PUBLICATIONS MATHÉMATIQUES D'ORSAY

Groupe de travail
en théorie analytique et élémentaire
des nombres

1987 - 1988

89 - 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

1987 - 1988

89 - 01

Université de PARIS-SUD

Département de Mathématiques

Bâtiment 425

91405 ORSAY France

Pour toute commande s'adresser à :

Bibliothèque Mathématique d'Orsay
Bâtiment 425
Université Paris-Sud Centre d'Orsay
91405 ORSAY CEDEX

Tél. : (1) 69 41 70 51

N° d'impression 1028
1er trimestre 1989

TABLE DES MATIERES

DÉCOMPTE DE BLOCS DANS LES ENTIERS ET FONCTION ZÉTA DE HURWITZ

J.–P. ALLOUCHE
(résumé d'un travail en commun avec J. SHALLIT)

Notons $s_B(n)$ la somme des chiffres de l'entier n écrit en base B, il est connu que la série

$$\sum_0^\infty (s_2(n))/n(n+1)$$

converge vers 2 Log 2, et J. Shallit a montré que l'on a plus généralement :

$$\sum_0^\infty (s_B(n))/n(n+1) = (B/(B-1)) \text{ Log } B.$$

Nous montrons que, si l'on note $a_{w,B}(n)$ le nombre d'occurences du bloc w dans le développement de n en base B, alors la série

$$\sum_1^\infty a_{w,B}(n)(n^{-s} - (n+1)^{-s})$$

s'exprime pour $\text{Re } s > 1$ au moyen des fonctions $\zeta(s,x)$ de Hurwitz en des valeurs rationnelles de x.

Rappelons que

$$\zeta(s,x) = \sum_1^\infty (n+x)^{-s}, \quad \text{Re } s > 1, \quad x \in]0,1],$$

(en particulier $\zeta(s,1)$ n'est autre que la fonction zêta de Riemann). Cette fonction

admet un prolongement méromorphe au plan complexe tout entier, avec un seul pôle en $s = 1$, et l'on a au voisinage de ce point :

$$\zeta(s,x) = (s-1)^{-1} - (\Gamma'(x)/\Gamma(x)) + o(1).$$

Le théorème de Gauss permet de calculer $(\Gamma'(x)/\Gamma(x))$ en un point x rationnel à l'aide de fonctions élémentaires, ce qui nous donne les valeurs des séries

$$\sum_1^\infty a_{w,B}(n)/n(n+1).$$

Nous obtenons par exemple :

$$\sum_1^\infty a_{0,2}(n)/n(n+1) = 2 - 2\mathrm{Log}\ 2$$

$$\sum_1^\infty a_{11,2}(n)/n(n+1) = (3/2)\mathrm{Log}\ 2 - \pi/4.$$

Jean–Paul ALLOUCHE
C.N.R.S. U.A. 226
Math. et Info.
351 Cours de la Libération
33405 TALENCE CEDEX

ON A VARIANT OF THE PJATECKIJ–ŠAPIRO PRIME NUMBER PROBLEM

A. BALOG

1. Let c be a positive constant and let us define

$$\pi_c(x) = \sum_{\substack{n \le x \\ [n^c] \text{ is prime}}} 1.$$

It is an elementary exercise to prove, using the prime number theorem, that

$$\pi_c(x) \sim \frac{x}{c \log x} \tag{1}$$

for $0 < c \le 1$. It is also clear that (1) cannot hold for integers > 1 but one still expects (1) to hold for any $c > 1$, c is not an integer. In fact Pjateckij–Šapiro [10] showed that (1) does indeed hold for $1 < c < 12/11$ and this range of c was somewhat extended by several authors [6], [3], [7]. On the other hand Deshouillers [1] proved that

$$\pi_c(x) \longrightarrow \infty \tag{2}$$

for almost all positive c (in the sense of Lebesgue measure), however, this result provides no specific value of c. His proof actually gave a little more, i.e. for almost all positive c we have

$$\limsup_{x \longrightarrow \infty} \pi_c(x) / \frac{x}{c \log x} \ge 1. \tag{3}$$

Let us quote here Heath–Brown's interpretation of these results. "While the problem of representing primes by linear polynomials is completely solved by Dirichlet's Theorem on primes in arithmetic progressions, it is not known if there exists any quadratic polynomial that takes infinitely many prime values. One can therefore look on the investigation of (1) for $1 < c < 2$ as an intermediate problem for 'polynomials of degree c'; the admissible range for c provides a measure of our progress towards the quadratic case".

For an analogue of (1) one can expect that

$$\Pi_c(x) =: \sum_{\substack{p \le x \\ [p^c] \text{ is prime}}} 1 \sim \frac{x}{c \log^2 x}. \tag{4}$$

Here and thoroughout the paper p denotes primes. This is not trivial even for $0 < c < 1$. Indeed the problem of representing primes by linear polynomials in prime variable is still not solved. One can therefore look on the investigation of (4) for $0 < c < 1$ as a progress towards the twin prime conjecture. Our present object is to prove results of this kind.

THEOREM 1. (4) *holds for* $0 < c < 5/6$.

THEOREM 2. ***For almost all positive*** c ***(in the sense of Lebesgue measure)***

$$\limsup_{x \to \infty} \Pi_c(x) / \frac{x}{c \log^2 x} \ge 1. \tag{5}$$

The proof of Theorem 1 is essentially the adaptation of classical methods of "primes in almost all short intervals" while the proof of Theorem 2 is an upper bound sieve combined with Deshouillers' argument. Using lower bound sieves one can get a variety of results having less interest. We state here without proof the general shape of these results.

THEOREM 3. ***There are a constant*** $C > 1$ ***and integers*** $r \ge 1$, $s \ge 1$ ***such that for any*** $0 < c < C$, c ***is not an integer, we have***

$$\sum_{\substack{P_r \le x \\ [P_r^c] \text{ is } P_s}} 1 \gg \frac{x}{\log^2 x}.$$

As usual P_r denotes an integer having at most r prime factors. One can take, for example

r	s
1	9
9	1
2	5
5	2

in the above theorem.

ACKNOWLEDGEMENT : The author is very grateful to Dr. D.R. Heath–Brown and Dr. I.Z. Ruzsa for their valuable comments.

2. Thoroughout the paper we use the standard notations of number theory. We use also the convention that

$$q \simeq Q \text{ means } Q < q \leq Q + \frac{Q}{\log Q},$$

where q denotes a prime number, although we will make use of not more than q is taken from a quite dense set of integers.

First we prove Theorem 1 so in this section we assume that $0 < c < 1$. It is trivial , that

$$[p^c] = q \Leftrightarrow p^c - 1 < q \leq p^c \Leftrightarrow q^{1/c} \leq p < (q+1)^{1/c}. \tag{6}$$

Using an easy splitting up argument we arrive at

$$\Pi_c(x) = \sum_{\{Q\}} \sum_{q \simeq Q} (\pi(q^{1/c}(1 + \frac{1}{cQ})) - \pi(q^{1/c})) + 0(\frac{x}{\log^3 x}), \tag{7}$$

where the first summation runs over a certain set of integers Q satisfying

$$Q \leq x^c \, ; \quad \sum_{\{Q\}} 1 \ll \log^2 x.$$

We can change the function $\pi(y)$ into $\psi(y)$ by partial summation and by easy estimates. Thus we are interested in getting asymptotic formula for

$$\sum_{q \simeq Q} (\psi(q^{1/c}(1 + \frac{1}{cQ})) - \psi(q^{1/c}))$$

for any fixed $Q \leq x^c$. Each term in this sum can be expressed by the zeros of the Riemann ζ–function via the well–known explicit formula for $\psi(y)$, (see for example Landau [8]). Let $Q^{1/c} \leq y \leq 2Q^{1/c}$, then

$$\psi(y) = y - \sum_{|\gamma| \leq T} \frac{y^\rho}{\rho} + O(\frac{Q^{1/c} \log^2 Q}{T} + \frac{Q^{1/c} \log T}{T} + \log Q) \tag{8}$$

uniformly for Q, $T \geq 3$. Here the summation runs over all non–trivial zeros $\rho = \beta + i\gamma$ of $\zeta(s)$ counted according to multiplicity. The almost optimal choice of T is given by $T = Q \log^5 Q$ which takes the error term being small enough. We define

$$F(s) = \sum_{q \sim Q} q^{-s}; \quad C(s) = \frac{(1 + \frac{1}{cQ})^s - 1}{s}.$$

The first term on the right hand side of (8) gives the expected main term while we treat with the contribution of the zeros as an error term. Writing all these things into (7) we get

$$\Pi_c(x) = \frac{x}{c \log^2 x} + O(\frac{x}{\log^3 x}) + O(\sum_{\{Q\}} |R(Q)|), \tag{9}$$

then we are ready after proving

$$R =: R(Q) =: \sum_{|\gamma| \leq T} F(-\frac{\rho}{c}) C(\rho) \ll \frac{x}{\log^5 x} \tag{10}$$

uniformly in $Q \leq x^c$. We have the trivial bound $|C(\rho)| \ll 1/Q$. Using the standard zero spacing process, the Cauchy–Schwartz inequality and the discrete mean–value theorem (see [9] Theorem 7.6) we get that

$$R \ll \log^3 Q \sup_{\sigma} N(\sigma, T)^{\frac{1}{2}} T^{\frac{1}{2}} Q^{-\frac{1}{2}+\sigma/c}, \tag{11}$$

where $N(\sigma, T)$ denotes the number of zeros $\rho = \beta + i\gamma$ of $\zeta(s)$ satisfying $\beta \geq \sigma$ and $|\gamma| \leq T$, while the supremum is taken over

$$\frac{1}{\log^{2/3} T \, \log\log^{1/3} T} \ll 1 - \sigma \leq \frac{1}{2} \tag{12}$$

by Vinogradov's zero–free region ([9] Corollary 11.4). From Huxley's zero–density theorem ([4]) we have

$$N(\sigma, T) \ll T^{12(1-\sigma)/5} \log^{44} T. \tag{13}$$

Remembering that $T = Q\log^5 Q$ and $Q \leq x^c$ and writing (11) and (13) into (10) we arrive at

$$R \ll x\log^{31} x \sup_{\sigma} x^{-(1-6c/5)(1-\sigma)}. \tag{14}$$

(12) and the condition $c < 5/6$ provide that the supremum in (14) is less than any negative power of $\log x$. This proves (10) and thus Theorem 1. Note that we chose not the simplest proof of our theorem but we prepared the way for a possible improvement. We hope to return for this in a subsequent paper.

3. Next we turn to the proof of Theorem 2. As the first step we give an upper bound for $\Pi_c(x)$ having independent interest.

LEMMA 1. ***For any*** $C > 0$ ***and*** $\epsilon > 0$ ***we have a*** $K = K(C,\epsilon)$ ***such that***

$$\Pi_c(x) \leq K\frac{x}{\log^2 x}$$

uniformly in $x \geq 3$, $0 < c \leq C$ ***and*** $\|c\| \geq \epsilon$.

The last condition says that c is separated from the integers by ϵ. In the proof of the above lemma we use the simplest Selberg's upper bound sieve and the van der Corput's exponential sum estimate.

For a given $D \geq 3$ there are sieving weights λ_d (see [2] Chapter 3) satisfying

$$\lambda_1 = 1\ ;\ |\lambda_d| \leq 1\ ;\ \lambda_d = 0 \text{ for } d > D\ ;\ \Sigma\frac{\lambda_d\lambda_{d'}}{[d,d']} \leq \frac{1}{\log D}, \tag{15}$$

where $[d,d']$ is the smallest common multiple. The advantage of this weights is the fact that

$$(\sum_{d|n}\lambda_d)^2 \begin{cases} \geq 0 \text{ trivially} \\ = 1 \text{ if } n > D \text{ is a prime.} \end{cases} \tag{16}$$

For later use it is convenient to introduce the function

$$\rho_d = \sum_{d=[d',d'']} \lambda_{d'}\lambda_{d''} \begin{cases} = 0 & \text{if } d > D^2, \\ \ll 3^{\nu(d)}, \end{cases}$$

where $\nu(d)$ is the number of prime factors of d. From (6) and (16) we have our starting formula

$$\Pi_c(x) = \sum_{p \le x} \sum_{p^{c-1} < q \le p^c} 1 \le \sum_{m \le x} \sum_{m^{c-1} < n \le m^c} \Big(\sum_{d \mid m} \lambda_d\Big)^2 \Big(\sum_{t \mid n} \lambda_t\Big)^2 + D + D^{1/c} =$$

(17)

$$= \sum_{d \le D^2} \sum_{t \le D^2} \rho_d \rho_t \sum_{\substack{m \le x/d \\ m^c d^{c-1} < nt \le m^c d^c}} \sum_{n \le x^c/t} 1 + D + D^{1/c}.$$

In estimating the inner double sum we use the Fourier–expansion of the fractional part function. We have (see, for example [5] Lemma 6)

$$(18) \qquad \sum_{\substack{m \le x/d \\ m^c d^{c-1} < nt \le m^c d^c}} \sum_{n \le x^c/t} 1 = \frac{x}{dt} + 0\Big(\frac{1}{t}\Big) + 0\Big(\sum_{h \le H} \frac{1}{h} \Big| \sum_{m \le \frac{x}{d}} e\Big(\frac{m^c d^c h}{t}\Big)\Big|\Big) + 0\Big(\frac{x}{dH}\Big).$$

There are more ways to bound the exponential sum in (18). As we are not interested in using the sharpest one we appeal to van der Corput's method (see [11] Theorem 5.13)

$$(19) \qquad \sum_{M < m \le 2M} e(\Delta m^c) \ll M\Lambda^{1/2K-2} + M^{1-2/K}\Lambda^{-1/2K-2},$$

with an absolute implied constant, where

$$k \ge 2; \quad K = 2^{k-1}; \quad \Lambda = \Delta c(c-1)\ldots(c-k+1)M^{c-k}.$$

A possible good choice of k is $k = [c+2]$. In this case the constant defined in Λ can be bounded from both direction by means of C and ϵ. Note that this is the point of the proof where we have to separate c from integers. Choosing $H = D^2\log^6 D$ and writing (18) and (19) into (17) and using crude estimates we get that

$$\Pi_c(x) \leq x \sum_d \sum_t \frac{\rho_d \rho_t}{dt} + 0\left(\frac{x}{\log^3 D}\right) + 0(x^{1-2^{-[C+2]}} D^6)$$

where the implied constants depend on C and ϵ only. It is clear that we can choose D as a small power of x and than (15) implies the lemma.

Note that proving the explicite results listed after the statement of Theorem 3 requires more careful analysis and first of all more effective sieve results. The essence, however, is the same.

4. For the proof of Theorem 2 we consider $\Pi_c(x)$ as a function of c. Let us introduce the next abbreviation

$$F_x(c) = \Pi_c(x) / \frac{x}{c \log^2 x}.$$

Lemma 1 asserts that the function family $F_x(c)$ is uniformly bounded in any fixed closed interval containing no integers. On the other hand we show that $F_x(c)$ is 1 in average. This is in fact an elementary consequence of the prime number theorem.

$$\int_a^b F_x(c)\, dc = \frac{\log^2 x}{x} \int_a^b \sum_{p \leq x} \sum_{p^{c-1} < q \leq p^c} c\, dc = \frac{\log^2 x}{x} \sum_{p \leq x} \sum_q \int_A^B c\, dc,$$

where $A = \max(a, \log q/\log p)$ and $B = \min(b, \log(q+1)/\log p)$. By simple calculation we have

$$\begin{aligned} \int_a^b F_x(c)\, dc &= \frac{\log^2 x}{x} \sum_{p \leq x} \sum_{p^a < q \leq p^b} \frac{\log q}{q \log^2 p} + 0\left(\frac{\log^2 x}{x} \sum_{p \leq x} \frac{1}{p^a \log p}\right) = \\ &= (b-a) + 0\left(\frac{1}{\log x}\right), \quad x \to \infty, \end{aligned} \tag{20}$$

where the implied constant depends on a and b.

We are interested in the measure of the set

$$S^* = \{0 < c : \limsup F_x(c) < 1\} = \bigcup_{N=1}^{\infty} \bigcup_{k=3}^{\infty} \bigcup_{n=3}^{\infty} \bigcup_{X=3}^{\infty} \bigcap_{x \geq X}^{\infty} S(J,n,x),$$

where

$$J = [N-1+\tfrac{1}{k}, N-\tfrac{1}{k}];\ S(J,n,x) = S = \{c \in J : F_x(c) \le 1-\tfrac{1}{n}\}.$$

S is Lebesgue measurable. We will show that for any fixed J and n and for large enough x the Lebesgue measure of S is zero. These implies that S^* is also measurable with measure zero. This proves Theorem 2.

From now on $|.|$ is the Lebesgue measure. Let us suppose that $|S| = \ell$. Than for any $\epsilon > 0$ there is a finite covering $\{I\}$ of S such that

$$S \subseteq \bigcup_{\{I\}} I \subseteq J;\ \sum_{\{I\}} |I| < \ell + \epsilon. \tag{21}$$

Let us fix an $I \subseteq J$. We have from (20) that for large enough x

$$\int_I F_x(c)\, dc \ge |I|\,(1-\tfrac{1}{2n}),$$

while from Lemma 1 we have

$$\int_I F_x(c)\, dc \le |S|\,(1-\tfrac{1}{n}) + (|I| - |S \cap I|)K,$$

where K depends on J alone. Comparing the upper and lower bounds we get

$$|S \cap I| \le |I|\,(1-\frac{1}{2nK-2n+2}).$$

Writing this into (21) we arrive at

$$\ell = |S| = |\bigcup_{\{I\}} (S \cap I)| \le \sum_{\{I\}} |S \cap I| \le (1-\frac{1}{2nK-2n+2}) \sum_{\{I\}} |I|$$
$$\le (1-\frac{1}{2nK-2n+2})(\ell+\epsilon)$$

and this is impossible except $\ell = 0$. Theorem 2 is proved.

BIBLIOGRAPHY

[1] J.–M. Deshouillers.– Nombres premiers de la forme $[n^c]$, C.R. Acad. Sci. Paris Sér. A–B 282 (3) (1976), A131–A133.

[2] H. Halberstam, H.E. Richert.– Sieve methods, Academic Press 1974.

[3] D.R. Heath–Brown.– The Pjateckij–Šapiro prime number theorem, J. Number Theory 16 (1983), 242–266.

[4] M.N. Huxley.– On the difference between consecutive primes, Invent. Math. 15 (1972), 164–170.

[5] H. Iwaniec.– On the Brun–Titchmarsh theorem, J. Math. Soc. Japan 34 (1982), 95–123.

[6] G.A. Kolesnik.– The distribution of primes in sequences of the form $[n^c]$, Mat. Zametki 2 (1972), 117–128.

[7] G.A. Kolesnik.– Primes of the form $[n^c]$, Pacific J. Math. 118 (1985), 437–447.

[8] E. Landau.– Über einige Summen, die von den Nullstellen der Riemann'schen Zetafunktion abhängen, Acta. Math. 35 (1912), 271–294.

[9] H.L. Montgomery.– Topics in multiplicative number theory, Springer 1971.

[10] I.I. Pjateckij–Šapiro.– On the distribution of prime numbers in sequences of the form $[f(n)]$, Mat. Sb. 33 (1953), 559–566.

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

Antal BALOG
Mathematical Institute of the
Hungarian Academy of Sciences
Budapest, Reáltanoda u. 13–15
1053 HUNGARY

ENSEMBLES NORMAUX ET SUITES BORNEES

J.–P. BOREL

Le but de cet exposé est d'établir quelques propriétés des ensembles normaux associés à une suite bornée. Essentiellement, deux résultats sont présentés :

– si $\Gamma_k = \bigcup_{i=1}^{k} \gamma_i \mathbb{Z}^*$, une méthode permettant de relier l'amplitude minimale $M(\Gamma_k)$ des intervalles dans lesquels on peut trouver une suite Λ telle que $\Gamma_k = B(\Lambda)$. Si on note $d(\Gamma_k)$ la densité asymptotique de Γ_k dans $\mathbb{Z}$, on montre que $d(\Gamma_k) \leq M(\Gamma_k)$, et le cas où il y a égalité est caractérisé. Dans le cas général, un encadrement de $M(\Gamma_k)$ est donné, dépendant de certaines propriétés de polynômes particuliers.

– si A est une partie normale de $\mathbb{R}$ associée à une suite bornée, il en est de même pour toute partie A' de A telle que $0 \notin A'$ et $k A' \subset A$ pour tout k entier non nul (ces propriétés sont clairement nécessaires). Une estimation quantitative est obtenue, qui est la majoration $M(A') \leq 2\, M(A)$.

1. Les ensembles b–normaux

1.1. On rappelle qu'une partie A de $\mathbb{R}$ est dite normale s'il existe une suite $\Lambda = (\lambda_n)_{n \geq 1}$ de nombres réels telle que A est l'ensemble normal associé $B(\Lambda)$, c'est–à–dire :

$$A = B(\Lambda) := \{x \in \mathbb{R} / x\,\Lambda \text{ est équirépartie modulo } 1\}$$

ce qui revient donc à dire, en utilisant le critère de Weyl (cf. [KUI]) :

$$(1) \qquad x \in A \Leftrightarrow \forall\, k \in \mathbb{N}^*,\ \lim_{n \to \infty} \frac{1}{N} \sum_{n=1}^{N} e^{2\pi i k \lambda_n} = 0.$$

1.2. Les parties normales de $\mathbb{R}$ ont été caractérisées par G. Rauzy [RAU 2] : A est normale si et seulement si elle vérifie (en notant $\mathbb{Z}^* = \mathbb{Z} - \{0\}$) :

(2) $$\begin{cases} 0 \notin A \\ \forall\, k \in \mathbb{Z}^*,\ kA \subset A \end{cases}$$

et

$$A \text{ est élémentaire}$$

(cette dernière condition est assez générale : par exemple, toute réunion dénombrable de fermés disjoints est élémentaire).

1.3. Il est intéressant de regarder les ensembles normaux associés à certains types de suites. Je considèrerai ici les ensembles *b–normaux*, c'est–à–dire normaux associés à une suite Λ bornée. Une translation de la suite Λ ne modifiant pas $B(\Lambda)$, il suffit de considérer les suites à valeurs dans un intervalle de la forme $[0,M]$. Les deux problèmes, qui ne sont pas complètement résolus ici, sont alors :

[pb 1] Trouver une condition (contenant donc (2)) nécessaire et suffisante pour que $A \subset \mathbb{R}$ soit un ensemble b–normal.

[pb 2] Si A est b–normal, calculer la quantité $M(A)$ définie par :
$M(A) := \inf \{M > 0 / \exists\ \Lambda : \mathbb{N}^* \longrightarrow [0,M],\ A = B(\Lambda)\}$

1.4. Des réponses partielles à ces problèmes sont déjà connues. Elles proviennent des travaux menés dans les années 1968–70 pour résoudre le problème de la caractérisation des ensembles normaux.

Toute partie de la forme $A = \bigcup_{k=1}^{\infty} \gamma_k \mathbb{Z}^*$ avec $\sum_{k=1}^{\infty} \frac{1}{\gamma_k} < \infty$ est b–normale, Mendès–France [MEN].

Une intersection dénombrable (ou finie...) de parties b–normales est b–normale, c'est une conséquence immédiate du principe de mixage de Dress [DRE 1].

Une partie A de $\mathbb{Z}$ est b–normale si et seulement si elle vérifie (2). Alors, $M(A) \leq 1$ et la borne inférieure est atteinte, Dress et Mendès–France [DRE 2].

2. Les résultats obtenus

2.1. Le premier résultat précise des conditions nécessaires qualitatives et quantitatives (i.e. concernant $M(A)$ pour que A soit b–normal. La partie qualitative est déjà connue, mais je n'en connais pas de rédaction publiée [LIA].

Pour cela, introduisons les quantités suivantes associées à une partie A de $\mathbb{R}$:

$$\mathscr{A}(X) := \frac{1}{2X} \sum_{\substack{a \in A \\ |a| < X}} 1 \qquad (X > 0)$$

$$s(A) := \sup_{X>0} \mathscr{A}(X) \qquad \underline{d}(A) := \liminf_{X \to \infty} \mathscr{A}(X)$$

$$\bar{d}(A) := \limsup_{X \to \infty} \mathscr{A}(X) \qquad d(A) := \lim_{X \to \infty} \mathscr{A}(X)$$

(ces quantités pouvant valoir $+\infty$, voire ne pas exister pour $d(A)$).

THEOREME 1. *Si A est b–normal, il est discret dans $\mathbb{R}$ et on a les inégalités :*

$$\underline{d}(A) \leq M(A) \tag{3}$$

$$s(A) \leq \frac{9}{5} M(A) \tag{4}$$

Par exemple, $\mathbb{Q}$ est normal [RAU 1] mais n'est pas b–normal.

2.2. La famille des ensembles b–normaux est stable pour certaines opérations ensemblistes.

THEOREME 2. *Soit A un ensemble b–normal, et A' une partie de A qui vérifie les conditions nécessaires* (2). *Alors A' est b–normal, et l'on a :*

$$M(A') \leq 2M(A).$$

Dans certains *bons* cas, qui seront précisés ultérieurement, le facteur 2 peut être supprimé.

COROLLAIRE 1. *La famille des ensembles b–normaux est stable par intersection quelconque et par réunion finie.*

(les intersections quelconques se ramènent en fait à des intersections dénombrables, puisque d'après le théorème 1 les ensembles b–normaux sont dénombrables !).

COROLLAIRE 2. *Soit* $(A_n)_{n\geq 1}$ *une suite croissante d'ensembles* b*–normaux. Alors* $A = \lim_{n\to\infty}\uparrow A_n$ *est* b*–normal si et seulement si la suite* $(M(A_n))_{n\geq 1}$ *est bornée. On a alors :*

$$\frac{1}{2}\sup_{n\geq 1} M(A_n) \leq M(A) \leq 2 \liminf_{n\to\infty} M(A_n).$$

2.3. D'après (2) et le théorème 1, tout ensemble b–normal A peut se mettre sous la forme *canonique* suivante :

$$(5)\quad A = \bigcup_{j\in J} \gamma_j \mathbb{Z}^* \qquad \text{avec} \quad \text{soit} \quad J = \{1,2,\ldots,k\} \text{ et } 0<\gamma_1<\gamma_2<\ldots<\gamma_k$$

$$\text{soit} \quad J = \mathbb{N}^*$$

$$\text{et} \quad 0<\gamma_1<\gamma_2<\ldots<\gamma_k<\ldots \lim_{j\to\infty} \gamma_j = +\infty$$

$$\text{et} \qquad \text{avec} \quad \gamma_j \notin \bigcup_{i<j} \gamma_i \mathbb{Z}^* \text{ pour tout } j\in J.$$

Considérons le cas simple $\Gamma_k = \bigcup_{j=1}^{k} \gamma_j \mathbb{Z}^*$, où les γ_j sont des entiers premiers entre eux dans leur ensemble et vérifiant les propriétés de (5). Posons alors :

$$m := ppcm\,(\gamma_1,\gamma_2,\ldots,\gamma_k) \text{ et } \zeta := e^{2\pi i/m}$$

$$\Delta_k := \Gamma_k \cap \{1,2,\ldots,m-1\}$$

et

$$P := \prod_{a\in\Delta_k} (X-\zeta^a) \qquad (\text{donc } P\in \mathbb{R}[X]).$$

THEOREME 3. 1) *Si* P *a tous ses coefficients positifs, on a :*

$$M(\Gamma_k) = d(\Gamma_k)$$

2) *Dans tous les cas, on a les inégalités :*

$$(6)\qquad d(\Gamma_k) + \frac{\delta P}{m} \leq M(\Gamma_k) \leq d(\Gamma_k) + \frac{\delta^* P}{m} \leq 1.$$

Les quantités δP et $\delta^* P$ sont définies pour tout polynôme $P\in\mathbb{R}[X]$ par, en notant $S\geq 0$ le fait que le polynôme S a tous ses coefficients (de X^0 à $X^{d^0 S}$) positifs :

$$\delta P := \min \{d^0Q,\ Q \in \mathbb{R}[X]/Q \neq 0 \text{ et } PQ \geq 0\}$$

$$\delta^* P := \min \{d^0Q,\ Q \in \mathbb{R}[X] \text{ et } \begin{cases} PQ \geq 0 \\ \forall\, z \in \mathbb{C},\ |z| = 1 \text{ et } \begin{matrix} Q(z)=0 \\ P(z) \neq 0 \end{matrix} \end{cases} \Rightarrow \exists\ n \geq 1,\ PQ(z^n) \neq 0$$

Une caractérisation des polynômes P pour lesquels ces quantités existent, ainsi que des encadrements de δP et $\delta^* P$ en fonction de d^0P et des arguments des racines de P peuvent être trouvées dans [BOR 3]. Par un argument de petite variation des racines des polynômes Q, et donc de petites variations des coefficients du polynôme PQ, il est facile de voir que :

$$\delta^* P \leq \delta^+ P := \min \{d^0Q,\ Q \in \mathbb{R}[X] \text{ et } PQ > 0\}$$

où $S > 0$ signifie que les coefficients du polynôme S, de X^0 à X^{d^0S}, sont strictement positifs.

3. Mesures de répartition

3.1. Soit $\Lambda = (\lambda_n)_{n \geq 1}$ une suite à valeurs dans $[0,M]$, et soit π_N la probabilité :

$$\pi_N := \frac{1}{N} \sum_{n=1}^{N} \delta_{\lambda_n}.$$

Je dirai qu'une probabilité μ est *adhérente à la suite* Λ, noté $\mu \in \Pi$, si μ est limite au sens de la convergence étroite d'une sous–suite $(\pi_{N_k})_{k \geq 1}$. Π est donc non vide, et toute mesure $\mu \in \Pi$ a son support contenu dans l'intervalle $[0,M]$. Lorsque Π n'a qu'un élément μ, je dirai que la suite Λ est μ–répartie.

