BULLETIN DE LA S. M. F.

JEAN-PAUL ALLOUCHE

Somme des chiffres et transcendance

Bulletin de la S. M. F., tome 110 (1982), p. 279-285

<http://www.numdam.org/item?id=BSMF_1982__110__279_0>

Bull. Soc. math. France,
110, 1982, p. 279-285.

SOMME DES CHIFFRES ET TRANSCENDANCE

PAR

JEAN-PAUL ALLOUCHE (*)

RÉSUMÉ. — Soit p un nombre premier, et soit $s_p(n)$ la somme des chiffres de n en base p.

Nous montrons les résultats suivants :

Si R est un polynôme de $\mathbb{Q}[X]$ tel que $R(\mathbb{N}) \subset \mathbb{N}$, la série formelle $\sum_{n=0}^{\infty} s_p(R(n)) X^n$ est algébrique sur $\mathbb{F}_p(X)$ si et seulement si le degré de R est inférieur ou égal à 1.

Soit $s_p^{(k)}$ la k-ième itérée de s_p, alors la série formelle $\sum_{n=0}^{\infty} s_p^{(k)}(n) X^n$ est algébrique sur $\mathbb{F}_p(X)$ si et seulement si $k=0$ ou $k=1$.

ABSTRACT. — Let p be a prime number, and let $s_p(n)$ be the sum of the digits of n in the p-adic expansion.

We prove the following results:

Let R be a polynomial in $\mathbb{Q}[X]$ with $R(\mathbb{N}) \subset \mathbb{N}$. The formal series $\sum_{n=0}^{\infty} s_p(R(n)) X^n$ is algebraic over $\mathbb{F}_p(X)$ if and only if the degree of R is less than or equal to 1.

Let $s_p^{(k)}$ be the iterates of s_p, then the formal series $\sum_{n=0}^{\infty} s_p^{(k)}(n) X^n$ is algebraic over $\mathbb{F}_p(X)$ if and only if $k=0$ or $k=1$.

Soit (ε_n) une suite d'entiers relatifs; on note $\mathbb{F}_p$ le corps à p éléments et $\mathbb{F}_p(X)$ le corps des fractions rationnelles à coefficients dans $\mathbb{F}_p$. Un résultat récent de G. CHRISTOL, T. KAMAE, M. MENDÈS FRANCE et G. RAUZY (*cf.* [1]) établit notamment que la série formelle $\sum_0^{\infty} \varepsilon_n X^n$ est algébrique sur $\mathbb{F}_p(X)$ si et seulement si la suite (ε_n) réduite modulo p est reconnue par un p-automate, ce qui est une façon de concevoir que la suite (ε_n) n'est pas répartie « au hasard ».

Si l'on note $s_p(n)$ la somme des chiffres de n en base p, on sait que la série $\sum_0^{\infty} s_p(n) X^n$ est algébrique sur $\mathbb{F}_p(X)$; il est donc intéressant de considérer par exemple les séries $\sum_0^{\infty} s_p(n^2) X^n$ ou $\sum_0^{\infty} s_p^{(2)}(n) X^n$, (où $s_p^{(k)}$ est la k-ième itérée de s_p) : les suites $(s_p(n^2))$ et $(s_p^{(2)}(n))$ sont données par des calculs effectifs (algorithmes?) mais nous allons démontrer qu'elles sont réparties « au

(*) Texte reçu le 17 mars 1981, révisé le 25 janvier 1982.

Jean-Paul ALLOUCHE, Laboratoire associé au C.N.R.S. n° 226, U.E.R. de Mathématiques et d'Informatique, Université de Bordeaux-I, 33405 Talence Cedex.

