le théorème de Bombieri-Vinogradov implique que, pour $\epsilon > 0$ et $A > 0$, on a

$$\sum_{\substack{q \leq x^{1/2-\epsilon} \\ (a,q)=1}} |\pi(x;q,a) - \frac{\ell i\, x}{\varphi(q)\log x}| = O_{\epsilon,A}\left(\frac{x}{(\log x)^A}\right)$$

La limite de validité en $x^{1/2-\epsilon}$ est conséquence de l'utilisation du grand crible. Dès que le module q des progressions arthmétiques est de l'ordre de grandeur de $x^{1/2}$, l'hypothèse de Riemann généralisée elle-même ne conduit (apparemment) qu'à une trivialité.

Nous nous proposons de démontrer ici le théorème suivant :

Théorème : Pour tout $A > 0$, il existe x_0 tel que, pour tout $x > x_0$, on ait

$$0{,}85 \frac{x}{\varphi(q)\log x} \leq \pi(x;q,1) \leq 1{,}48 \frac{x}{\varphi(q)\log x}$$

pour tout q de $[Q,2Q[$ avec au plus $Q(\log x)^{-A}$ exceptions, si $x^{\frac{1}{2}} \leq Q \leq x^{\frac{1}{2}+10^{-100}}$.

Dans cet énoncé, on peut remplacer $\pi(x;q,1)$ par $\pi(x;q,a)$ à condition de n'examiner que les q de $[Q,2Q[$ vérifiant $(q,a)=1$.

D'autre part, au prix de calculs plus compliqués que ceux exposés, j'ai pu remplacer 0,85 par 0,8607 et 1,48 par 1,4567.

Pour faciliter la démonstration, on ne présentera la preuve que pour $Q = x^{1/2}$, en suivant pas à pas la arguments, on se convaincra qu'on peut augmenter légèrement l'exposant 1/2, de 10^{-100} par exemple.

Même avec une valeur numérique moins performante, la minoration du théorème n'était pas encore connue. En revanche, des majorations du type de celle du théorème peuvent être trouvées dans [3], [6], [13] par exemple. Mais c'est le travail de Fouvry [7], qui est à l'origine de ce nouveau résultat. En effet, s'affranchissant pour la première fois de l'utilisation du crible, il montre la majoration du théorème avec la valeur numérique 1,73, qui est inférieure à la valeur 2 imposée par le phénomène de parité du crible. C'est cette avancée qui lui permet de montrer, par le critère d'Adleman et Heath-Brown, que le premier cas du grand théorème de Fermat est vrai pour une infinité d'exposants premiers.

Par rapport au travail de Fouvry, le nouvel ingrédient qui apparaît dans la démonstration du théorème est le lemme 8, issu du récent article de Friedlander et Iwaniec [9] basé sur la conjecture de Weil, démontrée par Deligne, et la méthode de Burgess.

II - Applications.

Il semble que la minoration donnée dans le théorème 1 soit de meilleure qualité que la majoration. Nous illustrons cela en reprenant un résultat de Balog [1] : (on note $P^+(n)$ le plus grand facteur premier de n).

Soit $$\pi(x,y) = |\{p \leq x,\ P^+(p-1) \leq y\}|$$

alors $$\pi(x,y) \gg \frac{x}{\log^2 x} \quad \text{pour} \quad y \geq x^{0,347}.$$

La constante 0,347 est ici une valeur approchée de $(2\exp((1+1,73)^{-1}))^{-1}$ où 1,73 est la borne de majoration donnée par Fouvry [7]. Placer ici la nouvelle borne 1,48 conduit au résultat :

$$\pi(x,y) \gg \frac{x}{\log^2 x} \quad \text{pour} \quad y \geq x^{0,335}.$$

On va, pour utiliser la borne de minoration du théorème 1, modifier un peu la méthode de Balog et obtenir le corollaire suivant :

Corollaire : On a $\pi(x,y) \gg \frac{x}{\log^2 x}$ pour $y \geq x^{0,316}$.

On reprend la fonction introduite dans [1],

$$g(p)= |\{p-1=mn;\ P^+(mn) \leq y\ N_1 < n \leq N_2\}|$$

avec maintenant le choix $N_1= x^{1/2-2\delta}$ $N_2= x^{1/2-\delta}$. Suivant la démonstration pas à pas, on retrouve

$$g(p)= |\{p-1=mn;\ P^+(m) \leq y,\ N_1< n \leq N_2\}| - |\{p-1=mp'\ell;\ y < p',\ N_1 < p'\ell \leq N_2\}|$$

d'où $\sum_{p\leq x} g(p) \geq S_1-S_2$, S_1 et S_2 inchangées; avec le même choix $T_1= N_1m+1$, $T_2= \min(x,N_2m+1)$ on a, pour δ choisi très petit,

$$S_1 = \sum_{\substack{m \\ P^+(m)\leq y \\ T_2>T_1}} (\pi(T_2,m,1) - \pi(T_1,m,1))$$

$$\geq 0,85\ \frac{x}{\log x} \sum_{x/N_2<m<x/N_1} \frac{1}{\varphi(m)} + O\left(\frac{x}{\log x}\right)$$

$$\geq 0,85\ \left(1-\log\left(\frac{\log x}{2 \log y}\right)\right)\frac{x}{\log x} \sum_{N_1<n\leq N_2} \frac{1}{\varphi(n)} + O\left(\frac{x}{\log x}\right)$$

puis

$$S_2 = \sum_{y<p'} \sum_{\frac{N_1}{p'}<\ell<\frac{N_2}{p'}} \pi(x;\ell p',1)$$

$$\leq \log\left(\frac{\log x}{2 \log y}\right)\ \frac{x}{\log x} \sum_{N_1<n\leq N_2} \frac{1}{\varphi(n)} + O\left(\frac{x}{\log x}\right)$$

par le théorème de Bombieri-Vinogradov. la condition $S_1-S_2 \gg x$ se traduit alors par

$$\frac{\log y}{\log x} > (2\exp((1+\frac{1}{0,85})^{-1}))^{-1} \simeq 0,316.$$

Toutefois, par une technique différente, Fouvry et Grupp [8] obtiennent, sur un domaine de validité moins étendu pour y, le bon ordre de grandeur de la fonction $\pi(x,y)$, c'est-à-dire

$$\pi(x,y) >>_\epsilon \frac{x}{\log x} \quad \underline{\text{pour}} \quad y \geq x^{\delta_0-\epsilon}$$

<u>avec</u> $\delta_0 = 3/(7\exp(0,3)) \simeq 0,3175.$

III - <u>Lemmes</u>.

Nous rappelons les notations du crible, si $\mathcal{A}$ est une suite finie d'entiers. $\mathcal{P}$ un ensemble de nombres premiers, pour $z\geq 2$ et $d\in\mathbb{N}^*$, on note

$$P(z) = \prod_{p\in\mathcal{P}, p<z} p$$

$$\mathcal{A}_d = \{a\in\mathcal{A},\ a\equiv 0[d]\}$$

$$S(\mathcal{A},\mathcal{P},z) = |\{a\in\mathcal{A},\ (a,P(z))=1\}|.$$

Le premier lemme s'obtient par itération de l'identité de Buchstab [11].

<u>Lemme 1</u> : Pour $R\in\mathbb{N}^*$, on a

$$S(\mathcal{A},\mathcal{P},z) + \sum_{r=1}^{R}(-1)^{r+1} \sum_{\substack{D_r\leq p_r<\ldots<p_1<z \\ p_i<D_i(1\leq i\leq r-1)}} S(\mathcal{A}_{p_1\ldots p_r},\mathcal{P},p_r)$$

$$= |\mathcal{A}| + \sum_{r=1}^{R-1}(-1)^r \sum_{\substack{p_r<\ldots<p_1<z \\ p_i<D_i(1\leq i\leq r)}} |\mathcal{A}_{p_1,\ldots,p_r}| + (-1)^R \sum_{\substack{p_r<\ldots<p_1<z \\ p_i<D_i(1\leq i\leq R)}} S(\mathcal{A}_{p_1,\ldots,p_R},\mathcal{P},p_R)$$

où les p_i sont des éléments de $\mathcal{P}$ et les D_i des fonctions arbitraires de $p_1,\ldots p_{i-1}$.

Supposons alors l'égalité :

$$|\mathcal{A}_d| = \frac{\omega(d)}{d} X + r(\mathcal{A},d)$$

avec $X > 1$ et $\omega(d)$ multiplicative. Notant alors

$$V(z)= \prod_{\substack{p<z \\ p\in\mathcal{P}}} (1 - \frac{\omega(p)}{p})$$

on fait l'hypothèse suivante :

(H) $$\frac{V(w)}{V(z)} \leq \frac{\log z}{\log w} (1 + \frac{K}{\log w}) \quad \text{pour} \quad 2 \leq w \leq z.$$

On peut énoncer le lemme suivant :

Lemme 2 (lemme fondamental) [12] : Pour $D \geq 2$, $z \geq 2$, notons $s= \frac{\log D}{\log z}$. Sous l'hypothèse (H), il existe deux suites de coefficients μ_d^+ et μ_d^- telles que

$$\mu_1^{\pm}=1, \ |\mu_d^{\pm}| \leq 1 \ d>D \Longrightarrow \mu_d^{\pm}=0, \ \mu^- * 1 \leq \mu * 1 \leq \mu^+ * 1$$

et vérifiant, si $s \geq 3$

$$\sum_{d|P(z)} \mu_d^+ \frac{\omega(d)}{d} \leq V(z) \ \{1+O(e^{-s}(1+e^{\sqrt{K}}(\log D)^{-1/3}))\}$$

$$\sum_{d|P(z)} \mu_d^- \frac{\omega(d)}{d} \geq V(z) \ \{1+O(e^{-s}(1+e^{\sqrt{K}}(\log D)^{-1/3}))\}$$

Les énoncés qui vont maintenant être exposés concernent la répartition de certaines suites dans les progressions arithmétiques. Par convention, $\alpha_m, \beta_n, \ldots$ auront leur support inclus dans des intervalles $[M,M'[$, $[N,N'[, \ldots$ avec $M' \leq 2M$, $N' \leq 2N, \ldots$ et les constantes intervenant dans les majorations seront indépendantes de M', $N', \ldots$

On dit qu'une suite ζ_k est *d'ordre* K si on a l'égalité $|\zeta_k| \leq \tau_K(k)$. Toutes les suites $(\alpha_m), (\beta_n), \ldots$ qui apparaîtront sont des suites d'ordre K. On note K_1 une fonction de K uniquement.

par convention,

1_n est la fonction caractéristique des entiers de $[N,N'[$

1_p est la fonction caractéristique des nombres premiers de $[P,P'[$.

Introduisons enfin, pour (α_m) donnée, l'erreur à étudier :

$$E((\alpha_m);q,a)= \sum_{m\equiv a[q]} \alpha_m - \frac{1}{\varphi(q)} \sum_{(m,q)=1} \alpha_m .$$

Le lemme suivant donne une majoration triviale qui sera fréquemment utilisée pour la mise en forme des expressions rencontrées.

Lemme 3 : On a, pour $K > a$ et $K > 2$

$$\sum_q \tau_q E(\zeta_k;q,a)= O_{K,a}(K^{1/2}(\sum_{\zeta_k\neq 0} 1)^{1/2} (\log QK)^{K_1})$$

et si $Q \leq K^{0,9}$

$$\sum_q \tau_q E(\zeta_k;q,a)= O_{K,a}(KQ^{-1/2}(\sum_{\tau_q\neq 0} 1)^{1/2} (\log KQ)^{K_1})$$

Ces deux majorations sont conséquences de l'inégalité de Hölder, la seconde nécessitant le résultat classique de Shiu [14].

Ecrivons maintenant l'hypothèse :

Pour tout $B > 0$, on a, pour $d \geq 1$, $q \geq 1$, $(q,a)=1$

$$(S.W) \qquad \sum_{\substack{k\equiv a[q]\\(k,q)=1}} \zeta_k = \frac{1}{\varphi(q)} \sum_{(k,qd)=1} \zeta_k + O_{K,B}(K \tau_K(d)(\log K)^{-B}).$$

On a alors les lemmes suivants :

Lemme 4 : [7]

Si (β_n) vérifie (S.W), si il existe $\epsilon > 0$ tel que

$$\max((MN)^{\epsilon}, (MN)^{\epsilon-1}Q^2) \leq N \leq (MN)^{5/6-\epsilon} Q^{-4/3}$$

alors, pour tout $A > 0$, pour tout $a \leq (\log MN)^A$

$$\sum_{(q,a)=1} \tau_q E(\alpha_m * \beta_n;q,a) = O_{\epsilon,A,K}(MN(\log MN)^{-A}).$$

Lemme 5 : [7]

Si (β_n) vérifie (S.W), si $p \leq \exp(\log^{1/4}(LMN))$ et $p|\ell n$ impliquent $\beta_n \lambda_\ell = 0$, si il existe $\epsilon > 0$ tel que

$$L^3N^2 \leq (LMN)^{1-\epsilon}Q,\ L^2N^5+L^3N^4 \leq (LMN)^{2-\epsilon} \quad \text{et} \quad LN \geq (LMN)^{\epsilon}Q$$

alors, pour tout $A > 0$, pour tout $a \leq \log(LMN)^A$

$$\sum_{(q,a)=1} \tau_q\, E(\alpha_m * \beta_n * \lambda_\ell; q, a) = O_{\epsilon,A,K}(LMN(\log LMN)^{-A}).$$

Lemme 6 : [7]

Si $(\beta_n) = (1_p)$, si il existe $\epsilon > 0$ tel que

$$Q(MP)^{\epsilon} \leq P \leq (MP)^{-\epsilon}\min((MP)^{3/8}Q^{3/8},\ (MP)^{4/7})$$

alors, pour tout $A > 0$,

$$\sum_{(q,a)=1} \tau_q\, (E(\alpha_m * 1_p; q, 1)) = O_{\epsilon,A,K}(MP(\log MP)^{-A})$$

Lemme 7 :

Si il existe $\epsilon > 0$ tel que

$$(LMN)^{1-\epsilon} \geq \max(MQ,\ L^3N^{-1}Q,\ QM^{1/2}L,\ ML^2N^{-2}Q)$$

alors, pour tout $A > 0$, pour tout $a \leq (\log LMN)^A$

$$\sum_q \tau_q\, E(\alpha_m * 1_n * \lambda_\ell; q, a) = O_{\epsilon,A,K}(LMN(\log LMN)^{-A}).$$

Ce lemme fait la synthèse des lemmes 6 et 7 de [7]. Nous donnons maintenant le lemme nouveau, provenant du travail de Friedlander et Iwaniec [9]. La première partie est leur proposition 1, la seconde s'obtient en sommant sur m et q.

Lemme 8 : Pour $N_1, N_2, N_3, Q, M \geq 1$, pour tout $\epsilon > 0$, pour $q \in \mathbb{N}^*$ quelconque, on a

$$E(1_{n_1} * 1_{n_2} * 1_{n_3}; q, 1) = O_\epsilon((N_1N_2N_3)^{46/75+\epsilon}q^{-7/30})$$

$$\sum_q \tau_q\, E(\alpha_m * 1_{n_1} * 1_{n_2} * 1_{n_3}; q, 1) = O_{K,\epsilon}(M^{1+\epsilon}(N_1N_2N_3)^{46/75+\epsilon}Q^{23/30+\epsilon}).$$

Nous aurons enfin besoin de deux derniers lemmes :

Lemme 9 : (Identité de Heath-Brown) [10]

Pour tout $J \geq 1$, $y > 0$ et $n < 2y^J$, on a

$$\Lambda(n) = \sum_{j=1}^{J} (-1)^j C_J^j \sum_{\substack{m,m'\ldots m^{(j)} \\ m^{(i)} \leq y}} \mu(m)\ldots\mu(m^{(j)}) \sum_{\substack{v,v'\ldots v^{(j)} \\ m\ldots m^{(j)} v\ldots v^{(j)}=n}} \log v.$$

Lemme 10 : [7]

Soit $w(u)$ définie par
$$w(u) = 0 \quad \text{sur} \quad [0,1[$$
$$w(u) = 1 \quad \text{sur} \quad [1,2[$$
$$w'(u) = \frac{w(u-1)}{u-1} \quad \text{pour} \quad u \geq 2$$

alors, pour tout $\epsilon > 0$ et $u \in]0,5]$, on a, pour $x > x_0(\epsilon)$

$$|\{n \leq x,\ p|n \Longrightarrow p > x^{1/u}\}| \leq (w(u) \pm \epsilon) \frac{x}{\log x}.$$

IV - Application du lemme fondamental.

Pour étudier la fonction $\pi(x;q,1)$, il est naturel de chercher les nombres premiers dans la suite

$$\mathcal{A}^{(q)} = \{qk+1 \leq x\}$$

à laquelle nous associons la suite

$$\mathcal{B}^{(q)} = \{n \leq x,\ (n,q)=1\}.$$

Cependant, il est plus commode, comme souvent en théorie du crible, de travailler avec des suites légèrement criblées; c'est pourquoi nous utiliserons

$$\tilde{\mathcal{A}}^{(q)} = \{a \in \mathcal{A}^{(a)},\ p|a \Rightarrow p > x^{\eta^2}\}$$

et
$$\tilde{\mathcal{B}}^{(q)} = \{b \in \mathcal{B}^{(q)},\ p|b \Rightarrow p > x^{\eta^2}\}$$

où η est positif et très petit.

Appliquons le lemme 1 aux suites $\tilde{\mathcal{A}}^{(q)}$ et $\tilde{\mathcal{B}}^{(q)}$, en choisissant $R = [\eta^{-2}+1]$, de sorte que le dernier terme de la partie droite des égalités est nul. Les D_i seront fixés par la suite, vérifiant seulement

(4.1) $$D_1,\dots,D_r < x^{1-3\eta}.$$

On choisit enfin $\mathcal{P}= \mathcal{P}^{(q)}= \{p \mid q\}$ et $z= \sqrt{x}$.

Ainsi nous avons les égalités

$$S(\tilde{\mathcal{A}}^{(q)},\mathcal{P}^{(q)},z)= \pi(x;q,1) + 1$$

$$S(\tilde{\mathcal{B}}^{(q)},\mathcal{P}^{(q)},z)= \pi(x) + O(\frac{x^{1/2}}{\log x}).$$

L'expression

$$S(\tilde{\mathcal{A}}^{(q)},\mathcal{P}^{(q)},z)- \frac{1}{\varphi(q)} S(\tilde{\mathcal{B}}^{(q)},\mathcal{P}^{(q)},z)= \pi(x;q,1)- \frac{\pi(x)}{\varphi(q)} + O(1)$$

devient

(4.2)

$$\pi(x;q,1) - \frac{\pi(x)}{\varphi(q)} + O(1) + \sum_{r=1}^{R} (-1)^{r-1}(S_r^{(q)}- \frac{T_r^{(q)}}{\varphi q})= \sum_d \lambda_d^{(q)}(|\tilde{\mathcal{A}}_d^{(q)}|- \frac{|\tilde{\mathcal{B}}_d^{(q)}|}{\varphi(q)})$$

avec les notations

$$\lambda_d^{(q)}= \begin{cases} 1 & \text{si } d=1 \\ (-1)^{\ell} & \text{si } d= p_1\dots p_\ell,\ x^{\eta^2} \le p_\ell <\dots< p_1< z, p_i < D_i \quad (1\le i\le \ell) \\ 0 & \text{sinon} \end{cases}$$

$$S_r^{(q)}= \sum_{\substack{D_r\le p_r<\dots<p_1<z \\ p_\ell<D_\ell\ (1\le \ell<r)}} S(\tilde{\mathcal{A}}^{(q)}_{p_1\dots p_r}, \mathcal{P}^{(q)}, p_r)$$

$T_r^{(q)}$ représentant la même somme que $S_r^{(q)}$ portant sur $\tilde{\mathcal{B}}^{(q)}$, et tous les p_i étant éléments de $\mathcal{P}^{(q)}$. Notons que (4.1) entraîne

$$d \ge x^{1-3\eta} \Longrightarrow \lambda_d^{(q)}= 0.$$

Pour $\omega(d)= \omega^{(q)}(d)=1$ si $(d,q)=1$ et 0 sinon, on a, pour $2 \le w \le z$, l'inégalité :

$$\frac{V^{(q)}(w)}{V^{(q)}(z)} = \prod_{\substack{w\le p<z \\ p\in\mathcal{P}^{(q)}}} (1 - \frac{\omega^{(q)}(p)}{p})^{-1} \le \prod_{w\le p<z} (1 - \frac{1}{p})^{-1}$$

et donc $$\frac{V^{(q)}(w)}{V^{(q)}(z)} \le \frac{\log z}{\log w} (1 + \frac{K}{\log w})$$

avec K constante absolue.

Pour $D= x^{\eta}$, introduisons les suites μ_r^+ et μ_r^- du lemme 2. Il vient l'inégalité

$$\sum_{\lambda_d^{(q)}=\pm 1} (\pm 1) \sum_{r\,|\,P^{(q)}(x^{\eta^2})} \mu_r^{\mp}(|\mathcal{A}_{dr}^{(q)}| - \frac{|\mathcal{B}_{dr}^{(q)}|}{\varphi(q)}) - \sum_{\lambda_d^{(q)}=\pm 1} \sum_{r\,|\,P^{(q)}(x^{\eta^2})} (\mu_r^+ - \mu_r^-) \frac{|\mathcal{B}_{dr}^{(q)}|}{\varphi(q)}$$

(4.3) $$\le \sum_d \lambda_d^{(q)} (|\tilde{\mathcal{A}}_d^{(q)}| - \frac{|\tilde{\mathcal{B}}_d^{(q)}|}{\varphi(q)})$$

$$\le \sum_{\lambda_d^{(q)}=\pm 1} (\pm 1) \sum_{r\,|\,P^{(q)}(x^{\eta^2})} \mu_r^{\pm}(|\mathcal{A}_{dr}^{(q)}| - \frac{|\mathcal{B}_{dr}^{(q)}|}{\varphi(q)} + \sum_{\lambda_d^{(q)}=\pm 1} \sum_{r\,|\,P^{(q)}(x^{\eta^2})} (\mu_r^+ - \mu_r^-) \frac{|\mathcal{B}_{dr}^{(q)}|}{\varphi(q)}$$

Un calcul facile nous donne :

$$|\mathcal{B}_{dr}^{(q)}| = \frac{\varphi(q)}{qdr} x + O(x^{\eta})$$

et le lemme 2 permet de conclure : pour $\epsilon > 0$ on peut prendre η assez petit pour que

(4.4) $$\sum_{\lambda_d^{(q)}=\pm 1} \sum_{r\,|\,P^{(q)}(x^{\eta^2})} (\mu_r^+ - \mu_r^-) \frac{|\mathcal{B}_{dr}^{(q)}|}{\varphi(q)} \le \frac{\epsilon\, x}{\varphi(q)\log x}.$$

Grâce à (4.3) et (4.4), pour démontrer que la partie droite de (4.2) est petite en moyenne, c'est-à-dire inférieure en valeur absolue à

$$\frac{2\,\epsilon\, x}{\varphi(q)\log x}$$

avec, pour tout $A > 0$, au plus $x^{1/2}(\log x)^{-A}$ exceptions pour q dans l'intervalle $[x^{1/2},2x^{1/2}[$, pour $x > x_0(\epsilon,A)$, il suffit de montrer la majoration

$$(4.5)\qquad \sum_{x^{1/2}\le q<2x^{1/2}} \tau_q\Big(\sum_{rp_1\dots p_\ell} \sum_{s\equiv 1[q]} \mu_r - \frac{1}{\varphi(q)} \sum_{(rp_1\dots p_\ell s,q)=1} \mu_r\Big) = O_{\eta,A}(x\,\mathcal{L}^{-A-2})$$

où les variables $r,p_1,\dots,p_\ell$ et s vérifient en outre les conditions $rp_1\dots p_\ell s \le x$, $x^{\eta^2} \le p_\ell \dots < p_1$, $p_i < D_i$ $(1\le i\le \ell)$, $r \le x^\eta$ et $r\,|\,P(x^{\eta^2})$, et où $\mathcal{L}= \log x$, ceci pour $\ell \le R$ et pour toutes suites τ_q et μ_r satisfaisant à $|\tau_q| \le 1$ et $|\mu_r| \le 1$.