3.2. Pour μ probabilité sur $\mathbb{R}$, soit $\mathcal{O}[\mu]$ l'ensemble des zéros de sa transformée de Fourier $\hat{\mu}(t) = \int e^{-2\pi i x t}\, d\mu(x)$, et soit :

$$B[\mu] := \bigcap_{k=1}^{\infty} \frac{1}{k}\, \mathcal{O}[\mu]$$

c'est–à–dire que l'on a :

$$t \in B[\mu] \Leftrightarrow \forall\ k \geq 1,\ \hat{\mu}(kt) = 0.$$

PROPOSITION 1. *Soit* Λ *à valeurs dans* $[0,M]$. *Alors* :

$$B(\Lambda) = \bigcap_{\pi \in \Pi} B[\mu].$$

Ce résultat se déduit simplement du critère de Weyl, on en trouvera la démonstration dans [BOR 2]. Comme $\hat{\mu}$ est ici une fonction entière, le principe des zéros isolés conduit immédiatement à $B(\Lambda)$ discret dans $\mathbb{R}$.

3.3. Soit μ adhérente à Λ, et posons $g(z) = e^{M\pi iz}\hat{\mu}(z)$. Soit $\mathscr{A}$ définie comme en 2.1 et associée à $A = B(\Lambda)$, $\mathscr{B}$ associée à $A = \mathcal{O}[\mu]$. On a alors :

$$|g(\rho e^{i\theta})| \leq e^{\pi RM|\sin\theta|}$$

et donc la formule de Jensen entraîne :

$$(7) \qquad \sum_{\substack{b \in B \\ |b| < X}} \operatorname{Log} \frac{X}{|b|} = \frac{1}{2\pi} \int_0^{2\pi} \operatorname{Log} |g(\rho e^{i\theta})| \, d\theta \leq 2\, R\, M$$

et donc :

$$\mathscr{A}(X) \leq \mathscr{B}(X) \leq \frac{1}{2X} \sum_{\substack{b \in B \\ |b| < eX}} \operatorname{Log} \frac{eX}{|b|} \leq e\, M$$

Donc $s(A) \leq e\, M(A)$. Les inégalités (3) et (4) du théorème 1 proviennent respectivement d'une sommation à la Abel, et d'une utilisation de la propriété (2) satisfaite par $B(\Lambda)$ associée à la majoration (7). Le lecteur trouvera la démonstration détaillée dans [BOR 1].

3.4. Soit μ une probabilité sur $\mathbb{R}$, de support borné. Il est alors très simple de construire des suites Λ qui soient μ–réparties. Cela conduit aux *caractérisations* peu pratiques des ensembles b–normaux qui suivent.

[car 1] $A \subset \mathbb{R}$ est b–normal si et seulement si A vérifie (2) et s'il existe une probabilité μ à support borné telle que $A \subset \mathcal{O}[\mu]$.

[car 2] $A \subset \mathbb{R}$ est b–normal si et seulement s'il existe une suite $(\mu_n)_{n\geq 1}$ de probabilités concentrées sur un même intervalle borné, et telles que :

$$A = \bigcap_{n=1}^{\infty} B[\mu_n].$$

4. Preuve du théorème 2

4.1. Il provient d'un résultat analogue concernant les ensembles $B[\mu]$. Il est cependant nécessaire d'introduire un type particulier de probabilités sur $\mathbb{R}$. Je dirai qu'une probabilité μ sur $\mathbb{R}$ est vraiment positive, noté $\mu \in \mathcal{P}_+$, lorsque :

μ est absolument continue, de densité f
il existe deux nombres réels $M_1 < M_2$ tels que :
$f(x) = 0$ si $x \notin [M_1, M_2]$
$\forall \eta > 0,\ \exists \epsilon > 0,\ x \in [M_1+\eta, M_2-\eta] \Rightarrow f(x) > \epsilon$

(si f est continue, cela signifie simplement que l'ensemble des x tels que $f(x) > 0$ est un intervalle ouvert borné).

4.2. Les propriétés suivantes des probabilités $\mu \in \mathcal{P}_+$ seront utilisées par la suite (leur démonstration, très simple, est laissée au lecteur).

(8) Soit λ_x la probabilité uniforme sur l'intervalle $[0,x]$, $x > 0$, et $\mu \in \mathcal{P}_+$. Alors $\mu * \lambda_x \in \mathcal{P}_+$.

(9) Soit μ une probabilité sur $\mathbb{R}$, de diamètre de son support $|\mu|$. Alors μ est facteur de convolution de $\mu' \in \mathcal{P}_+$, avec $|\mu'| \leq 2|\mu|$

(il suffit de prendre $\mu' = \mu * \lambda_{|\mu|}$. La facteur 2 qui apparaît ici est celui qui se retrouve dans le théorème 2. Pour une mesure μ explicetement donnée, il peut souvent être réduit).

4.3. Supposons que $t \in B[\mu]$, et que μ' est telle que l'on ait simultanément :

$$t \notin B[\mu'] \quad \text{et} \quad B[\mu'] \subset B[\mu].$$

L'ensemble $B[\mu']$ satisfait la propriété (2), donc il ne peut avoir d'élément de la forme t/k, $k \in \mathbb{Z}^*$. Je noterai $t/\mathbb{Z}^*$ cet ensemble.

THEOREME 4. *Soit* $\mu \in \mathcal{P}_+$, $t \in B[\mu]$ *et* $\eta > 0$ *donnés. Alors il existe* $\nu \in \mathcal{P}_+$ *vérifiant simultanément :*

$$B[\nu] = B[\mu] - t/\mathbb{I}^*$$
$$|\nu| \leq |\mu| + \eta$$

4.4. Je vais montrer ici que le théorème 4 entraîne le théorème 2.
Soit A un ensemble b–normal associé à Λ, suite à valeurs dans $[0,M(A)+\epsilon]$. Soit $\mu' \in \Pi$. Avec (9), il existe $\mu \in \mathcal{P}_+$ telle que :

$$\begin{cases} A' \subset A \subset B[\mu] \\ \mu([0,2M(A)+2\epsilon]) = 1. \end{cases}$$

Comme $B[\mu]$ est un ensemble dénombrable (cf. 3.2), on peut écrire $A' = B[\mu] - T$, où $T = \{t_n, n \in I\}$, I fini ou dénombrable. D'après le théorème 4, il existe ν_n telle que :

$$\begin{cases} B[\nu_n] = B[\mu] - t_n/\mathbb{I}^* & (n \in I) \\ \nu_n([O,2M(A)+3\epsilon]) = 1 \end{cases}$$

et comme A' vérifie (2), on a alors :

$$A' = \bigcap_{n \in I} B[\nu_n].$$

Soit $\Lambda^{(n)}$ une suite à valeurs dans $[0,2M(A)+3\epsilon]$ et ν_n–répartie. Le principe de mixage de Dress ([DRE 1]) permet de construire une suite, dont les valeurs sont les $\lambda_k^{(n)}$, et telle que :

$$\begin{aligned} B(\Lambda) &= \bigcap_{n \in I} B(\Lambda^{(n)}) \\ &= \bigcap_{n \in I} B[\nu_n] \\ &= A' \end{aligned}$$

et donc A' est b–normal, $M(A') \leq 2M(A) + 3\epsilon$. D'où le théorème 2.

La constante 2 apparaît donc pour garantir l'appartenance $\mu \in \mathcal{P}_+$. Si elle est connue à l'origine (par exemple s'il existe $\mu \in \mathcal{P}_+$ telle que $A \subset B[\mu]$ et $|\mu| = M(A)$), elle disparaît et on obtient donc $M(A') \leq M(A)$. Je pense que cette

dernière majoration est fausse dans le cas général, bien que la construction d'un contre–exemple semble un problème difficile.

5. Preuve du théorème 4

5.1. Le schéma de cette démonstration est le suivant : le théorème 4 est établi directement pour $\mu = \lambda_1$ (et ici avec $\eta = 0$), donc pour toutes les probabilités λ_x. Si $t \in B[\mu]$, $\lambda_{1/t}$ est presque un facteur de convolution de μ, on modifie un peu $\lambda_{1/t}$ par le théorème 4, et on retrouve ν en faisant la convolution.

Le lemme suivant est inspiré de la démonstration de Dress et Mendès–France dans [DRE 2].

LEMME 1. *Soit* $A = \mathbb{Z}^* - \{-1,+1\}$, C *une partie discrète de* $\mathbb{R} - \mathbb{Z}$ *et* $\epsilon > 0$ *donné. Alors il existe une fonction* g *continue sur* $[0,1]$, *nulle en dehors, positive, d'intégrale* 1, *telle que la probabilité* $\mu = g.\lambda_1$ *vérifie* :

$$A = B[\mu]$$
$$\mathcal{O}[\mu] = A \cup \{\pm t\} \quad \textit{avec } \frac{3}{4} < t < 1 \quad \textit{et } t \notin C$$
$$\forall\, x \in [0,1], \quad |1 - g(x)| \leq \epsilon.$$

Démonstration :

Il suffit de prendre $g(x) = 1 + 2a \cos 2\pi x$ sur $[0,1]$, avec a suffisament petit, ici :

$$a < \min\left(\frac{\epsilon}{2}, \frac{7}{18}, \frac{1-c^2}{2c^2}\right)$$

où c est le plus grand élément de C dans l'intervalle $]0,1[$. Le résultat provient du calcul de $\hat{\mu}$:

$$\hat{\mu}(-t) = \frac{e^{2\pi i t} - 1}{2\pi i}\left(\frac{1}{t} + a\,\frac{2t}{t^2-1}\right)$$

qui permet de préciser $\mathcal{O}[\mu]$, puis $B[\mu]$. □

5.2. Il est facile de montrer que si $t \in B[\mu]$, μ est nécessairement absolument continue, et que sa densité est (presque partout) majorée par t.

Pour $N \geq 1$, je poserai :

$$\omega_N = \frac{1}{N} \sum_{n=0}^{N-1} \delta_{\frac{n}{N}}.$$

On a donc $\lambda_1 = \omega_N * \lambda_{1/N}$ pour tout N, et la suite de probabilités $(\omega_N)_{N \geq 1}$ converge étroitement vers λ_1.

LEMME 2. *Soit μ une mesure finie absolument continue et à support borné, et telle que $1 \in B[\mu]$. Alors ω_2 est facteur de convolution de μ.*

Démonstration :

On peut se ramener par translation à $\mu([0,K]) = 1$, avec K entier. μ a une densité f et posons pour $0 \leq x \leq 1$ et $0 \leq m < 2K$:

$$f_m(x) = f(\frac{m}{2} + \frac{x}{2}).$$

La propriété $1 \in B[\mu]$ est alors équivalente à :

$$\sum_{k=0}^{K-1} f_{2k}(x) = \sum_{k=0}^{K-1} f_{2k+1}(x) = 1 \qquad (0 \leq x \leq 1) \tag{10}$$

Posons alors pour $0 \leq m \leq 2K-2$:

$$g_m(x) = 2 \sum_{k=0}^{m} (-1)^k f_{m-k}(x) \qquad (0 \leq x \leq 1)$$

et

$$g(x) = g_{[2x]}(\{2x\}) \qquad (0 \leq x \leq K - \frac{1}{2}).$$

On voit alors immédiatement que :

$$(12)\qquad \begin{cases} \text{si } 0 \le x \le \frac{1}{2}, & \frac{1}{2}\, g(x) = f(x) \\ \text{si } \frac{1}{2} \le x \le K - \frac{1}{2}, & \frac{1}{2}(g(x) + g(x - \frac{1}{2})) = f(x) \\ \text{si } K - \frac{1}{2} \le x \le K, & \frac{1}{2}\, g(x - \frac{1}{2}) = f(x) \end{cases}$$

la dernière propriété provenant de la relation $g_{2K-2}(x) = f_{2K-1}(x)$, qui vient de (10). Or (12) signifie exactement que l'on a $\mu = \omega_2 \star \nu$, où ν est la mesure de densité g sur l'intervalle $[0, K - \frac{1}{2}]$. □

LEMME 3. *Soit μ une probabilité à support borné, et telle que $1 \in B[\mu]$. Alors, pour tout $n \ge 1$, ω_{2^n} est facteur de convolution de μ.*

Démonstration : (récurrence sur n)

μ a une densité $f \le 1$. On peut lui appliquer le lemme 2. Les zéros de $\hat{\omega}_2$ étant les entiers impairs, on a donc $\mu = \omega_2 \star \nu_1$, avec ν_1 absolument continue et $2 \in B[\mu]$. Si $\tilde{\nu}_1$ est l'image de ν_1 par $x \longmapsto 2x$, on obtient $1 \in B[\tilde{\nu}_1]$, d'où :

$$\tilde{\nu}_1 = \omega_{2^n} \star \tilde{\nu}_2$$

si on suppose la propriété vraie au rang n. On a donc :

$$\nu_1 = \frac{1}{2^n} \sum_{k=0}^{2^n - 1} \delta_{\frac{k}{2^n}} \star \tilde{\nu}_2$$

et si ν_2 est la mesure image de $\tilde{\nu}_2$ par $x \longmapsto x/2$:

$$\nu_1 = \frac{1}{2^n} \sum_{k=0}^{2^n - 1} \delta_{\frac{k}{2^{n+1}}} \star \nu_2$$

d'où

$$\mu = \frac{1}{2}(\delta_0 \star \delta_{\frac{1}{2}}) \star \frac{1}{2^n} \sum_{k=0}^{2^n - 1} \delta_{\frac{k}{2^{n+1}}} \star \nu_2$$

$$= \omega_{2^{n+1}} \star \nu_2. \qquad \square$$

On a donc pu écrire $\mu = \omega_{2^n} \star \nu_n$, où ν_n est absolument continue, $|\nu_n| = |\mu| - |\omega_{2^n}| = |\mu| + 2^{-n} - 1$, la densité de ν_n étant majorée par $(2K)^n$, lorsque le support de μ est contenu dans $[0,K]$.

5.3. Montrons maintenant le théorème 4. D'après le principe utilisé plus haut, passage à la mesure image par $x \longmapsto tx$ et retour, il suffit de démontrer ce résultat avec $t = 1$.

Soit K l'ordre de zéro $t = 1$ de la fonction entière $\hat{\mu}$, et soit n tel que $K < \eta\, 2^n$. Je poserai :

$$\begin{aligned} \mu &:= \omega_{2^n} \star \mu_1 \\ \mu_2 &:= \lambda_1 \star \mu_1 = \lambda_{2^{-n}} \star \mu \\ \mu_3 &:= (g\,\lambda_1) \star \mu_1 \end{aligned}$$

ces trois probabilités étant construites de la façon suivante : μ_1 provient de l'application à μ du lemme 3. Cela donne μ_2. μ_3 provient du lemme 1, en choisissant les paramètres suivants :

$C = \frac{1}{2}(\mathcal{O}[\mu] - \mathbb{Z})$, qui est un ensemble discret dans $\mathbb{R}$;

$\epsilon = \frac{\epsilon'}{2X}$, où ϵ' est associé à η pour la mesure $\mu_2 \in \mathcal{P}_+$ ($\mu_2 \in \mathcal{P}_+$ d'après (8), en utilisant la seconde écriture $\mu_2 = \lambda_{2^{-n}} \star \mu$) et X est un majorant de $|f_1|$, où f_1 est densité de μ_1 (un tel majorant existe, d'après 5.2).

La démonstration du théorème utilise alors les deux lemmes techniques suivants :

LEMME 4. *μ_3 est un élément de $\mathcal{P}_+$.*

LEMME 5. *Si* $K = 1$, $B[\mu_3] = B[\mu] - 1/\mathbb{Z}^*$

Si $K > 1$, $B[\mu_3] = B[\mu]$ *et* 1 *est zéro d'ordre* $K-1$ *de* $\hat{\mu}_3$.

J'admettrai ici ces deux lemmes (voir [BOR 2] pour une démonstration). En itérant K fois cette méthode, on obtient donc une mesure $\nu \in \mathcal{P}_+$ telle que :

$$\begin{aligned} B[\nu] &= B[\mu] - 1/\mathbb{Z}^* \\ |\nu| &= |\mu| + K2^{-n} \leq |\mu| + \eta \end{aligned}$$

ce qui termine la preuve du théorème 4, et par là celle du théorème 2.

□

L'hypothèse $\mu \in \mathcal{P}_+$ est cruciale dans la démonstration du lemme 4, la densité f_3 de μ_3 vérifiant $|f_3 - f_2| < \frac{\epsilon}{2}$ sur certains intervalles. L'hypothèse $f_2 > \epsilon$ permet alors de conclure que f_3 est positive.

5.4. Le corollaire 1 se déduit facilement du théorème 2, en considérant $M = \liminf_{n \to \infty} M(A_n)$, et une probabilité μ adhérente à la suite $(\mu_k)_{k \geq 1}$ définie par :

$$\begin{cases} \mu_k \in \Pi^{(k)} \\ \Lambda^{(k)} \quad \text{suite à valeurs dans } [0, M(A_{n_k}) + \epsilon] \text{ telle que } B(\Lambda^{(k)}) = A_{n_k} \\ \lim_{k \to \infty} M(A_{n_k}) = M \end{cases}$$

On a alors $A \subset B[\mu]$, μ ayant son support contenu dans l'intervalle $[0, M+\epsilon]$.

6. Le cas $A \subset \mathbb{Z}$

6.1. Le cas simple $\Gamma_k = \bigcup_{j=1}^{k} \gamma_k \mathbb{Z}^*$ est régi par le théorème 3, que je ne montrerai pas ici (voir [BOR 1]). La démonstration se fait en deux étapes :

1er pas. Par une méthode de découpage, et en reprenant les notations définies en 2.3, si $\Gamma_k \subset B[\mu]$, on a aussi $\Gamma_k \subset B[\nu]$, où l'on pose :

$$\nu = \lambda_{1/m} * \sum_{n=0}^{m\,M} \mu([\tfrac{n}{m}, \tfrac{n+1}{m}[)\, \delta_{n/m}$$

si $\mu([0,M]) = 1$. Le polynôme $Q = \sum_{n=0}^{m\,M} \mu([\tfrac{n}{m}, \tfrac{n+1}{m}[)\, X^n$ est donc divisible par P, et comme $Q(1) = 1$, et $Q \geq 0$, on a donc

$$d^0 Q \geq d^0 P + \delta P.$$

D'où la minoration de M, et donc de $M(\Gamma_k)$.

2ème pas. Si $P \geq 0$, la mesure ν associée $\nu = \lambda_{1/m} * \prod_{a \in \Delta_k} (\delta_{1/m} - \zeta^a \delta_0)$ vérifie $\Gamma_k = B[\nu]$.

Dans le cas général, il en est de même pour ν associée à tout polynôme PQ où Q satisfait les conditions de δ^*. Si en particulier Q satisfait les conditions de δ^+ (i.e. $PQ > 0$), on a $\nu \in \mathcal{P}_+$. Cela donne la majoration de $M(\Gamma_k)$.

6.2. Un cas particulier intéressant est celui où les γ sont deux à deux premiers entre eux. On peut alors montrer que $\delta^+ P \leq d^0 P$ (voir [BOR 3]), et donc $M(\Gamma_k) \leq 2\, d(\Gamma_k)$. La démonstration de ce résultat utilise une formule de récurrence (sur k) sur les polynômes P associés aux Γ_k. A l'aide du théorème 2, on peut alors montrer le résultat suivant.

PROPOSITION 2. ***Soit*** $A = \bigcup_{j=1}^{\infty} \gamma_j \mathbb{Z}^*$, ***avec*** $\gamma_j = \frac{p_i}{q_j} \in \mathbb{Q}$, ***l'ensemble des*** p_j ***et*** q_j ***étant premiers entre deux à deux. Alors*** A ***est*** b***-normal si et seulement si*** $s(A)$ ***est fini.***

6.3. Posons $\gamma_j = T + j$, pour $1 \leq j \leq T$. L'ensemble Γ_T associé a été étudié par Erdös ([ERD] ou [HAL p. 256]), qui a montré que l'on a :

$$\lim_{T \to \infty} d(\Gamma_T) = 0. \tag{13}$$

Mais d'après (4), on a :

$$M(\Gamma_T) \geq \frac{5}{9}\, s(\Gamma_T) \geq \frac{5}{18}$$

ce qui, combiné avec (13), montre que l'hypothèse $M(A) \ll d(A)$, lorsque A vérifie (2), est fausse. Tenenbaum a donné l'ordre de grandeur de $d(\Gamma_T)$ lorsque T tend vers $+\infty$, [TEN 1], et a aussi obtenu dans [TEN 2] une estimation de la fonction $H(x,y,z)$, nombre d'entiers $n \leq x$ ayant au moins un diviseur entre y et z, qui permet de construire des Γ_k tels que :

$$\begin{cases} \lim_{k \to \infty} d(\Gamma_k) = 0 \\ \lim_{k \to \infty} s(\Gamma_k) = 1 \end{cases}.$$

7. Questions ouvertes.

7.1. La relation (4) conduit à la question suivante : si A vérifie (2) et si $s(A)$ est fini, est–ce que A est b–normal ? Cela entraînerait alors :

[conj. 1] A est b–normal si et seulement si il vérifie (2) et $s(A)$ est fini,

cette conjecture ayant une version quantitative plus forte :

[conj. 2] Si A vérifie (2), on a $M(A) \ll s(A)$.

7.2. Le théorème 2 permet alors de montrer que si [conj. 2] est vraie lorsque $A \subset \mathbb{Z}$, elle est encore vraie si $A \subset \mathbb{Q}$. Le passage de $\mathbb{Q}$ à $\mathbb{R}$ pose cependant un problème.

7.3. Une étude des cas simples conduit à émettre l'hypothèse suivante :

[conj. 3] Si A_1 et A_2 vérifient (2), et s'ils sont disjoints, alors $M(A_1 \cup A_2) = M(A_1) \cup M(A_2)$.

Il est à noter que si A_1 et A_2 sont disjoints et vérifient (2), ils sont indépendants sur $\mathbb{Q}$. Le théorème 1 combiné avec l'exemple dû à Erdös (cf. 6.3) entraîne alors :

[conj. 1] et [conj. 3] sont incompatibles.

7.4. On peut aussi s'intéresser à des ensembles b–normaux particuliers, par exemple :

A bien b–normal, i.e. il existe μ bornée telle que $A = B[\mu]$

A très bien b–normal, i.e. $A = B[\mu]$ avec μ à support dans $[O,M(A)]$.

Les Γ_k introduits en 2.3 sont très bien b–normaux. Mais je ne sais pas montrer que b–normal entraîne bien b–normal en général.

BIBLIOGRAPHIE

[BOR 1] J.–P. Borel.– Ensembles normaux associés aux suite bornées, pré–publication du département de Mathématiques de Limoges, 1988.

[BOR 2] J.–P. Borel.– Parties d'ensembles b–normaux, à paraître dans Manuscripta Mathématica.

[BOR 3] J.–P. Borel.– Polynômes à coefficient positifs multiples d'un polynôme donné, Colloque "50 ans de polynômes", 26–27 mai 1988, IHP Paris, à paraître.

[DRE 1] F. Dress.– Intersection d'ensembles normaux, J. of Number Theory 2 (1970), 352–353.

[DRE 2] F. Dress et M. Mendès France.– Caractérisation des ensembles normaux dans $\mathbb{Z}$, Acta Arith. 17 (1970), 115–120.

[ERD] P. Erdös.– Note on sequences of integers no one of which is divisible' by any other, J. london Math. Soc. 10 (1935), 126–128.

[HAL] H. Halberstam et K.F. Roth.– "Sequences", vol. I, Oxfort at the Clarendon Press, 1966.

[KUI] L. Kuipers et H. Niederreiter.– "Uniform distribution of Sequences", Wiley Interscience, 1974.

[LIA] P. Liardet et G. Rauzy.– Communication privée.

[MEN] M. Mendès France.– La réunion des ensembles normaux, J. of Number Theory 2 (1970), 345–351.

[RAU 1] G. Rauzy.– Normalité de $\mathbb{Q}^*$, Acta Arith. 19 (1971), 43–47.

[RAU 2] G. Rauzy.– Caractérisation des ensembles normaux, Bull. Soc. Math. France 98 (1970), 401–414.

[TEN 1] G. Tenenbaum.– Un problème de probabilité conditionnelle en Arithmétique, Acta Arith. 49 (1987).

[TEN 2] G. Tenenbaum.– Sur la probabilité qu'un entier possède un diviseur dans un intervalle donné, Compositio Math. 51 (1984), 243–263.

Jean–Pierre Borel
UFR des Sciences, Dep. de Math.
123, av. Albert Thomas
87060 LIMOGES CEDEX

MULTIPLICATIVE FUNCTIONS ON ARITHMETIC PROGRESSIONS

P.D.T.A. ELLIOTT

In this account g will be a multiplicative arithmetic function with values in the complex unit disc, $|g(n)| \leq 1$. There will be no other restrictions upon it, and that will be the novelty.

THEOREM 1. *There is a positive constant c so that*

$$\sum_{p \leq x^{\frac{1}{2}} \exp(-(\log x)^{23/24})}{}' \max_{(r,p)=1} \max_{y \leq x} \Big| \sum_{\substack{n \leq y \\ n \equiv r (\mathrm{mod}\ p)}} g(n) - \frac{1}{p-1} \sum_{\substack{n \leq y \\ (n,p)=1}} g(n) \Big|^2 \ll x^2 (\log x)^{-c}$$

where $'$ indicates that at most one prime modulus p is excluded from the summation.

The choice of a Dirichlet character $(\mathrm{mod}\ 5)$ for g shows that the exceptional modulus can actually occur. If the number of possible exceptional moduli is raised to $2\phi(a)$, then g may be restricted to the integers $n \equiv b (\mathrm{mod}\ a)$.

Small Moduli. These are taken care of by

THEOREM 2.

$$\sum_{\substack{n \leq y \\ n \equiv r (\mathrm{mod}\ D)}} g(n) = \frac{1}{\phi(D)} \sum_{\substack{n \leq y \\ (r,D)=1}} g(n) + O\Big(y \Big(\frac{\log\log x}{\log x}\Big)^{1/8} \frac{\log x}{\log y}\Big)$$

holds uniformly for $2 \leq y \leq x$, all $(r,D) = 1$ and all moduli D with the possible exception of the multiples of a single modulus $D_0 > 1$.

Once again the exceptional moduli may actually occur.

For complex $s = \sigma + i\tau$, $\sigma = \mathrm{Re}(s) > 1$, and Dirichlet character χ, define

$$G(s,\chi) = \sum_{n=1}^{\infty} g(n)\chi(n)n^{-s}, \quad G(s) = \sum_{n=1}^{\infty} g(n)n^{-s}.$$

Theorem 2 rests upon.

LEMMA A. *The inequality*

$$|\, G(\sigma+i\tau_1,\chi_1)|\, |G(\sigma+i\tau_2,\chi_2)| \le e^{40}(\zeta(\sigma)^3|L(\sigma+i(\tau_1-\tau_2),\chi_1\bar{\chi}_2)|\,)^{\frac{1}{2}}$$

holds uniformly for $\sigma > 1$, ***real*** τ_j. ***Here*** $L(s,\chi)$ ***and*** $\zeta(s)$ ***denote the usual L–function of Dirichlet and Riemann's zeta function, respectively.***

Proof. For complex numbers w_1, w_2 of absolute value 1,

$$2\mathrm{Re}(1-w_1\bar{w}_2) = |\,1-w_1\bar{w}_2|^{\,2} = |\,w_2-w_1|^{\,2} \le 2|\,w_2-1|^{\,2} + 2|\,1-w_2|^{\,2} = \sum_{j=1}^{2} 4(1-\mathrm{Re}\ w_j)$$

follows from the Cauchy–Schwarz inequality. If now z is a further complex number of modulus 1, then we note that each w_j can be replaced by zw_j without affecting the extreme left end of the chain. The desired result can now be obtained by considering Euler products.

LEMMA B. *Let* $T = (\log x)^7$, $\sigma_0 = 1 + (\log x)^{-1}$. *Then*

$$\sum_{n\le y} g(n) \ll \frac{\log x}{\log y} \max_{|\tau|\le T} (\zeta(\sigma_0)^{-1}|\, G(\sigma_0+i\tau)|\,)^{\frac{1}{2}}$$

uniformly for $2 \le y \le x$.

Results of this type may be derived with the methods of Wirsing and Halász. A proof of Theorem 2, including proofs of these two lemmas, is given in my 1988 paper in **Mathematika**. Lemma A represents a generalised Deuring–Heilbronn phenomenon. In fact the Dirichlet series corresponding to two multiplicative functions $h_j(n)$, $|\,h_j(n)| = 1$ for all n, can be linked in almost the same way to the Dirichlet series corresponding to their *Kronecker product* $h_1(n)\overline{h_2(n)}$. The present lemma A shows that for inequivalent characters χ_j at most one of the series

$G(s,\chi_j)$ can have a *pole* on the line $\sigma = 1$. The classical result rules out too many zeros, but an analogue may be derived by considering $\sum_{n=1}^{\infty} \mu(n)g(n)\chi(n)n^{-s} \approx G(s,\chi)^{-1}$ in place of $G(s,\chi)$.

The same arguments show that if for some constant θ, $0 < \theta < 1$,

$$M(x,D) = \max_{x^\theta \le y \le x} \max_{(r,D)=1} y^{-1} \Big| \sum_{\substack{n \le y \\ n \equiv r (\mathrm{mod}\ D)}} g(n) - \frac{1}{\phi(D)} \sum_{\substack{n \le y \\ (n,D)=1}} g(n) \Big|$$

$$M(x) = \max_{x^\theta \le y \le x} y^{-1} \Big| \sum_{n \le y} g(n) \Big|,$$

then

$$M(x)M(x,D) \ll \Big(\frac{\log\log x}{\log x}\Big)^{\frac{1}{4}},\ M(x,D_1)M(x,D_2) \ll \Big(\frac{\log\log x}{\log x}\Big)^{\frac{1}{4}}$$

provided $(D_1,D_2) = 1$. In a sense g can only be badly distributed within the residue classes of one modulus.

Large Moduli.

Three inequalities of Large Sieve type are used.