BULLETIN DE LA SOCIÉTÉ MATHÉMATIQUE DE FRANCE — 0037-9484/1982/279/$ 5.00

hasard » au sens défini plus haut; ces suites sont donc plus « complexes » que $(s_p(n))$; est-ce la raison qui rend difficile la question posée par Gelfond dans [2] : soit R un polynôme à coefficients entiers, estimer :

$$\text{card}\{n \leqslant x,\ s_p(R(n)) \equiv l \bmod m\}?$$

Cet article est divisé en cinq parties; dans la première nous étudions la série $\sum_0^\infty s_p(\alpha n+\beta) X^n$; dans la seconde nous établissons un lemme fondamental sur les propriétés de la suite (ε_n) lorsque la série $\sum_0^\infty \varepsilon_n X^n$ est algébrique sur $\mathbb{F}_p(X)$; dans la troisième, nous montrons que, si le polynôme R est dans $\mathbb{N}[X]$ et si la série $\sum_0^\infty s_p(R(n)) X^n$ est algébrique sur $\mathbb{F}_p(X)$, alors « il y a des relations » entre les sommes des chiffres de certains coefficients de R; dans la quatrième nous montrons que la série $\sum_0^\infty s_p(R(n)) X^n$ est algébrique sur $\mathbb{F}_p(X)$ si et seulement si le degré du polynôme R est inférieur ou égal à 1; dans la dernière partie nous prouvons que la série $\sum s_p^{(k)}(n) X^n$ est algébrique sur $\mathbb{F}_p(X)$ si et seulement si k vaut 0 ou 1.

1. Étude de la série formelle $\sum_0^\infty s_p(\alpha n+\beta) X^n$

1. THÉORÈME. — *Soit $(\varepsilon(n))$ une suite d'entiers telle que la série $\sum_0^\infty \varepsilon(n) X^n$ soit algébrique sur $\mathbb{F}_p(X)$, alors quels que soient α et β dans $\mathbb{N}$, la série $\sum_0^\infty \varepsilon(\alpha n+\beta) X^n$ est algébrique sur $\mathbb{F}_p(X)$.*

Posons d'abord $\beta = \alpha q + r$, $0 \leqslant r < \alpha$, alors $\varepsilon(\alpha n+\beta) = \varepsilon(\alpha(n+q)+r)$, et la série $\sum_0^\infty \varepsilon(\alpha n+\beta) X^n$ est algébrique sur $\mathbb{F}_p(X)$ si et seulement si la série $\sum \varepsilon(\alpha n+r) X^n$ l'est.

On pose alors $t(n) = \varepsilon(\alpha n+r)$ et on va montrer qu'il y a un nombre fini de suites $t_{k,a}(n) = t(p^k n+a)$ modulo p, $k \geqslant 0$, $0 \leqslant a \leqslant p^k-1$ ce qui montrera le résultat d'après [1].

Écrivons $\alpha a + r = up^k + v$, avec $0 \leqslant v \leqslant p^k-1$, alors u appartient à $[0, \alpha-1]$ et $t(p^k n+a) = \varepsilon((\alpha n+u)\, p^k+v) = \varepsilon_{k,v}(\alpha n+u)$. Comme la série $\sum_0^\infty \varepsilon(n) X^n$ est algébrique sur $\mathbb{F}_p(X)$, il y a un nombre fini de suites $\varepsilon_{k,v}(n) = \varepsilon(p^k n+v)$ modulo p, $k \geqslant 0$, $0 \leqslant v \leqslant p^k-1$, donc un nombre fini de suites $\varepsilon_{k,v}(\alpha n+u)$ pour chaque u.

Mais u décrit $[0, \alpha-1]$, il y a donc un nombre fini de suites $t_{k,a}(n)$ modulo p, $0 \leqslant k$, $0 \leqslant a \leqslant p^k-1$ et le résultat est démontré.

2. COROLLAIRE. — *Soit $R(X)$ un polynôme du premier degré dans $\mathbb{Q}[X]$, $R(X) = aX+b$, tel que $R(\mathbb{N})$ soit contenu dans $\mathbb{N}$, alors la série $\sum_0^\infty s_p(R(n)) X^n$ est algébrique sur le corps $\mathbb{F}_p(X)$.*

