ACTA ARITHMETICA
XXX (1976)

Factorisation de fractions rationnelles et de suites récurrentes

par

JEAN BERSTEL (Paris)

1. Introduction. Il est bien connu [8] qu'un polynôme f à coefficients entiers est puissance k-ième d'un polynôme h de même nature si $f(n)$ est puissance k-ième d'un entier pour tout $n \in \boldsymbol{N}$, et même s'il en est ainsi pour tout $n \in \boldsymbol{A}$, où $\boldsymbol{A}$ est une partie de $\boldsymbol{N}$ qui rencontre toute progression arithmétique [5]. D'autres propriétés analogues de factorisation, pour la composition des fonctions, de polynômes ou de fractions rationnelles peuvent se formuler comme solutions partielles au problème général suivant:

(*) „Etant donnés deux ensembles de nombres E et F, deux applications f et g de E dans F, et une partie A de E, déterminer les conséquences de l'inclusion $f(A) \subset g(E)$".

L'inclusion $f(A) \subset g(E)$ équivaut à l'existence d'une application h de A dans E telle que $f_{|A} = g \circ h$. La conclusion cherchée est que h est „de même nature" que f ou g, et de déterminer la forme particulière que prend alors la fonction f. Ainsi, si $f \in \boldsymbol{Z}[t]$, et $g = t^k$, le théorème rappelé ci-dessus affirme que h est également un polynôme, et donc que f se factorise en $f = h^k$.

On constate que des propriétés établies lorsque $A = \boldsymbol{N}$ ou $A = \boldsymbol{Z}$ restent souvent vraies lorsque A n'est plus qu'une partie de $\boldsymbol{N}$ ou de $\boldsymbol{Z}$ rencontrant toute progression arithmétique. Il en est ainsi pour l'énoncé cité plus haut; c'est le cas également du théorème de Fatou, comme l'a prouvé Rauzy [10]. Un tel ensemble A est appelé *arithmétiquement dense* par Rauzy.

Dans cet article, nous examinons le problème (*) dans quelques situations très particulières, qui trouvent leur application dans l'étude des propriétés arithmétiques de langages formels [3]; les ensembles E et F seront $\boldsymbol{N}$, $\boldsymbol{Z}$ ou $\boldsymbol{Q}$, et les fonctions f et g des polynômes, des fractions rationnelles ou ce que nous appelons des fonctions *adrationnelles*, c'est-à-dire des fonctions u telles que la suite $(u(n))_{n \geq 0}$ soit une suite récurrente.

Dans la deuxième section, nous rassemblons quelques propriétés des ensembles arithmétiquement denses. Les résultats que nous avons

obtenus se formulent en trois théorèmes principaux qui seront démontrés avec leurs corollaires dans les sections 3, 4 et 5. Ces théorèmes sont les suivants:

THÉORÈME I. *Soit* $\mathscr{A} = \sum u_n t^n$ *une série rationnelle à coefficients entiers et* $\boldsymbol{A} \subseteq \boldsymbol{N}$ *un ensemble arithmétiquement dense; s'il existe un entier* $k \geqslant 1$ *tel que, pour tout* $n \in \boldsymbol{A}$, u_n *a au plus* k *facteurs premiers distincts, alors* $\mathscr{A}$ *se met sous la forme:*

$$\mathscr{A} = P(t) + \sum_{m=0}^{M-1} \frac{a_m t^m}{1 - b_m t^M}, \tag{1}$$

où $P \in \boldsymbol{Z}[t]$, $M \geqslant 1$, *et* $a_m, b_m \in \boldsymbol{Z}$ *pour* $m = 0, \ldots, M-1$.

Ce théorème précise un résultat dû à G. Pólya [6]. Remarquons qu'il implique que, sauf si $\mathscr{A}$ se met sous la forme (1), le nombre de facteurs premiers divisant u_n, et donc la taille du plus grand facteur premier divisant u_n, ne sont pas bornés lorsque n tend vers l'infini. Dans le cadre du problème (*) on peut formuler les hypothèses du théorème I comme suit:

L'ensemble $\{u_n : n \geqslant 0\}$ est contenu dans l'ensemble $b(\boldsymbol{N})$ des valeurs d'une fonction b qui énumère les entiers ayant au plus k facteurs premiers distincts.