LEMMA C.

$$\sum_{D \le Q} \sideset{}{^*}\sum_{\chi (\mathrm{mod}\ D)} \Big| \sum_{n \le y} a_n \chi(n) \Big|^2 \ll (y + Q^2) \sum_{n \le y} |a_n|^2.$$

This is a standard version, with $*$ denoting summation over primitive characters.

LEMMA D.

$$\sum_{D \le Q} \sum_{\chi (\mathrm{mod}\ D)} \Big| \sum_{p \le y} a_p \chi(p) \Big|^2 \ll \Big(\frac{y}{\log y/Q} + Q^2\Big) \sum_{p \le Q} |a_p|^2$$

uniformly for $2 \le Q < y$.

Here p denotes a prime. A proof may be found in Chapter 6 of my 1984/85 Springer book on Arithmetic Functions and Integer Products.

LEMMA E.

$$\sum_{p\le Q}\max_{v-u\le H}\sum_{\chi(\mathrm{mod}\ p)}^{*}\Big|\sum_{u<n\le v}a_n\chi(n)\Big|^2\ll (H+Q^2\log Q)\sum_{n=-\infty}^{\infty}|a_n|^2$$

for all complex numbers a_n *for which the infinite sum converges,* $H\ge 0$, $Q\ge 2$.

A proof of this and other maximal variants of the large sieve may be found in my 1987 Proc. London Math. Soc. paper on Additive Functions on Arithmetic Progressions.

I next sketch the basic ideas when the range of summation in Theorem 1 is $p\le x^{\beta}$, for a fixed β, $0<\beta<1/2$. It will be enough to obtain an estimate

$$\sum_{p\le x^{\beta}}{}'\max_{y\le x}\frac{1}{p}\sum_{\chi(\mathrm{mod}\ p)}\Big|\sum_{n\le y}g(n)\chi(n)\log n\Big|\ll x(\log x)^{1-C_1}$$

for some positive absolute constant C_1.

I employ the representation

$$\sum_{n\le y}g(n)\chi(n)\log n=-\frac{1}{2\pi i}\int_{\sigma_0-i\infty}^{\sigma_0+i\infty}G'(s,\chi)\frac{y^s}{s}\,ds,$$

valid for positive non–integral y. In fact the integral need not converge absolutely, and it is convenient to integrate over intervals $x_1-\rho\le y\le x_1$ a few times, in order to introduce a power s^k into the denominator of the integrand. This introduces a factor $w(n)$ into the sum, where $w(n)=1$ if $n\le y-k\rho$, $0\le w(n)\le 1$ otherwise :

$$\sum_{n\le y}w(n)g(n)\log n=-\frac{1}{2\pi i}\int_{\sigma_0-i\infty}^{\sigma_0+i\infty}G'(s,\chi)K(s)\,ds,\quad y\le x. \tag{1}$$

Ultimately we choose ρ of the form $x(\log x)^{-C_2}$ for a (small) positive C_2, and apply a maximal version of the Large Sieve, Lemma E, to remove the $w(n)$. For such a value of ρ

$$\int_{-\infty}^{\infty}\sup_{y\le x}|K(\sigma_0+i\tau)|\,d\tau\ll x\ \mathrm{loglog}\ x,$$

a bound which we may regard as satisfactory.

Defining

$$G_p = G_p(s,\chi) = 1 + \frac{g(p)\chi(p)}{p^s} + \frac{g(p^2)\chi(p^2)}{p^{2s}} + \dots ,$$

we can replace $G(s,\chi)$ in (1) by the product of the G_p with $p \leq x$. If $0 < \epsilon < 1$, we can essentially confine ourselves to a range of prime moduli $y < q \leq y^{1+\varepsilon}$, where $y \leq x^{\beta}$. Define

$$L_1 = \prod_{p \leq y^{2+3\varepsilon}} G_p , \quad L_2 = \prod_{y^{2+3\varepsilon} < p \leq x} G_p .$$

Then

$$G'(s,\chi) = L_1' + L_1'(L_2 - 1) + L_1 L_2' , \tag{2}$$

and we are effectively reduced to giving an upper bound for

$$J_1 = \sum_{y < q \leq y^{1+3\varepsilon}} \frac{1}{q} \sideset{}{^*}\sum_{\chi (\mathrm{mod}\ q)} \Big| \int_{\sigma_0 - i\infty}^{\sigma_0 + i\infty} L_1' \frac{ds}{s^2} \Big| ,$$

and the two similar expressions J_2, J_3 derived from the decomposition (2). These are to be of the form $J_j \ll (\log x)^{1-C_1}$.

J_1 is estimated by moving the range of integration to the line $\sigma = 1/2$. This may be done *either* by using estimates for L_1 (and so by Cauchy's integral representation theorem, for L_1') obtained from Lemma C with the aid of high moment inequalities, *or* by modifying suitably a method of Gallagher. The use of high moments goes back to Linnik. In my original proof of a version of Theorem 1 I adapted relevant results of Wolke. The method of Gallagher cannot be directly adopted either, so that some complications anyway ensue.

To estimate J_2 we apply the Cauchy–Schwarz inequality, and again reduce ourselves to a more accessible inequality. Thus we need

$$\sum_{q \leq y^{1+\varepsilon}} \sideset{}{^*}\sum_{\chi (\mathrm{mod}\ q)} \int_{-\infty}^{\infty} |L_2(\sigma_0 + i\tau) - 1|^2 \frac{d\tau}{\tau^2} \ll (\log x)^{2-C_3} . \tag{3}$$

We employ Parseval's relation, in a manner familiar from the method of Halász, replacing the multiple sum at (3) by

$$\int_0^\infty \sum_{q \leq y^{1+\varepsilon}} \sum_{\chi (\mathrm{mod}\ q)}^{*} \Big| \sum_{n \leq e^u} g_2(n)\chi(n) \Big|^2 e^{-2\sigma_0 u}\, du.$$

Here

$$g_2(n) = \prod_{\substack{p^m \| n \\ y^{2+3\varepsilon} < p \leq x}} g(p^m),$$

so that $g_2(n) = 0$ if $1 < n \leq y^{2+3\varepsilon}$, and $|g_2(n)| \leq 1$ always. The integrand is estimated by Lemma C, and we gain for the multiple sum at (3) the upper bound

$$\int_0^\infty e^{2u(1-\sigma_0)}\, du \ll \frac{1}{\sigma_0 - 1} \ll \log x. \tag{4}$$

This has the desired form.

To estimated J_3 I employ the factorisation $L_2' = (L_2'/L_2).L_2$, and reduce to the consideration of an analogue of (3), the Dirichlet series $L_2(s)-1$ being replaced by L_2'/L_2. In place of Lemma C we apply Lemma D. In place of the bound

$$(e^u + y^{2+2\varepsilon}) \sum_{n \leq e^u} |g_2(n)|^2 \ll e^{2u}$$

employed in (4), we obtain essentially

$$(\frac{e^u}{u} + y^{2+2\varepsilon}) \sum_{p \leq e^u} (|g_2(p)| \log p)^2 \ll e^u \sum_{p \leq e^u} \log p \ll e^{2u}. \tag{5}$$

It is here that the restriction $g_2(p) = 0$ for $p \leq y^{2+3\varepsilon}$ becomes important, allowing e^u/u to dominate $y^{2+2\varepsilon}$.

This completes my sketch of the Theorem when the moduli run up to x^β. A detailed proof will appear in Mathematika.

To obtain the slightly longer range $p \leq x^{\frac{1}{2}} \exp(-(\log x)^{23/24})$ a more careful application is made of Lemma D, with the consequence that in the inequality (5), e^u/u is replaced by e^u/u^θ for a suitably chosen θ, $0 < \theta < 1$, and $y^{2+2\varepsilon}$ by $y^2 \exp((\log y)^\phi)$ for a suitably chosen ϕ, $0 < \phi < 1$. The exponent $23/24$ is not optimal.

No doubt Theorem 1 is valid with the summation over p running essentially up to x, say to $x(\log x)^{-C_4}$ for a suitable C_4.

Professor P.D.T.A. ELLIOTT
Department of Math. Box 426
University of Colorado
BOULDER, COLORADO 80309
U.S.A.

LE DIXIEME PROBLEME DE HILBERT

M. MARGENSTERN

Posé en 1900 par Hilbert lors de son adresse célèbre au Congrès des mathématiciens, ce problème demande une méthode générale de résolution des équations diophantiennes.

Que fallait–il comprendre par *méthode générale* ?

On devait attendre les années trente pour donner une formulation mathématique précise du problème. Grâce aux travaux de Gödel sur l'arithmétisation de théories formelles, l'étude des fonctions récursives connut un regain d'intérêt débouchant rapidement sur des résultats fondamentaux. Dans le même temps, des formalisations nouvelles de la notion d'algorithme apparaissaient : machines de Turing et λ–calcul, puis, beaucoup plus tard, systèmes de Post, algorithmes de Markov, machines à registres, etc... En 1936, la démonstration de l'équivalence des notions mathématiques d'algorithme alors connues permettait d'énoncer le dixième problème de Hilbert sous sa forme actuelle : existe–t–il un algorithme *formel** permettant de résoudre toute équation diophantienne donnée ?

Quelques définitions.

k étant un entier positif non nul, on appelle **fonction** de $\mathbb{N}^k$ dans $\mathbb{N}$ une *application* de $D \subset \mathbb{N}^k$ dans $\mathbb{N}$. D est le **domaine de définition** de la fonction. Si f est une fonction de $\mathbb{N}^k$ dans $\mathbb{N}$, on note *dom* f son domaine de définition. On note $\mathfrak{F}^k$ l'ensemble des fonctions de $\mathbb{N}^k$ dans $\mathbb{N}$.

On appelle **fonctions de base** les fonctions suivantes :

(i) $0 : x \longmapsto 0$ de $\mathbb{N}$ dans $\mathbb{N}$;

(ii) $S : x \longmapsto x + 1$ de $\mathbb{N}$ dans $\mathbb{N}$

(iii) $U_i^k : (x_1,\ldots x_k) \longmapsto x_i$ de $\mathbb{N}^k$ dans $\mathbb{N}$.

On appelle **schéma de composition** l'équation :

$$f(x_1,\ldots,x_k) \simeq h(g_1(x_1,\ldots,x_k),\ldots,g_n(x_1,\ldots,x_k))$$

---

* Quel que soit le sens de la formalisation choisie puiqu'elles sont équivalentes.

où f, h, $g_1,\ldots,g_n$ sont des fonctions, respectivement de $\mathbb{N}^k$ dans $\mathbb{N}$, de $\mathbb{N}^n$ dans $\mathbb{N}$ et de $\mathbb{N}^k$ dans $\mathbb{N}$;

schéma de récursion, le système d'équations :

$$f(x_1,\ldots,x_k,0) \simeq g(x_1,\ldots,x_k)$$
$$f(x_1,\ldots,x_k,Sy) \simeq h(x_1,\ldots,x_k,y,f(x_1,\ldots,x_k,y))$$

où f, g et h sont des fonctions, respectivement de $\mathbb{N}^{k+1}$ dans $\mathbb{N}$, de $\mathbb{N}^k$ dans $\mathbb{N}$ et de $\mathbb{N}^{k+2}$ dans $\mathbb{N}$;

et **schéma de minimisation**, l'équation :

$$f(x_1,\ldots,x_k) \simeq \mu y(gx_1,\ldots,x_k,y) = 0)$$

où f et g sont des fonctions, respectivement de $\mathbb{N}^k$ dans $\mathbb{N}$ et de $\mathbb{N}^{k+1}$ dans $\mathbb{N}$, l'expression $\mu y\, A(y)$, où A est une formule, désignant le plus petit des entiers naturels m pour lesquels $A(m)$ est vrai s'il existe de tels entiers, et n'étant pas définie s'il n'en existe pas.

On appelle **ensemble des fonctions récursives partielles** *le plus petit sous-ensemble de* $\bigcup_{k=1}^{\infty} \mathfrak{F}^k$ *contenant les fonctions de base et stable par rapport au schéma de composition, au schéma de récursion et au schéma de minimisation.*

On dit qu'une partie E de $\mathbb{N}^k$ est récursivement énumérable si et seulement si elle est le domaine de définition d'une fonction récursive partielle. On dit que E est diophantien si et seulement si il existe un polynôme $p \in \mathbb{Z}[X_1,\ldots,X_k, Y_1,\ldots,Y_n]$ tel que pour tout k–uple $a_1,\ldots,a_k$ on ait :

$$a_1,\ldots,a_k \in E \Leftrightarrow \exists y_1,\ldots,y_n \in \mathbb{N}\ (P(a_1,\ldots,a_k, y_1,\ldots,y_n) = 0).$$

Les $a_1,\ldots,a_k$ sont les *paramètres de* P et $y_1,\ldots,y_n$ ses *indéterminées.*

On dit que E est *exponentiellement diophantien* si et seulement si la condition ci–dessus est satisfaite pour un polynôme P *exponentiel* à coefficients dans $\mathbb{Z}$, c'est–à–dire comportant des monômes de la forme $ab^T Y_1^{U_1}\ldots Y_n^{U_n}$ où a, b $U\mathbb{N}$ et $T, U_i \in \mathbb{N} \cup \{X_1,\ldots,X_n\}$.

Le premier pas : le théorème de Davis–Putnam–Robinson.

En 1953, Davis conjecture que le dixième problème de Hilbert est impossible.

L'idée est simple : il *suffit* de démontrer qu'un ensemble récursivement énumérable est diophantien (la réciproque étant quasi immédiate). Or on sait depuis les années trente qu'il n'existe pas d'algorithme formel permettant de décider, pour tout ensemble récursivement énumérable E, si E est vide ou non.

Cette propriété ne sera établie qu'en 1970 par Matiyassévitch. En attendant, on franchit une étape importante : on établit qu'un ensemble récursivement énumérable est *exponentiellement diophantien.* Démontré tout d'abord à l'aide d'une conjecture (toujours ouverte) par Davis et Putnam, le théorème est établi dans toute sa généralité par Julia Robinson en 1961. On a remarqué en 1960 (Putnam) que si la conjecture de Davis est vraie, l'ensemble des nombres premiers est l'image positive d'un polynôme. Puis en 1969, J. Robinson démontre que la conjecture de Davis est vraie si et seulement si l'ensemble des nombres premiers est diophantien.

La démonstration du théorème de Davis–Putnam–Robinson est trop technique pour être esquissée ici. L'idée consiste à représenter le calcul de la valeur en un point d'une fonction récursive partielle (intuitivement, définie à partir d'un nombre fini d'équations correspondant aux schémas donnés ci–dessus) à l'aide de relations arithmétiques. En 1975, Matiyassevitch a considérablement simplifié la démonstration en se plaçant dans le cadre des machines de Turing*, le problème se ramenant alors à représenter le travail d'une telle machine par des suites finies d'entiers de la forme $r_0,\ldots,r_k$ assujetties à des relations arithmétiques simples dont *deux* sont exponentiellement diophantiennes :

$$r_1 + \ldots + r_k = 2^{r_0} - 1 \quad \text{et} \quad \binom{r_i + r_j}{r_i} \equiv 1 \bmod 2.$$

La réduction de l'exponentielle.

En 1970, Y. Matiyassévitch démontre, à la surprise générale**, la conjecture de Davis (cf. [5]). Précisément, il démontre qu'il existe un polynôme P à coefficients entiers tel que pour tous entiers a, b et c on ait :

$$a = b^c \Leftrightarrow \exists\ y_1,\ldots,y_n \in \mathbb{N}\ \ (P(a,b,c,y_1,\ldots,y_n) = 0).$$

---

* On peut encore simplifier la démonstration en remplaçant les machines de Turing par des machines à registre extrêmement simples (cf. [3]). Cependant le codage arithmétique est essentiellement le même que dans le cas des machines de Turing.

** On ne s'attendait pas à ce que ce résultat *tombe* si vite...

Il suffit en fait de démontrer cette propriété pour une *suite* dont la croissance est exponentielle. Matiyassévitch utilise à cet effet les solutions de l'équation de Pell :

$$x^2 - dy^2 = 1. \tag{1}$$

On prend $d = a^2 - 1$ avec $a > 1$. On sait que les solutions positives de (1) sont de la forme $x_a(n)$, $y_a(n)$ où ces nombres sont définis par :

$$(a + \sqrt{a^2-1})^n = x_a(n) + y_a(n)\sqrt{a^2-1}.$$

On démontre alors :

THEOREME (Matiyassévitch). *$c = y_a(b)$ avec $c, b > 0$, $a > 1$ si et seulement si il existe des entiers naturels d, e, f, g, h et i tels que :*

(i) $dfi = \square^*$

(ii) $d = (a^2-1)c^2 + 1$ *et* $c \geq b$

(iii) $f = (a^2-1)e^2 + 1$ *et* $2c^2d \mid e$

(iv) $f \mid g - a$ *et* $2cd \mid g - f$

(v) $i = (g^2-1)h^2 + 1$ *avec* $2c \mid h - b$ *et* $f \mid h - c$.

La *réduction* de l'exponentielle en découle par le résultat suivant :

THEOREME. *Soient x, y, $n \in \mathbb{N}$, y, $x > 0$ tels que $y = x^n$. Alors si $1 > 2ny$, $x^n = \langle \frac{y_{1x}(n+1)}{y_1(n+1)} \rangle$ où $\langle z \rangle$ est l'entier le plus proche de z en supposant $z \notin \mathbb{Z} + \frac{1}{2}$.*

D'où on déduit que $y = x^n$ si et seulement si il existe l, u, v et w entiers naturels tels que

$$\begin{aligned} l &= 2ny + 1 \\ 4(yv-u)^2 - v^2 + w + 1 &= 0 \\ u &= y_{1x}(n+1) \\ v &= y_1(n+1). \end{aligned}$$

---

* C'est–à–dire dfi est un carré.

Représentation des nombres premiers.

Les représentations des nombres premiers données dans la littérature se fondent sur le théorème de Wilson :

$$p \text{ est premier si et seulement si } p \mid (p-1)! + 1.$$

La représentation passe donc par celle de la factorielle elle-même réduite aux coefficients binomiaux :

On obtient facilement que

$$(*) \qquad \frac{n^k}{\binom{n}{k}} = k!\left(1 + \frac{1}{n-1}\right)\ldots\left(1 + \frac{k-1}{n+k-1}\right)$$

d'où on conclut que $k! = \left[\frac{n^k}{\binom{n}{k}}\right]$ pour n assez grand. Une majoration facile permet d'obtenir, à partir de (*) que, plus précisément,

$$k! = \left[\frac{n^k}{\binom{n}{k}}\right] \text{ dès que } n \geq 2k^{k+2},$$

d'où :

$f = k!$ si et seulement si il existe n, r, a et u tels que :

$$n = 2k^{k+2}$$
$$n^k = af + r$$
$$a = \binom{n}{k}$$
$$a = r + u + 1$$

Quant aux coefficients du binômes, on tire de

$$(x+1)^n = \sum_{j=0}^{n} \binom{n}{j} x^j$$

que pour $x > 2^n$, le membre de droite est l'écriture de $(x+1)^n$ en base x. D'où :

$a = \binom{n}{K}$ si et seulement si il existe x, y, b, c, u et v tels que

$$\begin{aligned} x &= 2^n + 1 \\ y &= x + 1 \\ y^n &= bx^{k+1} + ax^k + c \\ 2^n &= a + u \\ x^k &= c + v + 1 \end{aligned}$$

On trouve dans [4] un polynôme explicite dont l'image positive est l'ensemble des nombres premiers. Il comprend 26 variables et est de degré 25. Si on cherche à réduire le nombre des variables, le degré augmente fortement. Le meilleur résultat actuellement connu est dû à Matiyassevitch avec 10 variables et le degré 15905.

Représentation de l'hypothèse de Riemann.

Le théorème de Matiyassevitch permet d'exprimer un certain nombre de problèmes en terme de résolution d'une équation diophantienne.

Le principe général est le suivant :

Soit P un ensemble *récursif* de $\mathbb{N}$, c'est–à–dire récursivement énumérable ainsi que son complémentaire dans $\mathbb{N}$. Désignons par $F(n)$ la propriété $n \in P$. Il existe donc un polynôme Q à coefficients dans $\mathbb{Z}$ tel que :

$$\neg F(n) \Leftrightarrow \exists\, y_1,\ldots,y_m \in \mathbb{N}\ (Q(n,y_1,\ldots,y_m) = 0).$$

Donc, $\exists n \neg F(n) \Leftrightarrow \exists n y_1,\ldots,y_m \in \mathbb{N}\ (Q(n,y_1,\ldots,y_m) = 0)$

d'où $\forall n\, F(n) \Leftrightarrow \forall n y_1,\ldots,y_m \in \mathbb{N}\ (Q(n,y_1,\ldots,y_m) \neq 0).$

On obtient donc que la propriété $\forall n\, F(n)$ est vraie si et seulement si un certaine équation diophantienne n'a pas de solution.

Plusieurs conjectures célèbres peuvent s'énoncer ainsi et sont donc équivalentes à l'impossibilité de résoudre une certaine équation diophantienne.

Tel est le cas de l'hypothèse de Riemann (représentation dûe à H. Shapiro) :

THEOREME. *L'hypothèse de Riemann équivaut à :*

$$\left(\left(\sum_{k \leq \delta(n)} \frac{1}{k}\right) - \frac{n^2}{2}\right)^2 < 36n^3 \quad \textit{pour}\ n = 1, 2, 3 \ldots$$

où $\delta(x) = \prod_{n<x} \prod_{j \leq n} e^{\Lambda(j)}$, *avec* Λ, *la fonction de Von Mangoldt.*

On observe que, par définition de la fonction de Von Mangoldt, $\delta(x)$ est un entier pour tout x réel. Voir, par exemple, une démonstration du théorème dans [1].

Polynômes universels.

On démontre que les fonctions récursives partielles peuvent être énumérées par les entiers naturels. k étant fixé, on associe ainsi à tout élément n de $\mathbb{N}$ une fonction récursive partielle à k arguments notée φ_n. n est appelé *code* ou *numéro* de φ_n. Il importe de remarquer que toute fonction admet une infinité de codes et ceci, quel que soit le procédé de codage. On établit alors, propriété découverte dès les années trente, qu'il existe une fonction récursive partielle *universelle* U vérifiant la propriété suivante :

$$U(n,x_1,\ldots,x_k) \simeq \varphi_n(x_1,\ldots,x_k)$$

pour tout $x_1,\ldots,x_k \in \mathbb{N}^k$, U ne prenant pas de valeur si $x_1,\ldots,x_k \notin dom\ \varphi_n$.

Il résulte du théorème de Matiyassevitch qu'il existe des polynômes diphantiens universels. La recherche de polynômes *explicites* a conduit à une étude plus précises des relations entre le nombre des variables et le degré d'un polynôme universel. On appelle *couple universel* un couple d'entier (ν,δ) pour lequel il existe un polynôme diophantien universel de degré δ et à ν indéterminées. Ainsi Jones a obtenu les couples suivants [2] :

(58;4)	(28;20)	(21;96)	$(12;1,3.10^{44})$
(38;8)	(26;24)	(19;2668)	$(11;4,6.10^{44})$
(32;12)	(25;28)	$(14;2,0.10^{5})$	$(10;8,6.10^{44})$
(29;16)	(24;36)	$(13;6,6.10^{43})$	$(9;1,6.10^{45})$

Autres directions de recherche

D'autres directions de recherche ont été développées à partir du résultat de Matiyassévitch.

La première direction consiste à poser le dixième problème de Hilbert dans un autre cadre que celui des entiers naturels : on peut considérer des polynômes à coefficients dans un anneau $\mathcal{A}$ au lieu de $\mathbb{Z}$. Lorsque $\mathcal{A}$ est une extension de $\mathbb{Z}$, on a deux notions de polynômes diophantiens : les polynômes à coefficients dans $\mathbb{Z}$ sont alors dits *diophantiens purs*. Le problème a été étudié dans le cas des extensions quadratiques de $\mathbb{Q}$, des extensions de $\mathbb{Q}$ de degré fini, de l'anneau des entiers algébriques sur $\mathbb{Q}$, enfin, dans le cas de $\mathbb{Q}$ lui–même. Actuellement, on n'a de réponse que dans le cas des extension quadratiques : le dixième problème n'a pas de

solution dans ce cadre, que l'on considère les polynômes diophantiens ou seulement les diophantiens purs.

Une autre direction de recherche, plus récente, dévoloppée par Jones et Matiyassévitch (cf. par exemple [3]), est la représentation exponentiellement diophantienne. La simplification essentielle apportée à l'étude des polynômes diophantiens réside dans la propriété d'*unicité* des solutions qui peut être exigée dans ce cas. De façon plus précise, on dit qu'une partie E de $\mathbb{N}^k$ admet une représentation exponentiellement diophantienne *univoque*, s'il existe un polynôme diophantien exponentiel P tel que :

$$a_1,\ldots,a_k \in E \Leftrightarrow \exists!\ y_1,\ldots,y_n \in \mathbb{N}\ \ (P(a_1,\ldots,a_k\ ,\ y_1,\ldots,y_n) = 0).$$

Tout ensemble récursivement énumérable admet une représentation exponentiellement diophantienne *univoque*, ce qui n'est pas vrai avec des polynômes non exponentiels (cf. [6]). En particulier, il existe des polynômes diophantiens universels conservant cette propriété. Ceci permet d'obtenir des polynômes universels explicites plus *maniables* que les polynômes diophantiens universels *obtenus* par application du théorème de Matiyassévitch et de la propriété correspondante des fonctions récursives partielles (cf. [2]).

BIBLIOGRAPHIE

[1] M. Davis, Yu. V. Matiyassévitch, J. Robinson.– Proceed. Symposia in Pure Math. 28, 1976, 323–378.

[2] J.–P. Jones.– Universal Diophantine Equation.– J. Symbol. Logic. 47, 1982, 549–571.

[3] J.–P. Jones, Yu. V. Matiyassévitch.– Exponential Diophantine Representation of Recursively Enumerable Sets, Stud. Logic. and Found. Math. 107, 1982, 159–177.

[4] J.–P. Jones, D. Sato, H. Wada, D. Wiens.– Diophantine Representation of the Set of Prime Numbers, Amer. Math. Monthly. 83, 1976, 449–464.

[5] Yu. V. Matiyassévitch.– Les ensembles récursivements énumérables sont diophantiens, Comptes–Rendus Acad. Sc. URSS, 191, 1970, 279–282, (en russe).

[6] Yu. V. Matiyassévitch.– Indécidabilité algorithmique des équations exponentiellement diophantiennes à trois inconnues, in Recherches en théorie des algorithmes et en logique mathématique, Ed. Naouka, Moscou 1979, 69–77 (en russe).

L'article [1] est une revue très complète des recherches sur le Dixième problème de Hilbert jusqu'en 1976. On peut également consulter :

M. Margenstern.– le théorème de Matiyassévitch et résultats connexes, in Model Theory and Arithmetics, Lecture Notes 890, 1981, 198–241. qui contient la démonstration de la plupart des résultats cités.

Maurice MARGENSTERN
Département de Mathématiques
Bâtiment 425
Université Paris–Sud
91405 ORSAY CEDEX

COMMENTAIRE AUTOUR D'UN ARTICLE DE J. DELSARTE

J.–L. MAUCLAIRE

I – Une remarque préliminaire.

L'article en question, intitulé *essai sur l'application de la théorie des fonctions presque–périodique à l'Arithmétique*, est paru aux Annales Scientifiques de l'Ecole Normale Supérieure en 1945 [1]. Il est essentiel d'en citer les trois premières lignes : *l'examen de la plupart des développements formels ou asymptotiques que l'on rencontre en Théorie des Nombres, met en évidence le rôle important joué par les séries procédant suivant les sinus et les cosinus d'arcs qui sont des parties aliquotes de la circonférence*; puis l'auteur remarque qu'il y a un lien avec la presque–périodicité, et annonce que son article est une étude du problème : quand, à une fonction arithmétique, i.e. à une application $f : \mathbb{N}^* \longrightarrow \mathbb{C}$ où $\mathbb{N}^*$ dénote l'ensemble des entiers strictement positifs, est associée une série de Fourier du type énoncé précédemment, qui la définit et lui confère des propriétés de presque–périodicité ? La raison pour laquelle les trois première lignes sont mystérieuses est la suivante : à aucun moment, il n'est fait allusion à une seule référence bibliographique dans cet article; qui plus est, une note en bas de page dans le commentaire d'A. Weil [2] nous informe qu'il semble que ce soit seulement en 1947, à Princeton, que J. Delsarte ait pris connaissance des travaux de Ramanujan; d'ailleurs, dans l'article que nous allons examiner, la *somme de Ramanujan* $C_q(n)$ est appelée *indicateur d'ordre* n *de* q et notée $\Phi(q|\, n)$. Etant donné que le lien *sommes de Ramanujan - presque-périodicité* est mentionné dès 1940 aux U.S.A., avec les travaux d'Erdös, Wintner, Van Kampen, Kac etc [3] [4], et est alors particulièrement étudié, il semble difficile de tracer les idées de Delsarte à ce sujet dans le présent article, surprenant par son approche du problème. Il est en effet impossible de déterminer *les développements formels ou asymptotique* qui sont à l'origine de cette recherche très originale, et pour tout compliquer, il faut rappeler que son auteur est resté en France pendant la période 1939–1945; c'est dommage pour l'histoire de la pensée, mais le papier est là, que nous allons maintenant présenter en actualisant un peu les appellations.

II – **Ce que contient essentiellement le texte.**

A – **Partie formelle.**

a – Tout d'abord, la définition des *Sommes de Ramanujan* est donnée comme généralisation de celle de l'indicateur d'Euler. En effet, on peut considérer $C_q(n)$ comme la somme des puissances n–ièmes des racines primitives q–ièmes de l'unité, i.e. :

$$C_q(n) = \sum_{\substack{(h,q)=1 \\ 1 \leq h \leq q}} e^{2i\pi \frac{h}{q} n},$$

qui donne $\varphi(q)$ si n est divisible par q.