On donne maintenant les valeurs des D_i : pour la majoration

$$(4.6)\qquad \begin{aligned} &D_1= D_2= x^{1/3-3\eta} \qquad D_3= D_4= x^{1/6-\eta} \\ &\forall\, k \ge 5 \qquad D_k= \frac{x^{0,9}}{p_1\cdots p_{k-1}} \end{aligned}$$

pour la minoration

$$(4.7)\qquad \begin{aligned} &D_1= x^{1/2} \qquad D_2= D_3= \max\Big(x^{1/6-\eta}, \frac{x^{1/2-3\eta}}{p_1}\Big) \\ &D_4= x^{1/6-\eta}, \quad \forall k \ge 5 \quad D_k= \frac{x^{0,9}}{p_1\cdots p_{k-1}}\,. \end{aligned}$$

Il est nécessaire de rendre, dans (4.5), les variables de sommation $r,p_1,\dots,p_\ell,s$ indépendantes pour appliquer les lemmes 4, 5, 7 et 8. La technique employée est celle de [2], [5], et consiste à recouvrir, presque totalement, le domaine de sommation de $(r,p_1,\dots p_\ell,s)$ par des parallélotopes de la forme

$$[R,R\Delta[\times [P_1\Delta[\times \dots \times [P_\ell, P_\ell\Delta[\times [S,S\Delta[$$

avec

$$\Delta= 1 + \mathcal{L}^{-A_1} \qquad (A_1= A_1(A) \text{ suffisamment grand})$$

où R, S, $P_1\, x^{-\eta^2}, \ldots, P_\ell\, x^{-\eta^2}$ sont des puissances entières de Δ vérifiant

$$R\Delta \leq x^{\eta},\ x^{\eta^2} \leq P_1 < \ldots < P_\ell,\ P_i\Delta < D_i(P_1 \ldots P_{i-1}) \quad (1 \leq i \leq \ell)$$

$$\text{et} \quad RP_1 \ldots P_\ell S\, \Delta^{\ell+2} \leq x$$

Le nombre de tels parallélotopes est $O\ (\mathfrak{L}^{(A_1+1)(\ell+2)})$.

La contribution à la somme (4.5) de la partie non recouverte est négli geable, c'est-à-dire que la première partie du lemme 3 donne la majoration

$$O(x\, \mathfrak{L}^{-A+2})$$

car la suite ξ_k du lemme 3 est alors non nulle sur $O(x\, \mathfrak{L}^{-A_1'})$ entiers, où A_1' est une fonction de A_1 tendant vers l'infini avec A_1.

En conclusion, pour démontrer (4.5), il suffit, avec les notations du paragraphe III, d'établir l'estimation :

$$(4.8) \qquad \forall\, B > 0 \qquad \sum_q \gamma_q\, E(\mu_r * 1_{p_1} * \ldots * 1_{p_\ell} * 1_s; q, 1) = O\ (x\, \mathfrak{L}^{-B})$$

où

$$(4.9) \qquad R \leq x^{\eta}, x^{\eta^2} \leq P_\ell < \ldots < P_1,\ P_i < D_i\ (P_1 \ldots P_{i-1}) \quad (1 \leq i \leq \ell)$$
$$\text{et} \quad RP_1 \ \ldots \ P_\ell\, S \leq x.$$

V - <u>Démonstration de (4.8)</u>.

a) <u>Cas de la majoration</u>.

Les D_i ont les valeurs fixées en (4.6). Les cas $\ell=0$ et $\ell=1$ de (4.8) ne présentent aucune difficulté. Pour $\ell \geq 2$, on se ramène d'abord, en introduisant la fonction Λ, à montrer

$$(5.1) \quad \forall\, B > 0,\ \sum_q \gamma_q\, E(\mu_r * \Lambda(n_1) * \Lambda(n_2) * 1_{p_3} * \ldots * 1_{p_\ell} * 1_s; q, 1) = O(x \mathfrak{L}^{-B})$$

où $N_1 = P_1$, $N_2 = P_2$, et les P_i vérifiant toujours (4.9).

On applique le lemme 9 aux quantités $\Lambda(n_1)$ et $\Lambda(n_2)$ avec J=2 et $y= x^{1/6-\eta}$, les facteurs log sont éliminés par une sommation par parties; en conclusion, la démonstration de (5.1) se ramène à :

$$\forall\, B > 0 \quad \sum_q \gamma_q E(\mu_r * \mu(m_1) * \mu(m_1') * \mu(m_2) * \mu(m_2') * 1_{h_1} * 1_{h_1'} * 1_{h_2} * 1_{h_2'} * 1_{p_3} * \ldots * 1_{p_\ell} * 1_s ; q, 1)$$

$$= O_{B,\eta}(x\, \ell^{-B}) \tag{5.2}$$

avec $R \leq x^{\eta}$, $x^{\eta^2} \leq P_\ell < \ldots < P_3$ $M_1, M_1', M_2, M_2' < x^{1/6-\eta}$

$$P_i \leq M_i M_i' H_i H_i' \leq 2 P_i \qquad M_i' \leq M_i \qquad H_i' \leq H_i \qquad (i= 1,2)$$

$$P_i < D_i(P_1, \ldots, P_{i-1}) \quad (1 \leq i \leq \ell) \quad R M_1 M_1' M_2 M_2' H_1 H_1' H_2 H_2' P_3 \ldots P_\ell S \leq x \tag{5.3}$$

Remarquons que pour parvenir à (5.2) on a rendu indépendantes les variables, grâce au lemme 3, et que l'on retrouve les expressions j=1 du lemme 9 en faisant $M_1'= H_1'= 1$ ou $H_2'= M_2'= 1$.

La démonstration de (5.2) est alors une étude soigneuse des inégalités (5.3) pour construire des variables convenant à l'application des lemmes 4 ou 8. On note que la condition (S.W) est vérifiée par les suites $\mu(m)$, 1_n, 1_p et aussi par leurs produits de convolution.

On envisage deux situations :

* Si $M_1 M_1' M_2 M_2' H_1 H_1' H_2 H_2' P_3 \ldots P_\ell$ a un produit partiel dans l'intervalle $[x^{\eta}, x^{1/6-\eta}]$, alors, en prenant pour (β_n) la convolution des suites associées, le lemme 4 entraîne (5.2).

* Dans le cas contraire, les inégalités (5.3) impliquent qu'on a les relations

$$R \leq x^{\eta} \qquad M_1 M_1' M_2 M_2' H_1 H_2' P_3 \ldots P_\ell \leq x^{\eta}$$

et qu'on est ainsi ramené à étudier l'estimation

$$\forall\, B > 0 \quad \sum_q \gamma_q\, E(\mu_r' * 1_{h_1} * 1_{h_2} * 1_s ; q, 1) = O_{B,\eta}(x\, \ell^{-B}) \tag{5.4}$$

où μ_r', est d'ordre $\ell+5$ et $R' \leq x^{2\eta}$.

Le lemme 8 donne alors (5.4), ce qui termine la démonstration de (4.8) dans ce cas.

L'utilisation de ce lemme 8 a permis de porter la valeur de D_1 de $x^{3/10-3\eta}$ [7] à $x^{1/3-3\eta}$, ce qui est la cause véritable de l'amélioration de la borne 1,73 de [7].

b) Cas de la minoration.

Les D_i ont les valeurs fixées en (4.6). Nous supposons d'abord $\ell \geq 2$, et suivons une méthode proche de celle du paragraphe précédent. Nous appliquons le lemme 9 avec J=4 et $y = x^{1/6-\eta}$ à la variable $\Lambda(n_1)$ uniquement ; nous sommes alors conduits à montrer

$$\forall B > 0 \quad \sum_q \tau_q E(\mu_r * \mu(m_1) * \mu(m_1') * \mu(m_2'') * \mu(m_1''') * 1_{h_1} * 1_{h_1'} * 1_{h_1''} * 1_{h_1'''} * 1_{p_r} * \ldots 1_{p_\ell} * 1_s, q, 1)$$

$$= O_{B,\eta}(x, \mathcal{L}^{-B}) \tag{5.5}$$

où $R < x^\eta \quad M_1''' \leq M_1'' \leq M_1' \leq M_1 < y \quad H_1''' \leq H_1'' \leq H_1' \leq H_1$

$$P_1 < M_1 M_1' M_1'' M_1''' H_1 H_1' H_1'' H_1''' < 2P_1 \quad P_i < D_i(P_1, \ldots, P_{i-1}) \quad (1 \leq i \leq \ell)$$

$$\text{et} \quad R P_1 \ldots P_\ell S \leq x \tag{5.6}$$

Nous envisageons différents cas

* Si $M_1 M_1' M_1'' M_1''' H_1 H_1' H_1'' H_1''' P_2 \ldots P_\ell$ a un produit partiel dans $[x^\eta, x^{1/6-\eta}]$, le lemme 4 entraîne (5.5).

* Dans le cas contraire, l'étude des inégalités (5.6) prouve la situation suivante

$$M_1 M_1' M_1'' M_1''' H_1''' P_2 \ldots P_\ell \leq x^\eta$$

car l'inégalité

$$H_1''' \geq x^{1/6-\eta}$$

est impossible, ce qui nous ramène à démontrer

$$\forall B > 0, \ \sum_q \tau_q E(\mu_r' * 1_{h_1} * 1_{h_1'} * 1_{h_1''} * 1_{p_2} * 1_s; q, 1) = O_{B,\eta}(x \mathcal{L}^{-B}) \tag{5.7}$$

où μ_r', est d'ordre $\ell+5$ et $R' \leq x^{2\eta}$.

Comme nous avons déjà exclu la possibilité pour P_2 d'être dans $[x^{\eta}, x^{1/6-\eta}]$, supposons tout d'abord que

$$P_2 > x^{1/6-\eta}$$

alors, d'après la définition de D_2, on a

$$P_1 P_2 < x^{1/2-3\eta}.$$

Le lemme 7 prouve (5.7), les quantités N,M,L ayant respectivement pour valeurs S, $H_1 H_1' H_1'' P_2, R'$.

Supposons maintenant que l'on ait

$$P_2 < x^{\eta}.$$

L'expression (5.7) prend alors l'une des formes suivantes

$$(5.8) \qquad \forall\, B > 0,\ \sum_q \gamma_q\, E(\mu''_{r''} * 1_{h_1} * 1_{h_1'} * 1_{h_1''} * 1_s, q, 1) = O_{B,\eta}(x\, \mathcal{L}^{-B})$$

où $\mu''_{r''}$ est d'ordre $\ell+6$ et $R'' \leq x^{3\eta}$

ou

$$(5.9) \qquad \forall\, B > 0,\ \sum_q \gamma_q\, E(\mu'''_{r'''} * 1_{h_1} * 1_{h_1'} * 1_s; q, 1) = O_{B,\eta}(x\, \mathcal{L}^{-B})$$

où $\mu'''_{r'''}$ est d'ordre $\ell+7$ et $R''' \leq x^{4\eta}$

suivant que $H_1'' > x^{1/6-\eta}$ ou $H_1'' < x^{\eta}$.

Le lemme 8 prouve immédiatement (5.9).

Par application du lemme 7, avec les quantités N,M,L valant respectivement S, $H_1 H_1', R'' H_1''$, on prouve (5.8).

Le cas $\ell=1$ conduit uniquement à des expressions de la forme (5.8) et (5.9), qui ont été prouvées; le cas $\ell=0$ ne présente aucune difficulté. Nous avons donc montré (4.8) dans ce cas.

c) Conclusion.

En rassemblant les résultats (4.2), (4.3), (4.4) et (4.8) on a montré que l'inégalité suivante :

$\forall\, \epsilon > 0,\ \forall\, A > 0$

$$(5.10)\qquad \left|\pi(x,q,1) - \frac{\pi(x)}{\varphi(q)} + \sum_{r=1}^{R} (-1)^{r-1}\left(S_r^{(q)} - \frac{T_r^{(q)}}{\varphi(q)}\right)\right| \leq \frac{\epsilon\, x}{\varphi(q)\log x}$$

est vraie pour $x > x_0(\epsilon,A)$, pour tout q de $[x^{1/2},2x^{1/2}[$ avec au plus $x^{1/2}(\log x)^{-A}$ exceptions.

L'étude en moyenne des expressions

$$S_r^{(q)} - \frac{T_r^{(q)}}{\varphi(q)}$$

pour $r \geq 5$ ne pose pas de problème. En effet, pour prouver l'inégalité

$$\forall\, B > 0,\ \sum_q \gamma_q \left(S_r^{(q)} - \frac{T_r^{(q)}}{\varphi(q)}\right) = O_{B,\eta}(x\, \mathcal{L}^{-B})$$

on se ramène tout d'abord à

$$(5.11)\qquad \forall\, B > 0,\ \sum_q \gamma_q\, E(1_{p_1} * \ldots * 1_{p_r} * \zeta_s; q, 1) = O_{B,\eta}(x\, \mathcal{L}^{-B})$$

où ζ_s est la fonction caractéristique des entiers de $[S,S\Delta[$ dont les facteurs premiers sont plus grands que P_r et où

$$x^{\eta^2} \leq P_r < \ldots < P_1 \qquad P_i < D_i(P_1,\ldots,P_{i-1}) \quad (1 \leq i < r)$$
$$P_r \geq D_r(P_1,\ldots,P_{r-1}) \qquad P_1 \ldots P_r S \leq x$$

Le produit $P_1 \ldots P_r$ admet toujours un produit partiel dans $[x^{\eta},x^{1/6-\eta}]$, car, supposant le contraire, on trouve

$$P_3 \ldots P_r < x^{\eta}$$

d'où

$$P_1 P_2 P_3 \ldots P_r < x^{2/3-5\eta}.$$

Comme $P_r \geq D_r$ s'écrit

$$P_1P_2P_3\ldots P_r \geq x^{0,9}$$

on aboutit à une contradiction.

Le lemme 4 s'applique donc et prouve (5.11).

Remarquons que dans le cas de la majoration on a $S_2^{(q)} = T_2^{(q)} = S_4^{(q)} = T_4^{(q)} = 0$, ce qui donne, d'après (5.10)

$$\forall\, \epsilon > 0,\ \forall\, x > x_0(\epsilon,A)$$

$$\pi(x;q,1) - \frac{\pi(x)}{\varphi(q)} + S_1^{(q)} - \frac{T_1^{(q)}}{\varphi(q)} + S_3^{(q)} - \frac{T_3^{(q)}}{\varphi(q)} \leq \frac{\epsilon\, x}{\varphi(q)\log x}$$

pour tout q de $[x^{1/2}, 2x^{1/2}[$, avec au plus $x^{1/2}(\log x)^{-A}$ exceptions (5.12)

De même, dans le cas de la minoration

$$\forall\, \epsilon > 0,\ \forall\, A > 0,\ \forall\, x > x_0(\epsilon,A)$$

$$\pi(x;q,1) - \frac{\pi(x)}{\varphi(q)} - S_2^{(q)} + \frac{T_2^{(q)}}{\varphi(q)} - S_4^{(q)} + \frac{T_4^{(q)}}{\varphi(q)} \geq \frac{-\epsilon\, x}{\varphi(q)\log x}$$

pour tout q de $[x^{1/2}, 2x^{1/2}[$, avec au plus $x^{1/2}(\log x)^{-A}$ exceptions (5.13)

VI - La majoration : fin de la démonstration.

Nous ne pourrons pas montrer que la quantité

$$S_r^{(q)} - \frac{T_r^{(q)}}{\varphi(q)} \qquad (r = 1,3)$$

est petite en moyenne sur q, mais uniquement en donner des minorations.

a) Etude de $S_1^{(q)} - \dfrac{T_1^{(q)}}{\varphi(q)}$.

Soit

$$S_1^{*(q)} = \sum_{x^{7/16+\eta} < p_1 < x^{1/2-\eta}} S(\mathcal{A}_{p_1}^{(q)}, \mathcal{P}^{(q)}, p_1)$$

et soit

$$T_1^{*(q)} = \sum_{x^{7/16+\eta} < p_1 < x^{1/2-\eta}} S(\mathcal{B}_{p_1}^{(q)}, \mathcal{P}^{(q)}, p_1).$$

Pour démontrer l'estimation

$$\forall\, B > 0,\ \sum_q \tau_q \left(S_1^{*(q)} - \frac{T_1^{*(q)}}{\varphi(q)}\right) = O_{B,\eta}(x\, \mathcal{L}^{-B})$$

il suffit, après découpage des domaines de sommations, de vérifier l'égalité

$$\forall\, B > 0,\ \sum_q \tau_q E(1_{p_1} * 1_{p_2}; q, 1) = O_{B,\eta}(x\, \mathcal{L}^{-B})$$

$$\text{avec}\quad x^{7/16+\eta} < P_1 < x^{1/2-\eta} \qquad P_1 P_2 \leq x$$

qui est une conséquence directe du lemme 6.

Nous avons donc montré la minoration

$$\forall\, \epsilon > 0,\ \forall\, A > 0,\ \forall\, x > x_0(\epsilon, A)$$

$$S_1^{(q)} - \frac{T_1^{(q)}}{\varphi(q)} \geq -\frac{T_1^{(q)} - T_1^{*(q)}}{\varphi(q)} - \frac{\epsilon\, x}{\varphi(q) \log x}$$

pour tout q de $[x^{1/2}, 2x^{1/2}[$, avec au plus $x^{1/2}(\log x)^{-A}$ exceptions. Le théorème des nombres premiers donne un équivalent de $T_1^{(q)} - T_1^{*(q)}$, qui conduit à la minoration

$\forall\, \epsilon > 0,\ \forall\, A > 0,\ \forall\, x > x_0(\epsilon, A)$,

$$(6.1) \qquad S_1^{(q)} - \frac{T_1^{(q)}}{\varphi(q)} \geq -\left(\int_{1/3}^{7/16} \frac{dt}{t(1-t)} + 2\epsilon\right) \frac{x}{\varphi(q) \log x}$$

$$\geq -\left(\log \frac{14}{9} + 2\epsilon\right) \frac{x}{\varphi(q) \log x}$$

pour tout q de $[x^{1/2}, 2x^{1/2}[$, avec au plus $x^{1/2}(\log x)^{-A}$ exceptions

b) Etude de $S_3^{(q)} - \frac{T_3^{(q)}}{\varphi(q)}$.

Notre démarche est la même que précédemment : elle consiste à évaluer une sous-somme $S_3^{*(q)}$ de $S_3^{(q)}$ qui soit sensiblement égale, en moyenne sur q, à la somme correspondante $\frac{T_3^{*(q)}}{\varphi(q)}$.

Soit donc

$$S_3^{*(q)} = \sum_{p_1,p_2,p_3} S(\mathcal{A}_{p_1p_2p_3}^{(q)}, \mathcal{B}^{(q)}, p_3)$$

où la sommation est faite sur les triplets (p_1,p_2,p_3) vérifiant

$$x^{1/6+\eta} < p_3 < p_2 < p_1 < x^{1/3-3\eta}, \quad p_1^3 p_2^3 p_3 > x^{3/2+\eta},$$

$$p_1p_3 < x^{1/2-\eta} \quad \text{ou} \quad p_1p_3 > x^{1/2+\eta},$$

$$p_1^3p_2^3p_3^3 < x^{3/2-3\eta} \quad \text{ou} \quad p_1^3p_2^3p_3^3 > x^{3/2+\eta},$$

$$\text{et} \qquad p_1p_2p_3^2 \leq x$$

et soit $T_3^{*(q)}$ la somme associée portant sur $\mathcal{B}^{(q)}$.

On se propose de montrer que

$$(6.2) \qquad \forall B > 0, \ \sum_q \gamma_q \left(S_3^{*(q)} - \frac{T_3^{*(q)}}{\varphi(q)}\right) = O(x\,\mathcal{L}^{-B})$$

problème que l'on ramène tout d'abord au suivant :

$$\forall B > 0, \ \sum_q \gamma_q E(1_{p_1} * 1_{p_2} * 1_{p_3} * f_{p_3}(m);q,1) = O_{B,\eta}(x\,\mathcal{L}^{-B})$$

où $x^{1/6+\eta} < P_3 < P_2 < P_1 < x^{1/3-3\eta}$ $\quad P_1^3 P_2^3 P_3 > x^{3/2+\eta}$

$$P_1 P_3 < x^{1/2-\eta} \quad \text{ou} \quad P_1 P_3 > x^{1/2+\eta}$$
$$P_1^3 P_2 P_3^3 < x^{3/2-3\eta} \quad \text{ou} \quad P_1^3 P_2 P_3^3 > x^{3/2+\eta}$$

$$P_1 P_2 P_3^2 \leq x \qquad \text{et} \quad P_1 P_2 P_3 M \leq x$$

et où $f_{P_3}(m)$ est la fonction caractéristique des entiers de $[M,M'[$ dont tous les facteurs premiers sont supérieurs à P_3.

Nous envisageons trois situations

* si $P_1 P_3 \geq x^{1/2+\eta}$, on a $P_1^3 P_3^2 \leq x^{3/2-\eta}$ et $P_1^3 P_3^4 < x^{2-\eta}$

(car $P_1 P_3^3 \leq P_1 P_2 P_3^2 \leq x$) et on applique en (6.2) le lemme 5 avec les quantités N et L valant respectivement P_3 et P_1.

* si $P_1 P_3 < x^{1/2-\eta}$ et $P_1^3 P_2 P_3^3 > x^{3/2+\eta}$.