L'hypothèse $R(\mathbb{N})\subset\mathbb{N}$ implique facilement $a,\ b\in\mathbb{N}$. Il suffit donc de montrer que $\sum_0^\infty s_p(n)X^n$ est algébrique sur $\mathbb{F}_p(X)$; rappelons une démonstration de ce résultat :

$$\begin{aligned}G(X)=\sum_0^\infty s_p(n)X^n&=\sum_{a=0}^{p-1}\sum_{n=0}^{+\infty}s_p(pn+a)X^{pn+a}\\&=\sum_{a=0}^{p-1}X^a\sum_{n=0}^{+\infty}(s_p(n)+a)X^{np}\\&=\sum_{a=0}^{p-1}X^a(G(X))^p+\sum_{a=0}^{p-1}aX^a\sum_{n=0}^{\infty}X^{np},\end{aligned}$$

d'où :

$$G(X)=\frac{X^p-1}{X-1}G^p(X)+\frac{X}{(X-1)^2}.$$

2. Propriétés des coefficients d'une série algébrique

LEMME FONDAMENTAL. — *Soit $\varepsilon(n)$ une suite d'entiers telle que la série $\sum_0^\infty\varepsilon(n)X^n$ soit algébrique sur $\mathbb{F}_p(X)$; alors il existe $n_0\geqslant 1$ et $T\geqslant 1$ tels que :*

$$\forall j\geqslant 1,\quad \forall i\in[1,\ p],\quad \forall n\geqslant n_0,$$

$$\varepsilon\left(p^n\left(j-\frac{p-i}{p-1}\right)+\frac{p-i}{p-1}\right)\equiv\varepsilon\left(p^{n+T}\left(j-\frac{p-i}{p-1}\right)+\frac{p-i}{p-1}\right)\bmod p.$$

La série $\sum_0^{+\infty}\varepsilon(n)X^n$ étant algébrique sur $\mathbb{F}_p(X)$, il en est de même de la série $\sum_0^\infty\varepsilon(n+1)X^n$.

D'après [1] il existe un alphabet $Y=\{y_1,\ \ldots,\ y_l\}$, une p-substitution σ sur Y, une suite $(\lambda(n))_{n\geqslant 1}$ d'éléments de Y, et une application φ de Y vers $\{0,\ \ldots,\ p-1\}$ tels que :

- la suite $(\lambda(n))$ est point fixe de σ;
- pour $n\geqslant 1$, $\varphi(\lambda(n))\equiv\varepsilon(n)\bmod p$.

Il suffit donc de montrer le résultat annoncé dans le lemme pour la suite $\lambda(n)$ (avec $=$ au lieu de $\equiv$).

Soit $\sigma=(\tau_1,\ \ldots,\ \tau_p)$, où les τ_j sont des applications de Y dans lui-même, c'est-à-dire $\sigma(y)=\tau_1(y)\tau_2(y)\ldots\tau_p(y)$. Dire que $\lambda(n)$ est point fixe de σ, c'est dire que les suites $\lambda(1)\ldots\lambda(n)\ldots$ et $\tau_1(\lambda(1))\ldots\tau_p(\lambda(1))\tau_1(\lambda(2))\ldots\tau_p(\lambda(2))\ldots$ sont égales, autrement dit :

$$\forall i\in[1,\ p],\quad \forall j\geqslant 1,\qquad \tau_i(\lambda(j))=\lambda(i+p(j-1)),$$

d'où en itérant τ_i :

$$\tau_i^a(\lambda(j))=\lambda\left(p^a\left(j-\frac{p-i}{p-1}\right)+\frac{p-i}{p-1}\right).$$

Comme τ_i est une application de Y dans Y et que Y a l éléments, il existe une période $T_{ij} \leqslant l$ telle que :

$$\forall a \geqslant l+1, \qquad \tau_i^a(\lambda(j)) = \tau_i^{a+T_{i,j}}(\lambda(j)).$$

Posons $T = \text{ppcm}(1, 2, \ldots, l)$, alors T_{ij} divise T, d'où :

$$\forall i \in [1, p], \quad \forall j \geqslant 1, \quad \forall n \geqslant n_0 = l+1,$$
$$\lambda\left(p^n\left(j - \frac{p-i}{p-1}\right) + \frac{p-i}{p-1}\right) = \lambda\left(p^{n+T}\left(j - \frac{p-i}{p-1}\right) + \frac{p-i}{p-1}\right).$$