– On établit ensuite que

$$\sum_{d \mid q} C_d(n) = \begin{cases} q \text{ si } q \mid n \\ 0 \text{ sinon} \end{cases},$$

ceci par un calcul direct sur les racines de l'unité, et, si μ dénote la *fonction de Möbius*,

$$C_q(n) = \sum_{\substack{d \mid q \\ d \mid n}} d\, \mu\left(\frac{q}{d}\right)$$

en inversant la formule précédente par la *formule d'inversion de Möbius-Mertens.* Puis, en poussant les calculs, il est établi que

$$\frac{1}{q} C_q(n) = \frac{\mu(\nu)}{\nu} \times \frac{\varphi(\Delta)}{\Delta},$$

où $\nu = \frac{q}{(n,q)}$, $\Delta = \text{Max}\,\{d,\, d \mid q,\, d \mid n,\, (d,\nu) = 1\}$. De ceci, on déduit que

$C_q(n) \leq (n,q)$.
$C_q(n) = \mu(q)$ si $(n,q) = 1$
$C_q(n) C_{q'}(n) = C_{qq'}(n)$ si $(q,q') = 1$ etc...

b — Vient ensuite la présentation des développements formels des fonctions de la théorie des Nombres,

On va supposer, jusqu'au § B, que ce que l'on écrit a un sens :

On note $M(f)$ la moyenne arithmétique de f, i.e.

$$M(f) = \lim_{x \to +\infty} \frac{1}{x} \sum_{n \leq x} f(n).$$

Alors, si f et g sont deux fonctions arithmétiques, l'expression $<f,g> = M(f.\bar{g})$ peut être considérée comme un produit scalaire. Si α et β sont deux racines de l'unité, on a :

$$<\alpha,\beta> = \begin{matrix} 1 \text{ si } \alpha = \beta \\ 0 \text{ sinon} \end{matrix},$$

ce qui signifie que les racines de l'unité forment un système orthonormé. De cela, on déduit que

$$<C_q, C_{q'}> = \begin{matrix} 0 & \text{si } q \neq q' \\ \varphi(q) & \text{si } q = q' \end{matrix}.$$

c — Jusqu'à maintenant, nous sommes restés dans le domaine du ronronnement pour les initiés. Mais voici une surprise :

soit $f : \mathbb{N}^* \longrightarrow \mathbb{C}$. On définit :

$$F(n) = \sum_{d \mid n} \mu(d)\, f(\tfrac{n}{d}).$$

On sait que :

$$f(n) = \sum_{d \mid n} F(d).$$

Calculons $<f,\alpha>$, où α est une racine k–ième de l'unité. On a :

$$\sum_{m=1}^{n} f(m)\bar{\alpha}^m = \sum_{m=1}^{n} \Big(\sum_{d|n} F(d)\Big)\bar{\alpha}^m$$

$$= \sum_{d=1}^{n} F(d) \times \sum_{\substack{d|n \\ m\leq n}} \bar{\alpha}^m$$

$$= \sum_{d=1}^{n} F(d) \times (\bar{\alpha}^d + \bar{\alpha}^{2d} + \ldots + \bar{\alpha}^{d[\frac{n}{d}]})$$

si $k|d$, alors $\sum_{r=1}^{[\frac{n}{d}]} \bar{\alpha}^{dr} = [\frac{n}{d}]$; si $k \nmid d$, l'expression ci–dessus reste bornée. Donc, si l'on est dans un *bon cas*, quand $n \longrightarrow +\infty$, la *partie principale* de $\sum_{m=1}^{n} f(m)\bar{\alpha}^m$ sera

$$n \times \sum_{h=1}^{[\frac{n}{k}]} \frac{F(kh)}{kh}.$$

On prend donc :

$$<f,\alpha> = \sum_{h=1}^{+\infty} \frac{F(kh)}{kh}.$$

On associe alors à f la série

$$\sum_{k=1}^{+\infty} A_k\, C_k(n),$$

où $A_k = \sum_{h=1}^{+\infty} \frac{F(kh)}{kh}$.

On remarque ensuite qu'en inversant le sens des sommations, on a :

$$\begin{aligned}\sum_{k=1}^{+\infty} A_k\, C_k(n) &= \sum_{k=1}^{+\infty} \frac{F(kh)}{kh}\, C_k(n)\\ &= \sum_{d=1}^{+\infty} \frac{F(d)}{d} \Big(\sum_{k\mid d} C_k(n)\Big)\\ &= \sum_{d\mid n} F(d)\\ &= f(n).\end{aligned}$$

Donc, la série converge formellement vers f.

d – Cette construction peut se généraliser de la façon suivante :
Soit λ dans $\mathbb{N}^*$; on pose

$$A_\lambda(k) = \sum_{k\mid d^\lambda} \frac{F(d)}{d^\lambda},$$

et on considère :

$$\sum_{k=1}^{+\infty} A_\lambda(k)\, C_k(n).$$

Formellement, on a :

$$\begin{aligned}\sum_{k=1}^{+\infty} A_\lambda(k)\, C_k(n^\lambda) &= \sum_{k=1}^{+\infty} \sum_{k\mid d^\lambda} \frac{F(d)}{d^\lambda}\, C_k(n^\lambda)\\ &= \sum_{d=1}^{+\infty} \frac{F(d)}{d^\lambda} \sum_{k\mid d^\lambda} C_k(n^\lambda).\end{aligned}$$

Or, on sait que :

$$\sum_{k\mid d^\lambda} C_k(n^\lambda) = \begin{cases} 0 & \text{si } d^\lambda \nmid n^\lambda \\ d^\lambda & \text{si } d^\lambda \mid n^\lambda, \end{cases}$$

c'est–à–dire si $d\mid n$.

La somme formelle de la série est donc $\sum_{d\mid n} F(d)$, qui est égale à $f(n)$. De même, on a :

$$\sum_{k=1}^{+\infty} A_\lambda(k)\, C_k(n) = \sum_{d^\lambda \mid n} F(d).$$

On pose alors $[n]_\lambda = \mathrm{Max}\{d, d^\lambda \mid n\}$, et on remarque que $d^\lambda \mid n$ signifie $d \mid [n]_\lambda$. Par conséquent, la somme de la série est

$$\mathscr{F}_\lambda(n) = f([n]_\lambda).$$

e – Dernière généralisation.

On appelle *isomorphisme pour la divisibilité* une application $I : \mathbb{N}^* \longrightarrow \mathbb{N}^*$ telle que :

$$n \textit{ divise } m \text{ équivaut à } I(n) \textit{ divise } I(m).$$

exemple : si $n = p_1^{\alpha_1} \times \ldots \times p_r^{\alpha_r}$, alors, on peut définir $I(n)$ par :

$$I(n) = p_1^{f_1(\alpha_1)} \times \ldots \times p_r^{f_r(\alpha_r)},$$

où $f_1, \ldots, f_r, \ldots$ sont des applications $\mathbb{N}^* \longrightarrow \mathbb{N}^*$ strictement croissantes.

Si on a un tel isomorphisme, les calculs précédents peuvent encore s'effectuer. On pose :

$$A_I(k) = \sum_{k \mid I(d)} \frac{F(d)}{I(d)}.$$

Alors, la série $\sum_{k=1} A_I(k)\, C_k(n)$ converge formellement vers $\sum_{I(d) \mid n} F(d) = \mathscr{F}_I(n)$. En particulier, on a :

$$\mathscr{F}_I(I(n)) = f(n).$$

B – Quelques résultats généraux.

1 – Tout revient ici à légitimer les calculs. Trois cas sont considérés :

a – $$\sum_{n=1}^{+\infty} |F(r)| < +\infty.$$

Alors, on a :

THEOREME. *Les séries* $\sum_{k=1}^{+\infty} A_\lambda(k)\, C_k(n)$ *sont absolument convergentes par rapport à* n, *(*n *décrivant l'ensemble des entiers relatifs* $\mathbb{Z}$*), vers* $\mathscr{F}_\lambda(n)$ *i.e.* $f([n]_\lambda)$, *et par conséquent, les* $\mathscr{F}_\lambda$ *sont presque-périodiques au sens de Bohr.*

b – Il existe $A > 0$, $\sigma > 0$, tels que :

$$\sum_{k=1}^{n} |F(k)| \leq A\, n^\sigma.$$

Dans ce cas, on a le résultat :

THEOREME. *Si* $\lambda > \sigma$, $\sum_{k=1}^{+\infty} A_\lambda(k)\, C_k(n)$ *converge absolument, (*n *décrivant l'ensemble* $\mathbb{N}^*$ *des entiers strictement positifs), et a pour somme* $\mathscr{F}_\lambda(n)$.

Il est d'ailleurs remarqué que l'on ne peut ici rien en déduire en ce qui concerne une éventuelle presque–périodicité.

c – Par une utilisation simple du Théorème de Riesz–Fisher–Besicovitch, on peut montrer que :

THEOREME. *Si* $\lambda > 2\sigma$, *alors* $\mathscr{F}_\lambda$ *est presque-périodique au sens de Besicovitch sur* $\mathbb{Z}$.

N.B. : Il n'y a pas d'exposant autre que 2 de mentionné.

2 – a – Les résultats précédents sont appliqués à des fonctions classiques comme

$$S_\alpha(n) = \sum_{d|n} d^\alpha, \quad \alpha \in \mathbb{C}, \quad \frac{\varphi(n)}{n}, \quad n, \quad \mu(n).$$

Ici, les calculs effectifs sont développés, et le lecteur perspicace remarquera que toutes ces fonctions sont multiplicatives, ce qui rend possible le développement des sus–dits calculs, et l'application des résultats de II–B–a,b,c.

b – La partie relative aux entiers se clôt avec un très surprenant résultat, qui se déduit de α et β :

α – Pour que $\sum_{k=1}^{+\infty} A_I(k)\, C_k(n)$ converge absolument vers $\mathscr{F}_I(n)$, il suffit que I soit tel qu'il existe un α, $0 < \alpha < 1$, pour lequel $\sum_{k=1}^{+\infty} \frac{|F(k)|}{I(k)^\alpha} < +\infty$.

β – Si $\sum_{k=1}^{+\infty} \frac{|F(k)|}{I(k)^{\frac{1}{2}}} < +\infty$, alors $\mathscr{F}_I$ est presque–périodique au sens de Besicovitch pour l'exposant 2.

Le résultat surprenant, à première vue, est le suivant :

THEOREME. *Pour toute fonction arithmétique partout finie, il existe au moins un isomorphisme I tel que*

$$f(n) = \mathscr{F}_I(I(n)),$$

où $\mathscr{F}_I$ est presque–périodique au sens de Besicovitch et sa série de Fourier converge absolument, n appartenant à $\mathbb{N}^*$.

(La preuve revient à montrer l'existence de l'isomorphisme).

N.B. 1– Ce résultat a souvent été interprété de travers, la notion de finitude étant transformée en celle d'existence d'une borne uniforme.

2– Dans un article de 1986, A. Fuchs [5] retrouve, indépendamment comme le montre sa bibliographie, quelques–uns des résultats de Delsarte, en particulier, le précédent dans l'interprétation *borne uniforme*, par des méthodes similaires quoique simplifiées.

c – Cas des corps de nombres.

Delsarte étend alors les résultats précédents à des corps de nombres.

Soit K un *K*orps,

Z_K l'anneau des entiers,

I l'ensemble des *I*déaux entiers,

$N(.)$ la Norme d'un tel idéal.

Si $\underline{a}$ est un idéal entier, on considère $Z_K/\underline{a}$, qui est un groupe fini. On considère les caractères primitifs χ sur ce groupe additif, i.e. : les χ vérifiant $\chi^{-1}(1) = \underline{a}$, et l'on note leur nombre $\psi(\underline{a})$; par dénombrement, on a

$$\sum_{\underline{b} \mid \underline{a}} \psi(\underline{b}) = N(\underline{a}),$$

et par conséquent, la formule de Möbius donne

$$\psi(\underline{a}) = \varphi(\underline{a}),$$

où φ est l'indicateur d'Euler sur K.
On définit $C_{\underline{a}}(\alpha)$, $\alpha \in Z_K$, par :

$$C_{\underline{a}}(\alpha) = \sum_{\chi} \chi(\alpha),$$

où χ décrit l'ensemble des caractères primitifs mod $\underline{a}$. On a encore :

$$\sum_{\underline{b} \mid \underline{a}} C_{\underline{b}}(\alpha) = \begin{matrix} 0 & \text{si } \underline{a} \mid (\alpha) \\ N(\underline{a}) & \text{si } \underline{a} \mid (\alpha) \end{matrix}.$$

Par inversion, ceci donne :

$$C_{\underline{a}}(\alpha) = \sum_{\underline{d} \mid ((\alpha), \underline{a})} \mu\left(\frac{\underline{a}}{\underline{d}}\right) N(\underline{d}),$$

d'où l'on déduit des formules comme

$$C_{\underline{a}}(\alpha) \times C_{\underline{b}}(\alpha) = C_{\underline{a}.\underline{b}}(\alpha) \text{ si } (\underline{a}, \underline{b}) = 1 \text{ etc...}$$

On prolonge alors $C_{\underline{a}}$ à I tout entier par la formule

$$C_{\underline{a}}(\underline{b}) = \sum_{\underline{d} \mid (\underline{a},\underline{b})} N(\underline{d}).\mu(\frac{\underline{a}}{\underline{b}}).$$

On peut alors étendre sans difficultés particulières les résultats obtenus dans le cas des entiers ordinaires, ce que l'on ne vas pas faire ici.

En fin de cet article, Delsarte pose deux questions qui reviennent en fait à s'interroger sur les structures dissimulées derrière les constructions précédentes, et sur leur sens réel. C'est là l'origine des commentaires qui suivent et qui constituent la partie originale de cet exposé.

III – Commentaires.

1 – Sur la construction des *sommes de Ramanujan.*

On remarque que la construction des *sommes de Ramanujan*, que ce soit dans le cas des entiers ou celui d'un corps de nombres, est exactement la même. En effet, si l'on note :

P un idéal premier de K

K_{P} le complété P–adique de K

Z_{P} l'anneau des entiers de K_{P} , considéré comme groupe compact,

un caractère primitif de $Z_{\mathrm{K}}/\underline{a}$ s'identifie à un caractère primitif de

$$\prod_P (Z_{\mathrm{P}}/_{P^{v_{\mathrm{P}}(\underline{a})}Z_{\mathrm{P}}}),$$

où $v_{\mathrm{P}}(\underline{a})$ est l'exposant de P dans $\underline{a}$. Comme tous les caractères primitifs de $Z_{\mathrm{P}}/_{P^{v_{\mathrm{P}}(\underline{a})}Z_{\mathrm{P}}}$ se déduisent les uns des autres par les automorphismes de la forme $x \longmapsto h.x$ où $v_{\mathrm{P}}(h) = 0$, et $h \in Z_{\mathrm{P}}/_{P^{v_{\mathrm{P}}(\underline{a})}Z_{\mathrm{P}}}$, la construction des *sommes de Ramanujan* revient à celle d'une base de fonctions continues à valeurs réelles sur $\prod_P Z_{\mathrm{P}}$, invariantes par $\prod_P Z_{\mathrm{P}}^*$, où Z_{P}^* est le groupe des unités de Z_{P}, qui seront donc identifiées à des fonctions sur $\prod_P (Z_{\mathrm{P}}/Z_{\mathrm{P}}^*)$ considéré comme espace compact.

On va donc pouvoir rester dans le cas des entiers ordinaires sans que la généralité de l'exposé en souffre, les structures mises en oeuvre étant les mêmes.

2 — **Sur la construction de la série de Fourier.**

La méthode de Delsarte est très astucieuse; en effet, la *raison* pour laquelle seules les *sommes de Ramanujan* vont apparaître dans les séries de Fourier est essentiellement le choix effectué qui consiste à garder la *partie principale* dans le calcul formel des coefficients, ce qui revient à projeter une suite presque–périodique dans l'espace des suites limite–périodiques invariantes sous l'action de $\prod_P \mathbb{Z}_p^*$.

Signalons que l'on peut toujours décomposer une suite presque–périodique f en trois morceaux f_1, f_2, f_3 correspondant à des projections orthogonales deux à deux sur le groupe associé à f, et écrire :

$$f = f_1 + f_2 + f_3,$$

où le spectre de f_1 est irrationnel, celui de f_2 et f_3 est rationnel, mais f_2 n'a aucune composante invariante par $\prod_P \mathbb{Z}_p^*$, et f_3 est invariante par ce groupe, chacune de ces fonctions étant aussi presque–périodique au sens où f l'est, avec un même exposant que f dans le cas de la presque périodicité au sens de Besicovitch [6].

3 — **Sur la généralisation des constructions;** (on notera $*$ la convolution des fonctions arithmétiques). On commence comme Delsarte et l'on pose :

$$A_\lambda(k) = \sum_{k \mid d^\lambda} \frac{F(d)}{d^\lambda}, \quad \lambda \in \mathbb{N}^*.$$

Alors, formellement, on a :

$$\sum_{k=1}^{+\infty} A_\lambda(k)\, C_k(n) = \sum_{d^\lambda \mid n} F(d) = \mathcal{F}([n]_\lambda).$$

Plutôt que de nous lancer frénétiquement dans des calculs, regardons la série de Dirichlet associée à $\mathcal{F}([n]_\lambda)$. On a, formellement,

$$\sum_{n=1}^{+\infty} \frac{\mathscr{F}([n]_\lambda)}{n^s} = \sum_{n=1}^{+\infty} \Big(\sum_{d^\lambda \mid n} F(d) \Big) \times \frac{1}{n^s}$$

$$= \sum_{d=1}^{+\infty} F(d) \times \sum_{d^\lambda \mid n} \frac{1}{n^s}$$

$$= \zeta(s) \times \sum_{d=1}^{+\infty} \frac{F(d)}{d^{\lambda s}}$$

$$= \zeta(s) \times \sum_{d=1}^{+\infty} \frac{f * \mu(d)}{d^{\lambda s}}$$

$$= \zeta(s) \times \zeta(\lambda s)^{-1} \times \sum_{d=1}^{+\infty} \frac{f(d)}{d^{\lambda s}}.$$

On en déduit que l'*image de Möbius* $\Phi_\lambda(n)$ de $\mathscr{A}[n]_\lambda)$ a pour série de Dirichlet $\zeta^{-1}(\lambda s) \sum_{n=1}^{+\infty} \frac{f(n)}{n^{\lambda s}}$.

Si nous considérons un isomorphisme I général, on voit que la série de Dirichlet associée à $\mathscr{F}_I(n)$ s'écrit :

$$\sum_{n=1}^{+\infty} \frac{\mathscr{F}_I(n)}{n^s} = \sum_{n=1}^{+\infty} \frac{1}{n^s} \times \Big(\sum_{I(d) \mid n} F(d) \Big)$$

$$= \sum_{d=1}^{+\infty} F(d) \times \sum_{I(d) \mid n} \frac{1}{n^s}$$

$$= \zeta(s) \times \sum_{d=1}^{+\infty} \frac{F(d)}{I(d)^s}$$

$$= \zeta(s) \times \sum_{d=1}^{+\infty} \frac{f * \mu(d)}{I(d)^s}.$$

Par conséquent, l'*image de Möbius* $\Phi_I(n)$ de $\mathscr{F}_I(n)$ a pour série de Dirichlet

$$\sum_{n=1}^{+\infty} \frac{f * \mu(n)}{I(n)^s}.$$

Conclusion : L'introduction de l'isomorphisme pour la divisibilité permet de modifier très librement le rôle de la fonction initiale.

IV – **Sur les résultats généraux.**

a – Dans le B–1–a, où l'hypothèse est

$$\sum_{n=1}^{+\infty} |F(n)| < +\infty,$$

l'argument montre non seulement que f est presque–périodique au sens de Bohr, mais encore que f, étendue par continuité uniforme à $G = \prod_p \mathbb{Z}_p$, est dans l'algèbre $A(G)$ des fonctions continues sur G à série de Fourier absolument convergente. On peut en déduire, par le théorème de Wiener–Lévy [7], que si h est une fonction holomorphe dans un voisinage de $f(G)$, alors $h(f(t))$ a une série de Fourier absolument convergente et la suite $h(f(n)$ a donc cette propriété.

b – Les cas B–1–b, B–1–c, B–1–e sont très restrictifs; il est facile de faire beaucoup mieux. En effet, Wintner [8] a établi que si :

$$\sum_{n=1}^{+\infty} \frac{|\mu * f(n)|}{n} d(n) < +\infty ,$$

où $d(n)$ est le nombre de diviseurs de n, alors :

i) f est Besicovitch–presque–périodique

ii) la série de Fourier de f s'écrit

$$\sum_{q=1}^{+\infty} A_q \, C_q(n)$$

avec :

$$a_q = \sum_{m=1}^{+\infty} \frac{f * \mu(mq)}{mq}$$

iii) cette série converge absolument vers $f(n)$, pour tout n de $\mathbb{N}^*$.

N.B. : Delange [9] a raffiné sur ce thème en montrant que les conclusions restent vraies sous l'hypothèse plus faible

$$\sum_{n=1}^{+\infty} \frac{|f*\mu(n)|}{n} 2^{\omega(n)},$$

où $\omega(n)$ est le nombre de diviseurs premiers de n.

L'utilisation de l'un ou l'autre de ces résultats permet d'améliorer ceux obtenus par Delsarte. En effet, par exemple, on a d'après III–3,

$$\sum_{n=1}^{+\infty} \frac{\mathcal{F}_I * \mu(n)}{n^s} = \sum_{n=1}^{+\infty} \frac{F(d)}{I(d)^s}$$

et par conséquent, si

$$\sum_{d=1}^{+\infty} \frac{|F(d)| 2^{\omega(d)}}{I(d)} < +\infty ,$$

c'est gagné. Signalons que les conditions données par Delsarte entraînent, toutes, celles données par Wintner ou Delange dans [8] et [9].

V – Conclusion.

L'approche de Delsarte montre qu'étant donnée une suite quelconque, il existe toujours une suite f presque–périodique au sens de Besicovitch à série de Fourier absolument convergente *par blocs* telle que la suite figure dans l'ensemble des valeurs prises par f sur $\mathbb{N}^*$. Pour ce faire, on envoie les points *assez loin* dans $\prod_p (\mathbb{Z}_p/\mathbb{Z}_p^*)$, au moyen d'un *isomorphisme*, si besoin est, de façon à ce qu'ils soient privés de tout rôle, (quitte à ce qu'ils se retrouvent dans un ensemble de densité nulle). Bien entendu, on ne peut tirer de là aucune information sur leur distribution, ou sur une quelconque propriété de type probabiliste, même si on sait qu'ils sont valeurs d'une fonction presque–périodique au sens de Besicovitch.

Pour finir, signalons que les constructions précédentes se font intrinsèquement dans n'importe quel système de nombres A de Beurling vérifiant

$$\sum_{\substack{N(a)\leq x \\ a\in A}} 1 = L\,x + o(x),$$

où L est une constante strictement positive, $N(.)$ étant la norme définie sur le semi–groupe.

BIBLIOGRAPHIE

[1] **J. Delsarte.**– Essai sur l'application de la théorie des fonctions presque–périodiques à l'Arithmétique, Ann. Ec. Norm., 3, LXII, 185–204.

[2] **A. Weil.**– Oeuvres Scientifiques, Vol. 3 *L'oeuvre Mathématique de Delsarte*, Springer 1979, 229–247.

[3] **P. Erdõs et A. Wintner.**– Additive functions and almost periodicity (B^2), Amer. J. Math., 62, (1940), 635–645.

[4] **M. Kac, E.R. Van Kampen and A. Wintner.**– Ramanujan sums and almost periodic functions, Amer. J. Math., 62, (194à), 107–114.

[5] **A. Fuchs.**– Contribution à l'étude des densités asymptotique et analytique, Rend. Accad. Naz. Sci. Detta 40, V ser. Mem. Mat. 10 N° 1, (1986), 127–140.

[6] **J.–L. Mauclaire.**– Integration and Number Theory, Proc. Prospects of Math. Sci., World Sci. Pub., (1988), 97–125.

[7] **I. Katznelson.**– An introduction to Harmonic Analysis, J. Wiley and S. New York, (1968).

[8] **A. Wintner.**– Eratosthenian averages, Waverly press. 1943.

[9] **H. Delange.**– On Ramanujan expansions of certain arithmetical functions, Acta Arithmetica XXXI, (1976), 259–270.

J.–L. MAUCLAIRE
U.A. 212
Département de Mathématique–Informatique
Université Paris VII
2, place Jussieu
75251 PARIS

THE DISTRIBUTION OF PRIMES SATISFYING THE CONDITION $a^{p-1} \equiv 1 \pmod{p^2}$

(résumé)

Léo MURATA

For any fixed natural number $a \geq 2$, it is quite well known that a rational prime p with $(a,p) = 1$ satisfies the relation $a^{p-1} \equiv 1 \pmod{p}$ (Fermat's little theorem). But we know only little about the distribution of prime p which satisfies

$$(*) \qquad a^{p-1} \equiv 1 \pmod{p^2}.$$

To investigate these primes is an interesting problem not only from the view point of distribution of primes but also from the point of Fermat's last theorem. We know that (see, for example [4]) :

THEOREM A. *Let p be an odd prime. If (*) fails to hold at least one prime value of $a \leq 89$, then the first case of Fermat's last theorem is true for p, that is*

$$x^p + y^p = z^p, \ (xyz,p) = 1$$

has no non-trivial integral solution.

Now we introduce a few notations : for two distinct natural numbers a and b, we put

$$L_a(x) = \{p : p \text{ is an odd prime } \leq x, \ a^{p-1} \equiv 1 \pmod{p^2}\},$$
$$L_{a,b}(x) = L_a(x) \cap L_b(x).$$

Here we discuss about $|L_a(x)|$ and $|L_{a,b}(x)|$.

By means of electric computer, we have some numerical examples (see for example [1]), and these show that such primes exist very rarely :

$$L_2(3\times 10^9) = \{1093,3511\},$$
$$L_3(2^{30}) = \{11, 1006003\},$$
$$L_{12}(2^{28}) = \{2693, 123653\}, \text{ etc...}$$

From these numerical examples, it seems that there is not a big difference between the two cases, when a is equal to a prime number and when a is not equal to a prime. So, we consider an average of $|L_a(x)|$ for all natural values of a with $2 \leq a \leq x^4$. Our first theorem ([3]) states that, roughly speaking, $|L_a(x)|$ can be *normally* approximated by $\operatorname{loglog} x$:

THEOREM 1. *Let δ be an arbitrary fixed real number satisfying $\frac{1}{2} < \delta < 1$. We have,*

$$|L_a(x)| = \operatorname{loglog} x + \theta((\operatorname{loglog} x)^{\delta}) + O(1)$$

for all a such that $2 \leq a \leq x^4$ with at most $2x^4 (\operatorname{loglog} x)^{1-2\delta}$ exceptions of a, where, $f(x)$ being any positive valued function of x, $\theta(f(x))$ denotes a function of x with absolute value $\leq f(x)$.

The above-cited theorem A states that Fermat's last theorem Case I is true outside of the intersection of several $L_a(x)$'s, so it is reasonable to consider the behaviour of $|L_{a,b}(x)|$. Since we are motivated by Theorem A, we now have to limit a and b to be prime numbers. Then we have the following average type theorem :

THEOREM 2. *Let x be a sufficiently large real number, ϵ be a fixed real positive number, and we put*

$$\mathfrak{A} = \{(s,t); 2 < s < t < x^{2+\epsilon}, \ s \text{ and } t \text{ are primes}\}$$
$$\mathfrak{F}_0 = \{(s,t) \in \mathfrak{A} ; |L_{s,t}(x)| = 0\}.$$

Then we have

$$\lim_{x \to \infty} \frac{|\mathfrak{F}_0|}{|\mathfrak{A}|} = \frac{8}{\pi^2}.$$

From these two results, we can know only a tendency of $|L_a(x)|$ or $|L_{a,b}(x)|$ and cannot say anything about $|L_a(x)|$ with fixed a. In the next theorem, we give a result which gives an information for $|L_a(x)|$ with fixed a.

It is easy to see that, for an odd prime number p, the following two conditions are equivalent :

(i) $a^{p-1} \equiv 1 \pmod{p^2}$,

(ii) $[(Z/p^2 Z)^* : <a \pmod{p^2}>]$ is divisible by p, where $(Z/p^2 Z)^*$ denotes the multiplicative group consists of all inversible residue classes modulo p^2, $<a \pmod{p^2}>$ the cyclic subgroup generated by the residue class $a \pmod{p^2}$, and $[\ :\ ]$ the index of the subgroup.

So, when p is an element of $L_a(x)$, we can write

$$[(Z/p^2 Z)^* : <a \pmod{p^2}>] = p \times m_p, \quad m_p \text{ divides } p-1.$$

And we divide $L_a(x)$ into two parts :

$$L_a^{(1)}(x) = \{p : p \in L_a(x) \text{ and } m_p < \log\log x\},$$
$$L_a^{(2)}(x) = \{p : p \in L_a(x) \text{ and } m_p \geq \log\log x\}.$$

Concerning the second part, we have :

THEOREM 3. *Let a be a square free positive integer ≥ 2. We assume the truth of all Generalized Riemann Hypothesis, then we have*

$$|L_a^{(2)}(x)| = O(\pi(x) \frac{1}{\log_2 x})$$

where $\log_i$ denotes the i-th iteration of logarithmic function.

This result can be deduced from the following Theorem 4 which is a generalization of Hooley's work on Artin's conjecture for primitive roots.

THEOREM 4. *Let a be a square free natural number ≥ 2, $n \in N$, and we put*

$$N_a^{(n)}(x) = \{p : p \text{ is a prime number} \leq x,\ [(Z/pZ)^* : <a \pmod{p}>] = n\}.$$

Under the assumption of Generatized Riemann Hypothesis, we have

$$|N_a^{(n)}(x)| = C_a^{(n)} Li(x) + O((n \log_2 x + \log a) \frac{x}{(\log x)^2}),$$

where the values of the leading coefficients $C_a^{(n)}$ ***can be calculated exactly. Furthermore, for any positive real number*** z,

$$(**) \qquad \sum_{n \leq z} C_a^{(n)} = 1 + O(\frac{1}{z}).$$

It is rather easy to derive Theorem 3 from Theorem 4.