On applique le lemme 5, avec

$$N = P_2 \qquad L = x(P_1 P_2 P_3)^{-1}.$$

La condition $NL > x^{1/2+\eta}$ est vérifiée directement.
On a $N^2 L^3 = x^3 P_1^{-3} P_2^{-1} P_3^{-3}$, la condition

$$N^2 L^3 < x^{3/2+\epsilon} \quad \text{est vérifiée facilement.}$$

On a aussi

$$N^4 L^3 = x^3 (P_1 P_3)^{-3} P_2 < x^{3/2-\eta} P_2^2$$

par hypothèse. On a donc

$$N^4 L^3 < x^{2-\epsilon} \quad \text{pour} \quad P_2 \leq x^{1/4+\eta}.$$

Pour $P_2 \geq x^{1/4+\eta}$, on a

$$N^4L^3 < x^3P_2^{-2}P_3^{-3} < x^{5/2}P_2^{-2} < x^{2-\epsilon}.$$

La condition $N^5L^2 < x^{2-\epsilon}$ est toujours vérifiée, puisqu'on a

$$N^5L^2 = x^2P_1^{-2}P_2^3P_3^{-2} \leq x^2P_2P_3^{-2} \leq x^{2-\epsilon}$$

puisqu'on a toujours $P_2 < x^{1/3-\eta}$ et $P_3 > x^{1/6-\eta}$.

* si $P_1P_3 < x^{1/2-\eta}$ et si $P_1^3P_2P_3^3 < x^{3/2-3\eta}$, on trouve

$$P_2 \leq (P_1^3P_2)^{1/4} < (x^{3/2-3\eta}\, P_3^{-3})^{1/4} < x^{1/4-\eta}, \text{ puis}$$

$$P_2^2 < P_3^3, \text{ d'où } P_1P_2 \leq (P_1^3P_2P_3^3)^{1/3} \leq x^{1/2-\eta}$$

On applique le lemme 5 avec $N=P_3$ et $L=M$. La condition $NL > x^{1/2+\eta}$ est satisfaite d'après ce qui précède. La condition

$$N^2L^3 > x^{3/2+\epsilon} \text{ équivaut à } P_1^3P_2^3P_3 \geq x^{3/2-\epsilon}$$

qui est satisfaite par hypothèse.

L'autre condition

$N^5L^2 < x^{2-\epsilon}$ est une conséquence de l'inégalité triviale $P_3^3 \leq P_1^2P_2^2$. Quant à

$N^4L^3 < x^{2-\epsilon}$, elle se ramène à

$$P_3x^{1+\epsilon} < (P_1P_2)^3$$

mais, par hypothèse, $(P_1P_2)^3 > x^{3/2+\eta}P_3^{-1}$ ce qui termine la démonstration.

Nous avons terminé la vérification de (6.2), et par conséquent nous avons obtenu la minoration

$$\forall\, \epsilon > 0,\ \forall\, A > 0,\ \forall\, x > x_0(\epsilon, A)$$

$$S_3^{(q)} - \frac{T_3^{(q)}}{\varphi(q)} \geq -\frac{T_3^{(q)} - T_3^{*(q)}}{\varphi(q)} - \frac{\epsilon\, x}{\varphi(q)\log x}$$

pour tout q de $[x^{1/2}, 2x^{1/2}[$, avec au plus $x^{1/2}(\log x)^{-A}$ exceptions. Le théorème des nombres premiers et le lemme 10 donnent l'inégalité

$$\forall\, \epsilon > 0,$$

$$(6.3)\quad \left|T_3^{(q)} - T_3^{*(q)} - \frac{x}{\log x} \iiint_{\mathfrak{D}} \frac{w((1-t_1-t_2-t_3)/t_3)}{t_1 t_2 t_3 (1-t_1-t_2-t_3)}\, dt_1 dt_2 dt_3\right| \leq \frac{\epsilon\, x}{\log x}$$

pour η assez petit.

$\mathfrak{D}$ est le tétraèdre limité par les plans

$$t_1 = t_2,\ t_2 = t_3,\ t_3 = 1/6,\ 6t_1 + 6t_2 + 2t_3 = 3$$

$\mathfrak{D}$ a pour sommets (1/6,1/6,1/6), (5/18,1/6,1/6),(2/9,2/9,1/6) et (3/14,3/14,3/14), et son volume est 1/20412.

La fonction $(1-t_1-t_2-t_3)/t_3$ est, sur $\mathfrak{D}$, inférieure à 3, et ainsi sur $\mathfrak{D}$, on a

$$w((1-t_1-t_2-t_3)/t_3) \leq 1 + \log 2$$

La fonction $t_1 t_2 t_3 (1-t_1-t_2-t_3)$ est concave, elle prend son minimum sur $\mathfrak{D}$ en l'un de ses sommets, plus précisément en (1/6, 1/6,1/6).

L'intégrale triple de (6.3) est donc inférieure à $\frac{4(1+\log 2)}{189}$.

c) Conclusion.

Les calculs donnent $\log \frac{14}{9} + \frac{4(1+\log 2)}{189} < 0{,}48$. Grâce à l'expression (5.12), on en déduit la forme définitive de la majoration du théorème.

La clé de l'amélioration possible de ce résultat, signalée en I, consiste en la détermination d'un domaine de sommation un peu plus étendu pour $S_3^{*(q)}$, et d'un majoration plus fine de la fonction w, ceci au prix de calculs nettement plus compliqués.

VII - La minoration : fin de la démonstration.

a) Etude de $S_2^{(q)} - \frac{T_2^{(q)}}{\varphi(q)}$.

1) Découpage du domaine de sommation.

les quantités D_1 et D_2 ont les valeurs fixées en (4.6). La définition de $S_2^{*(q)}$ est délicate à formuler. Pour ce faire, sur le croquis, nous définissons 11 zones $\underline{A}, \underline{B}, \ldots, \underline{K}$ du plan des (t_1, t_2). Nous notons $\overset{\circ}{\underline{A}}, \overset{\circ}{\underline{B}}, \ldots, \overset{\circ}{\underline{K}}$, l'"intérieur" de chacun des polygones $\underline{A}, \underline{B}, \ldots, \underline{K}$, c'est-à-dire que $\overset{\circ}{\underline{A}}$ est le polygone intérieur à $\underline{A}$ tel que chacun de ses côtés soit parallèle et à la distance 100η du côté correspondant de $\underline{A}$; $\overset{\circ}{\underline{A}}$ est dessiné sur la figure.

Soit maintenant

$$S_2^{*(q)} = \sum_{p_1 p_2} S\left(\mathcal{A}_{p_1 p_2}^{(q)} ; \mathcal{B}^{(q)}, z\right)$$

où la sommation est faite sur les (p_1, p_2) tels que, en posant $t_i := \frac{\log p_i}{\log x}$, (t_1, t_2) appartienne à l'une des zones $\overset{\circ}{\underline{B}}, \overset{\circ}{\underline{C}}, \overset{\circ}{\underline{D}}, \overset{\circ}{\underline{E}}, \overset{\circ}{\underline{F}}, \overset{\circ}{\underline{G}}, \overset{\circ}{\underline{H}}, \overset{\circ}{\underline{I}}$.

On pose ensuite, pour $\underline{Z}$ l'une des zones $\underline{B}, \underline{C}, \underline{D}, \underline{E}, \underline{F}, \underline{G}, \underline{H}, \underline{I}$,

$$S_{2,\underline{Z}}^{*(q)} = \sum_{p_1, p_2} S\left(\mathcal{A}_{p_1 p_2}^{(q)} ; \mathcal{B}^{(q)}, z\right)$$

où $(t_1, t_2) \in \overset{\circ}{\underline{Z}}$, et on définit de même $T_{2,\underline{Z}}^{*(q)}$.

On veut donc montrer, pour $\underline{Z} = \underline{B}, \underline{C}, \underline{D}, \underline{E}, \underline{F}, \underline{G}, \underline{H}, \underline{I}$, l'estimation :

$$\forall B > 0, \ \sum_q \gamma_q \left(S_{2,\underline{Z}}^{*(q)} - \frac{T_{2,\underline{Z}}^{*(q)}}{\varphi(q)}\right) = O\left(x \mathcal{L}^{-B}\right)$$

qui se ramène, comme dans les parties précédentes, à

$$(7.1) \qquad \forall B > 0, \ \sum_q \gamma_q E(1_{\mathcal{P}_1} * 1_{\mathcal{P}_2} * f_{\mathcal{P}_2}(m); q, 1) = O(x \mathcal{L}^{-B})$$

pour $P_1P_2M \leq x$, et, avec maintenant $t_i = \frac{\log P_i}{\log x}$, pour la condition $(t_1,t_2) \in \overset{\circ}{\underline{Z}}$.

Le lemme 5 prouve directement (7.1) pour $\underline{Z} = \underline{B}$, par le choix des quantités N et L valant respectivement P_2 et P_1.

Le lemme 5 prouve encore (7.1) pour $\underline{Z} = \underline{C}$, par le choix des quantités N et L valant respectivement P_2 et M.

2) Réduction de la forme générale (7.1).

Remarquons que les zones $\overset{\circ}{\underline{D}}$, $\overset{\circ}{\underline{E}}$, $\overset{\circ}{\underline{F}}$, $\overset{\circ}{\underline{G}}$, sont situées dans la zone où $t_1 + 3t_2 \geq 1$; autrement dit, la variable m de (7.1) est une variable première, ce qui nous permet d'écrire (7.1) sous la forme

$$(7.2) \qquad \forall\, B > 0,\ \sum_q \gamma_q\, E(1_{P_1} * 1_{P_2} * 1_P; q, 1) = O(x\, \mathcal{L}^{-B})$$

avec $P_1P_2P \leq x$, $P > P_2$, $x^{1/2} > P_1 > P_2 > \max(x^{1/6-\eta}, x^{1/2-3\eta} P_1^{-1})$.

On utilise la même technique qu'en V a), c'est-à-dire l'application du lemme 9 à chacune des variables p_1, p_2, p. Le lemme 4 limite l'étude à quatre types d'estimations, (puisqu'on a toujours $P_2 < x^{1/3-\eta}$ sur ces régions) :

$$(7.3) \qquad \forall\, B > 0,\ \sum_q \gamma_q\, E(\alpha_r * 1_{n_1} * 1_{n_2} * 1_n; q, 1) = O\,(x\, \mathcal{L}^{-B})$$

où $R < x^{3\eta}$, $P_1 x^{-\eta} < N_1 < P_1$, $P_2 x^{-\eta} < N_2 < P_2$, $P x^{-\eta} < N < P$ et $R\, N_1 N_2 N \leq x$

$$(7.4) \qquad \forall\, B > 0,\ \sum_q \gamma_q\, E(\alpha_r * 1_{n_1} * 1_{n_1'} * 1_{n_2} * 1_n; q, 1) = O(x\, \mathcal{L}^{-B})$$

$$\text{où } R < x^{3\eta},\ P_1 x^{-\eta} < N_1 N_1' < P_1,\ P_2 x^{-\eta} < N_2 < P_2,\ P x^{-\eta} < N < P$$

$$N_1' > N_1 > x^{1/6-\eta} \quad \text{et} \quad R N_1 N_1' N_2 N \leq x$$

$$(7.5)\qquad \forall\, B > 0,\ \sum_q \gamma_q\, E(\alpha_r * 1_{n_1} * 1_{n_2} * 1_n * 1_{n'};q,1) = O\,(x\,\mathcal{L}^{-B})$$

$$\text{où}\quad R < x^{3\eta},\ P_1x^{-\eta} < N_1 < P_1,\ P_2x^{-\eta} < N_2 < P_2,\ Px^{-\eta} < NN' < P$$

$$N' > N > x^{1/6-\eta} \qquad \text{et} \qquad RN_1N_2NN' \le x$$

$$(7.6)\qquad \forall\, B > 0,\ \sum_q \gamma_q\, E(\alpha_r * 1_{n_1} * 1_{n_1'} * 1_{n_2} * 1_n * 1_{n'};q,1) = O\,(x\,\mathcal{L}^{-B})$$

$$\text{où}\quad R < x^{3\eta},\ P_1x^{-\eta} < N_1N_1' < P_1,\quad P_2x^{-\eta} < N_2 < P_2,\ Px^{-\eta} < NN' < P$$

$$N_1' > N_1 > x^{1/6-\eta},\ N' > N > x^{1/6-\eta} \quad \text{et} \quad RN_1N_1'N_2NN' \le x.$$

En revanche, en zone $\underline{H}$ et $\underline{I}$, la variable m de (7.1) est soit variable première, soit produit de deux variables premières. On sépare alors (7.1) en deux expressions

$$\forall\, B > 0,\ \sum_q \gamma_q\, E(1_{p_1} * 1_{p_2} * 1_p;q,1) = O\,(x\,\mathcal{L}^{-B})$$

$$\text{où}\quad P_1P_2P \le x,\ P > P_2,\ x^{1/2} > P_1 > P_2 > \max(x^{1/6-\eta}, x^{1/2-3\eta}\,P_1^{-1})$$

$$\forall\, B > 0,\ \sum_q \gamma_q\, E(1_{p_1} * 1_{p_2} * 1_p * 1_{p'};q,1) = O\,(x\,\mathcal{L}^{-B})$$

$$\text{où}\quad P_1P_2PP' \le x,\quad P' > P > P_2,\ x^{1/2} > P_1 > P_2 > \max\,(x^{1/6-\eta}, x^{1/2-3\eta}\,P_1^{-1}).$$

Par une nouvelle utilisation du lemme 9, ces sommes se ramènent aux types (7.3), (7.4), (7.5) et (7.6), avec les mêmes inégalités, grâce au lemme 4, puisqu'on a nécessairement sur les régions $\underline{H}$ et $\underline{I}$ l'inégalité $P < P' < x^{1/3-\eta}$.

3) Etude de (7.3).

L'estimation (7.3) est conséquence directe du lemme 8.

4) Etude de (7.4).

On applique le lemme 7 avec les variables M et N valant respectivement N_1N_2 et N, et (7.4) se trouve démontré pour $(t_1,t_2) \in \overset{\circ}{\underline{E}},\ \overset{\circ}{\underline{F}},\ \overset{\circ}{\underline{G}},\ \overset{\circ}{\underline{H}}$ et $\overset{\circ}{\underline{I}}$. On remarque ensuite que (7.4) n'apparaît pas dans le développement de (7.1) si $(t_1,t_2) \in \overset{\circ}{\underline{D}}$.

5) Etude de (7.5).

Le cas est symétrique : on applique le lemme 7 avec les variables M et N valant respectivement NN_2 et N_1, et (7.5) se trouve démontré pour $(t_1,t_2) \in \underset{\circ}{\underline{D}}, \underset{\circ}{\underline{F}}, \underset{\circ}{\underline{G}}, \underset{\circ}{\underline{H}}$ et $\underset{\circ}{\underline{I}}$. On remarque ensuite que (7.5) n'apparaît pas dans le développement de (7.1) si $(t_1,t_2) \in \underset{\circ}{\underline{E}}$.

6) Etude de (7.6).

Le cas (7.6) n'apparaît dans le développement de (7.1) que si $(t_1,t_2) \in \underset{\circ}{\underline{F}}, \underset{\circ}{\underline{G}}, \underset{\circ}{\underline{H}}, \underset{\circ}{\underline{I}}$. Sur $\underset{\circ}{\underline{F}}$, on applique le lemme 5 avec les variables N et L valant respectivement N et N_1N_1'. Sur $\underset{\circ}{\underline{G}}$, on applique le lemme 5 avec les variables N et L valant respectivement N_1 et NN'. Sur $\underset{\circ}{\underline{H}}$, on applique le lemme 5 avec les variables N et L valant respectivement N_1' et N_1N_2. Sur $\underset{\circ}{\underline{I}}$, on applique le lemme 5 avec les variables N et L valant respectivement N' et NN_2.

7) Etude de l'intégrale.

Nous avons en définitive montré que

$$\forall\, \epsilon > 0,\ \forall\, A > 0,\ \forall x > x_0(\epsilon,A)$$

$$S_2^{(q)} - \frac{T_2^{(q)}}{\varphi(q)} \geq -\frac{T_2^{(q)} - T_2^{*(q)}}{\varphi(q)} - \frac{\epsilon\, x}{\varphi(q)\log x}$$

pour tout q de $[x^{1/2}, 2x^{1/2}[$, avec au plus $x^{1/2}(\log x)^{-A}$ exceptions.

Par le théorème de nombres premiers et le lemme 10, on a, pour tout $\epsilon > 0$

$$\left|T_2^{(q)} - T_2^{*(q)} - \frac{x}{\log x}\iint_{\mathcal{D}} \frac{w((1-t_1-t_2)/t_2)}{t_1t_2(1-t_1-t_2)}\, dt_1dt_2\right| \leq \frac{\epsilon\, x}{\log x}$$

pour η assez petit, et où $\mathcal{D} = \underline{A} \cup (\underline{B}\setminus\underset{\circ}{\underline{B}}) \cup (\underline{C}\setminus\underset{\circ}{\underline{C}}), \ldots, \cup (\underline{I}\setminus\underset{\circ}{\underline{I}}) \cup \underline{J} \cup \underline{K}$.

L'intégrale sur $\underline{A}$ est nulle, car $w(\frac{1-t_1-t_2}{t_2})$ y vaut 0; l'intégrale sur $(\underline{B}\setminus\underset{\circ}{\underline{B}}) \cup \ldots \cup (\underline{I}\setminus\underset{\circ}{\underline{I}})$ peut être rendue plus petite que ϵ par le choix de η, et on obtient

$\forall\, \epsilon > 0$

$$|T_2^{(q)} - T_2^{*(q)} - \frac{x}{\log x} \iint_{\underline{J}\cup\underline{K}} \frac{w((1-t_1-t_2)/t_2)}{t_1 t_2 (1-t_1-t_2)}\, dt_1 dt_2| \leq \frac{\epsilon\, x}{\log x}$$

pour η assez petit.

On calcule alors

$$\iint_J \frac{w((1-t_1-t_2)/t_2)}{t_1 t_2 (1-t_1-t_2)}\, dt_1 dt_2 = \int_{3/16}^{3/14} \frac{\log((3-2t_2)(2t_2)/((1-3t_2)(3-4t_2)))}{t_2(1-t_2)}\, dt_2 + \int_{3/14}^{1/4} \frac{2\log((3-6t_2)/(1+2t_2))}{t_2(1-t_2)}\, dt_2$$

$$\text{d'où} \quad \iint_{\underline{J}} \frac{w((1-t_1-t_2)/t_2)}{t_1 t_2 (1-t_1-t_2)} \leq 0{,}02960 + 0{,}03776$$

Ensuite, en posant $u = \dfrac{1-t_1-t_2}{t_2}$ et $v = t_2$

$$\iint_{\underline{K}} \frac{w(\frac{1-t_1-t_2}{t_2})}{t_1 t_2 (1-t_1-t_2)}\, dt_1 dt_2 = \int_2^{7/3} \frac{(1+\log(u-1))\log((2u+3)/(3u-1))}{u}\, du + \int_{7/3}^{12/5} \frac{(1+\log(u-1))\log((4u+6)/(9u-9))}{u}\, du + \int_{12/5}^{13/5} \frac{(1+\log(u-1))\log((10-2u)/(3u-3))}{u}\, du$$

$$\text{d'où} \quad \iint_{\underline{K}} \frac{w((1-t_1-t_2)/t_2)}{t_1 t_2 (1-t_1-t_2)}\, dt_1 dt_2 \leq 0{,}05093 + 0{,}00848 + 0{,}01187.$$

On a donc :

$$\forall\, A > 0, \quad \forall\, x > x_0(A)$$

$$S_2^{(q)} - \frac{T_2^{(q)}}{\varphi(q)} \geq -(0{,}13864)\, \frac{x}{\varphi(q)\log x}$$

pour tout q de $[x^{1/2}, 2x^{1/2}[$, avec au plus $x^{1/2}(\log x)^{-A}$ exceptions.

b) Etude de $S_4^{(q)} - \frac{T_4^{(q)}}{\varphi(q)}$.

Soit $S_4^{*(q)}$ la sous-somme de $S_4^{(q)}$ définie par :

$$S_4^{*(q)} = \sum_{p_1,p_2,p_3,p_4} S(\mathcal{A}^{(q)}_{p_1p_2p_3p_4}, \mathcal{B}^{(q)}, p_4)$$

où $x^{1/6+\eta} < p_4 < p_3 < p_2 < p_1 < x^{1/2}$, $p_2 < \max(x^{1/6-\eta}, \frac{x^{1/2-3\eta}}{p_1})$

et $p_1 > x^{1/4+\eta}$ ou $p_2 < x^{3/16-\eta}$ ou $p_3 > x^{3/14+\eta}$

et soit $T_4^{*(q)}$ la somme associée portant sur $\mathcal{B}^{(q)}$.

On va montrer que

$$\forall B > 0,\ \sum_q \gamma_q \left(S_4^{*(q)} - \frac{T_4^{*(q)}}{\varphi(q)}\right) = O(x\, \mathcal{L}^{-B})$$

on ramène tout d'abord cette expression à la suivante :

$$\forall B > 0,\ \sum_q \gamma_q\, E(1_{p_1} * 1_{p_2} * 1_{p_3} * 1_{p_4} * f_{p_4}(m); q, 1) = O(x\, \mathcal{L}^{-B})$$

où $x^{1/6+\eta} < P_4 < P_3 < P_2 < P_1 < x^{1/2}$, $P_2 < \max(x^{1/6-\eta}, x^{1/2-3\eta} P_1^{-1})$

$P_1 > x^{1/4+\eta}$ ou $P_2 < x^{3/16-\eta}$ ou $P_3 > x^{3/14+\eta}$

et $P_1P_2P_3P_4M \leq x$.

On distingue alors trois cas :

* si $P_1 \geq x^{1/4+\eta}$, on a

$P_1P_2 \leq x^{1/2-3\eta}$, $x^{1/2+\eta} P_3P_4 \leq P_1^2P_2^2$, $P_1^4P_2^4P_3P_4 > x^{2+\eta}$ et

$P_1^5P_2^5P_3^3P_4^3 \geq x^{3+\eta}$; de plus $P_1P_2P_3P_4^3 > x$ donc la variable m est une variable première. On peut donc appliquer le lemme 5 avec les variables N et L valant respectivement M et P_3P_4.

* si $P_2 \leq x^{3/16-\eta}$, alors

$$P_2^2P_3^3P_4^3 \leq x^{3/2-8\eta}, \quad P_2^5P_3^2P_4^2 \leq x^{2-\eta}, \quad P_2^4P_3^3P_4^3 \leq x^{2-\eta} \quad \text{et} \quad P_2P_3P_4 \geq x^{1/2+3\eta}$$

on applique le lemme 5 avec les variables N et L valant respectivement P_2 et P_3P_4.

* $P_3 \geq x^{3/14+\eta}$, on a

$$P_1^3P_2^3P_3 \geq x^{3/2+7\eta}, \quad P_1P_2 \leq x^{1/2-\eta}, \quad P_3^3 \leq P_1^2P_2^2 \quad \text{et} \quad x^{1+\eta}P_3 < P_1^3P_2^3$$

et on peut appliquer le lemme 5, les variables N et L valant respectivement P_3 et P_4M.