3. Relations entre les coefficients de R si la série $\sum_0^\infty s_p(R(n))X^n$ est algébrique

PROPOSITION. — *Soit $R(X) = \sum_0^d a_k X^k$, avec $a_k \in \mathbb{N}$, $a_d \neq 0$, $d \geqslant 2$. Si la série $\sum_0^\infty s_p(R(n))X^n$ est algébrique sur $\mathbb{F}_p(X)$, alors :*

$$s_p(a_{d-1} + (d+1)a_d) \equiv s_p(a_{d-1} + da_d) + s_p(a_d) \bmod p.$$

Appliquons en effet le lemme de la deuxième partie, avec $i=1$, $j = p^{(d-1)n} + p^{(d-2)n} + 1$, à la suite $s_p(R(n))$: il existe $n_0 \geqslant 1$ et $T \geqslant 1$ tels que :

$$\forall n \geqslant n_0,$$
$$s_p(R(p^n(p^{(d-1)n} + p^{(d-2)n}) + 1)) \equiv s_p(R(p^{n+T}(p^{(d-1)n} + p^{(d-2)n}) + 1)).$$

Posons $b_{j,l} = \sum_{k=j}^d \binom{k}{j}\binom{j}{l} a_k$.

Calculons alors, pour $x = n$, puis $x = n+T$, la quantité :

$$\begin{aligned} f(x) &= s_p[R(p^x(p^{(d-1)n} + p^{(d-2)n}) + 1)] \\ &= s_p[R(p^{x+(d-2)n}(p^n+1) + 1)] \\ &= s_p\left[\sum_{k=0}^d a_k \sum_{j=0}^k \binom{k}{j} p^{(x+(d-2)n)j}(p^n+1)^j\right] \\ &= s_p\left[\sum_{k=0}^d \sum_{j=0}^k \sum_{l=0}^j a_k \binom{k}{j}\binom{j}{l} p^{(x+(d-2)n)j+nl}\right] \\ &= s_p[\textstyle\sum_{j=0}^d \sum_{l=0}^j b_{j,l}\, p^{(x+(d-2)n)j+nl)}] \end{aligned}$$

pour $x = n$:

$$f(n) = s_p(\textstyle\sum_{j=0}^d \sum_{l=0}^j b_{j,l}\, p^{(d-1)nj+nl}),$$

deux exposants $(d-1)nj+nl$ et $(d-1)nj'+nl'$, avec $(j, l)\neq(j', l')$ et par exemple $l'\geqslant l$, sont égaux si et seulement si $(d-1)(j-j')=l'-l$, avec $0\leqslant l\leqslant j\leqslant d$ et $0\leqslant l'\leqslant j'\leqslant d$, ce qui donne l'unique solution $l'=d-1$, $l=0$, $j'=d-1$, $j=d$. Par conséquent,

$$f(n)=s_p[\sum{}' b_{j,l}\,p^{(d-1)nj+nl}+(b_{d,0}+b_{d-1,d-1})\,p^{(d-1)dn}],$$

où $\sum'$ porte sur j, l avec $0\leqslant j\leqslant d, 0\leqslant l\leqslant j, (j, l)\notin\{(d, 0), (d-1, d-1)\}$ et où tous les exposants de p sont distincts donc diffèrent d'au moins n; si l'on choisit n assez grand ($p^n>2\sup_{j,l}(b_{j,l})$), on a donc :

$$\begin{aligned}f(n)&=\sum{}' s_p(b_{j,l})+s_p(b_{d,0}+b_{d-1,d-1})\\&=\sum_{j=0}^{d}\sum_{l=0}^{j}s_p(b_{j,l})-s_p(b_{d,0})-s_p(b_{d-1,d-1})+s_p(b_{d,0}+b_{d-1,d-1}).\end{aligned}$$

pour $x=n+T$:

$$f(n+T)=s_p[\sum_{j=0}^{d}\sum_{l=0}^{j}b_{j,l}\,p^{((d-1)n+T)j+nl}].$$

Ici, tous les exposants de p sont distincts si T est convenablement choisi et si n est suffisamment grand : en effet, si