We make remark that, for an odd prime number p, if p satisfies the condition $[(Z/pZ)^* : <a \pmod p>] = n$, then $[(Z/p^2 Z)^* : <a \pmod{p^2}>] = n$ or np. Therefore

$$|L_a^{(2)}(x)| \leq |\{p \leq x ; [(Z/pZ)^* : <a \pmod p>] \geq \log_2 x\}|,$$

and Theorem 4 gives the result of Theorem 3.

Theorem 3 gives a non-trivial upper bound for the second part of $|L_a(x)|$, but we do not succeed in obtaining any non-trivial upper bound for the first part until now. And probably, this might be very difficult problem.

I am grateful to Gérald Tenenbaum for his comment which enabled me to improve my original statement in (**).

BIBLIOGRAPHY

[1] J. Brillhart, J. Tonascia, P. Weinberger.– On the Fermat Quotient, Atkin–Birch (ed.) *Computers in Number Theory*, (1971) 213–222.

[2] C. Hooley.– On Artin's Conjecture, J. reine angew. Math., Vol. 225, (1967) 209–220.

[3] L. Murata.– An average Type Result on the Number of Primes Satisfying Generalized Wieferich Condition, Proc. Japan Acad., Vol. 57, (1981), 430–432.

[4] P. Ribenboim.– 13 Lectures on Fermat's Last Theorem, Springer–Verlag (1979).

Léo MURATA
Département de Mathématiques
Université de Nancy I
B.P. 239
54506 Vandoeuvre les Nancy Cedex

Meijigakuin University, Yokohama
Japon

ON A PROBABILISTIC METHODE IN ADDITIVE NUMBER THEORY

Imre Z. RUZSA(*)

1 – What is a random set ?

Probabilistic methods are widely applied in number theory. Their usage can be roughly split into two parts : properties of *random numbers*. The first is what is generally called *probabilistic number theory*, and it centers around multiplicative properties (like the number of prime divisors). The second, which we shall treat here, is not so systematized; its typical questions arise about additive properties of random sets. We shall consider only one such property in detail, that of being a basis.

Before doing that, we try to (partially) answer, or at least call the attention to three questions :

– what is a random set;

– why are random sets interesting; and

– what is the difference between a random and a deterministic construction.

We take a sequence (p_n) of real numbers, $0 \leq p_n \leq 1$. We build a random subset A of integers as follows : we select each n into A independently with probability

$$P(n \in A) = p_n$$

(think of throwing a coin). More exactly, let (ξ_n) be a sequence of independent $0-1$ valued random variables with the distribution

$$P(\xi_n = 1) = p_n ;$$

(*) Partially supported by the Hungarian National Foundation for Scientific Research Grant N° 1811

we define

$$A = \{n : \xi_n = 1\}.$$

We use $A(x)$ to denote the number of elements of A up to x. By the strong law of large numbers, whenever

$$\sum_{n=1}^{\infty} p_n = \infty ,$$

we have with probability 1

$$A(x) = \sum_{n \leq x} \xi_n \sim \sum_{n \leq x} p_n. \tag{1.1}$$

This A is now a random set; if we wanted to be very precise, we could write $A(\omega)$ to indicate the dependence on ω, an element of the underlying probability space. Instead of claiming that, say, A is a basis of order 2, we can only calculate the probability that A will be a basis of order 2. In most cases, the result will be 0 or 1.

As a limiting case, random sets include concrete sets. Say, if we put $p_n = 1$ if n is a prime, $p_n = 0$ if n is composite, then with probability 1, A will be the set of primes. To exclude this, let us tentatively call a random construction *purely probabilistic*, if p_n is a monotonic function of n.

Purely probabilistic constructions are quite often insufficient, and we have to include some numbers deterministically; or we may have a preconception how the numbers must look, say that they be all squares. In this case we can put, say, $p_n = 0$ if n is not a square, and some monotonic function of n if n is a square. In this way we can get a random set of squares, a sort of *semi-random* or *semi-probabilistic* construction.

The main importance of random constructions is that in this way we can often show the existence of sets which we cannot construct. Say, the thinnest additive complement to the set of primes (that is, a set A with the property that every large integer can be written in the form $a + p$, $a \in A$, p prime) was obtained probabilistically (Erdös, 1954). Similarly, the thinnest possible essential component is a random one (Ruzsa, 1987). The point is that our constructions tend to show regularities, and if a very irregular set is needed, then the best idea may be to turn

to a random set. It may also happen that a concrete set likely has the desired properties, just we cannot prove it, while a proof can be carried through for a similar random set. Say, the behaviour of the primes is remarkably similar to that of a random set with $p_n = 1/\log n$; and lots of properties that are conjectured for primes are easily established for this random set (like some forms of the Riemann hypothesis).

The border between random and deterministic construction is a rather fuzzy one. Let us illustrate this by an example. One can show that there is a set A of *primes* such that

$$A(x) = 0((x \log x)^{\frac{1}{2}}) \tag{1.2}$$

and $A + A$, the set of elements of the form $a_1 + a_2$, a_1, $a_2 \in A$, contains almost all even numbers. This is done by a probabilistic construction, by putting

$$p_n = c(\log n)^{3/2} n^{-1/2}$$

for prime n.

Alternatively, we could do the following. We show that with suitable constants n_0, c_1, c_2 for every $n > n_0$ there is a set $B \subset [n,4n]$ of primes such that

$$|B| < c_1(n \log n)^{\frac{1}{2}}, \tag{1.3}$$

and every even integer between $3n$ and $6n$ is the sum of two elements of B, with at most $c_2 n/(\log n)$ exceptions. This can be done by some random, or counting, argument. Apply it with $n = 2^k$, and from the resulting sets B let B_k be the first in some alphabetic ordering. This is a uniquely determined set, and consequently $A = \cup B_k$ is a deterministic example satisfying (1.2) and the required additive property. Of course, one feels that this is a sort of cheating.

For me, a real construction is when I *see* the set; or when to decide whether $n \in A$, we consider n separately, not together with other numbers. It looks like impossible to grasp this exactly. One could also argue that the time required to compute the set is a distinctive feature, and it deserves the name *construction* only if it goes fast. Personally I do not find this approach very charming either.

2 – Random bases

Let A be a set of nonnegative intergers. We use Ak to denote the set $A +...+ A$, that is, the set of all numbers in the form

$$a_1 +...+ a_k, \quad a_1,...,a_k \in A.$$

Recall that A is a *basis* of order k if all positive integers are in Ak; it is an *asymptotical basis* of order k, if all but finitely many positive integers are in Ak.

Let $A[k]$ be the set of integers representable as a sum of k *different* elements of A (this will be a technical convenience over the sums with no necessarily different terms).

Take a sequence (p_n) of real numbers, $0 \leq p_n \leq 1$. In the sequel A will be a random set of integers built with this sequence of probabilities, that is

$$P(n \in A) = p_n,$$

as described in the previous section. Put

$$\alpha_n = P(n \notin Ak). \tag{2.1}$$

$$\beta_n = P(n \notin A[k]). \tag{2.2}$$

We shall treat mainly $A[k]$ and β_n; in most cases, similar assertions will hold for Ak and α_n.

In terms of the ξ_n, let

$$\zeta_n = \Sigma\, \xi_{x_1} \cdots \xi_{x_k}$$

with the summation running over those k–tuples of subscripts that satisfy

$$x_1 +...+ x_k = n, \quad x_1 <...< x_k. \tag{2.3}$$

ζ_n is nothing else than the number of solutions of (2.3) with $x_i \in A$. We have now

$$\beta_n = P(\zeta_n = 0).$$

Let B be a fixed set of natural numbers. If

$$\sum_{n\in B} \beta_n < \infty \,, \tag{2.4}$$

then by the Borel–Cantelli lemma, with probability 1 only a finite number of the events $n\in B$, $n\notin A[k]$ occurs. In particular, if $B = \{\text{all natural numbers}\}$, this means that A is an asymptotic basis of order k with probability 1.

2.1.– Problem. Assume that $\sum_B \beta_n = \infty$ (or $\sum_B \alpha_n = \infty$). Does it follow that with probability 1 the set $B\setminus A[k]$, or, respectively, the set $B\setminus Ak$ is infinite?

If the events $n\in A[k]$ were independent, then condition (2.4) would be necessary as well. We show that they are always positively correlated, that is, dependent in the *wrong* direction.

2.2.– THEOREM. ***Let*** A ***be a random set of integers,*** $B_1,\ldots,B_m$ ***fixed finite sets. We have***

$$P(\cup B_i \subset A[k]) \geq \Pi P(B_i \subset A[k]), \tag{2.5}$$

$$P(\cup B_i \cap A[k]=\emptyset) \geq \Pi P(B_i \cap A[k]=\emptyset). \tag{2.6}$$

The analogous inequalities also hold for Ak ***in the place of*** $A[k]$.

In fact, these are part of a more general phenomenon. For a fixed set B, the condition $B\subset Ak$ defines a class of sets A; and the only important feature is that this class is *monotonic*, if a set A belongs to it then so does every set containing A.

2.3. DEFINITION. ***Let*** T ***be a class of subsets of a fixed set*** S. ***We call*** T ***increasing, if*** $A\subset B\subset S$, $A\in T$ ***implies*** $B\in T$; ***decreasing, if*** $A\subset B\subset S$, $B\in T$ ***implies*** $A\in T$.

2.4. THEOREM. ***Let*** $T_1,\ldots,T_k$ ***be classes of subsets of the set*** $\{1,2,\ldots,n\}$, ***either all increasing or all decreasing. Let*** A ***be a random subet of*** $\{1,2,\ldots,n\}$. ***We have***

$$P(A\in \cap\, T_i) \geq \Pi P(A\in T_i). \tag{2.7}$$

This is a result from Ruzsa (1976), where it is formulated in terms of the random variables ξ_j.

2.5. COROLLARY. *Let* A *be a random set of integers,* B *a fixed set. We have*

$$P(B \subset Ak) \geq \prod_{j=1}^{\infty} (1-\alpha_j), \tag{2.8}$$

$$P(B \subset A[k]) \geq \prod_{j=1}^{\infty} (1-\beta_j). \tag{2.9}$$

In particular, this shows that the events $Ak \supset B$ or $A[k] \supset B$ have a positive probability if $\sum_{n \in B} \alpha_n < \infty$ or $\sum_{n \in B} \beta_n < \infty$, respectively and no α_n (or β_n) is $=1$. This can also easily be deduced from the Borel–Cantelli lemma.

The next question is, how to calculate or estimate β_n. The event $n \in A[k]$ is the union of the events $Y_{\underline{x}}$, associated with each sequence $\underline{x} = (x_1,\ldots,x_k)$ of solutions of (2.3), which happen if $x_j \in A$ for all $1 \leq j \leq k$. We have clearly

$$P(Y_{\underline{x}}) = \prod p_{x_j}.$$

If $k = 2$, any number $0 \leq x \leq n$ is appearing in at most one vector $\underline{x}$ (in exactly one unless $n = 2x$), thus the events $Y_{\underline{x}}$ are independent and we conclude

$$\beta_n = \prod_{j < n/2} (1 - p_j p_{n-j}). \qquad (k = 2) \tag{2.10}$$

If $k > 2$, they are not independent, and the analogous equality does not hold. But, since the events $Y_{\underline{x}}$ arise from increasing classes, from Theorem 2.4 we can infer an inequality.

2.6. STATEMENT. *We have*

$$\beta_n \geq \prod (1 - p_{x_1} \ldots p_{x_k}), \tag{2.11}$$

with equality for $k = 2$, *where the product is taken over all solutions of* (2.3).

A similar inequality holds for α_n.

Now consider the case of a purely probabilistic set and $k = 2$. Try

$$p_n = c(\log n)\sqrt{n}. \tag{2.12}$$

A simple computation shows

$$\alpha_n \sim \beta_n \sim n^{-\rho c^2},$$

where

$$\rho = \int_0^{1/2} \frac{1}{(x(1-x))^{\frac{1}{2}}}\, dx = \pi/2.$$

Consequently for $c > \rho^{-1/2}$ A will be an asymptotical basis of order 2 with probability 1. For $c < \rho^{-1/2}$ we find $\Sigma \alpha_n = \Sigma \beta_n = \infty$. Indeed, in this case one can prove that with probability 1, A will not be an asymptotical basis. For this one has to establish the *quasi-independence* of the events $n \in Ak$.

Observe that under (2.12) one has with probability 1

$$A(x) \sim 2c(\log x)\sqrt{x}.$$

It is also well known that the thinnest possible bases have

$$A(x) = 0(\sqrt{x}),$$

(see Halberstam–Roth, 1966), thus we did not break a record with our random set.

These random sets do have, however, some remarkable properties. Let $r_A(n)$ denote the number of solutions of $n = a + a'$, a, $a' \in A$. Erdös (1956) established that for $c > c_0$ the above random set satisfies

$$c_1 \log n < r_A(n) < c_2 \log n \tag{2.13}$$

with positive constants c_1, c_2 for all $n > n_0(A)$ with probability 1. So far no concrete set is known with this property, neither is the answer known to any of the following related problems.

2.7. **Problem.** Is there a set A for which

$$r_A(n)/\log n \longrightarrow c \neq 0\ ?$$

2.8. Problem. Is there a set A for which r_A has such a stability property but with a smaller order of magnitude, say which satisfies (2.13) with the $\log n$ replaced by

$$(\log n)/\log\log n\ ?$$

I suggest the answer to both problems would be negative, thus random sets are from this aspect optimal.

3.– Semirandom bases

We start with a set S of nonnegative integers, a fixed integer $k>2$ and another set B of integers satisfying $B\subset Sk$. We take a sequence (q_n) of real numbers, $0\le q_n\le 1$. We build a random subset A of S as follows : we select every $n\in S$ into A independently with probability

$$P(n\in A)=p_n=\begin{cases}0 & \text{if } n\notin S\\ q_n & \text{if } n\in S\end{cases}.$$

The problem is, how to estimate β_n.

For our set S, let $r_k(n)$ denote the number of solutions of the equation

$$n=x_1+\ldots+x_k\ ,\quad x_1<\ldots<x_k\ ,\quad x_1,\ldots,x_k\in S. \tag{3.1}$$

Let $R_k(n)$ be the number of solutions without the requirement $x_1<\ldots<x_k$. Typically R_k is studied rather than r_k, but results on R_k permit us to estimate r_k. For instance, let $S_k(n)$ be the number of solutions of the equation

$$n=x_1+\ldots+x_{k-1}+2x_k\ ,\qquad x_1,\ldots,x_k\in S.$$

We have

$$r_k(n)\ge\frac{1}{k!}R_k(n)-\frac{1}{2(k-2)!}S_{k-1}(n). \tag{3.2}$$

The case $k=2$ is much simpler than that of greater k. β_n can be estimated by (2.10). This was applied by Erdös and Nathanson (1981) to extract a thin basis of order 2 from the set of integers with at most 2 prime factors, and essentially it dates back to Erdös (1956).

Assume that the sequence (p_n) is decreasing. By (2.10) we have

$$\beta_n \leq (1-p_n^2)^{r_2(n)} \leq \exp - p_n^2 r_2(n). \tag{3.3}$$

THEOREM 3.1. *Let (p_n) be a decreasing sequence of real numbers, $0 \leq p_n \leq 1$. Let S, B be sets of integers, and let A be a random subset of S, satisfying*

$$P(n \in A) = p_n \qquad (n \in S).$$

If for some $c > 1$ we have

$$p_n > c\left(\frac{\log n}{r_2(n)}\right)^{\frac{1}{2}} \tag{3.4}$$

for all $n > n_0$, $n \in B$, where $r_2(n)$ is defined above, then with probability 1 the set $A[2]$ contains all but a finite number of elements of B.

Proof: In view of (3.3), (3.4) yields $\beta_n < n^{-c}$ for $n > n_0$, thus $\Sigma \beta_n < \infty$. □

Assume now $k > 2$, fix n and write $r = r_k(n)$. The number n has r representations by elements of S, say

$$n = x_i^{(1)} + \ldots + x_i^{(k)}, \qquad i = 1,\ldots,r.$$

Let X_i $(i = 1,\ldots,r)$ be the event

$$x_i^{(j)} \in A \text{ for every } 1 \leq j \leq k.$$

In contrast to the case $k = 2$, the representation are generally not disjoint, hence the events X_i are not independent. Say, if we consider the representations of integers by three primes, then any representation of an even number must begin with 2. This observation indicates that, in contrast to (3.3), the assumption that r is big is alone insufficient to yield an estimate of β_n. Our auxiliary assumption will be that r_t is small for $t < k$, which will exclude these degenerate cases.

Erdös and Nathanson (1981) estimate β_n by selecting an independent subset of the events X_i. This yields a sharp result for three squares. Its applicability is, however, limited by the fact that no matter how big r is, at most n/k of the representations can be disjoint.

For simplicity, first we consider the case when all the numbers p_j are equal,

$$p_0 = p_1 = \ldots = p_n = p.$$

Then clearly we have

$$P(X_i) = p^k.$$

If the events X_i were independent, then we would have

$$\beta_n = \Pi\,(1 - P(X_i)) = (1 - p^k)^r. \tag{3.5}$$

This becomes smaller than $1 - \epsilon$ at about $p = r^{-1/k}$, and smaller than $1/n$ to ensure convergence at about $(r^{-1}\log n)^{1/k}$.

STATEMENT 3.2. *Let S, B, k, n, r be as above and let A be a random subset of S with*

$$P(m \in A) = p_m = p = r^{-1/k} \qquad (m \in S,\ m \leq n).$$

Assume that with a constant $C > 0$ we have for all $2 \leq t \leq k-1$ either

$$\sum_{m \leq n} r_t(m)^2 r_{k-t}(n-m) \leq C r^{1+t/k}, \tag{3.6}$$

or

$$\binom{k}{t} \max_{m \leq n} r_t(m) \leq C r^{t/k}. \tag{3.7}$$

Then we have

$$\beta_n \leq 1 - \frac{1}{2 + C(k-2)}. \tag{3.8}$$

Remark : (3.7) is *almost* stronger than (3.6); from (3.7) one can deduce (3.6) with $r^* = r + s_k(n)$ in the place of r.

We do not give a detailed proof of this statement. It is based on a second–moment method. If there are s numbers that occur both as an $x_i^{(\,)}$ and an $x_j^{(\,)}$, then we have

$$P(X_i \cap X_j) = p^{2k-s}.$$

Write $t = k - s$ and let L_t denote the number of pairs (i,j) such that in the i'th and j'th representations there are s common terms and t *individual* terms in each. We obtain

(3.9) $$P(\cup X_i) = 1 - \beta_n \geq \frac{r^2 p^{2k}}{\sum L_t p^{k+t}}.$$

We estimate L_t by the numbers $r_j(m)$ to arrive finally at (3.8).

This method gives a poor estimate for larger p. Even if we knew $L_t = 0$ for $1 \leq t \leq k-1$, for $p = vr^{-1/k}$ (3.9) only yields

$$\beta_n \leq (1 + v^k)^{-1}.$$

We use another method to show that β_n decreases at least exponentially. (3.5) would give an estimate of order $\exp -v^k$. We are unable to go as far, but we shall obtain $\exp -cv$.

Let $k \geq 2$ and $n \geq 1$ be fixed integers, $\underline{p} = (p_0,...,p_n)$ a sequence of $n+1$ real numbers, $0 \leq p_i \leq 1$. We define a random set $A \subset \{0,...,n\}$ by $P(k \in A) = p_k$. Let $\beta(\underline{p})$ denote the probability that the equation

$$n = x_1 +...+ x_k\,, \quad x_1 <...< x_k\,, \quad x_i \in A$$

has no solution (this is the β_n of the previous sections, but now we want to emphasize the dependence on the probabilities).

The novelty in my approach is the following sequence of simple inequalities. Exceptionally we include the detailed proofs.

LEMMA 3.3. *Assume that the sequences p, q, q' satisfy*

(3.10) $$p_j = q_j + q'_j - q_j q'_j$$

for all $j = 0,...,n$. Then we have

(3.11) $$\beta(\underline{p}) \leq \beta(\underline{q})\beta(\underline{q}').$$

Proof : Let (η_j), (η'_j), $j = 0,...,n$ be a collection of $2(n+1)$ independent $0-1$ valued random variables with the distribution

$$P(\eta_j = 1) = q_j, \quad P(\eta'_j = 1) = q'_j\,.$$

Put $\xi_j = \max(\eta_j, \eta'_j)$. Clearly

$$P(\zeta_j = 1) = p_j .$$

By definition,

(3.12) $$\xi_{x_1} \dots \xi_{x_n} = 0$$

implies that

$$\eta_{x_1} \dots \eta_{x_n} = \eta'_{x_1} \dots \eta'_{x_n} = 0.$$

Hence if S is the event that (3.12) holds whenever $n = x_1 + \dots + x_k$, $x_1 < \dots < x_k$, and T, T' are the corresponding events defined by the variables η_j and η'_j, then $S \subset T \cap T'$. By independence we infer

$$\beta(\underline{p}) = P(S) \leq P(T)P(T') = \beta(\underline{q})\beta(\underline{q}').$$ □

LEMMA 3.4. *If $p_j \geq q_j$ for all j, then we have*

$$\beta(\underline{p}) \leq \beta(\underline{q}).$$

Proof : If $p_j \geq q_j$, then we can find numbers q'_j satisfying (3.10), and then (3.11) yields

$$\beta(\underline{p}) \leq \beta(\underline{q})\beta(\underline{q}') \leq \beta(\underline{q}).$$ □

LEMMA 3.5. *Assume that the sequences $\underline{p}$, $\underline{q}$, $\underline{q}'$ satisfy*

(3.13) $$p_j \geq q_j + q'_j - q_j q_j$$

for all $j = 0,\dots,n$. Then we have

$$\beta(\underline{p}) \leq \beta(\underline{q})\beta(\underline{q}').$$

Proof : A combination of Lemmas 3.3 and 3.4. □

COROLLARY 3.6. *Let* $v \geq 1$ *be an integer. Assume that*

$$1 - p_j \leq (1 - q_j)^v \tag{3.14}$$

holds for all j. *Then*

$$\beta(\underline{p}) \leq \beta(\underline{q})^v.$$

Proof : Observe that (3.13) can be rewritten as

$$(1 - p_j) \leq (1 - q_j)(1 - q_j').$$

The Corollary now follows from Lemma 3.5 via a simple induction. □

COROLLARY 3.7. *If*

$$p_j \geq \min(1, vq_j) \tag{3.15}$$

for all j *with an integer* v, *then*

$$\beta(\underline{p}) \leq \beta(\underline{q})^v.$$

Proof : (3.15) implies (3.14). □

Problem : Does Corollary 3.6 or 3.7 hold for fractional values of v ?

We can use this inequality to generalize Statement 3.2.

STATEMENT 3.8. *Let* S, B *be sets of integers* $k \geq 2$, $n \geq 1$, $v \geq 1$ *integers,* $0 \leq p_j \leq 1$, *real numbers,* $r = r_k(n)$. *Let* A *be a random subset of* S *with*

$$P(m \in A) = p_m \geq \min(1, vr^{1/k}) \quad (m \in S,\ m \leq n).$$

Assume that with a constant $C > 0$ *for all* $2 \leq t \leq k-1$ *either* (3.6) *or* (3.7) *is satisfied* (*not necessarily always the same*). *Then we have*

$$\beta_n = P(n \notin A[k]) \leq \exp - \frac{v}{2 + C(k-2)}. \tag{3.16}$$

The proof applies Statement 3.2 and Corollary 3.7.

This enables us to present an analogue of Theorem 3.1 for the case $k > 2$.

THEOREM 3.9. *Let* (p_n) *be a decreasing sequence of real numbers,* $0 \leq p_n \leq 1$. *Let* S, B *be sets of integers, and let* A *be a random subset of* S, *satisfying*

$$P(n \in A) = p_n \quad (n \in S).$$

Assume that either (3.6) *or* (3.7) (*not necessarily always the same*) *holds for all* $2 \leq t \leq k-1$, $m \leq n$, $n \in B$. *If with some* $c > 2 + C(k-2)$ *we have*

$$p_n > \frac{c \log n}{r_k(n)^{1/k}},$$

for all $n > n_0$, $n \in B$, *then with probability* 1 *the set* $A[k]$ *contains all but a finite number of elements of* B.

Proof : We apply Statement 3.8 with $v = [c \log n]$. Inequality (3.16) yields $\beta_n = 0(n^{-d})$ with some $d > 1$, consequently $\Sigma \beta_n < \infty$ and an appeal to the Borel–Cantelli lemma concludes the proof. □

4.– Applications to economical bases

Every basis of order k must satisfy

$$A(x) \geq x^{1/k}, \tag{4.1}$$

so must every asymptotical basis for $x > x_0$. On the other hand, there exist bases of order k such that

$$A(x) = 0(x^{1/k}). \tag{4.2}$$

For more information see, for instance, Halberstam–Roth (1966).

Let us call a basis (asymptotical basis) of order k *economical*, if $A(x) = 0(x^{1/k+\epsilon})$ holds for every $\epsilon > 0$.

Many arithmetical important sets are known to form a basis, but they are mostly very uneconomical. Say, the squares form a basis of order 4 (Langrange's theorem), the primes form an asymptotical basis of order 4 (Vinogradov's theorem) and probably of order 3 (Goldbach's conjecture) and they have a lot more elements

than (4.1–2) would indicate. We are going to investigate the problem whether a smaller subset of these sets also suffices to form a basis, perhaps an economical one, of the same order.

For sets that do not form a basis of order k, it is interesting to decide which numbers are in Ak (say, for primes it depends on parity and for two or three squares it also depends on arithmetical properties). Or, if we cannot establish that Ak contains every number, sometimes one can prove that it contains *most*; a typical example is Davenport's theorem that almost all integers can be written as a sum of four cubes. For this sort of problem we can also ask whether a smaller subset suffices.

These problems were extensively investigated for squares. Härtter and Zöllner (1977) found a subset A of the set Q of squares which is a basis of order 4 and an infinity of squares are missing from A, but still with $A(x) \sim x^{\frac{1}{2}}$. Erdös and Nathanson (1981) improved this to $A(x) = 0(x^{3/8+\epsilon})$; Choi, Erdös and Nathanson (1980) to $A(x) = 0(x^{1/3})$; finally, Zöllner (1985) achieved $A(x) = 0(x^{1/4+\epsilon})$.

Erdös and Nathanson (1981) also found a set

$$A \subset Q,\ A(x) = 0(x^{1/3+\epsilon}),\ A3 = Q3, \tag{4.3}$$

that is, every number not in the form $4^k(8m+7)$ is representable as a sum of three squares from A; (some of thes results actually are stated for the corresponding finite problem; this seems to be only a technical convenience).

We have commented on Erdös and Nathanson's probabilistic method in the previous section. Zöllner's is based on Erdös and Nathanson's result. With our method we find some sharper results.

We start with the set $S = Q = \{n^2\}$ of squares.

THEOREM 4.1. *For any $k \geq 4$ there is a set $A \subset Q$ that is a basis of order k and satisfies*

$$A(x) = 0(x^{1/k} \log x).$$

We do not give a complete proof, just point out some ideas.

We start with the case $k = 4$. For R_4 there is a well–known formula

$$R_4(n) = \sum_{d \mid n,\ 4 \nmid d} d.$$

In particular, we have $R_4(n) \geq n$ if $4 \nmid n$. It is now easy to estimate r_4 and to find

$$r_4(n) > n/25 \quad (n > n_0\,,\ 4 \nmid n).$$

An application of Theorem 3.9 gives us a set A_1 for which $A_1 4$ contains every number not divisible by 4. Finally, our set A can be constructed by putting

$$A = \{2^m a : a \in A_1,\ m = 0,1,\ldots\}.$$

The case $k \geq 5$ is even simpler. It is easily seen that

$$r_k(n) \gg n^{(k/2)-1}$$

for all n, thus a direct application of Theorem 3.9 suffices.

The case $k = 3$ is more difficult. Let $\rho(n)$ denote the number of *primitive* solutions of the equation

$$n = x^2 + y^2 + z^2,$$

that is, those satisfying $(x,y,z) = 1$. Clearly we have

$$R_3(n) = \sum_{d^2 \mid n} \rho(n/d^2).$$

For ρ we have the expression, due to Gauss and Dirichlet

$$\rho(n) = g_n \sqrt{n} L(1,\chi),$$

where g_n depends on the residue of n modulo 8 ($g_n = 0$ if $n \equiv 0, 4, 7$, $=16/\pi$ if $n \equiv 3$ and $24/\pi$ otherwise) and χ denotes the character

$$\chi(m) = \left(\frac{-4\,n}{m}\right).$$

We know $L(1,\chi) \gg n^{-\epsilon}$ (Siegel's theorem); this is sufficient to ensure $R_3 \gg n^{1/2-\epsilon}$ and to find a set A satisfying (4.3). Assuming that there are no Siegel–roots, we would have $L(1,\chi) \gg 1/\log n$, and we could improve (4.3) to

$$A(x) = 0(x^{1/3}(\log x)^{4/3}).$$

Unconditionally I can prove the existence of a set $A \subset Q$ satisfying

$$A(x) = 0(x^{1/3}(\log x)^2)$$

such that $A3$ contains every *squarefree* number $\not\equiv 7$ (modulo 8). To this I use Landau's theorem that even if there are Siegel–roots, they are very rare.

The case $k = 2$ also raises some problems.

CONJECTURE 4.2. *If* $A \subset Q$ *and* $A2 = Q2$, *then* $A = Q$.

Since primes have only one representation (if at all), this would follow from

CONJECTURE 4.3. *For every natural number* n *there is another integer* m *such that* $n^2 + m^2$ *is a prime.*

In this form this seems to be hopeless. On the other hand, with sieve methods one can easily prove that such an m exists for a set of integers n with a positive lower density. Consequently, if $A \subset Q$, $A2 = Q2$ then $A(x) \gg \sqrt{x}$. Perhaps one could improve this to $(A(x)/\sqrt{x}) \to 1$.