En définitive, on obtient

$$\forall\, \epsilon > 0, \quad \forall\, A > 0, \quad \forall\, x > x_0(\epsilon,A)$$

$$S_4^{(q)} - \frac{T_4^{(q)}}{\varphi(q)} \geq -\frac{T_4^{(q)} - T_4^{*(q)}}{\varphi(q)} - \frac{\epsilon\, x}{\varphi(q)\log x}$$

pour tout q de $[x^{1/2}, 2x^{1/2}[$, avec au plus $x^{1/2}(\log x)^{-A}$ exceptions. Le théorème des nombres premiers et le lemme 10 fournissent l'inégalité

$\forall\, \epsilon > 0$

$$\left|T_4^{(q)} - T_4^{*(q)} - \frac{x}{\log x}\iiiint_{\mathcal{D}} \frac{w((1-t_1-t_2-t_3-t_4)/t_4)}{t_1t_2t_3t_4(1-t_1-t_2-t_3-t_4)}\, dt_1dt_2dt_3dt_3\right| \leq \frac{\epsilon\, x}{\log x}$$

pour η assez petit, et où

$$\mathcal{D} = \{1/6 < t_4 < t_3 < t_2 < t_1 < 1/4, \quad t_3 < 3/14, \quad t_2 > 3/16\}$$

on remarque que

$$\mathcal{D} \subset \{1/6 < t_4 < t_3 < t_2 < t_1 < 1/4, \ t_1 > 3/16\}$$

donc

$$\mathrm{Vol}(\mathcal{D}) \leq \frac{1}{24}\left(\frac{1}{4} - \frac{1}{6}\right)^4 - \frac{1}{24}\left(\frac{3}{16} - \frac{1}{6}\right)^4$$

de plus la fonction à intégrer sur $\mathcal{D}$ vaut

$$\frac{1}{t_1t_2t_3t_4(1-t_1-t_2-t_3-t_4)}$$

elle est donc convexe, et un calcul simple montre qu'elle est inférieure à 4880 en les neuf sommets de $\mathcal{D}$, qui est convexe.

Il vient donc

$$\iiiint_{\mathcal{D}} \frac{w((1-t_1-t_2-t_3-t_4)/t_4)}{t_1t_2t_3t_4(1-t_1-t_2-t_3-t_4)}\, dt_1dt_2dt_3dt_4 \leq 0,00977.$$

c) Conclusion.

Rassemblant les résultats numériques de a) et b), on a

$$0,13864 + 0,00977 \leq 0,14841 \leq 0,15.$$

Grâce à l'expression (5.13) on en déduit la forme définitive de la minoration du théorème.

Comme à la fin du VI, la clé de l'amélioration signalée au I réside dans une utilisation plus systématique du lemme 5 dans le b).

BIBLIOGRAPHIE

[1] Balog.- p+a without large prime factors, Sém. de Théorie des Nombres (1983-84) Université de Bordeaux I.

[2] Bombieri, Friedlander, Iwaniec.- Primes in arithmetical progressions to large moduli, Acta Math. 156 (1986), 203-251.

[3] Deshouillers, Iwaniec.- On the Brun-Titchmarsh theorem on average, Janos Bolyai Conf. (1981) (à paraître).

[4] Fouvry.- Répartition des suites dans les progressions arithmétiques, Résultats du type Bombieri-Vinogradov avec exposant supérieur à 1/2, Thèse de doctorat d'Etat (1981), Université de Bordeaux I.

[5] Fouvry.- Autour du théorème de Bombieri-Vinogradov, Acta Mathematica 152 (1984), 219-244.

[6] Fouvry.- Sur le théorème de Brun-Titchmarsh, Acta Arithmetica 42 (1984), 417-424.

[7] Fouvry.- Théorème de Brun-Titchmarsh, Applicatioon au théorème de Fermat, Invent. Math. 79 (1985), 383-407.

[8] Fouvry, Grupp.- On the switching principle in sieve theory, J. für reine angew. Math. 370 (1986), 101-126.

[9] Friedlander, Iwaniec.- Incomplete Kloosterman Sums and a Divisor Problem, Annals of Math. 121 (1985), 319-350.

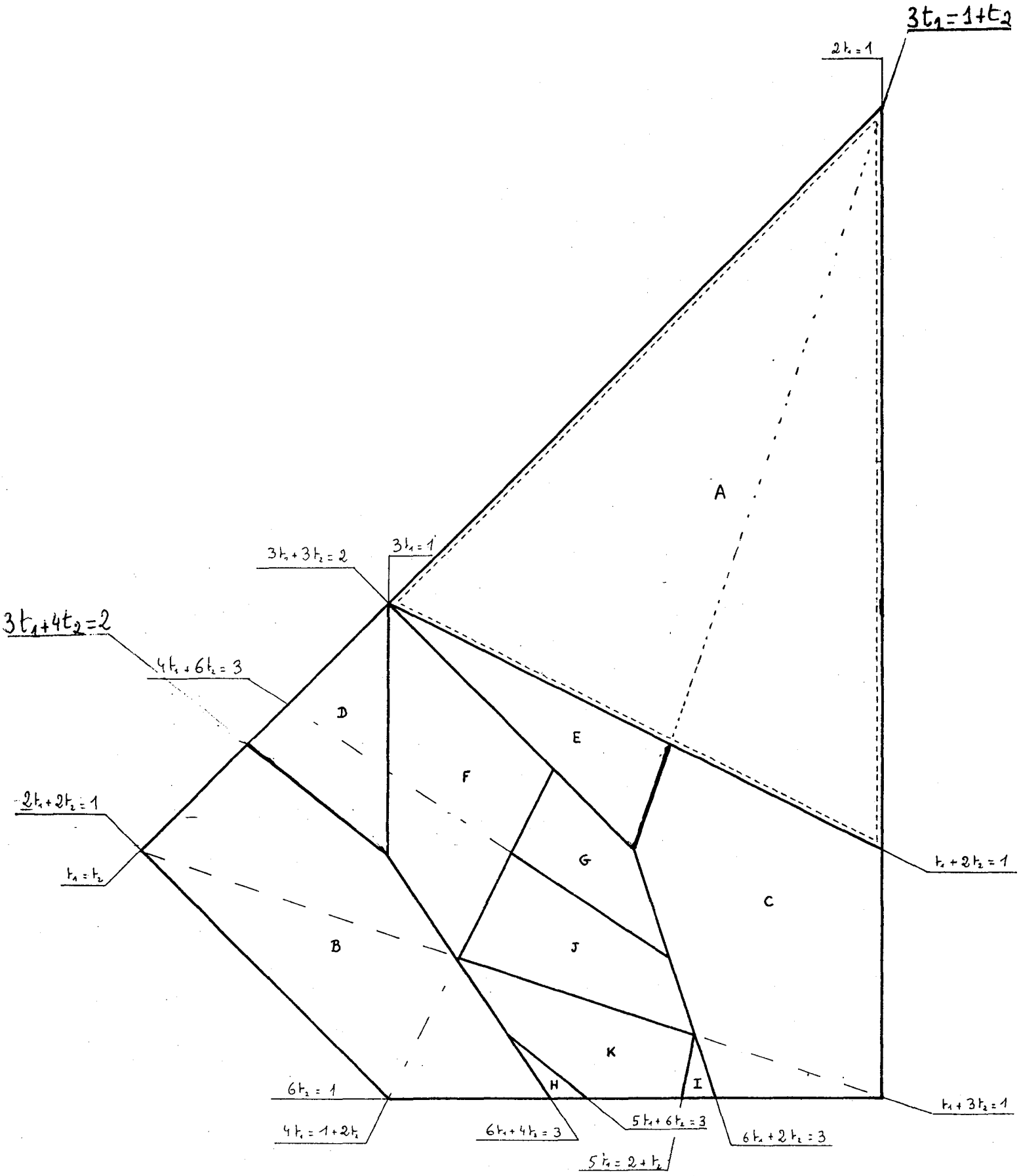
[10] Heath-Brown.- Prime numbers in short intervals and a generalized Vaughan identity, Can. J. of Math. 34 (1982), 1365-1377.

[11] Iwaniec.- The half dimensional sieve, Acta Arithmetica 29 (1976), 69-95.

[12] Iwaniec.- A new form of the error term in the linear sieve, Acta Arithmetica 37 (1980), 307-321.

[13] Iwaniec.- On the Brun-Titchmarsh theorem, J. Math. Soc. of Japan 34 (1982), 95-123.

[14] Shiu.- A Brun-Titchmarsh theorem for multiplicative functions, J. reine angew. Math. 313 (1980), 161-170.


$3t_1 = 1 + t_2$
$2t_1 = 1$
A
$3t_1 + 3t_2 = 2$
$3t_1 = 1$
$3t_1 + 4t_2 = 2$
$4t_1 + 6t_2 = 3$
D
E
F
$2t_1 + 2t_2 = 1$
$t_1 = t_2$
G
$t_1 + 2t_2 = 1$
C
B
J
K
H
I
$6t_2 = 1$
$t_1 + 3t_2 = 1$
$4t_1 = 1 + 2t_2$
$6t_1 + 4t_2 = 3$
$5t_1 + 6t_2 = 3$
$6t_1 + 2t_2 = 3$
$5t_2 = 2 + t_1$

NOTE.

Durant la publication de cet article, Bombieri, Friedlander et Iwaniec ont donné les premiers résultats en moyenne sur la fonction $\pi(x;q,a)$ $(q \geq x^{1/2})$:

Soit $a \neq 0$ $x,y \geq 3$, $y \leq x^{1/2}$ et $Q^2 \leq xy$, on a, pour une certaine constante absolue B l'égalité

$$\sum_{\substack{q \sim Q \\ (q,a)=1}} |\pi(x;q,a) - \frac{\mathrm{li}\, x}{\varphi(q)}| \ll \frac{x}{\log x} \left(\frac{\log y}{\log x}\right)^2 (\mathrm{loglog}\, x)^B.$$

Par une meilleure utilisation de la Kloostermanie et de la dispersion, ils ont trouvé de nouveaux résultats sur la répartition en moyenne de certaines formes trilinéaires. L'auteur est convaincu qu'en injectant ces nouveaux ingrédients dans la méthode qu'il vient de développer, il réduirait à néant les sommes qu'il ne peut que minoreer par 0, parvenant ainsi à l'énoncé typique.

$\forall\, \epsilon > 0$, $\forall\, A > 0$ il existe $x_0(\epsilon,A)$ tel que pour $x > x_0(\epsilon,A)$ on ait l'encadrement

$$(1-\epsilon) \frac{x}{\varphi(q)\log x} \leq \pi(x;q,1) \leq (1+\epsilon) \frac{x}{\varphi(q)\log x}$$

pour tout q de $[x^{1/2}, 2x^{1/2}]$ avec au plus $x^{1/2}(\log x)^{-A}$ exceptions.

texte reçu le 30/04/86

B. Rousselet
39, rue au Maire
75003 PARIS

SUBSTITUTIONS, AUTOMATES ET SERIES FORMELLES RELATIFS AUX SUITES A MULTI-INDICES

Olivier SALON

L'objet de cette note est de généraliser aux suites à multi-indices le théorème de Christol, Kamae, Mendès France et Rauzy ([4], voir aussi [1] et [5]) qui établit l'équivalence entre les notions de suite p-automatique, de suite engendrée par p-substitution et de série formelle algébrique sur le corps des fractions rationnelles sur un corps fini.

Le détail des démonstrations se trouve en [9].

Pour faciliter la présentation, nous ne considérerons que des suites doubles $u= (u(m,n))_{(m,n)\in\mathbb{N}^2}$.

A. (p,p)-substitutions.

Pour généraliser l'opération concaténation, nous utiliserons la juxtaposition de tableaux; considérons par exemple l'ensemble $A= \{a,b,c,d\}$ et l'application σ définie par :

$$\sigma(a)= \begin{pmatrix} a & b \\ b & c \end{pmatrix} \quad \sigma(b)= \begin{pmatrix} c & a \\ a & c \end{pmatrix} \quad \sigma(c)= \begin{pmatrix} c & d \\ d & a \end{pmatrix} \quad \sigma(d)= \begin{pmatrix} a & c \\ c & a \end{pmatrix}.$$

On définit alors $\sigma^2(a)= \sigma(\sigma(a))$ comme étant le tableau $\sigma^2(a)= \left[\begin{array}{c|c} \sigma(a) & \sigma(b) \\ \hline \sigma(b) & \sigma(c) \end{array}\right]$ et l'on pose, pour tout entier i strictement positif : $\sigma^i(a)= \sigma(\sigma^{i-1}(a))$.

Ainsi, $\sigma^2(a)= \begin{bmatrix} a & b & c & a \\ b & c & a & c \\ c & a & c & d \\ a & c & d & a \end{bmatrix}$, $\sigma^3(a)= \begin{bmatrix} a & b & c & a & c & d & a & b \\ b & c & a & c & d & a & b & c \\ c & a & c & d & a & b & c & d \\ a & c & d & a & b & c & d & a \\ c & d & a & b & c & d & a & c \\ d & a & b & c & d & a & c & a \\ a & b & c & d & a & c & a & b \\ b & c & d & a & c & a & b & c \end{bmatrix}$, etc...

et, puisque $\sigma(a)$ "commence par a", la suite $(\sigma^n(a))_{n\in\mathbb{N}}$ converge vers un tableau $\ell= (\ell(m,n))_{(m,n)\in\mathbb{N}^2}$ qui est un point fixe de σ (par construction, $\sigma(\ell)=\ell$).

Ainsi définie, σ est appelée (2,2)-substitution, ((p,p)-substitution lorsque l'image par σ d'un élément quelconque de A est un tableau d'éléments de A à p lignes et p colonnes) et, si τ

est une application de A dans un ensemble fini non vide T, la suite $(\tau(\ell(m,n)))_{(m,n)\in\mathbb{N}^2}$ est dite suite engendrée par (2,2)-substitution.

Par exemple, si l'on considère l'application τ_0 de A dans {0,1} définie par $\tau_0(a)=0$, $\tau_0(b)=1$, $\tau_0(c)=1$, $\tau_0(d)=0$, et si l'on revient à la suite ℓ définie ci-dessus, alors la suite $u=(\tau_0(\ell(m,n)))_{(m,n)\in\mathbb{N}^2}$ est une suite engendrée par (2,2)-substitution. Si l'on note $\epsilon=(\epsilon(k))_{k\in\mathbb{N}}$ la suite de Thue-Morse, alors on montre en fait que, pour tout couple d'entiers (m,n), $\tau_0(\ell(m,n))$ n'est autre que $\epsilon(m+n)$.

B. p-automates.

Pour définir un p-automate (p est un entier supérieur ou égal à 2) on se donne un ensemble fini non vide A (ensemble d'états) et un état initial a dans A, p^2 applications de A dans A, notées (i,j), avec $0\leq i,j\leq p-1$ - (i,j)(x) étant noté (i,j).x - , un ensemble fini non vide T et une application τ de A dans T.

Pour m et n entiers naturels, on définit (m,n).a de la façon suivante : ayant écrit m et n en base p et désigné par h+1 le plus grand nombre de chiffres dans les écritures de m et n en base p, quitte à rajouter des zéros à gauche dans l'une de ces écritures, ($m=\sum_{i=0}^{i=h} e_i(m)p^i$ et $n=\sum_{i=0}^{i=h} e_i(n)p^i$, avec $(e_h(m),e_h(n))\neq(0,0)$), on pose

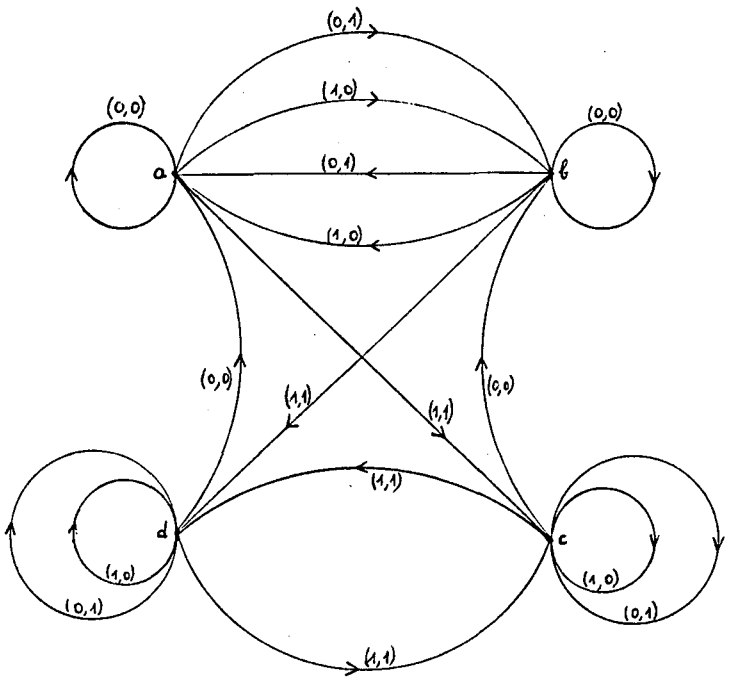
$$(m,n).a=(e_h(m),e_n(n)).[.....[(e_1(m),e_1(n)).[(e_0(m),e_0(n)).a]].....].$$

Une suite double $t=(t(m,n))_{(m,n)\in\mathbb{N}^2}$ est dite p-automatique lorsqu'il existe un p-automate tel que, pour tout couple (m,n) d'entiers, $t(m,n)=\tau((m,n).a)$.

Par exemple, si A= {a,b,c,d} et T= {0,1}, si l'on pose :

(0,0).a=a	(0,0).b=b	(0,0).c=b	(0,0).d=a
(1,0).a=b	(1,0).b=a	(1,0).c=c	(1,0).d=d
(0,1).a=b	(0,1).b=a	(0,1).c=c	(0,1).d=d
(1,1).a=c	(1,1).b=d	(1,1).c=d	(1,1).d=c

et si $\tau(a)=0$, $\tau(b)=1$, $\tau(c)=1$, $\tau(d)=0$, le 2-automate ainsi défini peut être schématisé comme suit :



et la suite $(\tau((m,n).a))_{(m,n)\in\mathbb{N}^2}$ est encore la suite $(\epsilon(m+n))_{(m,n)\in\mathbb{N}^2}$ définie au A., qui est donc 2-automatique.

Définition. Pour une suite double t à valeurs dans un ensemble fini non vide T, on note $N_p(t)$ et l'on appelle p-noyau de t l'ensemble $\{t(p^a m+r, p^a n+s);\ a\in\mathbb{N},\ 0\leq r,s\leq p^a-1\}$.

C. Et voici le théorème.

Le fait que la suite $\epsilon = (\epsilon(m+n))_{(m,n)\in\mathbb{N}^2}$ soit à la fois 2-automatique et engendrée par une (2,2)-substitution n'est pas l'effet d'un pur hasard ! En effet,

Théorème : Pour un entier p supérieur ou égal à 2 et une suite double $t = (t(m,n))$ à valeurs dans un ensemble fini non vide T, les trois propositions suivantes sont équivalentes :

(i) la suite t est engendrée par (p,p)-substitution

(ii) la suite t est p-automatique

(iii) le p-noyau $N_p(t)$ de la suite t est fini.

Si de surcroît l'entier p est premier, ces propositions sont équivalentes à :

(iv) il existe un corps fini K de caractéristique p et une injection I de T dans K tels que la série formelle $\sum_{(m,n)\in\mathbb{N}^2} I(t(m,n))X^m Y^n$ soit algébrique sur K(X,Y).

D. Application.

Une conséquence immédiate du théorème précédent est une aisée démonstration d'un cas particulier important d'un théorème prouvé par Deligne ([6]), puis par Denef et Lipschitz ([7]) :

Soit $u=(u(m,n))$ une suite double à valeurs dans un corps fini K, telle que la série formelle $F(X,Y)=\sum_{(m,n)\in\mathbb{N}^2} u(m,n)X^mY^n$ soit algébrique sur $K(X,Y)$; alors la série formelle "diagonale" $G(X)=\sum_{n\in\mathbb{N}} u(n,n)X^n$ est algébrique sur $K(X)$.

E. Exemples.

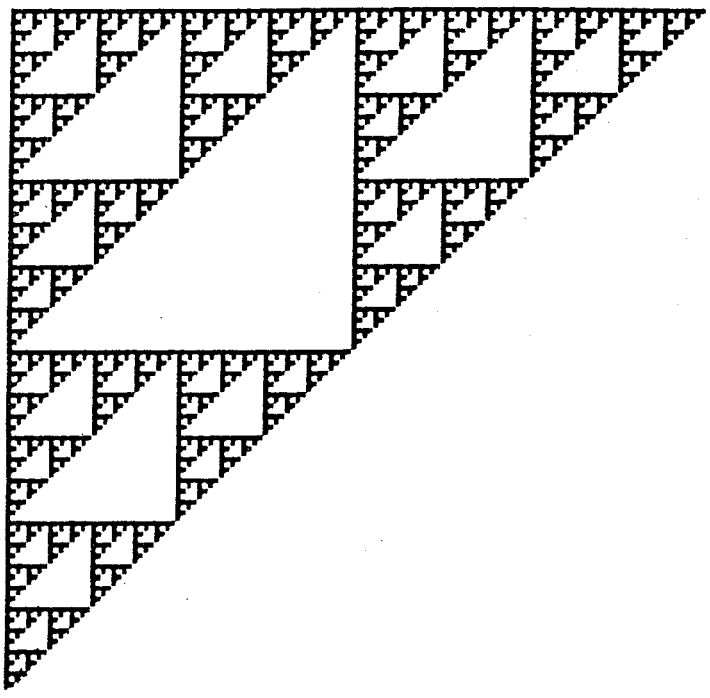
1. Pliages de mouchoir. (d'après une idée et des formules de J.-P. Allouche et M. Mendès France [2]).

Si l'on plie un mouchoir moitié inférieure sur moitié supérieure, puis moitié droite sur moitié gauche, et si l'on itère ce procédé sans jamais s'arrêter, le tableau (comportant une infinité de lignes et de colonnes) des intersections des plis ainsi formés, affectées des coefficients +1 ou -1 selon leur nature est le tableau des valeurs d'une suite double 2-automatique.

2. Triangle de Pascal.

Dans [3], A. Černy et J. Gruska montrent que la suite $v=(\overline{\binom{m+n}{m}})_{(m,n)\in\mathbb{N}^2}$, formée des coefficients binomiaux réduits modulo 2 est une suite 2-automatique.

Nous reproduisons ici le dessin de la partie des valeurs de cette suite correspondant à $0\leq m,n\leq 127$ qu'en obtient J. Shallit ([10]) en noircissant d'un petit carré les cases où $\overline{\binom{m+n}{m}}=1$.



3. Entiers de Gauss.

Rappelons (voir [8]) que pour tout couple d'entiers (relatifs) (m,n), il existe une unique famille $(\alpha_k)_{k\in\mathbb{N}}$ de 0 ou de 1, nulle à partir d'un certain rang, telle que $m+ni=\sum_{k=0}^{\infty}\alpha_k(-1+i)^k$. Notons alors $S(m,n)=\sum_{k=0}^{\infty}\alpha_k$ la somme réduite modulo 2 des chiffres composant $m+ni$ dans "la base" $(-1+i)^k$, $k\in\mathbb{N}$.

On peut montrer que la suite $S=(S(m,n))_{(m,n)\in\mathbb{N}^2}$ est une suite double à la fois 2-automatique et 5-automatique. En fait, la série formelle associée $\sum_{(m,n)\in\mathbb{N}^2}S(m,n)X^mY^n$ est non seulement algébrique sur $\mathbb{Z}/2\mathbb{Z}(X,Y)$, mais même rationnelle.