THÉORÈME II. *Soit* $f \in \boldsymbol{Q}(t)$ *une fraction rationnelle et* $\boldsymbol{A} \subset \boldsymbol{Z}$ *un ensemble arithmétiquement dense;*

(1) *si* $f(\boldsymbol{A}) \subset g(\boldsymbol{Z})$ *pour* $g \in \boldsymbol{Q}(t)$, *il existe* $r \in \boldsymbol{Q}[t]$ *tel que* $f = g \circ r$;

(2) *si* $f(\boldsymbol{A}) \subset p(\boldsymbol{Q})$ *pour* $p \in \boldsymbol{Q}[t]$, *il existe* $g \in \boldsymbol{Q}(t)$ *tel que* $f = p \circ g$.

Ceci représente l'extension aux fractions rationnelles du théorème cité plus haut.

THÉORÈME III. *Soit* $\mathscr{A} = \sum_{n=0}^{\infty} u(n) t^n$ *une série rationnelle à coefficients entiers ayant un pôle unique simple sur son cercle de convergence; si* $u(\boldsymbol{N}) \subset p(\boldsymbol{Z})$, *pour* $p \in \boldsymbol{C}[t]$, *il existe une série rationnelle* $\mathscr{B} = \sum_{n=0}^{\infty} v(n) t^n$ *à coefficients entiers telle que* $u = p \circ v$.

Ce théorème est une généralisation d'un résultat dû à Ch. Pisot (cf. [1]).

2. Ensembles arithmétiquement denses. Un sous-ensemble $\boldsymbol{A}$ de $\boldsymbol{N}$ (resp. de $\boldsymbol{Z}$) est *arithmétiquement dense* dans $\boldsymbol{N}$ (resp. $\boldsymbol{Z}$) s'il rencontre toute progression arithmétique infinie de $\boldsymbol{N}$ (resp. de $\boldsymbol{Z}$), donc si pour tout $a, b \in \boldsymbol{N}$ (resp. $a, b \in \boldsymbol{Z}$), $a \neq 0$, on a $\boldsymbol{A} \cap (a\boldsymbol{N} + b) \neq \emptyset$ (resp. $\boldsymbol{A} \cap (a\boldsymbol{Z} + b) \neq \emptyset$). La terminologie, due à Rauzy, est justifiée par la remarque suivante [10]:

Si l'on munit $\boldsymbol{N}$ (resp. $\boldsymbol{Z}$) de la topologie la moins fine qui rende continue les injections canoniques de $\boldsymbol{N}$ (resp. $\boldsymbol{Z}$) dans $\boldsymbol{Q}_p$, pour tout nombre premier p, alors les progressions arithmétiques $a\boldsymbol{N} + b$ (resp. $a\boldsymbol{Z} + b$) pour tout $a \in \boldsymbol{N} \setminus \{0\}$, forment une base de voisinages de b, et par conséquent $\boldsymbol{A}$ est un sous-ensemble dense de $\boldsymbol{N}$ (resp. $\boldsymbol{Z}$) pour cette topologie.

Si $\boldsymbol{A}$ est arithmétiquement dense dans $\boldsymbol{N}$, il l'est également dans $\boldsymbol{Z}$, mais si $\boldsymbol{A}$ est arithmétiquement dense dans $\boldsymbol{Z}$, alors $\boldsymbol{A} \cap \boldsymbol{N}$ ne l'est pas nécessairement dans $\boldsymbol{N}$. Donnons, pour références ultérieures, quelques propriétés élémentaires des ensembles arithmétiquement denses:

PROPOSITION 2.1. *Soit* $\boldsymbol{A}$ *arithmétiquement dense dans* $\boldsymbol{Z}$.

(i) *L'ensemble* $\boldsymbol{A}$ *rencontre toute progression arithmétique une infinité de fois*;

(ii) *Soient* $a, b \in \boldsymbol{Z}$, $a \neq 0$; *l'ensemble* $\boldsymbol{B} = \{k \in \boldsymbol{Z} \mid ak + b \in \boldsymbol{A}\}$ *est arithmétiquement dense*;

(iii) *L'ensemble* $\boldsymbol{B} = \{n \in \boldsymbol{A} \mid |n| \geqslant N_0 \text{ fixé}\}$ *est arithmétiquement dense.*

Preuve. (i) est due à [5], et (iii) en découle immédiatement. Pour vérifier (ii), soit $a'\boldsymbol{Z} + b'$ ($a' \neq 0$) une progression arithmétique et soit $x \in \boldsymbol{A} \cap (aa'\boldsymbol{Z} + ab' + b)$. Il existe $k' \in \boldsymbol{Z}$ tel que $x = a(a'k' + b') + b \in \boldsymbol{A}$, donc $k = a'k' + b' \in \boldsymbol{B}$ et $\boldsymbol{B} \cap (a'\boldsymbol{Z} + b')$ n'est pas vide.