Next we consider the set $S = \{n^3\}$.

Cubes are known to form a basis of order 9 and an asymptotic basis of order 7. We shall consider a slightly different aspect. H. Davenport proved that almost all integers (in the sense of asymptotical density) can be written as a sum of four cubes. This can be achieved with a thin subset.

THEOREM 4.4. *There is a set* A *of cubes such that*

$$A(x) = 0(x^{1/4}\log x)$$

and the set H *of positive integers not in* $A4$ *satisfies*

$$H(x) = 0(x(\log x)^{-\gamma})$$

with a constant $\gamma > 0$.

Our proof is based on a sharp effective form of Davenport's theorem, due to Vaughan (1986). It asserts

$$\sum_{n\leq x} (R_4(n) - \rho(n))^2 \ll x^{5/3}(\log x)^{-\gamma}. \tag{4.4}$$

where $\gamma > 0$ and the function $\rho(n)$ is the *expected* main term of R_4.

For the number of integers not representable as a sum of four cubes Vaughan gives the estimate $0(x^{103/115+\epsilon})$, much sharper than what (4.4) implies. We cannot get so far with our economical subset.

Now let $d \geq 3$ be fixed, and consider the set $S = \{n^d\}$. The set of d'th powers is not only known to form a basis, but for large j (for $j > (2+\epsilon)d\log d$) asymptotical formulas are known for $R_j(n)$ (see Vaughan, 1981).

THEOREM 4.5. *Assume that with an integer J for every $j \geq J$ we have*

$$c_j n^{j/d-1} \leq R_j(n) \leq C_j n^{j/d-1}$$

for $n > n_0(j)$; moreover, if $d = 3$ then suppose $J \geq 5$. Then for every

$$k > k^* = d(J-d+2)/2 \tag{4.5}$$

there is a set $A \subset S$ which is an asymptotical basis of order k and satisfies

$$A(x) = 0(x^{1/k}\log x). \tag{4.6}$$

We have $J > d$, thus always $k^* > J$. Typically k^* is of order dJ. SInce upper estimates of R_j are often available for $j < J$, it may be possible to improve this value of k^*, perhaps even to $k^* = J$. For cubes the best known value is $J = 8$ (Vaughan, 1986) which works for $k \geq 11$. I can improve this to 9 but presently I am unable to go down to 8.

Finally put $S = P = \{\text{primes}\}$.

THEOREM 4.6. *For every $k \geq 3$ there is a set $A \subset P$ such that Ak contains all sufficiently large integers $n \equiv k \pmod 2$ and which satisfies*

$$A(x) = 0(x^{1/k}\log x). \tag{4.7}$$

THEOREM 4.7. *There is a set* $A \subset P$ *which satisfies*

$$A(x) = 0((x\log x)^{1/2}) \tag{4.8}$$

such that $A2$ *contains almost all even integers; in fact, the number of even numbers up to* x *not in* $2A$ *is*

$$0(x(\log x)^{-\alpha}) \tag{4.9}$$

with every α.

The proof is based on Theorem 3.1 and an average asymptotical formula for $r_2(n)$, see Vaughan (1981), Theorem 3.7.

For the number of even integers that are not sums of two primes Montgomery and Vaughan (1975) found the sharp estimate $x^{1-\delta}$. Their result does not immediately lead to an improvement in ours, because if there are Siegel–roots, then this is reflected in the behaviour of $r_2(n)$ in the form of a second main term, which often makes it smaller than we need.

5.– Final remarks

We have found results that differ from the general lower bound in a logarithmic factor. One can ask whether this is necessary. For sake of definiteness, let $S = Q$ be the set of squares and $k = 5$. Is there an $A \subset Q$ which is a basis of order 5 and satisfies

$$A(x) = 0(x^{1/5}) \ ?$$

If we take a *random* subset of Q with this density, that is, we put

$$p_n = cn^{-0.3} \quad (n \in Q),$$

then it is easy to see that with probability 1, A will not be an asymptotical basis. What if we add a slowly increasing factor, say

$$p_n = cn^{-0.3} \log\log n \ ? \tag{5.1}$$

Heuristically one would expect that the border is at about $((\log n)/n)^{0.3}$.

Is there a $0-1$ law for this sort of problem ? Can one assert that, under reasonable restrictions, the probability that a random set is an asymptotical basis of order k is either 0 or 1 ? There must be a restriction to exclude degenerate cases like the primes, where the selection or omission of the number 2 may affect the basic property.

Let A be a random set, with defining probabilities p_n. Let $A_1,...,A_k$ be independent random sets with the same sequence (p_n). Assume that A is an asymptotical basis with probability 1. Does it follow that, with probability 1, $A_1 +...+ A_k$ contains all but a finite number of positive integers ?

A detailed account of the results of sections 3–4 will appear elsewhere, under the title *Economical bases.*

BIBLIOGRAPHY

[1] S.L. Choi, P. Erdös, M.B. Nathanson.– Lagrange's theorem with $n^{1/3}$ squares, Proc. Amer. Math. Soc. 79 (1980), 203–205.

[2] P. Erdös.– Some results on additive number theory, Proc. Amer. Math. Soc. 5 (1954), 847–853.

[3] P. Erdös.– Problems and results in additive number theory, Coll. Théorie des Nombres (CBRM), Bruxelles 1956, 127–137.

[4] P. Erdös, M.B. Nathanson.– Lagrange's theorem and thin subsequences of squares, in : J. Gani, V.K. Rohatgi (eds), *Contributions to probability*, Acad. press, New York 1981, 3–9.

[5] H. Halberstam, K.F. Roth.– Sequences, Oxford, Clarendon Press 1966.

[6] E. Härtter, J. Zöllner.– Darstellungen natürlicher Zahlen als Summe und als Differenz von Quadraten, Norske Vidensk. Selsk. Skr. Trondheim 1 (1977), 1–8.

[7] H.L. Montgomery, R.C. Vaughan.– The exceptional set in Goldbach's problem, Acta Arithmetica 27 (1975), 353–370.

[8] I.Z. Ruzsa.– Probabilistic generalization of a number–theoretical inequality, Amer. Math. Monthly 83 (1976), 723–725.

[9] I.Z. Ruzsa.– Essential components, Proc. London Math. Soc. 54 (1987), 38–56.

[10] R.C. Vaughan.– The Hardy–Littlewood method, Cambridge Univ. Press., Cambridge, U.K. 1981.

[11] R.C. Vaughan.– On Waring's problem for cubes, J. Reine Angew. Math. 365 (1986), 122–170.

[12] J. Zöllner.– Über eine Vermutung von Choi, Erdös and Nathanson, Acta Arithmetica 45 (1985), 211–213.

Imre Z. RUZSA
Mathematical Institute of the
Hungarian Academy of Sciences
Budapest, Pf. 127
H–1364 HUNGARY

ON A LIMITING FRACTAL MEASURE DEFINED BY CONJUGATE ALGEBRAIC INTEGERS

A.M. DAVIE and C.J. SMYTH

Let $\alpha \neq 0$ be an algebraic integer of degree d, and $M(\alpha)$ be its Mahler measure :

$$M(\alpha) = \prod_{i=1}^{d} \max(1, |\alpha^{(i)}|) \geq 1, \tag{1}$$

where $\alpha = \alpha^{(1)}, \alpha^{(2)}, \ldots, \alpha^{(d)}$ are the conjugates of α. Recently Langevin [2] has proved that $\theta < \pi$ if α and all its conjugates lie in a sector $-\theta \leq \arg \alpha_i \leq \theta$ of the complex plane, and if α is not a root of unity then

$$M(\alpha) \geq C(\theta)^d$$

where $C(\theta) > 1$ is independent of d.

For $\theta = 0$ (i.e. α totally positive) much more can be said about the values of the heights $\Omega(\alpha) := M(\alpha)^{1/d}$. From a result of Schinzel [5], a totally positive $\alpha \neq 0$ has $\Omega(\alpha) \geq \frac{1}{2}(1+\sqrt{5})$, with $C(0) = \frac{1}{2}(1+\sqrt{5})$ being the largest possible choice for $C(0)$. Indeed [6, 7] the set $\{\Omega(\alpha) : \alpha \neq 1, \ \alpha \text{ totally positive}\}$ has the following structure :

It consists of four discrete points $\Omega(\beta_1^2) = 1.618$, $\Omega(\beta_2^2) = 1.684$, $\Omega(\beta_3^2) = 1.710$, $\Omega(\gamma_6^2) = 1.713$, where β_n and γ_n are defined below. These are the only values less than 1.717. Between 1.717 and $\ell = \lim_{n \to \infty} \Omega(\beta_n^2) = 1.7273..$, the structure of the set is as yet undetermined (but thought to be discrete), while the set is dense on (ℓ, ∞) [6].

Let $H(x) = x - 1/x$. Then β_n is defined by $\beta_0 = 1$ and $H(\beta_n) = \beta_{n-1}$, with $\beta_n > 0$. Clearly β_n is totally real, and one can show that both β_n and β_n^2 have degree 2^n over the rationals. The number γ_n is defined as the largest positive root of $H^n(x) = x$, where $H^n(x) = H(H^{n-1}(x))$. As $\gamma_6 = 2\cos(\pi/7)$, $\deg \gamma_6 = 3$. However the degree of γ_n is not known in general (see [6, p. 148]).

The purpose of this note is to discuss two questions concerning the β_n:

1. Monotonicity of the sequence $\{\Omega(\beta_n^2)\}$,
2. Distribution of the conjugates of β_n on the real line.

1. Monotonicity of the sequence $\{\Omega(\beta_n^2)\}$.

It was shown in [6] that $\ell = \lim_{n\to\infty} \Omega(\beta_n^2)$ exists. Furthermore, computation revealed that $\Omega(\beta_1^2) < \Omega(\beta_2^2) < \ldots < \Omega(\beta_{15}^2)$. However, elaboration of the method used to prove that the limit exists, which involved explicit consideration of the conjugates of β_n, did not give a proof of monoticity. However, Terry Lyons suggested using the **analytic** definition of $M(\beta_n)$ for the proof (instead of (1)), and this idea readily produces the required monoticity, as we now show.

THEOREM 1. *The sequence $\Omega(\beta_n^2)$ is monotonic increasing.*

Clearly, this results helps to elucidate the structure of the set $\{\Omega(\alpha)\}$ in $(1.717, \ell)$. There are, however, other numbers in this interval, e.g. (some) numbers $\Omega(\gamma_n^2)$ and possibly some of the form $\Omega(\gamma_{nk+}^2)$, where $H^k(\gamma_{nk+}) = \gamma_n$ (γ_{nk+} is a preperiodic point of H), and of the form $\Omega(\gamma_{nk-}^2)$, where $H^k(\gamma_{nk-}) = 1/\gamma_n$.

Proof: We define polynomials P_n, Q_n by

$$H^n(x) = P_n(x^2)/x\, Q_n(x^2), \tag{2}$$

where P_n is monic of degree 2^{n-1} with $P_1(y) = y - 1$, and Q_n has degree $2^{n-1} - 1$ with $Q(y) = 1$. Then [6] for $n = 1,2,\ldots$

$$P_{n+1}(y) = P_n^2(y) - yQ_n^2(y) \tag{3}$$

$$Q_{n+1}(y) = P_1(y)P_2(n)\ldots P_n(y) = Q_n(y)P_n(y). \tag{4}$$

Further, since $H^{n+1}(\beta_n) = 0$, P_{n+1} is the minimal polynomial of β_n^2. Next, we need to note that

$$H^{n+1}(x) = H^n(H(x)) = \frac{P_n((x-\frac{1}{x})^2)}{(1-\frac{1}{x})Q_n((x-\frac{1}{x})^2)}$$

so that

$$P_{n+1}(x^2) = x^{2^n} P_n(x-\tfrac{1}{x})^2). \tag{5}$$

Also, from (3) with $y = (x-\frac{1}{x})^2)$

$$P_{n+1}((x-\tfrac{1}{x})^2) = P_n^2((x-\tfrac{1}{x})^2) - (x-\tfrac{1}{x})^2 Q_n((x-\tfrac{1}{x})^2). \tag{6}$$

Then we use Mahler's original definition of Mahler measure to write $M(\alpha)$ as

$$M(\alpha) = \exp(\int_0^1 \log|\, P_\alpha(e^{2\pi i\theta})|\, \mathrm{d}\theta, \tag{7}$$

where P_α is the minimal polynomial of α (equality of (1) and (7) follows from Jensen's Theorem). Hence

$$\begin{aligned}
\log \Omega(\beta_n^2) &= 2^{-n} \int_0^1 \log|\, P_{n+1}(e^{2\pi i\theta})|\; d\theta \\
&= 2^{-n} \int_0^1 \log|\, P_n(-4\sin^2 2\pi\theta)|\; d\theta \text{ using (5). But then} \\
\log \Omega(\beta_{n+1}^2) &= 2^{-(n+1)} \int_0^1 \log|\, P_{n+1}(-4\sin^2 2\pi\theta)|\; d\theta \\
&= 2^{-(n+1)} \int_0^1 \log|\, P_n^2(-4\sin^2 2\pi\theta) + 4\sin^2 2\pi\theta\, Q_n^2(-4\sin^2 2\pi\theta)|\; d\theta
\end{aligned}$$

using (6),

$$= \log \Omega(\beta_n^2) + 2^{-(n+1)} \int_0^1 \log|\, 1 + 4\sin^2 2\pi\theta . Q_n^2(-4\sin^2 2\pi\theta)/P_n^2(-4\sin^2 2\pi\theta)|\; d\theta$$

$> \log \Omega(\beta_n^2)$, as required.

2. Distribution of the conjugates of β_n on the real line

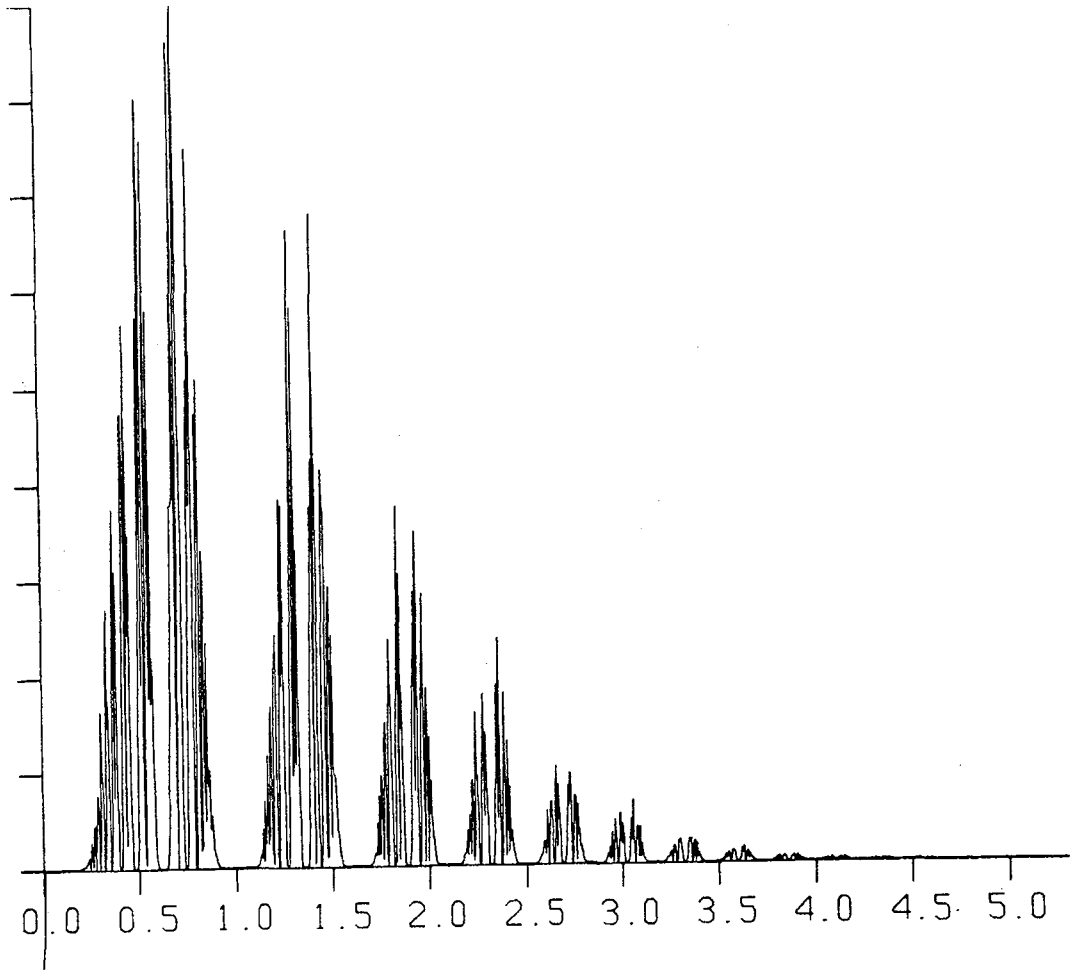

figure 1

Figure 1 shows the density distribution of the moduli of 32768 of the conjugates of β_{15}. The largest conjugate modulus is β_{15} itself, $= 5.4297...$ (when arranged in increasing order of modulus, the conjugates of β_n simply alternate in sign, so their signs are of no importance). As Figure 1 shows, the distribution of conjugate moduli is very irregular, and the aim of this discussion is to explain and quantify that irregularity.

This work arose from a discussion with Pierre Moussa, who pointed out the connection with fractal measures on Julia sets.

Our main result in this section is

THEOREM 2. *As* $n \to \infty$, *the limiting distribution of the conjugates of* β_n *define a probability measure* μ *of Hausdorff dimension* 0.800611138269168784.

We also investigate this limiting measure μ in the neighbourhoods of (a) conjugates of β_n, (b) fixed ponts of H^n. We show that μ behaves very differently in the two kinds of neighbourhoods :

THEOREM 3. *(a) For any conjugate a of β_n, and sufficiently small neighbourhood I of a, there is a constant $c_1 = c_1(a,n) > 0$ independent of I such that*

$$\mu(I) < \exp(-c_1/|I|^2).$$

(b) For any fixed point b of H^n and sufficiently small neighbourhood I of b, there is a constant $c_2 = c_2(b,n) > 0$ independent of I such that

$$\mu(I) > |I|^{c_2}.$$

The best values of c_1 and c_2 can be calculated explicitly in the limit as $|I| \to 0$ (see equations (12) and (13)).

First, we discuss the existence of the limiting measure μ. To define it, we take μ_n to be the atomic measure with weight 2^{-n} at each of the 2^n conjugates of β_n. Then the weak limit $\mu = \lim_{n\to\infty} \mu_n$ defined by $\mu(A) = \lim_{n\to\infty} \mu_n(A)$ for any Borel set $A \subseteq \mathbb{R}$ exists and is H–**invariant** in the sense that, when $H|_A$ is one–to–one then

$$\mu(H(A)) = 2\mu(A). \tag{8}$$

The existence of $\mu = \lim_{n\to\infty} \mu_n$ follows from the following more general result. For any C^1 function $f : \mathbb{C} \cup \{\infty\} \to \mathbb{C} \cup \{\infty\}$, its **Julia Set** J_f is the closure of the set of repelling fixed points of iterates of f (i.e. the closure of the set of points x with $f^n(x) = x$ but $|(f^n)'(x)| > 1$, for some n). Then a result of Freire, Lopes and Mañé [1] and Ljublich [3] states that there is a measure μ_f defined on J_f which is f–invariant ($\mu_f(f(A)) = (\deg f).\mu(A)$ if $f|_A$ is one–to–one) and is the weak limit of the atomic measures which have weight d^{-n} at each of the d^n pre–images under f^n of a single point (note $\deg f = \max(\deg p_1, \deg p_2)$, where $f = p_1/p_2$).

For our map $H(x) = x - 1/x$,

$$|\, Im\, H(x+iy)| \;= |\, y(1+(x^2+y^2)^{-1})| \;> |\, y| \quad \text{if} \quad y \neq 0.$$

Hence H^n has no non–real fixed points. Also $H'(x) = 1 + x^{-2} > 1$ on $\mathbb{R}$ for all n. Hence all fixed points of H^n are real and repelling, so that $J_H \subseteq \mathbb{R}$. It is not difficult to show that in fact $J_H = \mathbb{R}$. Thus the measure μ above is just the measure μ_f for $f = H$.

Before proving Theorem 2, we recall the definition of Hausdorff dimension for sets and measures. For $s > 0$ and a set $E \subset \mathbb{R}$, define

$$H^s(E) = \lim_{\delta \to 0} \inf\{\Sigma |\, U_i|^{\,s} : E \subseteq \bigcup_i U_i,\ U_i \text{ open intervals of length } < \delta\},$$

the infimum being taken over all such countable collections of intervals $\{U_i\}$. Then it is well–known that there is a unique number $h\dim E$ (*the Hausdorff dimension of* E) such that

$$H^s(E) = \begin{cases} \infty & \text{if } s < h\dim E \\ 0 & \text{if } s > h\dim E \end{cases}.$$

For a probability measure μ on $\mathbb{R}$, $h\dim \mu$ is defined as

$$h\dim \mu = \inf_E \{h\dim E : \mu(E) = 1\}.$$

Manning [4] showed that for a rational function f of degree d, the invariant measure μ_f defined on J_f above has

$$h\dim \mu_f = \log d / \int_{J_f} \log |\, f'| \, d\mu_f .$$

For $f = H$, this formula gives

$$h\dim \mu = \log 2/S$$

say, where

(9) $$S = \lim_{n\to\infty} \int_{-\infty}^{\infty} \log(1+x^{-2})\, d\mu_n .$$

Now $\int_{-\infty}^{\infty} \log(1+x^{-2})\, d\mu_n = 2^{-n} \sum_i \log(1+(\beta_n^{(i)})^{-2})$ where the $\beta_n^{(i)}$ are the conjugates of β_n

$$= 2^{-n} \sum_i \log(1+(\beta_n^{(i)})^2)$$

as $\beta_n^{(i)-1}$ is always a conjugate of $\beta_n^{(i)}$

$$= 2^{-n} \log \prod_i (1+(\beta_n^{(i)})^2)$$

(10) $$= 2^{-n} \log(P_{n+1}(-1)),$$

where P_{n+1} is as in (2). Now from (3) and (4)

$$P_{n+1}(-1) = P_n^2(-1) + [P_1(-1)P_2(-1)\ldots P_{n-1}(-1)]^2$$

(11) $$2^{-n}\log P_{n+1}(-1) = 2 \log P_n(-1) + 2^{-n} \log v_n ,$$

where $v_0 = 2$, $v_n = 1 + [P_1(-1)\ldots P_{n-1}(-1)/P_n(-1)]^2$ for $n \geq 1$. Then an easy calculation shows that

$$v_{n+1} = 1 + (v_n - 1)/v_n^2$$

and so from (11), on telescoping the series, and then using (9) and (10) we have

$$S = \lim_{n\to\infty} 2^{-n}\log P_{n+1}(-1) = \sum_{n=0}^{\infty} 2^{-n}\log v_n .$$

Summing 100 terms of the series gives

$$S = 0.865772592245533240$$

and hence $h\dim \mu = \log 2/S$ as in the statement of Theorem 2.

We now prove Theorem 3. Firstly, choosing $a : H^{n+1}(a) = 0$, we want to estimate $\mu((a-\delta, a+\delta))$. Put $c = c(n,a) = (H^{n+1})'(a)$. Then for δ sufficiently

small $H^{n+1}(a \pm \delta) = \pm \delta c + 0(\delta^2)$, the 0 terms depending on a and n. Hence $H^{n+2}(a \pm \delta) = \pm 1/\delta c + 0(1)$, and since $H^{n+2}(a) = \infty$ it follows from (8) that

$$\mu((a-\delta,a+\delta)) = 2^{-(n+2)}\mu((-\infty, -1/\delta|\, c| + 0(1)) \cup (1/\delta|\, c| + 0(1), \infty)).$$

As was shown in [6, Lemma 7 and p. 145],

$$\mu((-\infty, -\beta_k) \cup (\beta_k, \infty)) = 2^{-(k+1)}$$

and $\beta_k < \sqrt{2k+1}$, so that

$$\mu((-\infty, -\sqrt{2k+1} \cup (\beta_k, \infty)) < 2^{-(k+1)} < 2^{-k}.$$

Hence choosing k such that

$$\sqrt{2k+1} < \max(|\, H^{n+2}(a-\delta)|\,, |\, H^{n+2}(a+\delta)|\,),$$

i.e. $k = 1/(2c^2\delta^2) + 0(1/\delta)$

we obtain

$$\mu((a-\delta,a+\delta)) < 2^{-(n+1/(2c^2\delta^2)+0(1/\delta))} \tag{12}$$

as required. This result explains why the graph of Figure 1 is essentially zero around 0, 1, $\frac{1}{2}(1+\sqrt{5})$ etc.

In order to estimate μ in a neighbourhood of a fixed point of H^n, we must first describe these points. Put

$$H^{-1}(x) = \tfrac{1}{2}(x + \sqrt{x^2+4}).$$

Then the preimages of x under H are $H^{-1}(x)$ and $-(H^{-1}(x))^{-1}$. So define, for $\epsilon = \pm 1$,

$$\epsilon(x) = \epsilon(H^{-1}(x))^{\epsilon}$$

and hence for $\epsilon_k, \epsilon_{k-1}, \ldots, \epsilon_2, \epsilon_1 = \pm 1$

$$\epsilon_k\epsilon_{k-1} \cdots \epsilon_1(x) = \epsilon_k(\epsilon_{k-1} \cdots \epsilon_1(x)).$$

The 2^n-2 finite fixed points b of H^n are defined by the 2^n-2 sequences $\epsilon_n\epsilon_{n-1} \cdots \epsilon_1$, where the ϵ_i are not all equal, in the following way : take $b_0 = 1$ and put $b_{\ell+1} = \epsilon_n\epsilon_{n-1} \cdots \epsilon_1(b_\ell)$. Then $b = \lim_{h\to\infty} b_\ell$ satisfies $H^n b = b$.

Next, for a particular choice of $\epsilon_n\epsilon_{n-1}\cdots\epsilon_1$ not all equal and corresponding $b = H^n b$ we define a sequence $I_0 \supset I_1 \supset I_2 \cdots$ of neighbourhoods of b, as follows : choose I_0 to be whichever of the intervals $(-\infty,0)$ or $(0,\infty)$ which contains b. Then since $H^{-1}(0,\infty) = (1,\infty)$, $H^{-1}(-\infty,0) = (0,1)$, $-(H^{-1})^{-1}(0,\infty) = (-1,0)$, $-(H^{-1})^{-1}(-\infty,0) = (-\infty,-1)$ and not all the ϵ_i are equal, it follows that one of the intervals $\epsilon_1(I_0)$, $\epsilon_2\epsilon_1(I_0),\ldots,\epsilon_n\epsilon_{n-1}\cdots\epsilon_1(I_0)$ belongs to either $(0,1)$ or $(-1,0)$, and so is finite, of length less than 1.

So if we put

$$I_{\ell+1} = \epsilon_n\,\epsilon_{n-1}\cdots\epsilon_1(I_\ell) \quad (\ell = 0,1,2,\ldots)$$

then, since $|\epsilon'(x)| < 1$, $|I_1| < 1$. Also, since $b \in I_{\ell+1} \cap I_\ell$, and $|(\epsilon_n\,\epsilon_{n-1}\cdots\epsilon_1)'(x)| < 1$, $I_\ell \supset I_{\ell+1}$.

We now estimate the length $|I_\ell|$ of I_ℓ. Since $H^n(I_{\ell+1}) = I_\ell$, $|I_{\ell+1}|\,|(H^n)'(\xi_{\ell+1})| = |I_\ell|$ for some $\xi_{\ell+1} \in I_{\ell+1}$. If

$$m_n = \min_{\xi\in I_1} |(H^n)'(\xi)|$$

then $m_n > 1$ and so $|I_{\ell+1}| < m_n^{-1}|I_\ell|$, $|I_\ell| \le m_n^{-\ell+1}$ $(\ell = 1,2,\ldots)$. Hence putting $B = (H^n)'(b)$,

$$(H^n)'(\xi) = B + 0(m_n^{-\ell})$$

$$\sum_{i=1}^{\ell} \log|(H^n)'(\xi_i)| = \ell \log B + 0(1)$$

$$\log|I_\ell| = \log|I_1| + 0(1) - \ell \log B.$$

To compute $\mu(I_\ell)$, note that $\mu(I_0) = \frac{1}{2}$, and so by (8),

$$\mu(I_\ell) = 2^{-n\ell-1},$$

and thus

$$\mu(I_\ell) = |I_\ell|^{-n(\log 2/\log B + 0(1/\ell))} \tag{13}$$

which is a stronger form of Theorem 3(b).

Examples : We calculate explicitly values of the constant $c_2 = c_2(b,n)$ such that $\mu(I_\ell) = |I_\ell|^{c_2+0(1/\ell)}$, for some small values of n.

1. $n = 2$, $\epsilon_2\epsilon_1 = 1 - 1$, $b = 1/\sqrt{2}$, $c_2 = \log 2/\log 3 = 0.6309$.

2. $n = 4$, $\epsilon_4\epsilon_3\epsilon_2\epsilon_1 = -11 - 11$, $b = \sqrt{1+\sqrt{2}}$, $c_2 = 2 \log 2/\log 7 = 0.7124$

3(a) $n = 6$, $\epsilon_6\epsilon_5\epsilon_4\epsilon_3\epsilon_2\epsilon_1 = -1\ -11\ -1\ -11$, $b = 1/(2\cos 5\pi/18)$, $c_2 = 3 \log 2/\log 19 = 0.7062$

3(b) $n = 6$, $\epsilon_6\epsilon_5\epsilon_4\epsilon_3\epsilon_2\epsilon_1 = 11\ -111\ -1$, $b = 2 \cos 4\pi/7$, $c_2 = 3 \log 2/\log 13 = 0.8107$.

These examples show that not only is c_2 not monotonic in n, but it really does depend on b too.

It can be shown that, as $n \to \infty$, c_2 averaged over all fixed points b of H^n tends to $h\dim \mu = 0.8006\ldots$.