BIBLIOGRAPHIE

[1] J.-P. Allouche.- Automates finis en Théorie des Nombres, Expo. Math., 1987, à paraître.

[2] J.-P. Allouche et M. Mendès France.- Communication privée.

[3] A. Černy et J. Gruska.- Modular Trellises, in the book of L, G. Rozenberg et A. Salomaa, Springer Verlag, p. 45-61.

[4] G. Christol, T. Kamae, M. Mendès France et G. Rauzy.- Suites algébri-ques, Automates et Substitutions, Bull. Soc. Math. France, 108, 1980, p. 401-419.

[5] A. Cobham.- Uniform Tag Sequences, Mathem. Syst. Theory, 6, 1972, p. 164-192.

[6] P. Deligne.- Intégration sur un cycle évanescent, Invent. Math., 76, 1984, p. 129-143.

[7] J. Denef et L. Lipschitz.- Algebraic power series and diagonals, Preprint.

[8] D.E. Knuth.- The art of Computer Programming. Vol. 2, Seminumerical Algorithms, Addison-Wesley, Reading MA, 1981.

[9] O. Salon.- Suite automatiques à multi-indices, Séminaire de Théorie des Nombres, Bordeaux, 1986-87, exposé 4.

[10] J. Shallit.- Communication privée; voir aussi Appendix to the Paper of Salon, Séminaire de Théorie des Nombres, Bordeaux, 1986-87, exposé 4.

O. Salon
40, rue Maurice Ripoche
75014 PARIS

Generalised Salem Numbers

C.J. Smyth

This talk is concerned with some far from new, but unpublished results on generalised Salem numbers which were contained in my M.A. thesis [11]. Only the results from the thesis concerning generalised Pisot numbers were published [12].

I emphasise that the generalisations I am going to talk about are certainly not the only ones. Pisot, many of his students, and others have defined many alternative generalisations, which preserve different properties of the original Pisot numbers (see e.g. [6]).

Let S be the set of Pisot numbers :

$$S= \{\theta : \theta \text{ an algebraic integer } \in \mathbb{R},\ \theta>1, \text{ and all other conjugates } \theta_i \text{ of } \theta \text{ lie in } |z| < 1\}$$

and T be the set of Salem numbers :

$$T= \{\theta : \theta \text{ an algebraic integer } \in \mathbb{R},\ \theta>1 \text{ and all other conjugates } \theta_i \text{ of } \theta \text{ lie in } |z|\leq 1, \text{ with at least one on } |z|=1\}.$$

Then [Pisot, 5] for $\theta>1$, $\theta\in S$ iff $\sum_{n=1}^{\infty} \|\lambda\theta^n\|^2 < \infty$ for some real λ. Such a λ must then belong to $\mathbb{Q}(\theta)$.

For $\theta>1$, θ algebraic, $\theta\notin S$, a similar result is as follows : $\theta\in T$ iff for each $\epsilon>0$, $\|\lambda\theta^n\| < \epsilon$ $(n= 1,2,\ldots)$ for some real λ. For ϵ sufficiently small such a λ must belong to $\mathbb{Q}(\theta)$ (this result is well-known, though I don't know a source for it). [Here $\|\ \|$= distance from nearest integer].

Generalisations of S and T. We need these to discuss generalisations of the above properties of S and T. Let K be an algebraic number field, q a (equivalence class of) valuation on K, K_q the completion of K at q, Λ a finite extension of K_q, and $\tilde{K}_q$ an algebraic closure of K_q. Define $S_q(K,\Lambda)$ by

$$S_q(K,\Lambda)= \{\theta\in\Lambda : \theta \text{ algebraic over } K,\ |\theta|_q>1, \text{ and all zeros in } \tilde{K}_p \text{ of the minimal polynomial of } \theta \text{ over } K, \text{ except } \theta \text{ and its conjugates over}$$

$$K_q \text{ of } p=q, \text{ lie in } \begin{cases} |z|_p < 1 & \text{for } p \text{ an infinite valuation} \\ |z|_p \leq 1 & \text{for } p \text{ a finite valuation}\}. \end{cases}$$

Collecting ideas of Senge [10], Bertrandias [1] and others, I showed that $S_q(K,\Lambda)$ is closed in Λ, for K a J-field. This generalised result of Salem [7] for $S_\infty(\mathbb{Q},\mathbb{R}) = SU(-S)$, Kelly [4] for $S_\infty(\mathbb{Q},\mathbb{C})$, Chabauty [2] for $S_p(\mathbb{Q},\mathbb{Q}_p)$, p finite, and Grandet [3] for $S_\infty(\mathbb{Q}(\sqrt{-D}),\mathbb{C}) = \{\theta \in S_\infty(\mathbb{Q},\mathbb{C}) : \sqrt{-d} \in \mathbb{Q}(\theta)\}$.

Now $S_\infty(\mathbb{Q},\mathbb{R}) \cap \mathbb{Z} = \{\pm 2, \pm 3, \pm 4, \ldots\}$. More generally, define $S_q(K) = S_q(K,K_q) \cap K$. Then, like rational integers, $S_q(K)$ is discrete in K_q in the sense that there is a $c_q > 0$ such that if $\theta \neq \theta' \in S_q(K)$ then $|\theta - \theta'|_q > c_q$.

Denote by U the exceptional set

$$U = \begin{cases} \{0, \pm 1\} & \text{if } K=\mathbb{Q} \text{ and } q \text{ is infinite} \\ \{0\} \cup \{\text{absolute units of } K\} & \text{if } K \text{ quadratic imaginary, } q \text{ infinite} \\ \{0\} & \text{otherwise} \end{cases}$$

(U is the set of θ in K which satisfies all the conditions for belonging to $S_q(K)$ except the condition $|\theta|_q > 1$).

Then there is another constant $C_q = C_q(K,q)$ such that for any α in K_q, there is some $\theta \in S_q(K) \cup U$ with $|\alpha - \theta|_q < C_q$. Furthermore, generalising Pisot's results for S, we have

Theorem 1. An element $\theta \in \Lambda$, $|\theta|_q > 1$ belongs to $S_q(K,\Lambda)$ iff there is a $\lambda \in \Lambda$ and a sequence $\{c_n\}$ in $S_q(K) \cup U$ satisfying

$$\sum_{n=0}^{\infty} |\mathrm{tr}_q \lambda \theta^n - c_n|_q^2 < \infty \quad \text{if } q \text{ infinite}$$

$$|\mathrm{tr}_q \lambda \theta^n - c_n|_q \leq 1 \quad (n = 0,1,\ldots) \quad \text{if } q \text{ finite}$$

$$\sum_{n=0}^{\infty} |c_n|_p^2 < \infty \quad \text{for each infinite valuation } p \text{ (except } q).$$

Further, such a λ must belong to $K(\theta)$.

Here tr_q denotes $tr_{K_q(\theta)/K_q}$.

Generalisations of T. We define analogously $T_q(K,\Lambda)$ by $T_q(K,\Lambda)= \{\theta\in\Lambda$: algebraic over K, $|\theta|_q > 1$ and all zeros in $\tilde{K}_p$ of the minimal polynomial of θ, including θ and its conjugates over K_p if p=q, lie in $|z|_p \leq 1$ for each valuation p, with at least on conjugate on $|z|_p=1$ for some infinite p$\}$.

The corresponding generalisation of the result for T is as follows :

Theorem 2. An algebraic element $\theta\in\Lambda\setminus S_q(K,\Lambda)$, with $|\theta|_q > 1$, belongs to $T_q(K,\Lambda)$ iff for each $\epsilon>0$ there is a $\lambda\in\Lambda$ and a sequence $\{c_n\}$ in $S_q(K)\cup U$ with

$$|tr_q\lambda\theta^n - c_n|_q < \epsilon \quad (n=0,1,\ldots) \text{ if } q \text{ infinite}$$
$$|tr_q\lambda\theta^n - c_n|_q \leq 1 \quad (n=0,1,\ldots) \text{ if } q \text{ finite}$$
$$|c_n|_p < \epsilon \quad (n=0,1,\ldots) \text{ for } p \text{ infinite, excluding } q.$$

If ϵ is sufficiently small, such a λ must belong to $K(\theta)$.

We next look at inclusions among the $S_q(K,\Lambda)$ and $T_q(K,\Lambda)$.

Theorem 3. Let K be a subfield of L. Then

$$S_q(L,\Lambda)= \{\theta\in S_q(K,\Lambda) : [L_q(\theta):K_q(\theta) = [L(\theta):K(\theta)]\}$$
$$T_q(L,\Lambda)= \{\theta\in S_q(K,\Lambda) : [L_q(\theta):K_q(\theta) = [L(\theta):K(\theta)]\}.$$

While the sets $S_q(K)$ are non-empty and discrete, the sets

$$T_q(K)= T_q(K,K_q) \cap K$$

may be empty. For q real, Salem [8] proved that if $T_q(K)$ was non-empty, then it was of the form

$$T_q(K)= \{\pm\sigma^n : n= 1,2,\ldots\} \qquad (q \text{ real})$$

for some σ in $T_q(K)$. For q complex, Samet [9] generalised this : if $T_q(K)$ non-empty then

$T_q(K) = \{\epsilon\sigma^n : n=1,2,\ldots$ for certain roots of unity $\epsilon \in K\}$, (q complex)

for some σ in $T_q(K)$. However, there is one exceptional class of fields K overlooked by Samet for which his result is not valid. If $[K:\mathbb{Q}]=6$, the lowest degree for which $T_q(K)$, q complex, can be non-empty, then $T_q(K)$ may have two (complex) conjugate generators. Let us look at an example :

Example : Let $x^3 - x - 1 = (x-\alpha)(x-\alpha_2)(x-\bar{\alpha}_2)$, α real, and $K = \mathbb{Q}(\alpha,\alpha_2)$, $[K:\mathbb{Q}]=6$, K normal over $\mathbb{Q}$. Put $\theta = \alpha/\alpha_2$. Then, applying the six automorphisms of $K/\mathbb{Q}$, we obtain the following table of conjugates of θ and $\bar{\theta}$:

(1)	$(2\bar{2})$	$(1\bar{2})$	(12)	$(12\bar{2})$	$(1\bar{2}2)$
θ	$\bar{\theta}$	$\bar{\alpha}_2/\alpha_2$	θ^{-1}	$\alpha_2/\bar{\alpha}_2$	$\bar{\theta}^{-1}$
$\bar{\theta}$	θ	$\bar{\theta}^{-1}$	$\alpha_2/\bar{\alpha}_2$	θ^{-1}	$\bar{\alpha}_2/\alpha_2$

(here e.g. $(1\bar{2})$ denotes the automorphisms $\alpha \longrightarrow \bar{\alpha}_2, \bar{\alpha}_2 \longrightarrow \alpha$). We see that θ and $\bar{\theta}$ both belong to $T_\infty(K)$. They cannot, however, both be of the form $\epsilon_1\sigma^{n_1}$, $\epsilon_2\sigma^{n_2}$, $n_1,n_2>0$ as $|(1\bar{2})\theta|=1$ and $|(1\bar{2})\bar{\theta}|<1$ while $|(1\bar{2})(\epsilon_1\sigma^{n_1})|$ and $|(1\bar{2})\,\epsilon_2\sigma^{n_2}|$ are either both >1, both $=1$ or both <1.

The reason for this exceptional case is that for $[K:\mathbb{Q}]=6$, q complex and $\theta,\theta' \in T_q(K)$, θ/θ' need not have any conjugates of modulus one. The following lemma is needed for proving Samet's result : it does not hold for $[K:\mathbb{Q}]=6$.

Lemma. Let $T_q(K)$ be non-empty. Then, except when $[K:\mathbb{Q}]=6$, K is totally complex and q is infinite, there is a unique valuation q' of K such that $|\theta|_{q'} = |\theta|_q^{-1} < 1$.

Proof. It is easy to show that for $\theta \in T_q(K)$, $|\theta|_p=1$ for all valuations $p \neq q, q_\theta$, where $|\theta|_{q_\theta} = |\theta|_q^{-1} < 1$. Further, all infinite valuations except possibly q and q_θ must be complex. We must show that q_θ is independent of θ.

Take, θ, $\phi \in T_q(K)$. If there is a p infinite with $p \neq q, q_\theta, q_\phi$ then $|\theta|_p = |\phi|_p = |\theta\phi|_p = 1$. Hence there is an automorphism taking

$\theta\phi \longrightarrow (\theta\phi)^{-1}$, and so a valuation q' with $|\theta\phi|_{q'} = |\theta\phi|_q^{-1}$. This is possible only if $q_\theta = q_\phi$.

It is only in the exceptional case in the statement of the lemma that such a p does not exist : if q is non-complex, then so are q_θ and q_ϕ, so p must exist, by the definition of $T_q(K)$. If on the other hand q is complex and $[K:\mathbb{Q}]>6$ then K has at least four complex valuations.

Using this lemma it is now almost trivial to prove the following generalisation of Salem's result :

Theorem 4. Let $T_q(K)$ be non-empty, and K,q not the exceptional case of the previous lemma. Then there is an element $\sigma \in T_q(K)$ such that

$$T_q(K) = \{\epsilon\sigma^n \ (n=1,2,\ldots), \ \epsilon \text{ a root of unity in } K\}.$$

For the proof one takes (following Salem and Samet) $\sigma \in T_q(K)$ with $|\sigma|_q$ minimal, $\theta \in T_q(K)$ and n such that $|\sigma|_q^n \leq |\theta|_q < |\sigma|_q^{n+1}$, and shows that $\theta\sigma^{-n}$ is a root of unity in K.

For example, when $K = \mathbb{Q}(\sqrt{-1})$, $q_o \equiv 1 \pmod 4$ a rational prime, $q = a+ib$ where $q_o = a^2+b^2$ then

$$T_q(K) = \{(\sqrt{-1})^k (\tfrac{\bar{q}}{q})^n : k = 0,1,\ldots ; \ n = 1,2,\ldots\}.$$

BIBLIOGRAPHY

[1] F. Bertrandias.- Ensembles remarquables d'adèles algébriques, Bull. Soc. Math. France , Mémoire N° 4 (1965).

[2] C. Chabauty.- Sur la répartition modulo un de certaines suites p-adiques, C.R. Acad. Sci. Paris, 231 (1950), 465-466.

[3] M. Grandet.- Sur les dérives d'un ensemble d'entiers algébriques, C.R. Acad. Sci. Paris, 254 (1962), 2905-2906.

[4] J.B. Kelly.- A closed set of algebraic integers, Amer. J. Math., 72 (1950), 565-572.

[5] C. Pisot.- La répartition modulo 1 et les nombres algébriques, Annali di Pisa, ser. 2,7 (1938), 205-248.

[6] C. Pisot.- Quelques aspects de la théorie des entiers algébriques, Montréal 1963.

[7] R. Salem.- A remarkable class of algebraic integers. Proof of a conjecture of Vijayaraghavan, Cuke Math. J., 11 (1944), 103-108.

[8] R. Salem.- Power Series with integer coefficients, Duke Math. J., 12 (1945), 153-172.

[9] P.A. Samet.- Algebraic integers with two conjugates outside the unit circle, Proc. Camb. Phil. Soc., 49 (1953), 421-436.

[10] H.G. Senge.- Closed sets of algebraic numbers, Duke Math. J., 34 (1967), 307-323.

[11] C.J. Smyth.- Remarkable sets of algebraic numbers, M.A. thesis, University of Adelaide, (1970).

[12] C.J. Smyth.- Closed sets of algebraic numbers in complete fields, Mathematika 17, (1970), 199-205.

Congruences for Symmetric Functions

C.J. Smyth

We start by giving a proof of Fermat's Theorem : $a^p \equiv a \bmod p$, for $a \in \mathbb{Z}$, and p prime. This proof, very similar to [3, p. 8], goes as follows : write $1-az$ as

$$1-az = \prod_{k=1}^{\infty} (1-z^k)^{b_k} \tag{1}$$

where the equality is purely formal. Then it is clear inductively that $b_1, b_2, \ldots$, obtained successively by formal expansion, must be integers. Taking logs, differentiating and multiplying by $-z$ we get

$$\frac{az}{1-az} = \sum_{k=1}^{\infty} \frac{kb_k z^k}{1-z^k},$$

a Lambert series. Expanding $z^k/(1-z^k)$ and interchanging summations we have

$$\sum_{n=1}^{\infty} a^n z^n = \sum_{n=1}^{\infty} \Big(\sum_{k \mid n} kb_k \Big) z^n.$$

Then $b_1 = a$ and for $n = p$ prime, $a^p = b_1 + pb_p \equiv a \pmod p$, giving Fermat's Theorem. However, we obtain more : for general n,

$$a^n = \sum_{k \mid n} kb_k, \quad \sum_{\ell \mid n} a^{\ell} \mu\left(\frac{n}{\ell}\right) = nb_n \quad \text{so that}$$

$$\sum_{\ell \mid n} a^{\ell} \mu\left(\frac{n}{\ell}\right) \equiv 0 \pmod n \tag{2}$$

a generalisation due to Lucas, Pellet, etc. $\simeq$ 1880 (see [2] for references).

Another Generalisation. Schönemann (1839) showed that

$$s_p(\alpha) \equiv s_1(\alpha) \pmod p \tag{3}$$

where $s_\ell = s_\ell(\alpha) = a_1^\ell + a_2^\ell + \ldots + a_d^\ell$, and α is an algebraic integer with conjugates $\alpha = \alpha_1, \alpha_2, \ldots, \alpha_d$. In fact both (2) and (3) can be generalised to obtain

$$(4) \qquad \sum_{\ell \mid n} s_\ell \mu(\tfrac{n}{\ell}) \equiv 0 \pmod{n}.$$

This is readily proved by putting $\prod_{i=1}^{d} (1-\alpha_i z) = \prod_{k=1}^{\infty} (1-z^k)^{b_k}$ and proceeding as in the above proof of (1). For another proof of (4), extending Petersen's proof of Fermat's Theorem, which used a colouring argument, see [2].

Improving the Congruence Modulus. We now look for congruences like (4), which are monic with constant coefficients, i.e. of the form

$$(5) \qquad s_n + \sum_{i=1}^{n-1} c_i s_{n-i} \equiv 0 \pmod{N(n)}$$

where the integers c_i are independent of α. How big can $N(n)$ be ? Can the modulus n in (4) be improved ? The answer is no in general, as if α has minimal polynomial

$$x^d - a_1 x^{d-1} - a_2 x^{d-2} - \ldots - a_d$$

then (Newton's identities)

$$(6) \qquad s_n = a_1 s_{n-1} + a_2 s_{n-2} + \ldots + a_{n-1} s_1 + n a_n,$$

taking $a_k = 0$ for $k > d$. So if $s_1, s_2, \ldots, s_{n-1}$ are given, $a_1, a_2, \ldots, a_{n-1}$ are determined, and so s_n is determined mod n by (6), say $s_n = f(s_1, \ldots, s_{n-1}) \bmod n$. Conversely to any s_n satisfying this congruence there is a value of a_n making (6) true, so that we can make two choices $s_n' = f(s_1, \ldots, s_{n-1})$ and $s_n'' = s_n' + n$ for s_n. Then both sequences $s_1, \ldots, s_{n-1}, s_n'$ and $s_1, \ldots, s_{n-1}, s_n''$ must satisfy (5). Subtracting these, we get $n \equiv 0 \pmod{N(n)}$, so that $N(n) \leq n$ (note in passing that (6) does not give a congruence of the type (5), since its coefficients $a_1, a_2, \ldots$ depend on α).

Congruences for fixed degree. When we fix the degree d of α, however, better congruences may exist. We look first at the simple case $d=1$. Here the congruence (2) to modulus n can be vastly improved by

$$a(a-1) \ldots (a-n+1) = \binom{a}{n} n! \equiv 0 \pmod{n!}$$

Furthermore n! is the largest possible modulus. For if

$$a^n + \sum_{i=1}^{n-1} c_i a^{n-i} \equiv 0 (\mathrm{mod}\ N(n))$$

then, replacing a by a+1, subtracting the two congruences and dividing by n we obtain $N(n)/\gcd(n,N(n)) \leq N(n-1)$, $N(n) \leq nN(n-1)$. Since $N(1)=1$ this gives $N(n) \leq n!$.

For d=2, however, the situation is considerably more complicated, and seems to be essentially as difficult in the case of general fixed d. Let $x^2-ax-b = (x-\alpha)(x-\alpha')$, $a,b \in \mathbb{Z}$, $s_n = \alpha^n+\alpha'^n$, and $S^{(2)}$ be the $\mathbb{Z}$-module generated by all linear or quadratic sequences $\{s_n\}$. Any congruence which is satisfied by such a sequence is also satisfied by any sequence in $S^{(2)}$. Further, suppose that $0 \ldots 0 M_n^{(2)} \ldots \in S^{(2)}$, with $M_n^{(2)} > 0$ the n^{th} term of the sequence. Then if (5) holds for $\{s_n\}$ linear or quadratic, we have $M_n^{(2)} \equiv 0 (\mathrm{mod}\ N(n))$, $N(n) \mid M_n^{(2)}$.

For convenience we truncate sequences in $S^{(2)}$ to n terms, and denote this set of n-tuples by $S_n^{(2)}$. It would be nice to have an explicit basis for $S_n^{(2)}$. The following theorem shows how to obtain the next best thing : a finite spanning set for $S_n^{(2)}$. For any fixed n, row reduction of this set will of course yield a basis.

For the theorem it is useful to write $A_1, \ldots, A_n \in S_n^{(2)}$ as a polynomial $A_1 x + A_2 x^2 + \ldots + A_n x^n$.

Theorem. The set $S_n^{(2)}$ is spanned by the vectors (= polynomials) c_{ij} where $i \geq 0$, $j \geq 0$, $i+2j \leq n$, defined as follows :

let P be a rectangular lattice-path from (i,j) to (0,0) with both i and j non-increasing. Put $\lambda_{ij} = (1-ix-jx^2)^{-1}$, and λ_P the power series for $\prod_{(i_o,j_o)\ \text{on P}} \lambda_{i_o j_o}$, truncated to terms in $x^{n-(1+2j)}$. Then

$$c_{ij} = i!j!\, x^{i+2j} \sum_{\text{all P}} \epsilon_P \lambda_P \tag{7}$$

where

$$\epsilon_P = \begin{cases} 2 & \text{if } P \text{ hits } j\text{-axis first} \\ 1 & \text{if } P \text{ hits } i\text{-axis first.} \end{cases}$$

The theorem gives a spanning set of $\simeq n^2/4$ sequences for $S_n^{(2)}$.