PROPOSITION 2.2. *Si la densité supérieure* d^+P *de* $P \subset \boldsymbol{N}$ *est* 1, *alors* P *est arithmétiquement dense.*

Supposons en effet que

$$d^+P = \limsup_{n \to \infty} \frac{1}{n} \operatorname{Card}(P \cap \{0, \ldots, n\}) = 1,$$

et que P ne rencontre pas la progression arithmétique $a\boldsymbol{N} + b$ ($a \geqslant 1$, $b \geqslant 0$). Alors

$$\operatorname{Card}(P \cap \{0, \ldots, ak + b\}) \leqslant (a-1)k + b, \quad \text{et} \quad d^+P \leqslant 1 - 1/a.$$

Des *exemples* d'ensembles arithmétiquement denses $\boldsymbol{A} \subset \boldsymbol{N}$ sont fournis par les ensembles de fréquence > 1 de Rauzy [9]. Rappelons que $\boldsymbol{J} \subset \boldsymbol{N}$ est de fréquence > 1 s'il existe un nombre réel $q > 1$, tel que, pour tout $x_0 > 0$, il existe $x \in \boldsymbol{N}$, avec $x \geqslant x_0$, et $\{n : x \leqslant n < qx\} \subset \boldsymbol{J}$. Les ensembles de fréquence > 1 ne forment pas tous les ensembles arithmétiquement denses de $\boldsymbol{N}$, comme le montre l'ensemble $\boldsymbol{A} = \bigcup_{n \in \boldsymbol{N}} \{n^3, n^3+1, \ldots, n^3+n-1\}$. L'ensemble $\boldsymbol{A}$ contient des intervalles arbitrairement grands de nombres entiers consécutifs, et est par conséquent arithmétiquement dense, sans être de fréquence > 1. Cet exemple montre de plus qu'il existe des ensembles arithmétiquement denses de densité asymptotique nulle. Une autre différence importante entre les ensembles arithmétiquement denses de $\boldsymbol{N}$ et les ensembles de fréquence > 1 est que les

premiers peuvent ne pas contenir de nombres premiers, alors que les derniers en contiennent nécessairement une infinité. L'ensemble $\boldsymbol{B}$ des entiers qui possèdent un facteur carré est en effet arithmétiquement dense, comme on le voit immédiatement.

3. Une généralisation d'un théorème de G. Pólya. Soit K un corps commutatif; une série formelle $\mathscr{A} = \sum_{n=0}^{\infty} a_n t^n$ à coefficients dans K est rationnelle s'il existe $q_1, \ldots, q_h \in K$, $q_h \neq 0$, $n_0 \in \boldsymbol{N}$ tels que:

$$a_{n+h} = q_1 a_{n+h-1} + \ldots + q_h a_n \tag{3.1}$$

pour tout $n \geqslant n_0$. La série $\mathscr{A}$ représente alors le développement en série de Taylor, autour de l'origine, d'une fraction rationnelle P/Q, où $P, Q \in K[t]$, et

$$Q(t) = 1 - q_1 t - \ldots - q_h t^h.$$

La relation (3.1) est la plus courte relation de récurrence que vérifie la suite $(a_n)_{n\geqslant 0}$ si, et seulement si P et Q sont sans facteur commun; elle est vérifiée pour tout $n \geqslant 0$ si, et seulement si degré $P <$ degré Q. Une série rationnelle pour laquelle cette condition est vérifiée est dite normalisée.

Appelons, avec B. Benzaghou [1], *suite de Pólya* une suite $(a_n)_{n\geqslant 0}$ de nombres rationnels telle que la valuation p-adique v_p est triviale sur $(a_n)_{n\geqslant 0}$ pour tous les nombres premiers à l'exception d'un nombre fini d'entre eux. G. Pólya [5] a caractérisé les séries rationnelles normalisées $\mathscr{A} = \sum_{n=0}^{\infty} a_n t^n$ à coefficients rationnels dont la suite des coefficients $(a_n)_{n\geqslant 0}$ est une suite de Pólya (résultat généralisé aux séries à coefficients dans un corps de nombres quelconque par Benzaghou [1]), en prouvant que ces séries sont de la forme suivante, appelées *séries de Pólya*:

$$\mathscr{A} = \sum_{m=0}^{M-1} \frac{a_m t^m}{1 - b_m t^M}, \tag{3.2}$$

où $M \geqslant 1$, et $b_0, \ldots, b_{M-1} \in \boldsymbol{Q}$. Nous prouvons ici que la même conclusion reste vraie lorsque les hypothèses du théorème de Pólya sont affaiblies. D'une part, on ne considère qu'une sous-suite de la suite des coefficients dont les indices forment un ensemble $\boldsymbol{A}$ arithmétiquement dense dans $\boldsymbol{N}$ (suivant ainsi la démarche de Rauzy [10] pour le théorème de Fatou), et d'autre part, on n'impose de restrictions que sur le nombre de facteurs premiers distincts des coefficients a_n ($n \in \boldsymbol{N}$).

Appelons *suite de Pólya généralisée* une suite $(a_n)_{n\geqslant 0}$ de nombres rationnels pour laquelle il existe un ensemble arithmétiquement dense $\boldsymbol{A}$ de $\boldsymbol{N}$, et un entier $k \geqslant 0$ vérifiant: pour tout $n \in \boldsymbol{A}$ tel que $a_n \neq 0$, on a $v_p(a_n) \neq 0$ pour au plus k nombres premiers. Toute suite de Pólya est une suite de Pólya généralisée. La suite $(a_n)_{n\geqslant 0}$ définie par $a_p = p$ si p est premier, $a_n = 0$ sinon, est une suite de Pólya généralisée sans être une suite de Pólya.

THÉORÈME I. *Soit $\mathscr{A} = \sum_{n=0}^{\infty} a_n t^n$ une série rationnelle normalisée à coefficients rationnels; la suite $(a_n)_{n\geqslant 0}$ est une suite de Pólya généralisée si, et seulement si $\mathscr{A}$ est une série de Pólya.*

Remarque. On construit facilement des exemples montrant que ce résultat est le meilleur possible en ce sens que si l'ensemble $\boldsymbol{A}$ a une intersection vide avec au moins une progression arithmétique $b\boldsymbol{N}+c$, la conclusion du théorème cesse d'être vraie.

Dans l'exemple d'application suivant, $d(m)$ est le nombre de diviseurs de $m \in \boldsymbol{Z} \setminus \{0\}$, et $d(0) = 0$.

COROLLAIRE 3.1. *Soit $\mathscr{A} = \sum_{n=0}^{\infty} a_n t^n$ une série rationnelle à coefficients entiers; pour que les coefficients de $\mathscr{A}$ soient bornés, il faut et il suffit qu'il existe un ensemble arithmétiquement dense $\boldsymbol{A} \subset \boldsymbol{N}$ tel que $\sup\{d(a_n) \mid n \in \boldsymbol{A}\} < \infty$.*

Preuve. Si $\sup\{d(a_n) \mid n \in \boldsymbol{A}\} = k < \infty$, chaque a_n ($n \in \boldsymbol{A}$) a au plus k facteurs premiers. On peut d'autre part supposer la série $\mathscr{A}$ normalisée en modifiant au plus un nombre fini de coefficients; par le théorème, la série $\mathscr{A}$ est donc de la forme (3.2), et comme chacune des séries $\dfrac{a_m t^m}{1 - b_m t^M}$ contient une infinité de coefficients d'indice dans $\boldsymbol{A}$ (par 2.1), on a $b_m \in \{-1, 0, 1\}$, ce qui prouve le corollaire.

Pour la preuve du théorème I, on utilise les trois lemmes suivants dus à Benzaghou [1].

LEMME 3.2. *Soit $\mathscr{A} = \sum_{n=0}^{\infty} a_n t^n$ une série rationnelle à coefficients complexes. Il existe un entier M tel que pour tout $m = 0, \ldots, M-1$, les séries $\mathscr{A}_{m,M} = \sum_{r=0}^{\infty} a_{m+rM} t^r$ satisfont la propriété* (S) *suivante:*

(S) *aucun des quotients α_i/α_j de deux pôles α_i et α_j de $\mathscr{A}_{m,M}$ n'est une racine de l'unité autre que* 1.

LEMME 3.3. *Soit $\mathscr{A} = \sum_{n=0}^{\infty} a_n t^n$ une série rationnelle à coefficients rationnels, et p un nombre premier. Il existe $N \geqslant 1$, et un entier n_0 tels que l'application $r \to v_p(a_{n_0+rN})$ soit une fonction affine.*

Autrement dit, il existe λ, μ tels que $v_p(a_{n_0+rN}) = \lambda r + \mu$.