REFERENCES

[1] A. Freire, A. Lopes and R. Mañé.– An invariant measure for rational maps, Bol. Soc. Bras. Mat. 14, N° 1 (1983), 45–62.

[2] M. Langevin.– Méthode de Fekete–Szego et problème de Lehmer, C.R. Acad. Sci. Paris, 301 (1985), 436–466.

[3] M.J. Ljubich.– Entropy properties of rational endomorphisms of the Riemann sphere, Ergod. Th. and Dynam. Sys. 3 (1983), 351–385.

[4] A. Manning.– The dimension of the maximal measure for a polynomial map, Ann. Math. 119 (1984), 425–430.

[5] A. Schinzel.– On the product of the conjugates outside the unit circle of an algebraic number, Acta Arith. 24 (1973), 385–399.

[6] C.J. Smyth.– On the measure of totally real algebraic integers, J. Austral. Math. Soc. (Series A), 30 (1980), 137–149.

[7] C.J. Smyth.– On the measure of totally real algebraic integers II, Math. of Computation 37 (1980), 205–208.

A.M. Davie and C.J. Smyth
Department of Mathematics
University of Edinburgh
James Clerk Maxwell Building,
The King's Buildings,
Mayfield Road,
Edinburgh EH9 3JZ
SCOTLAND

SUR LE THEOREME DE DIRICHLET CONCERNANT L'APPROXIMATION DIOPHANTIENNE

(résumé)

P. THURNHEER

i) Soient $\alpha_1,\ldots,\alpha_n$, $n \geq 2$, des nombres réels donnés. Il s'ensuit du théorème de Dirichlet qu'il existe une infinité de points entiers $(g_1,\ldots,g_{n+1}) \in \mathbb{Z}^{n+1}$ tels que

$$|\alpha_1 g_1 + \ldots + \alpha_n g_n + g_{n+1}| \leq (\max_{1 \leq \nu \leq n} |g_\nu|)^{-n}.$$

On montre qu'essentiellement cet énoncé reste vrai si on restreint le choix des points approximants $(g_1,\ldots,g_{n+1})$ à certains sous–ensembles de $\mathbb{R}^{n+1}$. En d'autres termes, on montre que l'effectivité dans le théorème de Dirichlet peut être remplacée par une condition concernant la position des points approximants.

ii) Soient ϵ,δ deux nombres arbitraires positifs. Pour $\mathscr{H} = (\xi_1,\ldots,\xi_{n+1}) \in \mathbb{R}^{n+1}$ on note

$$p(\mathscr{H}) = (\sum_{\nu=1}^{n-1} \xi_\nu^2)^{\frac{1}{2}}$$

et

$$\Phi(w) = \{\mathscr{H} \in \mathbb{R}^{n+1} \mid |\xi_n| \leq (1+\epsilon)p(\mathscr{H})^w\} \cup \{\mathscr{H} \in \mathbb{R}^{n+1} \mid p(\mathscr{H}) \leq 1\};$$

$$\Psi = \{\mathscr{H} \in \mathbb{R}^{n+1} \mid |\xi_n| \leq \epsilon\, p(\mathscr{H})\}.$$

THEOREME.

a) *Si*

$$w = w(n) = 1 + \frac{1}{n} + \frac{1}{n^2},$$

alors il existe une infinité de points entiers $\mathcal{G} = (g_1,\ldots g_{n+1})$ *tels que*

$$(1) \qquad \mathcal{G} \in \Phi(w) \quad et \quad |\alpha_1 g_1 + \ldots + \alpha_n g_n + g_{n+1}| < (1+\delta)(\max_{1 \leq \nu \leq n} |g_\nu|)^{-n}.$$

b) *Si*

$$v = v(n) = \frac{1}{2}(n-1+\sqrt{n^2+2n-3})$$

et si les nombres $1, \alpha_1,\ldots \alpha_n$ *sont* $\mathbb{Q}$*–linéairement indépendants, alors il existe une infinité de points entiers* $\mathcal{G} = (g_1,\ldots,g_{n+1})$ *tels que*

$$(2) \qquad \mathcal{G} \in \Psi \quad et \quad |\alpha_1 g_1 + \ldots + \alpha_n g_n + g_{n+1}| < \delta(\max_{1 \leq \nu \leq n} |g_\nu|)^{-v}.$$

iii) Remarques

1) Le théorème reste vrai si on remplace $\Phi(w)$ et Ψ par leurs images sous une transformation des coordonnées Λ linéaire régulière – mais sinon arbitraire – à condition qu'à la place de $1 + \delta$ dans (1) on introduise une borne $d(\Lambda)$ dépendante de Λ.

2) Pour $n = 2$ la partie b) du théorème a été démontrée par W.M. Schmidt [1] qui s'était intéressé à l'approximation diophantienne par des entiers positifs. Il a donné aussi un exemple [1, remark C] qui montre que le théorème b) est faux sans l'hypothèse sur la $\mathbb{Q}$–indépendance linéaire des nombres $1, \alpha_1,\ldots,\alpha_n$.

3) Pour $n \geq 3$ le résultat suivant de W.M. Schmidt [1] donne une 'borne inférieure' pour les domaines $\Phi(w)$ et Ψ, montrant que dans le théorème ceux–ci ne peuvent pas être remplacés par l'ensemble $\Theta = \{\mathcal{H} \in \mathbb{R}^{n+1} | \xi_\nu > 0, \quad \nu = 1,\ldots,n\}$, comme [1, remark F] :

étant donnés $n \geq 3$ *et* $\epsilon > 0$ *il existe des nombres* $\alpha_1,\ldots,\alpha_n$ *et* $c(\epsilon) > 0$ *tels que* $1, \alpha_1,\ldots,\alpha_n$ *sont* $\mathbb{Q}$*-linéairement indépendants et*

$$|\alpha_1 g_1 + \ldots + \alpha_n g_n + g_{n+1}| > c(\epsilon)(\max_{1 \leq \nu \leq n} |g_\nu|)^{-2-\epsilon}$$

pour tout point entier $\mathcal{G} = (g_1,\ldots,g_{n+1}) \in \Theta$.

4) W.M. Schmidt conjecture que (2) est vrai avec $n-\epsilon$ à la place de $v(n)$.
5) La démonstration du théorème ci–dessus est donnée dans [2]. On y trouve également un théorème qui entre autre caractérise une classe de n–tuplets $\alpha_1,\ldots,\alpha_n$ permettant une approximation diophantienne d'ordre $n-\epsilon$ par des points entiers dans Ψ.

P.S. : Une argumentation un peu plus raffinée – trouvée entretemps – permet de démontrer que seul la 'partie essentielle' de $\Phi(w)$ est nécessaire. En d'autres termes

> *le théorème* a) *reste vrai avec* $\{\mathcal{H} \in \mathbb{R}^{n+1} \mid \xi_n| \leq (1+\epsilon)p(\mathcal{H})^w\}$ *à la place de* $\Phi(w)$.

BIBLIOGRAPHY

[1] W.M. Schmidt.– Two questions in diophantine approximation, Monatshefte für Mathematik 82 (1976), 237–245.

[2] P. Thurnheer.– On Dirichlet's theorem concerning diophantine approximation, à paraître dans Acta Arithmetica Vol. 54.

P. THURNHEER
ETH–Z Hg G
Rämistr. 101
CH – 8092 ZÜRICH

COMMENT DEVINER LES RACINES ℓ–ièmes MODULO n EN RÉDUISANT DES RÉSEAUX

B. VALLÉE, M. GIRAULT, P. TOFFIN

Résumé : Dans de nombreux problèmes de théorie algorithmique des nombres, on est confronté à des équations ou des inéquations polynomiales modulo un nombre n. Quand ce nombre est une puissance d'un nombre premier, des algorithmes polynomiaux, déterministes ou probabilistes, permettent de résoudre ces problèmes. Il en est de même, via le théorème des restes chinois, quand on connaît la factorisation du module n. On peut alors se poser une question importante :

Est–ce que la résolution d'inéquations ou d'équations polynomiales modulo un nombre n est aussi difficile que la factorisation de n ?

Nous montrons ici que, même si on ne connaît pas la factorisation de n, on peut résoudre en temps polynomial probabiliste des équations ou des inéquations polynomiales modulo ce nombre n pourvu que l'on connaisse une bonne approximation des solutions.

Ce sont les réseaux que nous utilisons principalement, après une linéarisation préalable du problème; nous étudions une famille particulière de réseaux, qui généralise celle de Frieze et la solution de notre problème réside dans la régularité géométrique de ces réseaux.

Nos résultats sont à la fois algorithmiques et structurels.

D'une part, nous construisons un algorithme, fondé sur la réduction des réseaux, qui reconstruit les valeurs exactes des racines tronquées d'un polynôme, et nous étendons ainsi des résultats précédents, uniquement obtenus dans le cas linéaire par Frieze.

Cet algorithme a de nombreuses applications pratiques, puisque la sécurité de plusieurs schémas cryptographiques repose sur la difficulté de résoudre des équations polynomiales congruentes. Nous déduisons d'abord qu'il est facile de casser les versions de plus haut degré du récent système d'Okamoto, et nous étendons ainsi des attaques précédentes de Brickell et Shamir. Nous obtenons aussi de nouveaux résultats concernant la prédictibilité du générateur RSA.

D'autre part, nous établissons, pour tout ℓ, de nouveaux résultas concernant la répartition relative des puissances ℓ–ièmes et de leurs racines ℓ–ièmes, et nous prouvons, dans le cas $\ell = 2$, une propriété très naturelle sur cette répartition. Ces résultats peuvent être considérés commes des extensions d'un résultat précédent de Blum.

1.– INTRODUCTION.

Chacun sait que l'extraction des racines ℓ–ièmes modulo un nombre composé n est un problème difficile. En fait, la résolution de la congruence $x^2 \equiv y[n]$ est de même difficulté que la factorisation de n. Dans ce papier, nous étudions le problème de l'extraction des racines ℓ–ièmes : nous voulons résoudre la congruence

$$x^\ell \equiv y\ [n] \tag{1}$$

quand la factorisation de n est inconnue. Notre méthode s'applique à des modules n qui sont soit square–free soit presque square–free –un nombre est dit square–free s'il n'est divisible par aucun carré–, et nous supposons que nous disposons, au départ, d'approximations initiales sur les deux variables x_0 pour x et y_0 pour y.

Nous avons de bonnes raisons pour choisir un pareil cadre :

(i) Les modules square–free correspondent à des cas plus difficiles que des nombres hautement composés, plus faciles à factoriser.

(ii) Dans les applications à la cryptographie, on ne doit souvent deviner que certains bits sur chacune des variables de la congruence (1), tandis que les autres bits constituent la donnée du problème : c'est en particulier le cadre général dans les systèmes proposés par Okamoto, et aussi dans les problèmes de prédictibilité du générateur pseudo–aléatoire RSA.

(iii) L'étude de cette congruence dans le cas particulier $\ell = 2$ joue un rôle particulier important dans les algorithmes de factorisation entière fondés sur la recherche de congruences de carrés.

Nous décrivons ici une méthode générale, fondée sur l'étude géométrique d'une certaine famille de réseaux, et nous construisons un nouveau cadre qui permet d'étendre et d'unifier de précédents résultats de Frieze, Brickell, Shamir et Blum.

Notre papier est organisé comme suit :

D'abord, nous rappelons les problèmes déjà posés, leurs solutions partielles, et décrivons comment nos résultats résolvent des extensions de ces problèmes. Après quoi, nous introduisons notre outil principal, les réseaux, et expliquons comment leurs propriétés géométriques interviennent dans ce sujet. Enfin, nous revenons aux nombres, déduisons nos résultats et décrivons leurs applications cryptographiques. Nous terminons en posant quelques problèmes ouverts.

1.1.– Quelques définitions.

$Z(n)$ désigne l'anneau des entiers modulo n que nous identifions avec l'intervalle des entiers de longueur n centré en 0. Un entier $\ell \geq 2$ nous est donné et nous cherchons des racines ℓ–ièmes modulo n. Nous supposerons toujours que n et ℓ sont premiers entre eux.

Nous considèrerons souvent des modules n qui ne sont pas trop différents de modules square–free, et nous définissons, pour $\delta \in [0,1]$, un nombre n comme étant *δ–square–free* si

$$n = \prod_{i=1}^{f} p_i^{e_i}, \text{ et } \prod_{i=1}^{f} p_i^{e_i - 1} \leq n^{\delta} \quad (p_i \text{ sont les facteurs premiers distincts de } n)$$

Ces nombres sont des généralisations des nombres square–free (un nombre square–free est 0 square–free) et ont déjà été introduits par Frieze et al. [4].
Un nombre n est dit *δ–monosquare-free* si on peut l'écrire :

$$n = p^2 q \text{ avec un premier } p \text{ égal à } n^{\delta}, \text{ et premier avec les square–free } q.$$

Quand n est un nombre δ–monosquare–free et que ℓ est un entier supérieur ou égal à 2, certains éléments x_0 de $Z(n)$ jouent un rôle particulier; ce sont les x_0 tels que x_0 modulo pq est inférieur à $n^{(1-\delta)/2\ell}$. Un tel élément x_0 est appelé *facile.*
Nous utilisons des approximations d'un nombre x_0 de $Z(n)$, et nous considérons des espèces variées de voisinages de x_0. Si $|u|$ désigne, pour $u \in Z(n)$, le maximum de u et de $-u$, il est usuel de considérer les intervalles

$$I(a,x_0) = \{x \in Z(n) \,/\, x = x_0 + u, |u| \leq n^a\}.$$

Plus généralement, si a, a_1, a_2 sont trois réels de $[0,1]$, nous définissons les sous–ensembles

$$K(a_1,a_2,x_0) = \{x \in Z(n) \,/\, x = u_1 x_0 + u_2, |u_1| \leq n^{a_1}, |u_2| \leq n^{a_2}\}$$

formés de petits intervalles centrés autour de petits multiples de x_0. Nous distinguons dex cas particuliers importants : $a_1 = 0$ et aussi $a_1 = a_2 = a/2$ et nous posons

$$H(a,x_0) = K(0,a,x_0) \text{ and } J(a,x_0) = K(\tfrac{a}{2},\tfrac{a}{2},x_0).$$

Remarquons que $H(a,x_0)$ est la réunion des trois intervalles $I(a,x_0)$, $I(a,0)$ and $I(a,-x_0)$. Les intervalles I, et, plus généralement les sous–ensembles H, définissent ce que nous appelons des *approximations inhomogènes*, tandis que les sous–ensembles J définissent des *approximations homogènes.*

1.2.– Des problèmes naturels.

Au lieu de résoudre exactement l'équation $x^\ell \equiv y\ [n]$, qui est sûrement un problème difficile quand la factorisation de n est cachée, nous nous permettons des approximations sur l'une ou l'autre des deux variables x ou y. Les problèmes suivants arrivent alors de manière naturelle :

Problème 1 :

Etant donné un intervalle $I(b,y_0)$ de $Z(n)$, trouver x tel que x^ℓ appartienne à $I(b,y_0)$.

Problème 2 :

Etant donnés

(i) y_0 , une puissance ℓ–ième dans $Z(n)$ dont les racines ℓ–ièmes sont inconnues,

(ii) un sous–ensemble $H(a,x_0)$ –resp $J(a,x_0)$– dont on sait qu'il contient une racine ℓ–ième x de y_0 ,

trouver x.

1.3.– Nos principaux résultats.

En fait, le problème que nous résolvons est plus général; c'est une extension naturelle des deux problèmes précédents, puisqu'il permet une approximation simultanée sur les deux variables x and y.

Problème 3 :

Etant donnés deux sous–ensembles de $Z(n)$

(i) un intervalle $I(b,y_0)$

(ii) et un sous–ensemble $K(a_1,a_2,x_0)$

existe–t–il x dans $K(a_1,a_2,x_0)$ et y dans $I(b,y_0)$ tels que $x^\ell \equiv y[n]$?

Si oui, les trouver.

Avant d'établir notre principal résultat, qui résoud essentiellement ce dernier problème, nous donnons deux définitions :

Deux sous–ensembles A et B de $Z(n)$ sont dits ℓ–compatibles si et seulement si il existe une paire (x,y) de $A \times B$ telle que $x^\ell \equiv y$ modulo n; on dit qu'une telle paire est une *paire de compatibilité*. Deux paries de compatibilité (x,y) et (x',y') sont dites *jumelles* si et seulement si $x' = -x$ et $y' = y$; remarquons que, si A est symétrique par rapport à 0 et que si ℓ est pair, la compatibilité d'une paire entraîne la compatibilité de sa jumelle.

La définition suivante permet d'établir des conditions suffisantes de compatibilité.

Soient a_1, a_2, b, δ quatre réels de $[0,1]$, $\epsilon > 0$ un nombre réel et un entier $\ell > 2$; le triplet (a_1, a_2, b) satisfait les conditions $C(\ell,\delta,\epsilon)$ si et seulement si

$$(2) \qquad \frac{\ell(\ell+1)}{2} a_1 + \frac{\ell(\ell-1)}{2} a_2 + b = 1 - \delta - \ell\epsilon \text{ et } b \geq \ell a_2.$$

Nous pouvons maintenant donner notre premier résultat principal, qui décrit la répartition relative des racines et des puissances ℓ–ièmes; c'est un résultat d'unicité : si x et x_0 sont deux éléments de $Z(n)$ assez proches, leurs puissances ℓ–ièmes ne sont pas souvent très proches.

THEOREME 1. *Pour $\epsilon > 0$, pour $\ell \geq 2$, pour n δ-square-free, $n \geq n_0(\ell,\epsilon)$, pour un triplet (a_1,a_2,b) qui satisfait $C(\ell,\delta,\epsilon)$, il existe un ensemble exceptionnel $S(\epsilon)$ de $Z(n)$ vérifiant les deux conditions suivantes :*

(i) $|S(\epsilon)| \leq n^{1-\epsilon}$

(ii) pour tout x_0 n'appartenant pas à $S(\epsilon)$, pour tout y_0, les deux sous-ensembles $K(a_1,a_2,x_0)$ et $I(b,y_0)$ ont au plus une paire de compatibilité (pour un ℓ impair) et au plus deux paires de compatibilité (pour un ℓ pair).

Et notre second résultat, qui est effectif :

THEOREME 2. *Pour $\epsilon > 0$, pour $\ell \geq 2$, pour n δ–square-free, $n \geq n_0(\ell,\epsilon)$, pour un triplet a_1,a_2,b qui satisfait $C(\ell,\delta,\epsilon)$, il existe un algorithme polynomial probabiliste qui décide, en dehors de $S(\epsilon)$, de la ℓ-compatibilité des deux sous-ensembles $K(a_1,a_2,x_0)$ and $I(b,y_0)$.*

1.4.– Le cas particulier $\ell = 2$.

Le résultat suivant montre que les racines carrées d'un intervalle suffisamment grand ont une répartition presque régulière à l'intérieur de la totalité de $Z(n)$: c'est une extension d'un résultat précédent de Blum [2]. Plus précisément, nous obtenons ici un résultat d'existence :

THEOREME 3. *Pour tout* $\epsilon > 0$, *pour tout* $n \geq n_0(\ell,\epsilon)$, *pour toute paire* (a,b) *satisfaisant les deux conditions*

$$a + b = 1 + 4\epsilon \quad et \quad b \geq 2a$$

il existe un ensemble exceptionnel $S(\epsilon)$ *de* $Z(n)$ *vérifiant les deux conditions suivantes* :

(i) $|S(\epsilon)| \leq n^{1-\epsilon}$.

(ii) Pour tout x_0 *n'appartenant pas à* $S(\epsilon)$, *pour tout* y_0, *les deux intervallles* $I(a,x_0)$ *et* $I(b,y_0)$ *sont 2-compatibles.*

Il y a aussi une version effective de ce résultat.

1.5.– La prédictibilité du générateur pseudo–aléatoire RSA.

Adèle choisit un élément x_0 dans $Z(n)$ et calcule $y_0 \equiv x_0^{\ell}[n]$; puis elle cache les $[a \log_2 n]$ bits moins significatifs –i.e. ceux les plus à droite– de x_0 et les $[b \log_2 n]$ bits moins significatifs de y_0. Elle demande à Basile de deviner ces bits cachés.

Basile peut–il le faire ? Quelle proportion totale de bits peut–il deviner ?

Notre réponse est la suivante : Basile peut deviner une proportion totale de bits égale à

$$a + b = \frac{2}{\ell}(1 - \delta - \ell\epsilon) \quad \text{pourvu que} \quad b \geq \ell a.$$

Donc, pour $\ell = 2$, puisque δ peut être négligé, Basile devine presque autant de bits qu'il y a de bits visibles.

Nous donnons aussi une autre application aux générateurs pseudo–aléatoires. Le générateur RSA de degré ℓ et de module n est bien connu : on choisit x_0 dans $Z(n)$ et on définit la suite x_i par la relation

$$x_{i+1} \equiv x_i^{\ell} [n].$$

En cachant les $[a \log_2 n]$ bits moins significatifs de x_i, on obtient une suite tronquée. On peut se poser la question :

Comment choisir a pour que cette suite soit prédictible ?

Nous pouvons répondre :

Le générateur pseudo aléatoire obtenu par troncature des $[a \log_2 n]$ bits moins significatifs du générateur RSA associé à un degré ℓ et à un module n δ–square–free est prédictible à droite pourvu que

$$a < 2 \frac{1-\delta}{\ell(\ell+1)};$$

le terme correctif δ peut être négligé si $\ell = 2$.

1.6.– Application à l'attaque des schémas cryptographiques.

Le fait de casser les schémas d'Okamoto [6], [7], [8], repose sur l'existence d'algorithmes polynomiaux probabilistes qui résolvent, même si la factorisation du module n est cachée, les problèmes déjà posés; nous ne donnerons pas ici une description précise de ces schémas ni de leurs attaques [10], mais nous insisterons sur la relation entre nos résultats et ces attaques.

Ici, on suppose le module n particulier : $n = p^2 q$ où p et q sont des premiers distincts $(p < q)$; un tel nombre est $1/3$–monosquare–free.

La résolution du Problème 1 : l'attaque du schéma de signature.

Dans [6], Okamoto et Shiraaishi ont proposé un schéma de signature :

Etant donnée une fonction **one–way** h, on considère qu'une signature x est valide pour un message u si

$h(u) \leq x^{\ell}[n] \leq h(u) + O(n^b)$ avec $|x|$ pas trop petit et b beaucoup plus petit que 1.

Par suite, le Problème 1 est exactement le problème que nous devons résoudre pour casser ce schéma de signature. Ce problème a déjà été résolu avec $b = 2/3$ dans les deux cas particuliers suivants :

(1) si $\ell = 2$, même si la factorisation de n est cachée (Brickell [3])

(2) pour tout ℓ, pourvu que la factorisation de $n = p^2 q$ (p et q premiers distincts) soit connue (Okamoto [6], Brickell [3]). Il est facile de voir que ce résultat peut se généraliser quand q est un nombre quelconque premier avec le premier p.

Nous améliorons ici ce résultat précédent dans deux directions différentes sans utiliser la forme particulière du module :

(1) dans le cas $\ell = 2$, nous montrons qu'on peut trouver la racine carrée x dans presque n'importe quel intervalle $I(a,x_0)$ pourvu que a soit suffisamment grand : $a > 1/3$

(2) pour tout ℓ, pour tout n, nous montrons qu'on peut deviner x, pourvu qu'on en connaisse une estimation (qui dépend de ℓ et de δ, le degré de simplicité de n) même si la factorisation de n est cachée et ainsi attaquer le schéma de signature avec cette information supplémentaire.

La résolution du Problème 2 : les attaques des cryptosystèmes.

Dans [7], Okamoto a proposé un premier cryptosystème à clef publique :
La clef publique est la paire (n,x_0), où x_0 est un élément facile de $Z(n)$. A partir d'un message u, qui est petit comparé à n, le texte codé y est construit comme suit :

$$y \equiv (x_0 + u)^{\ell} [n].$$

Comme il est signalé dans [8], Shamir [9] a deux attaques pour casser ce système : la première convient pour une paire quelconque (n,x_0) tandis que la seconde attaque utilise la forme particulière de la clef publique.
Okamoto [8] a proposé alors un second cryptosystème : x_0 est le quotient connu modulo n de deux nombres faciles de $Z(n)$ tenus secrets. Un message (u_1,u_2), où les u_i sont petits comparés à n, permet de construire un texte codé y comme suit :

$$y \equiv (u_1 x_0 + u_2)^{\ell} [n].$$

La sécurité d'un tel cryptosystème est fondée sur la difficulté de résoudre le Problème 2, dans le cas inhomogène pour la première version [7], dans le cas homogène pour la seconde version [8].
Shamir [9] (1986) a déjà attaqué la première version –inhomogène– du cryptosystème d'Okamoto en résolvant le second problème dans le cas d'une approximation inhomogène. Il a d'ailleurs deux attaques différentes (expliquées dans [8]) selon que l'une ou l'autre des conditions suivantes est réalisée :

(1) pour toute paire (x_0,n), quand $\ell = 2$ et $a < 1/3$

(2) pour tout ℓ, pour tout n δ–monosquarefree, pour tout x_0 facile.

Le cas d'une approximation homogène, sous les mêmes conditions des paramètres, est une question ouverte d'Okamoto (1987); une réponse positive à cette question casserait d'ailleurs la seconde version du cryptosystème d'Okamoto.
Nous pouvons effectivement résoudre cette question ouverte.
De plus, nous avons une méthode plus générale qui étend la première attaque de Shamir et permet de résoudre le Problème 2.

THEOREME 4. ***Pour tout ℓ, pour les deux sortes d'approximation, il existe un algorithme polynomial probabiliste qui résoud le Problème 2 pourvu que n soit un nombre δ–square-free suffisamment grand et que***

$$a < 2\,\frac{1-\delta}{\ell(\ell+1)}\,.$$

Le terme correctif, qui dépend de δ, peut être négligé quand $\ell = 2$.

Nous déduisons le résultat suivant :

Toutes les versions des cryptosystèmes d'Okamoto peuvent être cassées.

Nous présentons aussi une seconde attaque qui utilise la forme particulière de la clef publique et généralise la deuxième attaque de Shamir.

2.– Comment arrivent les réseaux.

Ici, nous montrons comment les réseaux interviennent de manière naturelle dans notre problème; nous introduisons tout d'abord un premier réseau, puis un second afin de pouvoir travailler avec la norme sup. Nous rappelons certaines propriétés géométriques des réseaux liées à la longueur du premier minimum et finissons en analysant la taille du premier minimum dans notre famille particulière de réseaux.

2.1.– Le premier réseau.

On se donne un entier $\ell \geq 2$ et une paire (x_0,y_0) de deux éléments de $Z(n)$ et nous considérons la ℓ–ième compatibilité des deux sous–ensembles $K(a_1,a_2,x_0)$ and $I(b,y_0)$: nous cherchons un triplet (u_1,u_2,v) de petits entiers, solution de l'équation

$$(u_1x_0+u_2)^{\ell} \equiv y_0 + v\ [n]$$

En utilisant le développement du binôme, on obtient

$$x_0^\ell u_1^\ell + C_\ell^1 x_0^{\ell-1} u_1^{\ell-1} u_2 + \ldots + C_\ell^i x_0^{\ell-i} u_1^{\ell-i} u_2^i + \ldots + C_\ell^1 x_0 u_1 u_2^{\ell-1} + u_2^\ell - v \equiv y_0 \ [n].$$

Nous considérons le réseau $L(\ell,x_0)$ des vecteurs $w = (w_0,w_1,\ldots,w_\ell)$ de $Z^{\ell+1}$ vérifiant

$$\sum_{i=0}^{\ell-1} C_\ell^i x_0^{\ell-i} w_i - w_\ell \equiv 0 \ [n]$$

et nous devons trouver, dans $L(\ell,x_0)$, un point w qui est –au sens d'une norme un peu inusuelle– près du point $(0,0,\ldots,y_0)$. Le réseau $L(\ell,x_0)$ a la matrice suivante $(\ell+1,\ell+1)$

$$\begin{pmatrix} 1 & 0 & 0 & \ldots & 0 & 0 \\ 0 & 1 & 0 & \ldots & 0 & 0 \\ 0 & 0 & 1 & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & 1 & 0 \\ x_0^\ell & C_\ell^1 x_0^{\ell-1} & C_\ell^2 x_0^{\ell-2} & \ldots & C_\ell^{\ell-1} x_0 & n \end{pmatrix}.$$

Pour chacune des composantes, l'approximation imposée par le choix des voisinages est la suivante :

$$|w_i| \leq n^{(\ell-i)a_1 + ia_2} \text{ pour tout } i : 0 \leq i \leq \ell - 1 \text{ et aussi } |w_\ell - y_0| \leq 2n^b$$

(la dernière condition utilise l'hypothèse : $b \geq \ell a_2$).

2.2.– Le second réseau.

A part le cas particulier où $a_1 = a_2 = b/\ell$, ces approximations ne sont pas du même ordre de grandeur et nous utilisons un système de multiplicateurs

$$k = (k_0,k_1,\ldots,k_\ell)$$

qui **dilate–contracte** le réseau $L(\ell,x_0)$ en un autre, le réseau $M(\ell,x_0)$, où les approximations sont égalisées : on peut maintenant, dans un tel réseau, utiliser la norme sup. Le réseau $M(\ell,x_0)$ est l'ensemble des vecteurs $t=(t_0,t_1,\ldots,t_\ell)$ qui satisfont

$$t_i = k_i\, w_i \text{ pour tout } i : 0 \leq i \leq \ell \text{ avec } w=(w_0,w_1,\ldots,w_\ell) \in L(\ell,x_0).$$

Afin de préserver le déterminant du réseau, on peut supposer, après une éventuelle homothétie, que

$$\prod_{i=0}^{\ell} k_i = 1$$

Le système de multiplicateurs k et le réseau $M(\ell,x_0)$ associés seront dits simples. Si nous cherchons les rationnels k_i sous la forme $k_i = n^{c_i}$ pour tout $i : 0 \leq i \leq \ell$, nous avons les conditions d'égalisations des approximations :

(3) $\quad c_i + (\ell-i)a_1 + ia_2$ indépendant de $i : 0 \leq i \leq \ell-1$ et égal à $c_\ell + b$

et la condition de simplicité

$$\sum_{i=0}^{\ell} c_i = 0. \tag{4}$$

Maintenant, nous transportons notre problème dans le réseau simple $M(\ell,x_0)$ dont la matrice est la suivante :

$$\begin{pmatrix} k_0 & 0 & 0 & \ldots & 0 & 0 \\ 0 & k_1 & 0 & \ldots & 0 & 0 \\ 0 & 0 & k_2 & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & k_{\ell-1} & 0 \\ k_\ell x_0^\ell & k_\ell C_\ell^1 x_0^{\ell-1} & k_\ell C_\ell^2 x_0^{\ell-2} & \ldots & k_\ell C_\ell^{\ell-1} x_0 & k_\ell n \end{pmatrix}$$

et nous cherchons un point $t=(t_0,t_1,\ldots,t_\ell)$ de ce réseau proche du point $m=(0,\ldots,k_\ell y_0)$ pour la norme sup.