The idea of the proof is to write s_n as a function of a and b, in a different way. The standard formula, expressing s_n as a polynomial in a and b, is

$$s_n = a^n + \sum_{k=1}^{[\frac{n}{2}]} a^{n-2k} b^k \frac{n}{k}\binom{n-k-1}{k-1}, \tag{8}$$

readily obtained from the generating function $\Sigma s_n x^n = (ax+2bx^2)/(1-ax-bx^2)$. Instead, we write s_n as

$$s_n = \sum_{i+2j\le n} c_n(i,j) a_i b_j,$$

the range of i and j being clear from (8). Here

$$a_i = \binom{a}{i} i! = a(a-1)\ \dots\ (a-i+1)$$

(b_j similarly). Define

$$\Delta_a f(a,b) = f(a+1,b) - f(a,b)$$
$$\Delta_b f(a,b) = f(a,b+1) - f(a,b).$$

Then $\Delta_a a_i = i a_{i-1}$, $\Delta_a^i a_i = i!$, $\Delta_a^i a_\ell \big|_{a=0} = \begin{cases} i! & i=\ell \\ 0 & i\neq\ell \end{cases}$. Hence

$$\Delta_a^i \Delta_b^j \Big(\sum_{k=1}^{n} s_k x^k \Big)\Big|_{\substack{a=0 \\ b=0}} = \Big(\sum_{k=1}^{n} c_k(i,j) x^k \Big) i!\, j!$$

so that for any i,j the sequence (series) $c_{ij} = \sum_{k=1}^{n} c_k(i,j) x^k$ lies in $S^{(2)}$. Conversely, since $\{s_n\} = \{0,0,\dots\}$ if $a=b=0$ any $\sum_{k=1}^{n} s_k x^k$ can be

written as an integer linear combination of the c_{ij}, using Newton's two-dimensional interpolation formula. Thus the c_{ij} span $S_n^{(2)}$.

We can calculate a recurrence for the $c_k(i,j)$ as follows :

$$a \times a_i = a_{i+1} + ia_i$$

(and similarly for $b \times b_j$), so from $s_k = as_{k-1} + bs_{k-2}$ we get

$$c_k(i,j) = c_k(i-1,j) + ic_{k-1}(i,j) + c_{k-2}(i,j-1) + j\, c_{k-2}(i,j)$$

with $c_k(i,j)=0$ if $k<1$ or $i<0$ or $j<0$. Thus

$$c_{ij} = x\, c_{i-1j} + xic_{ij} + x^2 c_{ij-1} + x^2 jc_{ij},$$

$$c_{ij} = (xc_{i-1j} + x^2 c_{ij-1})/(1-ix-jx^2).$$

From this recurrence, and $c_1(1,0)=1$, $c_2(0,1)=2$ we readily obtain the result of the theorem.

The theorem can readily be generalised to $S^{(d)}$, generated by $\{s_n\}$ corresponding to α of degree $\leq d$. Returning to the case $d=2$, we can use the theorem to obtain an integer-reduced echelon form basis for $S_{12}^{(2)}$ (and thus also for all $n \leq 12$), using a computer. The basis is

$$\begin{bmatrix}
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
 & 2 & 0 & 2 & 0 & 2 & 0 & 2 & 0 & 2 & 0 & 2 \\
 & & 3 & 0 & 0 & 3 & 0 & 0 & 3 & 0 & 0 & 3 \\
 & & & 4 & 0 & 0 & 0 & 4 & 0 & 0 & 0 & 4 \\
 & & & & 10 & 0 & 28 & 0 & 0 & 60 & 198 & 0 \\
 & & & & & 6 & 0 & 0 & 0 & 0 & 0 & 6 \\
 & & & & & & 42 & 0 & 0 & 0 & 132 & 12 \\
 & & & & & & & 8 & 0 & 60 & 0 & 12 \\
 & & & & & & & & 36 & 0 & 0 & 0 \\
 & & & & & & & & & 120 & 0 & 0 \\
 & & & & & & & & & & 660 & 0 \\
 & & & & & & & & & & & 24
\end{bmatrix}$$

From this basis, it is easy to obtain congruences of the type (5) for sequences in $S^{(2)}$. Note that except for $n=10$, in every other case $N(n)$ is the leading term $M_n^{(2)}$ of the n^{th} row. The congruences are :

$s_2 \equiv s_1 \pmod 2$, $s_3 \equiv s_1 \pmod 3$, $s_4 \equiv s_2 \pmod 4$, $s_5 \equiv s_1 \pmod{10}$,
$s_6 \equiv s_3+s_2-s_1 \pmod 6$, $s_7 \equiv 7s_5-6s_1 \pmod{42}$, $s_8 \equiv s_4 \pmod 8$,
$s_9 \equiv s_3 \pmod{36}$, $s_{10} \equiv s_2 \pmod{60}$, $s_{11} \equiv -44s_7+11s_5+34s_1 \pmod{660}$,
$s_{12} \equiv 2s_{11}-s_{10}+9s_8-6s_7+s_6-2s_4-6s_2+4s_1 \pmod{24}$.

Applications : 1. Irreducibility testing : we can show that a polynomial cannot decompose into only linear or quadratic factors by finding one of the above congruences which the sequence $\{s_n\}$ for the polynomial does not satisfy.

Example : $x^4-2x^3-2x^2-x-1$ has $\{s_n\}= \{2, 8, 23, 68, 192, 551, 1577,\ldots\}$. But $s_7 \not\equiv 7s_5- 6s_1 \pmod{42}$, so the polynomial must be irreducible, or have an irreducible cubic factor. Since it clearly has no linear factor, it is irreducible.

2. Lehmer's Problem : Recent methods of attack on Lehmer's problem (see e.g. [1]) have used determinants of Vandermonde type. One part of the method is to bound the determinant from below in modulus. For this purpose it is useful to have determinants which can be shown to be large.

In this context it may be useful to consider determinants of the form

$$\begin{vmatrix} s_1^{(1)} & \cdots & s_1^{(n)} \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ s_n^{(1)} & \cdots & s_n^{(n)} \end{vmatrix}$$

where monic integral polynomials P_i have root powersum sequence $\{s_1^{(i)}, s_2^{(i)}, \ldots\}$, $(i= 1,\ldots,n)$. If d is the largest degree of any irreducible factor of any P_i, it follows that the modulus of the determinant is either 0 or $\geq \prod_{i=1}^{n} M_i^{(d)}$. Here the $M_i^{(d)}>0$ are the leading terms of the integer reduced echelon form of the basis for $S_n^{(d)}$, the generalisation of $S_n^{(2)}$ to α of degree $\leq d$.

3. Reducible polynomials differing by an integer. If two monic integral polynomials of degree n both have only linear factors over the rationals, and differ by a non-zero integer $k_1(n)$, then it is certainly known that $|k_1(n)| \geq (n-1)!$ (this can be proved without difficulty using (6)). One can

extend this to show that if the polynomials have only factors of degree $\leq d$, and differ only by a constant $k_d(n) \neq 0$ then $|k_d(n)| \geq M_n^{(d)}/n$.

BIBLIOGRAPHY

[1] D.C. Cantor and E.G. Straus.- On a conjecture of D.H. Lehmer, Acta Arithmetica, 42 (1982), 97-100.

[2] C.J. Smyth.- A coloring proof of a generalisation of Fermat's Little Theorem, 93 (1986), 469-471.

[3] J.H. van Lint.- Introduction to Coding Theory, Springer 1982.

C.J. Smyth
Department of Mathematics
James Clerk Maxwell Building
The King's Buildings
Mayfield road
Edinbourgh EH9 3JZ
ECOSSE

UN PROBLÈME CENTRAL

EN GÉOMETRIE ALGORITHMIQUE DES NOMBRES:

LA RÉDUCTION DES RÉSEAUX.

AUTOUR DE L'ALGORITHME DE LENSTRA LENSTRA LOVASZ.

BRIGITTE VALLÉE
Département de mathématiques,
Université de CAEN, 14032 CAEN, FRANCE

Nous rappelons d'abord le cadre général de la Géométrie des Nombres, tant classique que algorithmique; nous définissons les principaux problèmes des réseaux et insistons sur leurs nombreuses applications.

Nous définissons ensuite les différentes notions de réduction, et décrivons en particulier l'algorithme de Gauss qui résoud le problème parfaitement en dimension 2. Puis nous précisons la notion de réduction au sens de Lovász.

Nous décrivons alors l'algorithme dû à Lenstra, Lenstra et Lovász qui construit en temps polynomial une base réduite au sens de Lovász et nous analysons précisément sa complexité. Puis nous mentionnons d'autres algorithmes, qui en sont issus et qui permettent de calculer aisément dans les réseaux. Nous insistons enfin sur la simplicité de l'implémentation de tels algorithmes.

Nous continuons en présentant le champ d'application très étendu d'un tel algorithme, d'abord interne à la théorie des réseaux. Nous terminons en insistant sur des applications externes à la théorie et en décrivant des thèmes particuliers : théorie des nombres, algèbre ou cryptographie.

Cet exposé veut juste décrire l'ampleur des problèmes posés, aussi bien que la qualité des réponses qui y sont apportées. Pour plus de précisions, on pourra se reporter aux références bibliographiques finales.
Citons ici une référence générale, qui contient une bibliographie exhaustive sur le sujet et où la technique des problèmes est abordée de manière plus détaillée qu'ici:
R.Kannan, Algorithmic Geometry of Numbers, à paraitre dans Annual Reviews in Computer Science

1. LA PROBLÉMATIQUE GÉNÉRALE DES RÉSEAUX.

$\mathbf{R}^p$ est muni de sa structure euclidienne canonique, et on désigne par $|v|$ la norme de v et par $(u|v)$ le produit scalaire de u et v.
Par ailleurs $[r]$ désignera l'entier le plus proche du réel r.

Un réseau de $\mathbf{R}^p$ est l'ensemble des combinaisons linéaires à coefficients entiers de vecteurs linéairement indépendants de $\mathbf{R}^p$: c'est un sous-groupe discret de $\mathbf{R}^p$. Un tel réseau est dit *entier* s'il est inclus dans $\mathbf{Z}^p$.
Si $b = (b_1, b_2, ..., b_n)$ est un système de n vecteurs linéairement indépendants de $\mathbf{R}^p$, $(n \leq p)$, le réseau engendré par b, noté $L(b)$, est l'ensemble $\{ \sum_{i=1}^n \lambda_i b_i \ / \ \lambda_i \in \mathbf{Z}\}$ et n est appelé rang ou dimension du réseau.

1.1 . Le déterminant d'un réseau.

Soient $(b_1, b_2, ..., b_n)$ et $(c_1, c_2, ..., c_n)$ deux bases du même réseau, et B et C leurs matrices $(n \times p)$ dans la base canonique de $\mathbf{R}^p$. Il existe alors une matrice unimodulaire U (matrice entière dont le déterminant vaut ± 1) vérifiant $C = UB$.

On remarque donc que le volume n-dimensionnel du parallélépipède construit sur une base quelconque du réseau est indépendant de la base: c'est un invariant du réseau, noté $d(L)$, et appelé le déterminant du réseau. Cette quantité est aisément calculable: si $G(b)$ désigne la matrice de Gram de la base b, c'est-à-dire la matrice B^tB de coefficient général $(b_i|b_j)$, on a

$$\det G(b) = \det B^tB = d(L)^2.$$

1.2. Les minima successifs des réseaux.

Il existe, dans un réseau, d'autres objets qui ne dépendent que du réseau et non d'une base servant à le définir; en particulier, les minima successifs...
Pour un réseau L de $\mathbf{R}^p$ de dimension n , on définit le i-ème minimum $\Lambda_i(L)$ comme étant le plus petit réel positif t pour lequel il existe i vecteurs v linéairement indépendants de L vérifiant $|v|^2 \leq t$.
Il est clair que les n nombres $\Lambda_i(\mathrm{L})$ sont bien définis et vérifient

$$\Lambda_1(L) \leq \Lambda_2(L) \leq ... \leq \Lambda_n(L)$$

et qu'il existe un ensemble - non nécessairement unique - de n vecteurs linéairement indépendants de L , appelés aussi minima successifs, notés $\lambda_i(\mathrm{L})$ vérifiant

$$\text{pour tout } k, \ \ 1 \leq k \leq n, \ \ |\lambda_k(L)|^2 = \Lambda_k(L)$$

En particulier, $\lambda_1(\mathrm{L})$ désigne un plus court vecteur du réseau L.

1.3. Les problèmes théoriques et pratiques des réseaux.

Les problèmes *théoriques* sont de deux types:

1. .Chercher à relier les quantités intrinsèques d'un réseau, en particulier les $\Lambda_i(L)$ et $d(L)$.
2. Construire, exhiber, au moyen d'algorithmes, les objets intrinsèques d'un réseau, en particulier ses vecteurs minimaux successifs $\lambda_i(L)$.

Il arrive parfois que, en voulant résoudre le premier type de problèmes, on trouve une méthode explicite qui résolve du même coup le deuxième type de problèmes. Ce n'est jamais le cas pour cette théorie, et c'est ce qui explique la nécessité et l'importance de l'algorithmique en Géométrie des Nombres.

Dans la *pratique*, on se pose d'autres sortes de problèmes, qui sont essentiellement des problèmes de calcul dans les réseaux entiers; citons-en quelques uns:

- Soit un réseau L donné par une base b et un vecteur v. Décider si v appartient au réseau $L(b)$

- Soit un réseau L engendré par un système c de vecteurs non nécessairement indépendants. Trouver une base b du réseau.

- Soit $b = (b_1, b_2, ..., b_n)$ un système de n vecteurs de $\mathbf{Z}^p$. Trouver le réseau L des relations, c'est-à-dire l'ensemble

$$\{v = (v_1, v_2, ..., v_n) \in \mathbf{Z}^n \ / \ \sum_{i=1}^{n} v_i \, b_i = 0 \}$$

1.4. Les résultats théoriques non constructifs: Hermite, Minkowski.

Il existe [7] une constante γ_n, dite constante d'Hermite, ne dépendant que de n et vérifiant

$$\gamma_n = \text{Max} \{ \frac{\Lambda_1(L)}{d(L)^{2/n}} \ / \ L \text{ réseau de rang } n\}.$$

On a aussi:

$$\prod_{i=1}^{n} \Lambda_i(L) \leq \gamma_n^n d^2(L)$$

Les preuves de ces résultats sont non constructives, et la première majoration connue de γ_n, due à Hermite, assez peu fine. Puis Minkowski a obtenu les inégalités suivantes:

$$\gamma_n \leq \frac{2}{3}n \text{ si } n \text{ est pair et } \gamma_n \leq \frac{1}{2}n \text{ si } n \text{ est impair}$$

Seules les huit premières valeurs de cette constante sont exactement connues;

1.5. Les problèmes algorithmiques des réseaux.

Un réseau entier L étant donné par son rang n et une base b de longueur $M = \mathrm{Max}_i\ |b_i|$, on se pose les problèmes suivants:

1. *Déterminer $\lambda_1(L)$, un plus court vecteur du réseau L*
2. *Déterminer les $\lambda_i(L)$, une suite de vecteurs minimaux successifs du réseau L*

L'intérêt porté à ces problèmes est dû à la conjonction de trois facteurs:

- ce sont des problèmes **probablement difficiles**
- qui admettent pourtant des solutions **approchées**
- permettant la résolution d'autres problèmes, **variés et essentiels**, en théorie des nombres, en algèbre ou en cryptographie.

Nous allons décrire, tout au long de cet exposé, le second pôle, dans les sections 2. et 3., puis le troisième pôle, dans la section 4.
Attardons-nous un instant sur le premier pôle:

1.6. La difficulté probable de ces problèmes.

Le second problème est NP-dur en la donnée $(n, \log M)$ [17].

Rien n'est connu sur la "facilité" du premier: on ne connait, à l'heure actuelle, aucun algorithme polynomial en $(n, \log M)$ qui résolve ce problème, à première vue plus facile que le second. L'opinion courante semble croire aussi en sa "dureté", en vertu de trois arguments principaux:

- ce problème est NP-dur pour la norme sup [21];
- le problème non homogène associé, qui consiste à chercher le point d'un réseau L le plus proche d'un point donné de $\mathbf{Q}^n$, est lui aussi NP-dur, même pour la norme euclidienne [21];
- les inégalités de Minkowski ne sont pas plus fines pour le premier minimum que pour la moyenne géométrique des autres minima successifs.

Deux points de vue différents mais complémentaires peuvent alors exister pour tourner cette difficulté presque sûre :

1. Chercher des algorithmes *approchés* polynomiaux en la donnée $(n, \log M)$.
2. Chercher, *à dimension n fixée*, des algorithmes *exacts* polynomiaux en la donnée $\log M$.

Ce ne sont pas d'ailleurs des points de vue divergents.

Le second point de vue est fructueux en petite dimension: l'algorithme de Gauss est un algorithme polynomial qui trouve les deux minima d'un réseau de dimension 2. Sa complexité est parfaitement connue, et il admet une généralisation totale en dimension 3 [19].

Le premier point de vue utilise alors en procédures des algorithmes de ce type -exacts en petite dimension- pour construire en dimension supérieure des algorithmes approchés. On obtient alors ce qu'on appelle une *base réduite*: c'est une base formée de vecteurs *"assez courts" et "assez orthogonaux"* qui permet de bien décrire le réseau et donc

1. de donner une bonne approximation des objets intrinsèques du réseau.

2. de pouvoir calculer facilement dans le réseau.

Nous verrons, dans la section 4., comment une telle base peut résoudre, de manière étonnament satisfaisante, la totalité des problèmes algorithmiques, tant théoriques que pratiques.

Nous décrivons maintenant plus précisément les notions de réduction des réseaux.

2. LA RÉDUCTION DES RÉSEAUX.

On cherche donc une base formée de vecteurs assez orthogonaux; remarquons qu'un réseau ne possède pas, en général, de base orthogonale. Le procédé d'orthogonalisation de Gram-Schmidt associe bien à une base b d'un réseau de $\mathbf{R}^p$ une base orthogonale b^* du Q-espace vectoriel engendré par b mais cette dernière n'appartient pas, en général au réseau $L(b)$.

2.1. Le procédé d'orthogonalisation de Gram-Schmidt.

Il associe à un système ordonné $b = (b_1, b_2, ..., b_n)$ le système $b^* = (b_1^*, b_2^*, ..., b_n^*)$ et la matrice $m = (m_{ij})$ qui exprime le système b dans le système b^* définis comme suit:

(i) $b_1^* = b_1$

(ii) b_i^* est le projeté de b_i orthogonalement au sous-espace H_{i-1} engendré par les $i-1$ premiers vecteurs de b. On peut donc écrire

$$b_i = b_i^* + \sum_{j<i} m_{ij}\, b_j^* \quad \text{où} \quad m_{ij} \text{ est défini par la relation} \quad m_{ij} = \frac{(b_i|b_j^*)}{|b_j^*|^2}$$

ce qui permet de calculer facilement les b_i^* et les m_{ij} par récurrence sur i.

La matrice m est donc une matrice triangulaire inférieure, possédant une diagonale de 1.

Remarquons que $d(L)$ est égal au produit des longueurs des vecteurs b_i^*.

Si b est un système de n vecteurs de $\mathbf{Z}^p$, de longueur M, le calcul du couple (b^*, m) est polynomial en la taille de la donnée $(n, \log M)$. Soient L_i le réseau engendré par les i premiers vecteurs de b, et $d_i = d(L_i)^2$ le déterminant de Gram associé.

L'inégalité d'Hadamard permet d'affirmer que

$$d_i \leq \prod_{j=1}^{i} |b_j|^2 \leq M^{2i}$$

D'autre part, les rationnels intervenant dans b^* ou dans m ont comme dénominateurs les d_j; plus précisément:

$$|b_i^*|^2 = \frac{d_i}{d_{i-1}} \quad \text{pour tout} \quad i,\ 2 \leq i \leq n$$

$$d_{i-1}\, b_i^* \in \mathbf{Z}^p \quad \text{pour tout} \quad i,\ 2 \leq i \leq n$$

$$d_j\, m_{ij} \in \mathbf{Z} \quad \text{pour tout couple } (i,j)\ 1 \leq j < i \leq n$$

Remarquons donc que la quantité

$$D = \prod_{j=1}^{n-1} d_j$$

représente un dénominateur commun à tous les rationnels intervenant dans le couple (b^*, m).

2.2. Les défauts de longueur et d'orthogonalité.

Les deux conditions cherchées -vecteurs assez courts, vecteurs assez orthogonaux- sont heureusement compatibles en vertu des résultats de Hermite et Minkowski.
Soit $b = (b_1, b_2, ..., b_n)$ une base d'un réseau L; les deux paramètres suivants mesurent la qualité de la base :

Le rapport $\rho(b) = \frac{\prod_{i=1}^{n} |b_i|^2}{d(L)^2}$ s'appelle le *défaut d'orthogonalité* de la base b.

Le rapport $\mu_i(b) = \frac{|b_i|^2}{\Lambda_i(L)}$ s'appelle le i-ème *défaut de longueur* de la base b.

Les résultats de 1.4. permettent de montrer la liaison entre ces deux paramètres, qui vérifient la double inégalité:

$$\frac{1}{n!} \leq \frac{\rho(b)}{\prod_{i=1}^{n} \mu_i(b)} \leq \gamma_n^n$$

2.3. Les différentes notions de réduction.

Il n'existe pas *une* réduction; historiquement, il s'est dégagé principalement quatre notions de réduction.

Celle de Minkowski privilégie la recherche de vecteurs courts.

Les trois autres privilégient la recherche de vecteurs presque orthogonaux, et se décrivent donc aisément sur la matrice m, qui exprime le système b en fonction du système b^* : ce sont les réductions au sens de Korkhine-Zolotarev, Siegel et enfin Lovász.

On peut montrer que la réduction au sens de Siegel est la plus générale [7]: toute base réduite en un des trois autres sens est réduite au sens de Siegel. Cette réduction, qui est aussi la moins fine, est suffisante dans beaucoup d'applications, car la base réduite obtenue est d'assez bonne qualité.
Les deux premières réductions

celle de Minkowski, qui est la meilleure possible pour les défauts de longueur,

celle de Korkhine-Zolotarev qui parait être la meilleure pour le défaut d'orthogonalité,

ne peuvent pas être obtenues -semble-t-il- en temps polynomial. Nous privilégierons alors les deux autres réductions, et nous montrerons qu'elles sont obtenues "facilement".

Si ces réductions diffèrent quelque peu en dimension quelconque, elles coincident toutes en dimension 2 avec la célèbre réduction de Gauss que nous décrivons maintenant, avant de préciser ces différentes notions de réduction.

2.4. La réduction de Gauss en dimension 2. [4]

Les minima successifs forment toujours dans un réseau un système de vecteurs indépendants, mais n'engendrent pas en général le réseau; c'est cependant vrai en petite dimension:

> *Les minima successifs d'un réseau de dimension $n \leq 4$ forment une base du réseau, appelée base minimale du réseau : c'est alors la "meilleure base" du réseau.*

En dimension 2, l'algorithme de Gauss construit en temps polynomial une base minimale du réseau; il généralise, en dimension 2, l'algorithme d'Euclide centré :

$$a = bq + r \text{ avec } -\frac{b}{2} < r \leq +\frac{b}{2}$$

Algorithme de Gauss
Donnée: une base (u, v) d'un réseau L.
Résultat: une base minimale (u, v) du réseau L.
Répéter
1. **Echanger** éventuellement u et v pour que $|u| \leq |v|$
2. **Translater** v parallèlement à u de manière à le raccourcir au maximum:
   plus précisément, choisir dans l'ensemble $\{ \ w = \epsilon(v - mu) \ / \ \epsilon = \pm 1, \ \ m \in \mathbf{Z}\}$
   le vecteur w qui vérifie $0 \leq \frac{(w|u)}{(u|u)} \leq \frac{1}{2}$.
   ce dernier est aisément calculable en fonction de $r = \frac{(v|u)}{(u|u)}$:
   on choisit $m = [r]$ et $\epsilon = signe(r - m)$
jusqu'à ce que $|v| \geq |u|$

On peut modifier le test d'arrêt, en le changeant en un test moins fin:
Si t est un nombre réel vérifiant la double inégalité $1 < t \leq \sqrt{3}$, on obtient ainsi un algorithme dit de t-Gauss qui "tourne" un peu moins longtemps , mais qui possède une configuration de sortie comparable: le triangle construit sur la base contient les deux minima du réseau.