$$((d-1)n+T)\,j+nl=((d-1)n+T)\,j'+nl',$$

alors

$$T(j'-j)=n((l-l')+(d-1)(j-j')).$$

Si le coefficient de n est nul, alors $j'=j$ d'où $l'=l$; si le coefficient de n est non nul, alors :

$$|T(j'-j)-n((l-l')+(d-1)(j-j'))|\geqslant n-|T(j'-j)|\geqslant n-Td\geqslant T$$

pour n assez grand.

Les exposants de p diffèrent ainsi d'au moins T; comme on peut, quitte à remplacer T par un multiple, supposer $p^T>\sup_{j,l}(b_{j,l})$, on a donc :

$$f(n+T)=\sum_{j=0}^{d}\sum_{l=0}^{j}s_p(b_{j,l}).$$

Comme $f(n)\equiv f(n+T)\bmod p$ pour n assez grand, on a par conséquent :

$$s_p(a_{d-1}+(d+1)\,a_d)\equiv s_p(a_{d-1}+da_d)+s_p(a_d)\bmod p.$$

4. Fin de l'étude de la série $\sum_0^\infty s_p(R(n))\,X^n$

THÉORÈME. — *Soit R un polynôme de $\mathbb{Q}[X]$ tel que $R(\mathbb{N})\subset\mathbb{N}$, et $d^0\,R\geqslant 2$; alors la série $\sum_0^\infty s_p(R(n))\,X^n$ n'est pas algébrique sur $\mathbb{F}_p(X)$.*

L'hypothèse $R(\mathbb{N})\subset\mathbb{N}$ implique que le coefficient du terme de plus haut degré de R est positif et que le terme constant est entier positif.

Posons :

$$R(X)=\frac{1}{q}\sum_{k=0}^{d} a_k X^k,$$

où $a_k\in\mathbb{Z}$, $q\in\mathbb{N}^*$, $a_d\in\mathbb{N}^*$ et $a_0 q^{-1}\in\mathbb{N}$. Si la série formelle $\sum_0^\infty s_p(R(n))X^n$ est algébrique sur $\mathbb{F}_p(X)$, il en est de même, d'après la première partie, pour la série $\sum_0^\infty s_p(R(qun+qv))X^n$ quels que soient u et v dans $\mathbb{N}^*$.

On écrit alors :

$$\sum_0^\infty s_p(R(qun+qv))X^n=\sum_0^\infty s_p\left(R(qv)+\sum_{k=1}^{d}(qu)^k\left(\frac{R^{(k)}(qv)}{k!}\right)n^k\right)X^n;$$

Si v est choisi assez grand, les $q(R^{(k)}(qv)/k!)$ appartiennent à $\mathbb{N}$; on applique alors le résultat de la troisième partie, et l'on obtient : quel que soit v assez grand, $s_p(\alpha v+\beta)-s_p(\alpha v+\gamma)=\delta$, mod p, avec :

$$\alpha=da_d q^{d-1}u^{d-1},$$
$$\beta=a_{d-1}q^{d-2}u^{d-1}+(d+1)a_d q^{d-1}u^d,$$
$$\gamma=a_{d-1}q^{d-2}u^{d-1}+da_d q^{d-1}u^d,$$
$$\delta=s_p(a_d q^{d-1}u^d).$$

Fixons alors $u>d$, ceci entraîne $\beta-\gamma>\alpha$, d'où une contradiction grâce au lemme suivant :

LEMME. — *Soient* α, β, γ, *trois entiers, tels que* $\beta-\gamma>\alpha$. *Alors* $s_p(\alpha v+\beta)-s_p(\alpha v+\gamma)$ *n'est pas constant modulo p pour v voisin de* $+\infty$.