2.3.– Des questions.

Nous sommes maintenant conduits à quelques questions importantes :

1) Comment trouver ce point ? Nous conviendra–t–il ? Plus précisément,

2) Ce point t sera–t–il suffisamment proche de $m = (0,0,...,0,k_\ell y_0)$ pour que les approximations exigées puissent être satisfaites ?

3) Ce point provient–il de notre problème : est–ce un point **puissance**, i.e. existe–t–il un triplet (u_1,u_2,v) vérifiant

$$(5) \qquad t_i = k_i\, u_2^i u_1^{\ell-i} \text{ pour tout } i : 0 \leq i \geq \ell-1 \text{ et aussi } t_\ell = k_\ell(y_0+v-u_2^\ell).$$

2.4.– L'algorithme PointProche.

Etant donnés dans $\mathbf{Q}^{\ell+1}$ un réseau L de rang $\ell+1$ et un point m, le problème consistant à trouver le point du réseau L le plus proche de m est NP–dur. Mais, il y a beaucoup d'algorithmes polynomiaux, utilisant la réduction LLL, qui permettent de trouver un point du réseau t proche de m

i) à un facteur $F(\ell)$ près,

ii) au sens de la norme euclidienne.

Ces deux restrictions nous poseront d'ailleurs quelques problèmes.

Nous utilisons ici l'algorithme le plus simple [1], qui donne aussi la plus mauvaise estimation pour $F(\ell)$

$$F(\ell) = O(\ell 2^\ell)$$

mais il est adéquat, puisque, de toute façon, nous devrons choisir ℓ petit par rapport à n. Nous décrivons ce dernier algorithme :

1) Réduire, par l'algorithme LLL, la base de départ de L en un base $e = (e_0,e_1,...,e_\ell)$.

2) Exprimer m dans la base e : $m = \Sigma_{i=0}^{\ell} m_i e_i$

3) Définir $t = \Sigma_{i=0}^{\ell} t_i e_i$ où t_i est l'entier le plus proche de m_i.

2.5.– Les réseaux ϵ–réguliers.

Par définition, un réseau L de rang $\ell+1$, dont le déterminant n est δ–square–free, est dit ϵ–régulier si et seulement si son premier minimum, noté $\lambda_1(L)$ satisfait

$$\|\lambda_1(L)\|_\infty \geq n^{(1-\delta-2\epsilon)/(\ell+1)}.$$

En comparant cette borne avec la moyenne géométrique des longueurs euclidiennes des minima –qui est de l'ordre de $n^{1/(\ell+1)}$–, et en utilisant l'inégalité de Minkowski, nous remarquons que tous les minima ont à peu près la même longueur. Tout ceci explique la terminologie et permet d'obtenir les faits géométriques suivants, à la fois théoriques et pratiques.
Soit $B(m,r)$ une boule euclidienne de centre m et de rayon r dans $\mathbf{Q}^{\ell+1}$; on pose

$$r_0 = n^{(1-\delta-2\epsilon)/(\ell+1)} \qquad r_1 = \frac{r_0}{F(\ell)} = n^{(1-\delta-\ell\epsilon)/(\ell+1)}$$
$$r_2 = (2\gamma_\ell)^{\ell/2}\, n^{1/(\ell+1)+\delta+2\epsilon} \quad (\gamma_\ell \text{ est la constante d'Hermite}) \qquad r_3 = F(\ell) r_2$$

et on a :

Unicité : if $r < r_0/2$, cette boule contient **au plus** un point du réseau,

Unicité effective : et si, de plus $r < r_1$, l'algorithme **PointProche** donne les résultats suivants : ce point, s'il existe, ou la réponse *aucun* sinon.

Existence : si $r > r_2$, cette boule contient **au moins** un point du réseau,

Existence effective : et si, de plus $r > r_3$, l'algorithme **PointProche** trouve un point du réseau dans cette boule.

2.6.– L'ϵ–régularité des réseaux $M(\ell,x_0)$.

Nous prouvons ici que la plupart des réseaux $M(\ell,x_0)$ sont ϵ–réguliers. Plus précisément, nous établissons le résultat suivant :

PROPOSITION. ***Pour*** $\epsilon > 0$, ***pour*** $\ell \geq 2$, ***pour*** n ***δ–square-free,*** $n \geq n_0(\ell,\epsilon)$, ***pour tout système simple de multiplicateurs*** k, ***il existe un sous-ensemble exceptionnel*** $S(\epsilon)$ ***de*** $Z(n)$, ***dépendant de*** k, ***pour lequel ce qui suit est vrai :***

(i) $|S(\epsilon)| \leq n^{1-\epsilon}$

(ii) ***pour tout*** x_0 ***n'appartenant pas à*** $S(\epsilon)$, ***le réseau simple*** $M(\ell,x_0)$ ***est ϵ-régulier.***

Nous donnons ici un abrégé de la preuve, qui suit essentiellement celle de FHKFS [4] sans cependant nécessiter le passage au réseau réciproque.
Pour un R fixé, nous voulons estimer la taille de l'ensemble

$$U(R) = \{x_0 \in Z(n)/\exists t \in M(\ell,x_0), \|t\|_\infty \leq R\}$$

Si $x_0 \in U(R)$, alors il existe $w \in L(\ell,x_0)$ vérifiant

$$|w_i| \leq \frac{R}{k_i} \text{ pour tout } i : 0 \leq i \leq \ell$$

et aussi

$$w_\ell \equiv \sum_{i=0}^{\ell-1} w_i C_\ell^i x_0^{\ell-i} \ [n].$$

Notons que le pgcd d de $(w_0,w_1,\ldots,w_\ell)$ divise n. Si nous posons

$$s_i = C_\ell^i w_i \text{ pour } i : 0 \leq i \leq \ell-1 \text{ et } s_\ell = -w_\ell$$

nous avons également

$$d = \text{pgcd}(s_0,s_1,\ldots,s_\ell)$$

et

$$x_0 \text{ est une racine modulo } n \text{ du polynôme } Q(X) = \sum_{i=0}^{\ell} s_i X^{\ell-i}.$$

Pour un diviseur fixé d de n, il y a au plus

$$\prod_{i=0}^{\ell} \frac{3R}{k_i d} = \left(\frac{3R}{d}\right)^{\ell+1}$$

tels polynômes.

Nous estimons maintenant le nombre de racines x_0 modulo n d'un tel polynôme. Si nous remplaçons n par n', quotient de n par d et s_i par s'_i, quotient de s_i par d, nous obtenons

$$\sum_{i=0}^{\ell} s'_i x_0^{\ell-i} \equiv 0 \ [n'].$$

Chaque solution modulo n' se relève en au plus d solutions modulo n et, alors, par le théorème chinois, il y a au plus [4]

$$\ell^f n^\delta \text{ solutions modulo } n'$$

sauf dans le cas $\ell = 2$, où nous pouvons négliger le terme n^δ.
Nous déduisons donc :

$$| U(R)| \leq (3R)^{\ell+1} \ell^f n^\delta \sum_{d| n} \frac{1}{d^\ell}.$$

Pour simplifier plus, nous utilisons deux estimations : d'abord,

$$\sum_{d| n} \frac{1}{d^\ell} \leq 2 \text{ pour } \ell \geq 2$$

et aussi l'estimation bien connue pour le nombre f des facteurs premiers distincts de n : il existe une constante c_0 telle que pour n suffisamment grand, on a :

$$f \leq c_0 \frac{\log n}{\log\log n}.$$

Et donc :

$$\ell^f \leq n^\epsilon \text{ pourvu que } \log n \geq .\ell^{c_0/\epsilon}$$

et aussi :

$$| U(R)| \leq (3R)^{\ell+1} n^{\delta+\epsilon}.$$

Si nous choisissons $3R = n^{(1-\delta-2\epsilon)/(\ell+1)}$, nous avons $| U(R)| \leq n^{1-\epsilon}$. Nous définissons alors $S(\epsilon)$ comme étant égal à $U(R)$ pour cette valeur particulière de R et ceci termine la preuve.

2.7.– **Un autre ensemble exceptionnel $T(\epsilon)$.**
Trouver le vecteur le plus court du réseau $M(\ell,x_0)$ est presque certainement un problème NP–dur, et donc nous ne pouvons savoir si nous sommes ou non dans l'ensemble exceptionnel $S(\epsilon)$: nous considérons donc un autre ensemble exceptionnel $T(\epsilon)$ avec une définition plus algorithmique.
Définissant $e_0(x_0)$ comme le premier vecteur de la base réduite au sens de LLL de $L(\ell,x_0)$, nous savons que [5]

$$| e_0(x_0)| \leq 2^{\ell/2} | \lambda_1(x_0)| .$$

Si nous posons

$$T(\epsilon) = \{x_0 \in Z(n) / | e_0(x_0)| \leq 2^{\ell/2} n^{(1-\delta-2\epsilon)/(\ell+1)}\}$$

nous pouvons travailler avec $T(\epsilon)$ de la même manière que $S(\epsilon)$ pourvu que

$$2^{\ell(\ell+1)/2} \text{ soit petit par rapport à } n^{\epsilon}.$$

3.– DES RÉSEAUX AUX NOMBRES.

Nous revenons à notre problème original. D'abord nous choisissons un réseau convenable dans lequel nous pouvons appliquer les résultats de la section précédente : nous prouvons alors les principaux résultats (Théorèmes 1 et 2). Puis, nous précisons les conditions sur les paramètres et déterminons un choix optimal sur ceux–ci, que nous utilisons dans la preuve du Théorème 4. Après quoi, nous décrivons les particularités du cas $\ell = 2$ qui nous permettent de démontrer le Théorème 3. Nous terminons en donnant une autre attaque du second cryptosystème d'Okamoto qui étend la seconde attaque de Shamir.

3.1.– Le choix de la dilatation–contraction $M(\ell,x_0)$ du réseau $L(\ell,x_0)$.

Nous choisissons maintenant la dilatation–contraction afin de pouvoir utiliser la seconde conséquence de la ϵ–régularité des réseaux $M(\ell,x_0)$: une paire de compatibilité doit donner naissance à un vecteur t de $M(\ell,x_0)$ dans la boule $B(m,r_1)$.

Si nous revenons aux conditions d'égalisation (3), nous demandons

$$c_i + (\ell-i)a_1 + ia_2 = \frac{1-\delta-\ell\epsilon}{\ell+1} \text{ pour tout } i : 0 \leq i \leq \ell-1 \text{ et aussi } c_\ell + b = \frac{1-\delta-\ell\epsilon}{\ell+1}.$$

Puisque le système de multiplicateurs est simple, nous avons (4)

$$\sum_{i=0}^{\ell} c_i = 0 = 1 - \delta - \ell\epsilon - (a_1\frac{\ell(\ell+1)}{2} + a_2\frac{\ell(\ell-1)}{2} + b).$$

Les conditions (3) sont donc satisfaites si et seulement si le triplet (a_1,a_2,b) satisfait les conditions $C(\ell,\delta,\epsilon)$ définies en (2).

3.2.– Les preuves des principaux résultats, les Théorèmes 1 et 2.

Nous revenons maintenant au problème initial. Nous fixons ϵ et un triplet (a_1,a_2,b) qui satisfait $C(\ell,\delta,\epsilon)$; nous considérons $x_0 \notin S(\epsilon)$ et (x,y) une éventuelle paire de compatibilité pour les deux sous–ensembles $K(a_1,a_2,x_0)$ et $I(b,y_0)$.

Nous associons à cette paire le point **puissance** w de $L(\ell,x_0)$, puis le point correspondant t de notre réseau particulier $M(\ell,x_0)$ défini en (5).

Ce point puissance t est dans la boule euclidienne $B(m,r_0/2)$ qui contient au plus un point de $M(\ell,x_0)$, et a fortiori au plus un point puissance de $M(\ell,x_0)$. Chaque point puissance provient de deux triplets jumeaux (u_1,u_2,v) (si ℓ est pair) ou d'un triplet (si ℓ est impair) et nous calculons aisément ces triplets en extrayant, dans les entiers, des racines ℓ–ièmes en (5). Après quoi, il reste à vérifier que les approximations exigées sont bien vérifiées. Tout ceci démontre le principal résultat théorique (Théorème 1).

Maintenant, nous adoptons un point de vue constructif pour la preuve du Théorème 2 : Sous les mêmes hypothèses, le point puissance t, si il existe, est dans la boule euclidienne $B(m,r_1)$ et c'est le seul point du réseau dans la boule $B(m,r_0/2)$. L'algorithme **PointProche** trouve un point du réseau t' et nous avons donc $t = t'$: nous avons trouvé t. Si $B(m,r_1)$ ne contient aucun point puissance, il se peut que l'algorithme **PointProche** trouve un autre point du réseau, dont on vérifie qu'il ne peut convenir.

3.3.– Quelques précisions sur le choix de n, ℓ, ϵ.

A trois reprises, nous avons eu besoin de conditions sur ce triplet :

1) $c_0 \log \ell \leq \epsilon \operatorname{loglog} n$ (dans la preuve 2.6)

2) $\ell(\ell+1) \leq \epsilon \log n$ (dans la remarque 2.7)

3) $\log F(\ell) \leq \epsilon \log n$ avec $F(\ell) = O(\ell 2^{\ell})$ (dans la description 2.4).

Nous devrons donc choisir

$$\log \ell \leq K \epsilon \operatorname{loglog} n \quad \text{i.e.} \quad n \geq n_0(\ell,\epsilon)$$

où K est une constante à déterminer en fonction de c_0.

3.4.– Le choix optimal des paramètres.

Nous revenons aux sous–ensembles particuliers $H(a,x_0)$, $I(a,x_0)$ et $J(a,x_0)$, et les conditions $C(\ell,\delta,\epsilon)$ peuvent s'écrire dans chacun des cas :

$$\text{Cas inhomogène : } \frac{\ell(\ell-1)}{2}\,a + b = 1 - \delta - \ell\epsilon \text{ et } b \geq \ell a$$

$$\text{Cas homogène : } \frac{\ell^2}{2}\,a + b = 1 - \delta - \ell\epsilon \text{ et } b \geq \ell\frac{a}{2}.$$

Dans chacun de ces cas, le choix optimal (a_0,b_0) pour la paire (a,b) est le suivant :

$$\text{Pour les deux cas : } a_0 = 2\,\frac{1-\delta-\ell\epsilon}{\ell(\ell+1)}.$$

Dans le premier cas : $b_0 = 2\,\frac{1-\delta-\ell\epsilon}{\ell+1}$. Dans le second cas : $b_0 = \frac{1-\delta-\ell\epsilon}{\ell+1}$.

3.5.– Une description précise de l'Algorithme DevineRacine.

Entrée : Deux sous–ensembles $K(a_1,a_2,x_0)$ et $I(b,y_0)$ de $Z(n)$, un entier $\ell \geq 2$ et une estimation de δ.

Sortie : deux booléens *prudent* et *compatible* et, si elle existe une paire de compatibilité (x,y)

1) Déterminer $\epsilon > 0$ tel que le triplet (a_1,a_2,b) vérifie les conditions $C(\ell,\delta,\epsilon)$ et vérifier que le triplet (n,ℓ,ϵ) satisfait les conditions 3.3. Sinon, répondre "Mauvais choix des paramètres" et finir.

2) Choisir le système k des multiplicateurs comme en 3.1, calculer le point m et réduire le réseau associé $M(\ell,x_0)$; *prudent* := faux; *compatible* := faux;

3) Si x_0 est dans $T(\epsilon)$ alors *prudent* := vrai.

4) t := **PointProche** $(m,M(\ell,x_0))$

   si t est un point puissance alors

   calculer le triplet (u_1,u_2,v).

   Si les approximations demandées sont satisfaites, alors *compatible* := vrai.

5) Si non *compatible et prudent* alors répondre

   "Echec de l'algorithme";

Si non *compatible* et non *prudent* alors écrire

"Pas de paire de compatibilité";

Si *compatible* alors écrire $x = u_1x_0 + u_2, y = y_0 + v$.

3.6.– La répartition comparée des racines ℓ–ièmes et des puissances ℓ–ièmes.

Nous rappelons tout d'abord le résultat de Blum [2] :

Soit $n \geq 5$ un nombre square-free. La probabilité que $I(1/2,x_0)$ et $\{x_0^2\}$ aient une paire de compatibilité non triviale est inférieure à $n^{-\frac{1}{2}}48$ loglog n.

Ce résultat doit être comparé avec le nôtre :

Soient n un nombre δ-square-free, $\ell \geq 2$ un entier et $(0,a,b)$ un triplet qui vérifie $C(\ell,\delta,\epsilon)$. La probabilité que $I(a,x_0)$ et $I(b,x_0^\ell)$ aient une paire de compatibilité non triviale est inférieure à $n^{-\epsilon}$.

Nous pouvons interpréter l'ensemble exceptionnel $S(\epsilon)$.

Si le triplet $(0,a,b)$ vérifie $C(\ell,\delta,\epsilon)$, nous remarquons le fait suivant :

Si (x_1,y_1) est une paire de compatibilité de $I(a,x_0) \times I(b,x_0^\ell)$ et si x_0 est un élément de $S(\epsilon)$, alors tout l'intervalle $[x_0,x_1]$ est inclus dans $S(\epsilon)$.

$S(\epsilon)$ nous empêche de séparer les racines, car il contient des intervalles où les racines s'empilent.

3.7.– Le cas particulier $\ell = 2$; la preuve du Théorème 3.

Ici, nous considérons le cas de deux intervalles $I(a,x_0)$ et $I(b,y_0)$. L'étude de leur compatibilité nous emmène dans la section du réseau $L(\ell,x_0)$ par l'hyperplan $w_0 = 1$ que nous appelons $L'(\ell,x_0)$. L'analyse de la ϵ–régularité du réseau associé $M'(\ell,x_0)$, qui est de rang ℓ, est similaire à celle de 2.6 et nous pouvons négliger le terme ν^δ si $\ell = 2$. Dans ce cas particulier, nous voulons trouver de "petites" solutions (u,v) à l'équation :

$$2x_0\, u - (v - u^2) \equiv y_0 - x_0^2\ [n].$$

Le réseau $M'(\ell,x_0)$ a la matrice suivante

$$\begin{pmatrix} k_1 & 0 \\ 2k_2x_0 & k_2n \end{pmatrix}$$

et nous devons trouver, dans $M'(\ell,x_0)$, un point "proche" de $m = (0,k_2(y_0 - x_0^2))$.
Nous observons d'abord un fait tout à fait particulier au cas $\ell = 2$:
Etant donnés $w = (w_1,w_2)$ et (x_0,y_0), le système suivant a toujours une solution (u,v) :

$$w_1 = u, \quad w_2 = y_0 - x_0^2 + v - u^2.$$

Tous les points de $M'(\ell,x_0)$ proviennent du problème : ce sont des points puissance !
D'autre part, nous pouvons choisir les multiplicateurs (k_1,k_2) de sorte à pouvoir utiliser les troisième et quatrième conséquences de la ϵ–régularité de $M'(\ell,x_0)$: nous exigeons

$$c_1 = \frac{1}{2} - a + 2\epsilon, \quad c_2 = \frac{1}{2} - b + 2\epsilon, \quad \text{et} \quad c_1 + c_2 = 0,$$

ce qui est possible par hypothèse.
Nous obtenons ainsi un point t de $M'(\ell,x_0)$, puis un point w de $L'(\ell,x_0)$ et résolvons le système précédent afin d'obtenir notre paire de compatibilité. Tout ceci termine la preuve du Théorème 3.

3.8.– Une extension de la seconde attaque de Shamir.
Okamoto pensait que la première attaque de Shamir [9] ne pouvait s'appliquer au cas homogène. Nous avons montré ici comment une extension des idées de Shamir pouvait s'appliquer au cas homogène. De plus, nous montrons que la seconde attaque de Shamir [9] peut s'étendre, elle aussi, au cas homogène.
Nous rappelons le cadre général de la seconde version du cryptosystème d'Okamoto [8] : Nous considérons un nombre δ–monosquare et deux nombres faciles x_1 et x_2 ; nous posons $x_0 = x_1/x_2$ et nous voulons retrouver la solution x dans $J(a,x_0)$ de l'équation : $x^\ell \equiv y_0\ [n]$. Plus précisément, nous cherchons un x sous la forme $x = u_1x_0 + u_2$, avec les conditions $0 < u_1,u_2 < n^c$ et $c < (1-\delta)/2\ell$.
Nous connaissons les quantités n, y_0, $x_0 = x_1/x_2$. Les facteurs p et q de n et la paire (x_1,x_2) nous sont cachés. De plus, puisque les x_i sont faciles, nous avons

$$x_i = y_i + z_ipq\ [n] \text{ avec } 0 < y_i < \frac{1}{2}\, n^c \text{ et aussi } 0 < z_i < p.$$

La clef secrète est (y_1,y_2,z_1,z_2,p,q).

Nous montrons d'abord que nous pouvons aisément retrouver le triplet (y_1,y_2,p). En effet,

$$(x_2p)x_0 \equiv (x_1p)\ [n]$$

et puisque x_i est facile, tous les multiples $(x_ik)\ [n]$ de x_i (pour $0 < k < p$) sont de l'ordre de $n^{1-\delta}$ tandis que le multiple particulier $x_ip\ [n] = y_ip$ est inférieur à $n^{c+\delta}$.

Pourvu que $c + \delta$ soit inférieur à $(3/4)(1-\delta)$ et à $1/2$, ce dernier multiple est nettement inférieur aux autres et suffisamment petit pour pouvoir être trouvé au moyen de l'algorithme d'Euclide. Ces conditions sont équivalentes à

$$\delta < \min\left(\frac{3\ell-2}{7\ell-2}, \frac{\ell-1}{2\ell-1}\right),$$

La dernière inégalité est d'ailleurs toujours satisfaite dans le cas usuel où $\delta \leq \frac{1}{3}$.

Nous déterminons chacun de deux y_1p and y_2p et, après un calcul de pgcd, chacun des trois termes y_1,y_2,p.

Comme $pq = n/p$ est désormais connu, nous considérons notre équation modulo pq et nous obtenons

$$(u_1y_1 + u_2y_2)^\ell \equiv y_0y_1^\ell\ [pq].$$

Mais, d'après l'hypothèse d'Okamoto,

$$0 < u_1y_1 + u_2y_2 < n^{(1-\delta)/\ell}$$

et, en extrayant une racine ℓ–ième dans les entiers, nous retrouvons la valeur A de la quantité $u_1y_1 + u_2y_2$. Il reste à trouver de petites solutions à cette équation linéaire, ce qui peut être fait grâce à l'algorithme d'Euclide à cause de l'ordre de grandeur de u_1 et u_2.

4.– DES EXTENSIONS ET DES PROBLÈMES OUVERTS.

4.1.– Une extension facile.

Il est clair que notre méthode permet de trouver une racine x de n'importe quel polynôme P de degré ℓ, à condition que nous connaissions une approximation suffisante x_0 de cette racine. Nous devons résoudre

$$P(u_1x_0 + u_2) \equiv y_0 + v\ [n].$$

A la place de la formule du binôme, nous utilisons la formule de Taylor

$$x_0^{\ell}u_1^{\ell}\frac{P^{(\ell)}(u_2)}{\ell!} + x_0^{\ell-1}u_1^{\ell-1}\frac{P^{(\ell-1)}(u_2)}{(\ell-1)!}+\ldots+x_0^{\ell-i}u_1^{\ell-i}\frac{P^{(\ell-i)}(u_2)}{(\ell-i)!}+\ldots+P(u_2)-v\equiv y_0[n].$$

Nous remplaçons la condition (5) par la suivante

$$t_i = k_i\frac{P^{(\ell-i)}(u_2)}{(\ell-i)!}u_1^{\ell-i} \text{ pour tout } i : 0\leq i\leq \ell 61 \text{ et aussi } t_{\ell} = k_{\ell}(y_0+v-P(u_2))$$

et nous obtenons le même genre d'égalisations d'approximations.

Dans le cas des approximations homogènes, nous pouvons aussi travailler dans un autre réseau : à la place de $L(\ell,x_0)$, nous utilisons un réseau dont la matrice contient en dernière ligne :

$$(P(x_0),\ P'(x_0),\ \frac{P''(x_0)}{2!},\ldots,\ \frac{P^{(i)}(x_0)}{i!},\ldots,\ \frac{P^{(\ell-1)}(x_0)}{(\ell-1)!},\ n).$$

L'analyse des propriétés géométriques de ces réseaux, lorsque P est fixé et x_0 varie, est la même qu'en 2.6; à la place des polynômes Q précédents, nous étudions des polynômes Q qui peuvent s'écrire comme de petites combinaisons linéaires de

$$P(x_0),\ P'(x_0),\ \frac{P''(x_0)}{2!},\ldots,\ \frac{P^{(i)}(x_0)}{i!},\ldots,\ \frac{P^{(\ell-1)}(x_0)}{(\ell-1)!}.$$

4.2.– Des algorithmes quasi–uniformes pour trouver de petis résidus quadratiques et une application à la factorisation entière.

Trouver de petits résidus quadratiques modulo n, lorsque n est une grand nombre composé de factorisation inconnue, est presque certainement un problème difficile. Nous décrivons dans [11] des algorithmes polynomiaux, fondés sur des idées

similaires à celles que nous avons développés ici, qui trouvent, de manière quasi–uniforme, des éléments x de $Z(n)$ dont les carrés modulo n sont inférieurs à $n^{2/3}$. Nous décrivons aussi l'application de tels algorithmes à la factorisation entière : nous obtenons ainsi un algorithme de factorisation entière avec la meilleure borne prouvée à ce jour.

4.3.– Une question ouverte au sujet de $S(\epsilon)$.

Nous posons maintenant une question :

Peut–on décrire précisément l'ensemble exceptionnel $S(\epsilon)$? Cette description, dans le cadre de 3.7, est faite en [11]. Comment la généraliser ? L'étude de cet ensemble peut–elle nous aider dans la localisation des facteurs de n ?

4.4.– Une question ouverte à propos de la probabilité de compatibilité.

Soient a et b deux réels de [0,1]. Nous définissons

$$Q(a,b,\ell,) = \{(x_0,y_0) \in Z(n)^2 / I(a,x_0) \text{ et } I(b,y_0) \text{ sont } \ell\text{–compatibles}\}.$$

Quelle est la probabilité $q(a,b,\ell)$ de l'évènement $Q(a,b,\ell)$?

Dans le cas particulier $\ell = 2$, nous avons démontré que

$$q(a,b,\ell) \geq 1 - \frac{1}{n^{\epsilon}}$$

pourvu que

$$a + b = 1 + 4\epsilon \text{ et } b \geq 2a.$$

Un argument heuristique simple pourrait généraliser ce résultat au cas où la fonction $x \longmapsto x^{\ell}$ est "ℓ–to–one" pour chaque p_i (cette conditioon est équivalente à $p_i \equiv 1[\ell]$).

Nous posons donc une question ouverte :

Est-il vrai que $q(a,b,\ell)$ est peu différente de 1 quand $a + b$ est proche de 1 ?

BIBLIOGRAPHIE

[1] L. Babai.– On Lovasz's lattice reduction and the nearest lattice point problem, Combinatorica 6, 1–14.

[2] M. Blum.– How to exchange (secret) keys, ACM transactions on Computer systems, 1, 2, may 83, 175–193.

[3] E. Brickell, J. Delaurentis.– An attack on a signature scheme proposed by Okamoto and Shiraishi, Proc of Crypto'85, 1–4.

[4] A. Frieze, J. Hastad, R. Kannan, J.–C. Lagarias, A. Shamir.– Reconstructing truncated variables satisfying linear congruences, to appear in SIAM Journal of Computing.

[5] A.K. Lenstra, H.W. Lenstra, L. Lovasz.– Factoring polynomials with integer coefficients, Mathematische Annalen, 261, (1982), 513–534.

[6] T. Okamoto, A. Shiraishi.– A fast signature scheme based on quadratic inequalities, Proc. of the 1985 Symposium on Security and Privacy, April 1985, Oakland, CA.

[7] T. Okamoto.– Fast public–key cryptosystem using congruent polynomial equations, Electronics Letters, 1986, 22, 581–582.

[8] T. Okamoto.– Modification of a public–key cryptosystem, Electronics Letters, 1987, 23, 814–815.

[9] A. Shamir.– Private communications to Okamoto, August and October 1986 (cité dans Okamoto [8]).

[10] B. Vallée, M. Girault, Ph. Toffin.– How to break Okamoto's cryptosystems by reducing lattice bases, Proceedings of Eurocrypt'87, Lecture notes in Computer Science.

[11] B. Vallée.– Quasi–uniform algorithms for finding small quadratic residues and application to integer factorisation, ou Factorisation entière par génération quasi–uniforme de petits résidus quadratiques, Rapport de Recherche 1988–1 du Laboratoire A3L de l'Université de Caen.

Brigitte VALLÉE
Philippe TOFFIN
Département de Mathématiques
Université de Caen
14032 CAEN CEDEX

Marc GIRAULT
SEPT
BP 6243
14066 CAEN CEDEX