Algorithme de t-Gauss
Donnée: une base (u, v) d'un réseau L.
Résultat: une base "quasi-minimale" (u, v) du réseau L.
Répéter
1. **Echanger** éventuellement u et v pour que $|u| \leq |v|$
2. **Translater** v parallèlement à u
jusqu'à ce que $|v| \geq \frac{1}{t}|u|$

2.5. Etude de la complexité de l'algorithme de Gauss.

Soient $k(t)$ et k le nombre d'itérations des algorithmes de t-Gauss et de Gauss sur la même base (u, v) de départ de longueur M . Il est clair que l'on a $k(t) \leq \log_t M + 1$.
On peut montrer aussi que, pour $1 \leq t \leq \sqrt{2}$, on a : $k(t) \leq k \leq k(t) + 1$
ce qui démontre la complexité polynomiale -non tout à fait triviale- de l'algorithme de Gauss.
Une étude plus fine du plus mauvais cas de l'algorithme de Gauss [19] permet d'exhiber la meilleure borne possible: on obtient

$$k \leq \log_{1+\sqrt{2}} M + 3$$

qui est une borne similaire à celle obtenue dans l'algorithme d'Euclide centré [3].

2.6. L'effet de l'algorithme de Gauss sur l'orthogonalisée (u^*, v^*).

A la sortie de l'algorithme de Gauss, le vecteur v vérifie les deux conditions

i) $|v| \geq \frac{1}{t}|u|$

ii) $0 \leq (v|u) \leq \frac{1}{2}(u|u)$

La projection de v orthogonalement à u, égale par définition à v^*, vérifie donc

$|v^*|^2 \geq (\frac{1}{t^2} - \frac{1}{4})|u^*|^2$. Posant $s = \sqrt{\frac{4t^2}{4-t^2}}$, nous obtenons donc $|u^*| \leq s|v^*|$

Remarquons que si l'on a $1 \leq t \leq \sqrt{2}$, on a $\frac{4}{3} \leq s^2 \leq 4$

La valeur $s^2 = 2$, correspondant à la valeur $t = \frac{2}{\sqrt{3}}$ est usuellement choisie pour simplifier les calculs.

2.7. La propreté d'une base.

L'idée la plus simple, pour rapprocher b de b^* est de diminuer les coefficients de la matrice m sans modifier ni b^*, ni le réseau $L(b)$; cela justifie la définition suivante qui reprend les notations de 2.1:

Une base $b = (b_1, b_2, ..., b_n)$ est propre si la matrice m associée a tous ses coefficients m_{ij} pour $j < i$ inférieurs, en valeur absolue, à 1/2.

Explicitons géométriquement cette condition: chaque vecteur b_i se projette orthogonalement sur l'hyperplan H_{i-1} à l'intérieur du parallélépipède rectangle construit sur les b_j pour $j < i$ et défini comme étant l'ensemble des vecteurs

$$\sum_{j=1}^{i-1} \alpha_j \, b_j^* \quad \text{pour} \quad -\frac{1}{2} < \alpha_j \leq \frac{1}{2}$$

Il existe un algorithme Totalement-Propre qui, étant donné un système $b = (b_1, b_2, ..., b_n)$, le rend propre; c'est une succession d'appels à Propre(i) qui généralise la seconde étape de l'algorithme de Gauss ; cette procédure translate b_i parallèlement à chaque vecteur b_j pour $j < i$ et ne modifie donc ni b_i^* ni $L(b)$.

Algorithme Propre(i)
  Pour j allant de $i-1$ à 1 faire
    $r_j := \lfloor m_{ij} \rceil$;
    $b_i := b_i - r_j \, b_j$;
Algorithme Totalement-Propre
  Pour i allant de 2 à n faire Propre(i);

2.8. La réduction au sens de Siegel.

Le fait qu'une base soit propre ne garantit pas le fait qu'elle soit presque orthogonale; on sait seulement pour le moment que la projection de chaque b_{i+1} sur H_i est assez petite. Pour pouvoir minorer l'angle θ_{i+1} que forme b_{i+1} avec le plan H_i, on a besoin de minorer en plus la projection de b_{i+1} orthogonalement à H_i, c'est-à-dire la longueur de b^*_{i+1}. C'est l'objet de la condition de réduction de Siegel:

$$|b^*_{i+1}| \geq \frac{1}{s}\,|b^*_i| \quad \text{pour tout} \quad i,\ 1 \leq i \leq n-1$$

Une base propre qui vérifie de plus la condition de Siegel pour le paramètre s est dite s-réduite au sens de Siegel.

Cette condition est suffisante pour assurer la qualité de la base et pour majorer les défauts de longueur et d'orthogonalité. On obtient les résultats suivants:

$$|\sin(\theta_i)| \geq \frac{1}{s^{i-1}} \quad \text{et donc} \quad |b_i| \leq |b^*_i|\ s^{i-1} \quad \text{pour tout} \quad i,\ 1 \leq i \leq n-1$$

On en déduit

$$\rho(b) \leq s^{n(n-1)/2} \quad \text{et aussi} \quad s^{-2(i-1)} \leq \mu_i(b) \leq s^{2(n-1)} \quad \text{pour tout} \quad i,\ 1 \leq i \leq n$$

2.9. La réduction au sens de Lovász.

Il reste deux questions essentielles

i) Tout réseau admet-il une base réduite au sens de Siegel? Si oui, pour quelles valeurs du paramètre s ?

ii) Si oui, existe-t-il un algorithme qui, partant d'une base quelconque b de longueur M d'un réseau L de rang n, construise une base de Siegel du réseau L en un temps polynomial en la donnée $(n, \log M)$? Toutes ces questions vont recevoir des réponses positives. Puisque l'existence et la constructibilité d'une base propre ne posent pas de problèmes, la question se résume ainsi:

Comment assurer la condition de Siegel ?

On va chercher à assurer une condition un peu plus forte : la condition de Lovász. Nous avons remarqué en 2.6. que l'algorithme de Gauss -en dimension 2- permet d'obtenir une inégalité sur les orthogonalisés proche de la condition de Siegel ; précisons:

Soit P_i l'orthogonal de H_{i-1} dans H_{i+1} et B_i le système (la "boîte") formé par les projections u_i et v_i de b_i et de b_{i+1} sur P_i. Par définition de b^*_{i+1}, on a $b^*_{i+1} = v^*_i$. Si donc, on applique l'algorithme de t-Gauss aux systèmes B_i, nous obtenons, à la sortie, les trois conditions suivantes valables pour tout i, $1 \leq i \leq n-1$:

(les paramètres t et s sont liés, comme en 2.6., par la relation $s = \sqrt{\frac{4t^2}{4-t^2}}$)

i) $0 \leq (v_i|u_i) \leq \frac{1}{2}\,(u_i|u_i)$

ii) $|v_i| \geq \frac{1}{t} |u_i|$

iii) $|u_i^*| \leq s\, |v_i^*|$

La première condition est une condition de propreté; la troisième est la condition de Siegel; la deuxième est appelée condition de Lovász.

D'après le paragraphe 2.6. , on a : i) + ii) $\Rightarrow$ i) + iii) . Tout ceci justifie la définition suivante:

Une base propre qui vérifie la condition de Lovász pour le paramètre t est dite t-réduite au sens de Lovász.

et permet d'affirmer

Soient s et t deux paramètres liés par la relation $s = \sqrt{\frac{4t^2}{4-t^2}}$. *Une base t-réduite au sens de Lovász est s-réduite au sens de Siegel.*

3. L'ALGORITHME DE LENSTRA LENSTRA LOVASZ. [14]

Cet algorithme construit, à partir d'une base b de longueur M d'un réseau L de rang n , une base t-réduite au sens de Lovász , en un temps polynomial en la taille de la donnée $(n, \log_t M)$. Pour $t = 1$, cet algorithme se termine sans qu'on sache, à l'heure actuelle, préciser davantage sa complexité.

3.1. Les principales phases de l'algorithme.

Cet algorithme se compose de trois phases principales :

- une phase d'**initialisation**; elle consiste essentiellement à calculer le système b^* , la matrice m et la liste ℓ formée des éléments $\ell_i = |b_i^*|^2$, par la procédure d'orthogonalisation de Gram-Schmidt décrit en 2.1. Ces deux derniers objets vont d'ailleurs être essentiels tout au long de l'algorithme.

- des phases de **translation** des vecteurs b_i parallèlement à H_{i-1} qui s'effectuent par les procédures **Propre** décrites en 2.7. Rappelons aussi que ces phases ne modifient pas le système b^*.

- des phases d'**échange** des vecteurs b_i et b_{i+1} afin de réaliser,

  i) sur le système B_i la condition de t-Gauss.

  ii) et donc sur le système b^* la condition de s-Siegel.

Le triplet (b^*, m, ℓ) est modifié lors de cet échange et nous devons en recalculer une partie, au moyen d'une procédure **Nouvortho** que nous décrirons plus loin.

Le choix -**translater ou échanger**- se fait en effectuant le test de l'algorithme de t-Gauss sur le système B_i défini en 2.9 et ce choix est répercuté ensuite sur les vecteurs $(b_1, b_2, ..., b_n)$ de la base b.

Remarquons d'abord que les vecteurs u_i et v_i du système B_i se lisent sur la matrice m: ce sont les vecteurs-ligne de la boîte B_i visualisée par un encadré ci-dessous.

$$m = \begin{array}{c} \\ b_1 \\ b_2 \\ \vdots \\ b_i \\ b_{i+1} \\ \vdots \\ b_n \end{array} \begin{array}{c} \begin{array}{cccccccc} b_1^* & b_2^* & \dots & b_i^* & b_{i+1}^* & \dots & b_n^* \end{array} \\ \left(\begin{array}{cccccccc} 1 & 0 & \dots & 0 & 0 & \dots & 0 \\ m_{21} & 1 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ m_{i1} & m_{i2} & \dots & 1 & 0 & \dots & 0 \\ m_{i+1,1} & m_{i+1,2} & \dots & m_{i+1,i} & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ m_{n1} & m_{n2} & \dots & m_{ni} & m_{n,i+1} & \dots & 1 \end{array} \right) \end{array}$$

En particulier, nous avons les relations suivantes:

$$u_i = b_i^* \quad , \quad v_i = b_{i+1}^* + m_{i+1,i}\, b_i^*$$

et donc les trois quantités intervenant dans l'algorithme de Gauss effectué sur B_i se calculent facilement en fonction de la liste ℓ et de la matrice m:

$$|u_i|^2 = \ell_i \quad , \quad |v_i|^2 = \ell_{i+1} + m_{i+1,i}^2\, \ell_i \quad \text{et aussi} \quad \frac{(v_i|u_i)}{(u_i|u_i)} = m_{i+1,i}$$

3.2. La description générale de l'algorithme.

Algorithme LLL (t) ;

Donnée : une base b d'un réseau L de rang n de $\mathbf{R}^p$.
Résultat : une base b de L t-réduite au sens de Lovász.

Gram;

i:=1 ;

Tant que $i < n$ **répéter**

1. **Translater** v_i parallèlement à u_i et donc b_{i+1} parallèlement à b_i : calculer $r_i = [m_{i+1,i}]$ et **faire** $v_i := v_i - r_i\, u_i$; $b_{i+1} := b_{i+1} - r_i\, b_i$;
2. **Tester si** $|v_i|^2 \geq \frac{1}{t^2}\, |u_i|^2$;

   Si **oui** la boîte B_i est réduite au sens de t-Gauss; **faire:**
   Translater alors b_{i+1} parallèlement aux b_j pour $j < i$ au moyen de **Propre**$(i+1)$;
   Modifier l'indice $i := i + 1$;
   Sinon faire:
   Echanger b_i et b_{i+1} ;
   Recalculer par la procédure **Nouvortho** le triplet (b^*, m, ℓ) ;
   La boîte B_{i-1} n'est plus nécessairement réduite:
   Modifier éventuellement l'indice **si** $i \neq 1$ **alors** $i := i - 1$.

La variable i désigne un indice, l'indice courant de l'algorithme qui va varier de 1 à $n-1$: c'est le plus grand indice k pour lequel le système $(b_1, b_2, ..., b_k)$ est t-réduit au sens de Lovász. La matrice m_i formée par les i premières lignes et les i premières colonnes de la matrice m et la liste formée par les i premiers termes de la liste ℓ ont déjà les formes souhaitées.

On considère alors la $i+1$-ème ligne de m, représentant le vecteur b_{i+1} et on effectue les opérations de translation ou d'échange suivant le résultat du test de t-Gauss.

Il reste à préciser maintenant la procédure **Nouvortho**.

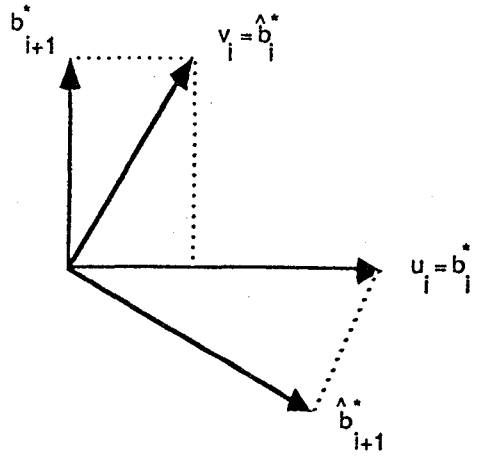
3.3. Les modifications de b^* de ℓ et de m effectuées dans la seconde étape.

Si le test est positif, ni b^* ni ℓ ne sont modifiés; seule la $i+1$-ème ligne de la matrice m est remplacée par une combinaison linéaire de la $i+1$-ème et des lignes précédentes, conformément à la description de la procédure Propre$(i+1)$.

Si, par contre, le test est négatif, l'échange des vecteurs b_i et b_{i+1} dans le système b modifie les deux vecteurs b_i^* et b_{i+1}^*. Soit $(\hat{b}, \hat{b}^*, \hat{m})$ le nouveau triplet recalculé par la procédure Nouvortho :

v_i est la projection de $\hat{b}_i = b_{i+1}$ sur P_i: on a donc $\hat{b}_i^* = v_i$; $\hat{b}_{i+1}^*$ est la projection de $\hat{b}_{i+1} = b_i$ orthogonalement à $\hat{H}_i$ et donc la projection de u_i sur l'orthogonal de v_i.

On a donc le schéma suivant:



On obtient aussi les formules suivantes:

$$\hat{m}_{i+1,i} = m_{i+1,i} \frac{|u_i|^2}{|v_i|^2}$$

et aussi

$$\begin{pmatrix} \hat{b}_{i+1}^* \\ \hat{b}_i^* \end{pmatrix} = \begin{pmatrix} -\hat{m}_{i+1,i} & 1 - \hat{m}_{i+1,i}\, m_{i+1,i} \\ 1 & m_{i+1,i} \end{pmatrix} \begin{pmatrix} b_{i+1}^* \\ b_i^* \end{pmatrix}$$

qui permettent de recalculer les colonnes i et $i+1$ de la matrice m et les éléments i et $i+1$ de la liste ℓ.

3.4. Le nombre de tours effectués par l'algorithme.

L'indice i est un indice qui croît si le test en 2. est positif et qui décroît si ce test est négatif. On va d'abord majorer en fonction de t, M, n le nombre de fois k_- où l'on passe en 2. avec un test négatif; Cette majoration suffira à conclure car le nombre de fois k_+ où l'on passe en 2. avec un test positif vérifie :

$$k_+ \leq k_- + n - 1$$

Le nombre total k d'itérations de l'algorithme vérifiera donc $k \leq n - 1 + 2k_-$

Comme dans l'étude de la complexité de la procédure **Gram** étudiée en 2.1 , c'est la quantité

$$D = \prod_{j=1}^{n-1} d_j$$

définie dans ce paragraphe 2.1 qui va jouer un rôle essentiel. Remarquons qu'à chaque passage en 2, avec un test négatif, avec une valeur de i donnée:

- les réseaux L_j pour $j < i$ et pour $j \geq i+1$ ne sont pas modifiés; donc les d_j correspondants non plus.

- par contre le réseau L_i est modifié et son déterminant de Gram également;

$$\text{On avait précédemment: } d_i = \prod_{j=1}^{i} |b_i^*|^2 = |u_i|^2 \prod_{j=1}^{i-1} |b_i^*|^2 \ ;$$

$$\text{On a maintenant : } \hat{d}_i = \prod_{j=1}^{i} |\hat{b}_i^*|^2 = |v_i|^2 \prod_{j=1}^{i-1} |b_i^*|^2$$

Le test de t-Gauss étant négatif, on en déduit que

$$\hat{d}_i < \frac{1}{t^2} d_i \quad \text{et donc} \quad \hat{D} < \frac{1}{t^2} D$$

D'autre part, lors de chaque passage en 2. avec un test positif, aucun des réseaux L_i n'est modifié. Donc D reste inchangé lors de cette étape. Finalement, D décroît tout au long de l'algorithme, et:

$$\text{au départ, on a : } D \leq M^{n(n-1)}$$

$$\text{à l'arrivée, on a : } D \geq 1$$

$$\text{On obtient donc la majoration : } k_- \leq \frac{n(n-1)}{2} \log_t M$$

3.5. La complexité de l'algorithme.

Il reste à borner la taille des nombres apparaissant dans l'algorithme en fonction de la donnée $(n, \log M)$.

Il est clair que les nombres ℓ_i sont des rationnels dont numérateur et dénominateur décroissent tout au long de l'algorithme.

Par contre, la situation est moins claire pour les entiers $|b_i|^2$ et la matrice rationnelle m : en effet, tout se passe bien en projection sur les plans P_i , mais, lors du relèvement, même choisi de manière minimale, on ne peut affirmer que la taille de ces quantités décroissent.

Un peu de technique, que nous ne développerons ici, permet de montrer que les majorations suivantes ont cours tout au long de l'algorithme -y compris dans les phases de relèvement- :

On a toujours :

$$|m_{ij}| \leq |b_i|\sqrt{d_j} \quad \text{et aussi} \quad |b_i| \leq \sqrt{2n}\, M^{2n}$$

et donc $|m_{ij}|$ est un rationnel dont la taille des numérateur et dénominateur est majorée par une quantité polynomiale en $(n, \log M)$.

Les opérations effectuées sur ces nombres sont très simples:

calcul d'entier le plus proche

élévations au carré.

On déduit de tout cela le théorème suivant:

L'algorithme de Lenstra, Lenstra, Lovász, associé au paramètre t construit, à partir d'une base b de longueur M d'un réseau L de rang n, une t-base réduite au sens de Lovász en un temps polynomial en $(n, \log_t M)$

3.6. Le cas particulier de la valeur $t = 1$ du paramètre t.

On ne sait pas prouver la complexité polynomiale de l'algorithme **LLL(1)** . On conjecture actuellement que le nombre d'itérations de cet algorithme est encore polynomial en la taille de la donnée $(n, \log M)$. Plusieurs arguments plaident en la faveur d'une telle conjecture :

- Lagarias et Odlyzko [12] ont effectué une étude expérimentale de cette conjecture : en pratique, le nombre d'itérations de **LLL(1)** ne dépasse pas trois fois le nombre d'itérations de **LLL**(t) pour $t^2 = 4/3$, valeur usuelle du paramètre.

- Nous avons montré en 2.5 combien le nombre d'itérations de l'algorithme de t-Gauss dépendait faussement du paramètre t.

3.7. La pratique de l'algorithme.

Le succès de l'algorithme provient aussi de la simplicité de sa mise en oeuvre : cet algorithme est plus simple à programmer qu'à comprendre , ce qui n'est pas si usuel pour un algorithme !

Toutes les opérations de l'algorithme s'effectuent sur le système b ou sur le triplet (b^*, m, ℓ); il est d'ailleurs facile de se convaincre que b^* doit être seulement calculé lors de l'initialisation, et qu'il n'est plus nécessaire, ni de le conserver, ni de le mettre à jour ensuite. On peut donc seulement travailler sur les trois données b, m, ℓ.

Nous montrons maintenant comment des adaptations simples de l'algorithme **LLL** permettent de résoudre de manière satisfaisante des problèmes pratiques de base:

- trouver une base d'un réseau
- trouver des relations linéaires entières.

On peut consulter [20] si on veut avoir une idée plus précise de la pratique de toutes ces implémentations.

3.8. La recherche d'une base d'un réseau donné par un système de générateurs. [5]
Soit $b = (b_1, b_2, ..., b_n)$ un système de générateurs d'un réseau L de rang s; on veut trouver une base du réseau, pour pouvoir calculer, par exemple, le déterminant du réseau. L'idée est de faire "comme si" les b_i étaient indépendants.

On peut en effet généraliser le procédé d'orthogonalisation de Gram-Schmidt à un système de vecteurs non indépendants : on obtient un couple (b^*, ℓ) et une liste I formée des indices i pour lesquels b_i^* et donc ℓ_i sont nuls. Par définition , le réseau L' est le réseau engendré par le système $\{b_i \ / i \notin I\}$ qui, par définition est un système de vecteurs indépendants : les deux réseaux L et L' engendrent bien sûr le même espace vectoriel de dimension s et L' est inclus dans L; on a donc $d(L) \leq d(L')$
L'idée est alors de faire décroître les indices $i \in I$, jusqu'à ce qu'ils soient tous au début; alors, les deux réseaux seront les mêmes et le système des vecteurs correspondant aux indices finaux sera le système cherché.

Pour cela, on utilise un algorithme **LLL** modifié, dont la structure générale est proche de celle de l'algorithme initial:

Algorithme Construction de base ;

Donnée : un système générateur $b = (b_1, b_2, ..., b_n)$ d'un réseau L de $\mathbf{R}^p$.
Résultat : une base b de L.

Gram;

$q := 0$;
Pour $i \in I$ **répéter**
$q := q + 1$;
Pour j allant de $i - 1$ à q **faire**

1. **Tant que** $\ell_j \neq 0$ faire
   Translater v_j parallèlement à u_j et donc b_{j+1} parallèlement à b_j ;
   Echanger b_j et b_{j+1} ;
   Recalculer par la procédure **Nouvortho** le triplet (b^*, m, ℓ) ;
2. **Translater** alors b_{j+1} parallèlement aux b_k pour $k < j$ au moyen de **Propre** $(j + 1)$.