Choisissons en effet un entier r tel que $\gamma/\alpha<r<\beta/\alpha$ (c'est possible car $\beta-\gamma>\alpha$), et soit $v=p^t-r$ où t tend vers $+\infty$. Alors, pour t assez grand, on a :

$$\begin{aligned}s_p(\alpha v+\beta)-s_p(\alpha v+\gamma)&\\ &\equiv s_p(\alpha p^t+\beta-\alpha r)-s_p(\alpha p^t+\gamma-\alpha r)\\ &\equiv s_p(\alpha)+s_p(\beta-\alpha r)-s_p(\alpha p^t+\gamma-\alpha r).\end{aligned}$$

Fixons un s tel que $p^s>\alpha r-\gamma$, on en déduit donc :

$$\begin{aligned}s_p(\alpha v+\beta)-s_p(\alpha v+\gamma)&\\ &\equiv s_p(\alpha)+s_p(\beta-\alpha r)-s_p((\alpha-1)p^t+p^t-p^s+p^s+\gamma-\alpha r)\\ &\equiv s_p(\alpha)+s_p(\beta-\alpha r)-s_p(\alpha-1)-s_p(p^t-p^s)-s_p(p^s+\gamma-\alpha r),\end{aligned}$$

on conclut en remarquant que :

$$s_p(p^t-p^s)=s_p(p^{t-s}-1)=s_p((p-1)(1+p+\ldots+p^{t-s-1}))=(p-1)(t-s),$$

n'est pas constant modulo p lorsque t tend vers $+\infty$.

5. Étude de la série $\sum_0^\infty s_p^{(k)}(n)X^n$

THÉORÈME. — *Pour $k\geqslant 2$, la série formelle $\sum_{n=0}^{+\infty} s_p^{(k)}(n)X^n$ n'est pas algébrique sur $\mathbb{F}_p(X)$.*

Il résulte du lemme ci-dessus que la suite $s_p(n)$ n'est pas périodique à partir d'un certain rang (sinon pour n assez grand, $s_p(n+T)-s_p(n)\equiv 0 \bmod p$ ce qui correspond avec les notations ci-dessus à $\alpha=1$, $\beta=T$, $\gamma=0$), donc la série $\sum_0^\infty s_p(n)X^n$ n'est pas dans $\mathbb{F}_p(X)$.

Pour montrer le résultat annoncé, il suffit donc de prouver que si la série $\sum_0^\infty s_p^{(k+1)}(n)X^n$ est algébrique sur $\mathbb{F}_p(X)$, alors $\sum_0^\infty s_p^{(k)}(n)X^n$ est dans $\mathbb{F}_p(X)$:

Si la série $\sum_0^\infty s_p^{(k+1)}(n)X^n$ est algébrique sur $\mathbb{F}_p(X)$, il en est de même de la série $\sum_0^\infty \varepsilon(n)X^n$, où $\varepsilon(0)=0$ et $\varepsilon(n)=s_p^{(k+1)}(n-1)$ pour $n\geqslant 1$.

On applique alors à la suite $\varepsilon(n)$ le lemme fondamental de la deuxième partie, avec $j=1$, $i=2$, donc il existe $T\geqslant 1$ tel que, pour n assez grand :

$$\varepsilon\left(\frac{p^n-1}{p-1}+1\right)\equiv\varepsilon\left(\frac{p^{n+T}-1}{p-1}+1\right)\bmod p,$$

soit :

$$s_p^{(k+1)}(1+p+\ldots+p^{n-1})\equiv s_p^{(k+1)}(1+p+\ldots+p^{n+T-1}),$$

d'où :

$$s_p^{(k)}(n)\equiv s_p^{(k)}(n+T),$$

pour n assez grand, ce qui signifie que la série $\sum_0^\infty s_p^{(k)}(n)X^n$ est dans $\mathbb{F}_p(X)$.

BIBLIOGRAPHIE

[1] CHRISTOL (G.), KAMAE (T.), MENDÈS FRANCE (M.) et RAUZY (G.). — Suites algébriques, automates et substitutions, *Bull. Soc. math. France*, vol. 108, fasc. 4, 1980, p. 401-420.

[2] GELFOND (A. O.). — Sur les nombres qui ont des propriétés additives et multiplicatives données, *Acta Arithmetica*, vol. XIII, 1968.