Dans la procédure d'initialisation, on calcule les colonnes d'indice $i \notin I$ de la matrice m comme précédemment. Les colonnes d'indice $i \in I$ seront, par convention, égales à celles de la matrice identité correspondante.

Puis on procède là aussi par une succession d'échanges et de translations, mais uniquement sur les boîtes B_i associées à un indice i vérifiant $i+1 \in I$; pour ces indices-là, ces boîtes contiennent deux vecteurs u_i et v_i colinéaires: puisqu' elles sont "aplaties", l'algorithme de Gauss y coincide avec l'algorithme d'Euclide centré, et la procédure Nouvortho est juste un échange entre u_i^* et v_i^*. L'étude de la complexité de cet algorithme est assez proche de l'algorithme classique. La quantité qui décroît ici tout au long de l'algorithme est la suivante :

$$D = \prod_{i \in I} 2^i \prod_{i \notin I} |b_i^*| = \prod_{i \in I} 2^i \quad d(L')$$

Lors de chaque boucle interne, la partie I n'est pas modifiée, mais L', lui, est modifié par l'échange des vecteurs de la boîte et $d(L')$ est divisé par 2. Lors d'une boucle externe, L' n'est pas modifié, mais, par contre, un indice $i \in I$ diminue d'une unité: c'est au tour de la première quantité d'être divisée par 2.

Nous remarquons donc que le nombre d'itérations de l'algorithme est une fonction linéaire de n - à s fixé, alors qu'il était quadratique dans l'algorithme classique..

3.9. La recherche d'une relation linéaire courte entre n vecteurs de $\mathbf{Z}^p$.

Soit $y = (y_1, y_2, ..., y_n)$ le système formé par ces vecteurs, Y la matrice dont les colonnes sont les y_i. Soit $x = (x_1, x_2, .., x_p)$ le système formé par les lignes de la matrice Y et L le réseau de $\mathbf{Z}^n$ engendré par x qu'on suppose de rang q ($q \leq p$). On veut construire un vecteur court de ce qu'on appelle *le réseau des relations* c'est-à-dire le réseau R des vecteurs $v = (v_1, v_2, ..., v_n)$ de $\mathbf{Z}^n$ vérifiant

$$\sum_{i=1}^{n} v_i \, y_i = 0 \quad \text{et donc} \quad (v|x_i) = 0 \text{ pour tout } \; i \;\; 1 \leq i \leq p$$

On procède de la manière suivante:

1. On construit une base $b = (b_1, b_2, ..., b_n)$ du réseau $\mathbf{Z}^n$ tel que les q premiers vecteurs de b engendrent le même $\mathbf{Q}$-sous-espace vectoriel H que x.

2. Les derniers $n - q$ vecteurs de la base $c = (c_1, c_2, ..., c_n)$ duale de la base b sont alors une base du réseau des relations.

3. Il reste alors à chercher un vecteur court de ce réseau.

Il est clair que **LLL** résoud l'étape 3. Il est vrai aussi qu'un algorithme assez semblable à celui du paragraphe précédent permet de résoudre la première étape:

Partant de la base canonique b de $\mathbf{Z}^n$, nous définissons

1. le système b^* formé par les vecteurs b_i^* , projections des vecteurs b_i orthogonalement aux sous-espaces $K_{i-1} = H + H_{i-1}$

2. le couple (m, ℓ) et la partie I correspondant dont le cardinal est q, dimension de H

Travaillant alors sur le triplet (b^*, m, ℓ), nous cherchons par une succession d'échanges et de translations à faire décroître les indices $i \in I$ jusqu'à ce que $I = \{1, 2, ..., q\}$: nous avons ainsi obtenu la base b cherchée.

4. LE CHAMP D'APPLICATIONS DE L'ALGORITHME

Il s'agit de montrer ici comment l'algorithme **LLL** permet de résoudre de manière satisfaisante les problèmes internes à la théorie des réseaux mais aussi beaucoup d'autres problèmes externes.

Les applications internes, les trois premières, permettent de résoudre polynomialement, à dimension fixée, les problèmes difficiles de la théorie.

Les applications externes, au moins les trois premières d'entre elles, sont si essentielles en algorithmique qu'elles ont été un moteur puissant pour l'élaboration même de l' algorithme **LLL**.

Les dernières ont agi après coup en utilisant l'algorithme déjà existant.

Cet exposé ne prétend pas à l'exhaustivité sur ce sujet : on veut juste donner un apercu de l'importance de l'utilisation de cet algorithme.

4.1. Le vecteur le plus court du réseau.

Le premier vecteur b_1 de la base réduite obtenue est assez court, grâce à la majoration du premier défaut de longueur d'une base s-réduite au sens de Siegel. Pour les valeurs usuelles des paramètres s et t, on obtient:

$$|b_1|^2 \leq 2^{n-1}\Lambda_1(L)$$

Nous verrons plus tard comment ce vecteur b_1 peut jouer, dans les applications, le même rôle que le vecteur $\lambda_1(L)$, même s'il est en général plus long que lui.

Ici, nous nous posons la question:

Comment, à partir d'une base réduite au sens de Siegel, trouver $\lambda_1(L)$?

Rappelons qu'à ce jour, aucun algorithme polynomial n'est connu pour résoudre ce problème. Il existe essentiellement trois algorithmes de complexité décroissante mais de complication croissante, tous trois essentiellement dûs à Kannan.

Le premier opère une simple mais longue recherche systématique. Exprimant $\lambda_1(L)$ dans la base b, sous la forme $\lambda_1(L) = \sum_{i=1}^{n} \beta_i b_i$, les formules de Cramer donnent :

$$\beta_i = \frac{\det\,(b_1, b_2, ..., b_{i-1}, \lambda_1(L), b_{i+1}, ..., b_n)}{\det\,(b_1, b_2, ..., b_n)}$$

En utilisant l'inégalité d'Hadamard, et la définition de $\lambda_1(L)$, on obtient: $|\beta_i| \leq \rho(b)$
Puisque b est de Siegel, le défaut d'orthogonalité $\rho(b)$ est majoré et on obtient:

$$\rho(b) \leq 2^{n(n-1)/4} \text{ on en déduit : } |\beta_i| \leq 2^{n(n-1)/4} \text{ pour tout } i \quad 1 \leq i \leq n$$

On en déduit donc un algorithme qui doit calculer la longueur de 2^{n^3} vecteurs du réseau.

Le second procède de manière récursive en utilisant un argument géométrique assez simple: la longueur $|b_n^*|$ mesure la distance entre deux hyperplans consécutifs du réseau parallèles à H_{n-1}.

Or, puisque b est s-réduite au sens de Siegel, ces hyperplans sont assez "espacés" et on a, d'après le paragraphe 2.8:

$$|b_n^*| \geq \frac{1}{s^{n-1}} |b_n| \quad \text{et donc} \quad |b_n^*| \geq \frac{1}{s^{n-1}} |\lambda_1(L)|$$

Par conséquent, $\lambda_1(L)$ ne peut se trouver que dans un petit nombre d'hyperplans de direction parallèle à H_{n-1} (ce "petit" nombre est de l'ordre de $2s^{n-1}$) : on projette successivement dans ce nombre fini d'hyperplans affines, et dans chacun d'eux on peut utiliser le même genre d'arguments car l'inégalité précédente est vraie quand on remplace n par $n-1$ et $\lambda_1(L)$ par ses projetés dans ces hyperplans.

On obtient ainsi un algorithme qui considère $2^n s^{n(n-1)/2}$ vecteurs du réseau. Du fait que cet algorithme est affine et non pas vectoriel comme les deux autres, nous y reviendrons dans le paragraphe suivant 4.2.

Le troisième [9] construit, à partir d'une base réduite au sens de Siegel, une base réduite au sens de Korkhine-Zolotarev : nous y reviendrons dans le paragraphe 4.3.

4.2. La recherche du vecteur d'un réseau L le plus proche d'un point donné N.

Rappelons que l'on sait que ce problème est NP-dur.

Il existe pour le résoudre un algorithme dû à Babai [1], qui reprend les mêmes principes que le second algorithme de la section précédente et qui a la même complexité.

Par contre, si on cherche seulement un point "assez" proche, on possède un algorithme polynomial, fondé aussi sur les mêmes principes, qui trouve un point A du réseau vérifiant

$$d(N,A) \leq 2^{(n-1)/2}\, d(N,L) \quad \text{où, par définition,} \quad d(N,L) = \text{Min } \{\, d(N,A)/A \in L \,\}$$

4.3. Les autres réductions.

Nous avions mentionné dans la section 1. les réductions au sens de Minkowski et au sens de Korkhine-Zolotarev.

Il est prouvé que la première réduction est NP-dure. Helfrich-Just [6] a cependant construit à dimension n fixée, un algorithme polynomial, qui, partant d'une base réduite au sens de Siegel, détermine une base réduite au sens de Minkowski. En dimension 3, on peut également construire un algorithme polynomial qui, généralisant exactement l'algorithme de Gauss, construit directement une base réduite au sens de Minkowski, sans réduction préalable au sens de Siegel [19].

La seconde réduction a la même "dureté" que la recherche du plus court vecteur. Bien que cette réduction soit donc moralement NP-dure, nous avons déjà mentionné en 4.1 qu'un algorithme de Kannan [9] résoud le problème polynomialement à dimension n fixée.

4.4. La factorisation des polynômes à coefficients entiers.

L'idée fondamentale est la suivante :

Etant donné un polynôme $f(X)$ à coefficients entiers de degré n et de longueur $M(f) = \max |f_i|$ et une assez bonne approximation d'une racine α de f on peut déterminer h le polynôme minimal du nombre algébrique α qui est par définition un facteur irréductible de f

L'approximation $\bar{\alpha}$ de α sera

soit complexe et obtenue par l'algorithme de Newton [10].

soit p-adique, et obtenue alors par l'algorithme de factorisation mod p dû à Berlekamp suivi d'un relèvement par le lemme de Hensel [14].

Si l'approximation est suffisamment fine, on peut alors appliquer un principe de séparation qui affirme : il existe δ dont la taille est polynomiale en la taille de la donnée $(n, \log M)$ tel que les deux propositions soient équivalentes :

i) g est multiple de h

ii) $|g(\bar{\alpha})| \leq \delta$ (la valeur absolue est archimédienne ou p-adique selon le cas envisagé)

Dans le premier cas, la proposition ii) incite donc à chercher un vecteur court du réseau L engendré par les lignes v_i ($0 \leq i \leq n$) de la matrice

$$A = \begin{pmatrix} C & 0 & 0 & \dots & 0 & 1 & 0 \\ 0 & C & 0 & \dots & 0 & \beta_1 & \gamma_1 \\ 0 & 0 & C & \dots & 0 & \beta_2 & \gamma_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & C & \beta_n & \gamma_n \end{pmatrix}$$

où $\beta_i = \Re(\bar{\alpha}^i)$, $\gamma_i = \Im(\bar{\alpha}^i)$, et C est une constante dépendant polynomialement de $M(f)$ et de δ.

On montre effectivement qu'en appliquant l'algorithme **LLL** à la matrice ci-dessus, on peut construire exactement h: le premier vecteur de la base réduite obtenue, mis sous la forme

$$v = \sum_{i=0}^{n} h_i \, v_i \quad \text{permet de construire le polynôme } h \text{ sous la forme } \quad h = \sum_{i=0}^{n} h_i \, X^i$$

Dans le second cas, on considère un nombre premier p, et on détermine, par l'algorithme de Berlekamp, un polynôme g de degré m vérifiant les deux propriétés:

i) g mod p est irréductible dans $\mathbf{F}_p[X]$

ii) g mod p divise f mod p dans $\mathbf{F}_p[X]$

On choisit alors une puissance de p, suffisamment grande, de la forme p^ℓ et on "relève" g mod p en un polynôme g mod p^ℓ. Nous cherchons alors un polynôme $h \in \mathbf{Z}[X]$ de degré inférieur ou égal à q tel que:

i) h soit un facteur irréductible de f dans $\mathbf{Z}[X]$

ii) h mod p^ℓ soit un multiple de g mod p^ℓ

Nous travaillons donc dans le réseau

$$L = \{\, \phi \in \mathbf{Z}[X] \text{ de degré } q \;/\; \phi = p^\ell b + ag \text{ pour } a \text{ et } b \in \mathbf{Z}[X] \,\}$$

qui admet la base $V = \{p^\ell, p^\ell X, p^\ell X^2, ..., p^\ell X^{m-1}, g, gX, gX^2, ..., gX^{q-m}\}$. Remarquons donc que si p^ℓ est suffisamment grand, en fonction de la hauteur de g qu'on sait borner en fonction de celle de f, les vecteurs courts de L auront, dans la base V, leurs m premières composantes nulles et seront donc des petits multiples de g.

4.5. Les approximations diophantiennes simultanées.

Le problème à résoudre est le suivant:

Soit $(\alpha_1, \alpha_2, ..., \alpha_n)$ un n-uplet de nombres réels. On cherche n nombres entiers $(p_1, p_2, ..., p_n)$ et un nombre entier q tels que les n nombres rationnels $(p_1/q, p_2/q, ..., p_n/q)$ soient de bonnes approximations des nombres donnés.

On connait une réponse à cette question, due à Dirichlet, fondée sur le théorème de Minkowski, et donc non constructive:

Pour tout n , pour tout n-uplet $(\alpha_1, \alpha_2, ..., \alpha_n)$, pour tout couple (ϵ, Q) vérifiant $\epsilon > 0$ et $Q \geq \epsilon^{-n}$, il existe des entiers $(p_1, p_2, ..., p_n)$ et un entier q vérifiant

$$0 < q \leq Q \text{ et } |q\alpha_i - p_i| \leq \epsilon \text{ pour tout } i, \quad 1 \leq i \leq n$$

Lagarias [11] a pu donner une version approchée mais constructive à ce théorème en appliquant l'algorithme **LLL** au réseau L engendré par les lignes v_i de la matrice

$$A = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & 0 \\ \alpha_1 & \alpha_2 & \alpha_3 & \dots & \alpha_n & \epsilon/Q \end{pmatrix}$$

Là encore, le premier vecteur de la base réduite obtenue , mis sous la forme

$$v = \sum_{i=1}^{n} p'_i\, v_i \; + q' v_{n+1} \text{ permet de construire une bonne approximation } \left(\frac{p'_1}{q'}, \frac{p'_2}{q'}, ..., \frac{p'_n}{q'}\right)$$

On peut préciser le théorème constructif obtenu avec l'algorithme **LLL** associé à la valeur usuelle du paramètre:

Pour tout n , pour tout n-uplet $(\alpha_1, \alpha_2, ..., \alpha_n)$, pour tout couple (ϵ, Q) vérifiant $\epsilon > 0$ et $Q \geq 2^{n^2} \epsilon^{-n}$, on peut construire des entiers $(p'_1, p'_2, ..., p'_n)$ et un entier q' vérifiant

$$0 < q' \leq Q \text{ et } |q\alpha_i - p_i| \leq \epsilon \text{ pour tout } i, \quad 1 \leq i \leq n$$

4.6. La programmation linéaire en nombres entiers.

Le problème principal est le suivant:

Etant donné un polytope P de $\mathbf{R}^n$ de volume non nul, déterminer les points de coordonnées entières situés à l'intérieur de ce polytope.

On sait que ce problème est NP-dur en général. Mais, là encore, on peut chercher un algorithme qui soit polynomial quand la dimension n est fixée. C'est la démarche de Lenstra [15], assez analogue au second algorithme de 4.1.

On commence par considérer que P est un ellipsoïde, puis on se ramènera à ce cas, en coïnçant un polytope entre deux ellipsoïdes.

Soit f la transformation linéaire qui transforme P en une sphère unité S. Soit L le transformé du réseau $\mathbf{Z}^n$ par f. Le problème est alors transformé en le suivant:

Déterminer les points de L situés à l'intérieur de S

On réduit le réseau L en lui appliquant l'algorithme **LLL**: on obtient ainsi une base $(b_1, b_2, ..., b_n)$. Puis, on procède de manière récursive, en bornant le nombre d'hyperplans affines parallèles à H_{n-1} qui rencontrent la sphère S.

4.7. L'attaque du système de Merkle-Hellmann.

Le système de cryptographie de Merkle-Hellmann est fondé sur la difficulté du problème dit du "sac à dos"

Soient n entiers positifs a_i -les paquets- et un entier M -le sac-, trouver un élément $X = (x_i)_{1 \leq i \leq n}$ élément de $\{0,1\}^n$ solution de l'équation

$$\sum_{i=1}^{n} a_i \, x_i \; = \; M$$

Quels paquets doit-on mettre pour pouvoir remplir exactement le sac ?

Ce problème est facile quand la suite a_i est super-croissante : $a_i \geq \sum_{j<i} a_j$.

On peut l'utiliser dans un système de cryptographie dont la clé publique est le système des a_i: étant donné un message formé du mot X, on le code en M. Alors, de deux choses l'une:

- si la suite n'est pas super-croissante, personne ne peut décoder
- si elle l'est, tout le monde pourra décoder!

On utilise une suite super-croissante, dont on cache la super-croissance en lui appliquant une transformation $a \rightarrow va \bmod u$: le couple (u, v^{-1}) sera alors la clé secrète qui permettra le décodage.

Cependant, ce système n'est pas sûr: Shamir [16] a montré qu'on pouvait retrouver cette clé et donc briser le code, en utilisant un vecteur assez court d'un réseau bien choisi.

4.8. La prédictibilité de la suite des bits produits par le générateur congruentiel linéaire.

Le générateur pseudo-aléatoire le plus célèbre est sans doute le générateur linéaire congruentiel: on choisit un module m et un multiplicateur a, premier avec m, et une donnée x_1 de départ; puis on considère la suite (x_i) définie par

$$x_{i+1} = a\ x_i \mod m$$

Stern [18] a montré, en améliorant les résultats de Frieze [2] que, même si aucun des paramètres n'est connu, la suite y_i formée par une proportion "assez grande" des bits dominants des x_i est prédictible et donc que le générateur n'est pas cryptographiquement sûr. On travaille dans les réseaux X et Y engendrés respectivement par les vecteurs

$$u_i = \begin{pmatrix} x_{i+1} - x_i \\ x_{i+2} - x_{i+1} \\ x_{i+3} - x_{i+2} \end{pmatrix} \quad \text{et} \quad v_i = \begin{pmatrix} y_{i+1} - y_i \\ y_{i+2} - y_{i+1} \\ y_{i+3} - y_{i+2} \end{pmatrix}$$

Les k premiers vecteurs v_i étant donnés, on trouve, par l'algorithme 3.9 une relation entière courte entre eux de la forme

$$\sum_{i=1}^{k} \lambda_i\ v_i = 0$$

On déduit de cela que le vecteur $\sum_{i=1}^{k} \lambda_i\ u_i$ est un vecteur tellement court du réseau X ... qu'il est nul.

Si k est bien choisi en fonction de la taille présumée des données, on construit ainsi un polynôme P défini par $P(t) = \sum_{i=1}^{k} \lambda_i\ t^i$ et vérifiant $P(a) \equiv 0 \mod m$

Si on réitère cette construction, on détermine ainsi une suite de ℓ polynômes P_j qui appartiennent tous à un réseau L de base

$$q_0(t) = m \ \text{ et } \ q_i(t) = t^i - a^i \ \text{ pour } i \ \ 1 \leq i \leq \ k$$

Le réseau L a comme déterminant le nombre m cherché. On peut, par l'algorithme 3.8, trouver le déterminant $\hat{m}$ du réseau engendré par les P_j . $\hat{m}$ est un multiple de m qui décroît rapidement quand ℓ augmente; on déduit donc la valeur de m puis ensuite une valeur très probable de a.

4.9. Quelque autres applications...

Kaltofen [8] a utilisé l'algorithme en dimension 4 pour écrire un algorithme qui détermine le pgcd de deux nombres d'un corps quadratique principal mais non euclidien.

Landau et Miller [13] ont utilisé l'algorithme pour résoudre algorithmiquement la solvabilité par radicaux d'une équation polynomiale.

5. RÉFÉRENCES BIBLIOGRAPHIQUES

[1] L.Babai, *On Lovász's lattice reduction and the nearest lattice point problem*, Combinatorica 5 (1985).

[2] A.Frieze, J.Hastad, R.Kannan, J.C.Lagarias, A.Shamir, *Reconstructing truncated integer variables satisfying linear congruences*, to appear in SIAM Journal on computing.

[3] A.Dupré, Journal de Mathématiques 11, (1846), pp 41-64.

[4] C.F.Gauss, *Recherches Arithmétiques*, Paris 1807, réimprimé par Blanchard, Paris 1953.

[5] J.Hastad, B.Just, J.C.Lagarias, C.P.Schnorr, *Polynomial time algorithms for finding integer relations among real numbers*, Proceedings of STACS , Lecture Notes in Computer Science (1986).

[6] B.Helfrich, *Algorithms to construct Minkowski and Hermite reduced bases*, Univ. Frankfurt Technical Report, à paraitre dans TCS (1985).

[7] J.W.S.Cassels, *Rational quadratic forms*, Academic Press (1978).

[8] E.Kaltofen, H.Rolletschek, *Arithmetic in quadratic fields with unique factorization*, Comptes-rendus de EUROCAL'85, Lectures notes in Computer Science 204, Springer-Verlag.

[9] R.Kannan, *Improved algorithms for integer programming and related lattice problem*, JACM (1983), pp 193-206.

[10] R.Kannan, H.W.Lenstra, L.Lovász, *Polynomial factorization and bits of algebraic and some transcendental numbers* , Carnegie-Mellon University (1984).

[11] J.C.Lagarias, *Computational complexity of simultaneous diophantina approximation problem*, 23rd IEEE Symp. FOCS (1982).

[12] J.C.Lagarias, A.Odlyzko, *Solving low-density subset sum problems*, 24th IEEE Symp. FOCS (1983).

[13] S.Landau, G.L.Miller, *Solvability by radicals is in polynomial time* 15th Annual ACM Symposium on Theory of Computing (1983).

[14] A.K.Lenstra, H.W.Lenstra, L.Lovász, *Factoring polynomial with rational coefficients*, Math. Annalen 261 (1982) pp 513-534.

[15] H.W.Lenstra, *Integer programming with a fixed number of variables*, Mathematics of Operations Research, Vol 8, Number 4, Nov (1983).

[16] A.Shamir, *A polynomial time algorithm for breaking the Merkle- Hellman cryptosystem* , 23rd IEEE Symp. FOCS (1982).

[17] J.Stern, Lectures Notes, University of Singapore (1986).

[18] J.Stern, *Secret linear congruential generators are not cryptographically secure* , 28th IEEE Symp. FOCS (1987).

[19] B.Vallée, *Une approche géométrique de la réduction des réseaux en petite dimension*, Thèse de doctorat de l'Université de Caen (1986).

[20] Ph.Toffin, B.Vallée, *Implémentation d'algorithmes de type* **LLL** , Prépublications de l'Université de Caen (1987).

[21] P. Van Emde Boas, *Another NP-complete partition problem and the complexity of computing short vectors in a lattice*, Rep MI, UVA 81-04, Amsterdam (1981).

N° d'impression 963
2ème trimestre 1988