LEMME 3.4. *Si la série* $\mathscr{A} = \sum_{n=0}^{\infty} a_n t^n$ *à coefficients complexes, satisfait à la propriété* (S), *et si, pour des entiers* r, T, *il existe une relation de la forme* $a_{r+nT} = cd^n$ *pour une infinité de* $n \in \boldsymbol{N}$, *alors* a_n *est de la forme* $a_n = c'd'^n$ *pour tout* $n \in \boldsymbol{N}$.

Preuve du théorème I. a). D'après le théorème d'Eisenstein, il existe un entier $q \in \boldsymbol{N} \setminus \{0\}$ tel que $b_n = q^{n+1} a_n \in \boldsymbol{Z}$ pour tout $n \geqslant 0$. Comme $(a_n)_{n\geqslant 0}$ est une suite de Pólya généralisée si, et seulement si $(b_n)_{n\geqslant 0}$ l'est, on peut supposer $\mathscr{A}$ à coefficients entiers. D'autre part, il suffit de prouver le théorème pour les séries $\mathscr{A}_{m,M}$ données par le lemme 3.2. Comme les ensembles $\boldsymbol{A}'_m = \{k: m + kM \in \boldsymbol{A}\}$ sont arithmétiquement denses par 2.1 (ii), on peut supposer que la série $\mathscr{A}$ vérifie la condition (S).

b). Notons, pour une suite $(u_n)_{n\in\boldsymbol{N}}$ d'entiers, et une partie $P \subset \boldsymbol{N}$, par $\mathscr{P}\{(u_n)_{n\in P}\}$ l'ensemble des nombres premiers divisant au moins un u_n avec $n \in P$. Nous allons vérifier qu'il existe un ensemble arithmétiquement dense $\boldsymbol{B} \subset \boldsymbol{N}$, et deux entiers r, T $(T \geqslant 1)$ tels que $\mathscr{P}\{(a_{r+nT})_{n\in\boldsymbol{B}}\}$ est fini.

Soit en effet k le plus petit entier k' tel que pour tout $n \in \boldsymbol{A}$, a_n a au plus k' facteurs premiers distincts. Si $k = 0$, l'assertion est trivialement vraie pour $r = 0$, $T = 1$, $\boldsymbol{B} = \boldsymbol{A}$.

Supposons donc $k > 0$, et supposons que la suite (a_n) vérifie la relation de récurrence (3.1). Dans l'ensemble J des indices $n \in \boldsymbol{A}$ tels que a_n a exactement k facteurs premiers distincts, soit r un de ceux pour lesquels le nombre s de facteurs premiers p ne divisant pas q_h soit maximal. Par construction, on a donc $\mathrm{Card}\{p \mid p \mid a_n \text{ et } p \nmid q_h\} \leqslant s$ pour tout $n \in J$. Soient $p_1, p_2, \ldots, p_s$ les facteurs premiers de a_r ne divisant pas q_h, et $P = p_1 p_2 \ldots p_s$. La suite $(a_n)_{n\in\boldsymbol{N}}$ est purement périodique mod P [4]. Soit T sa période, et soit $\boldsymbol{B} = \{n: r + nT \in \boldsymbol{A}\}$. Par 2.1 (ii), $\boldsymbol{B}$ est arithmétiquement dense dans $\boldsymbol{N}$. Pour tout $n \in \boldsymbol{B}$, $a_{r+nT} \equiv 0 (\mathrm{mod} P)$, et en vertu de la maximalité de s, a_{r+nT} ne possède comme facteurs premiers, en dehors de $p_1, p_2, \ldots, p_s$, au plus les facteurs premiers de q_h, qui sont en nombre fini. Les diviseurs premiers de la suite $(a_{r+nT})_{n\in\boldsymbol{B}}$ sont donc en nombre fini, ce qu'il fallait prouver.

c). Posons $\mathscr{P}\{(a_{r+nT})_{n\in\boldsymbol{B}}\} = \{p_1, \ldots, p_l\}$. La démonstration se fait par récurrence sur l. Si $l = 0$, alors $a_{r+nT} \in \{0, 1, -1\}$ pour tout $n \in \boldsymbol{B}$; par le lemme 3.4, on a $a_n = \varepsilon$ pour tout $n \in \boldsymbol{N}$ où $\varepsilon \in \{0, 1, -1\}$, et la série est bien de Pólya. Si $l > 0$, posons pour tout $n \in \boldsymbol{B}$,

$$a_{r+nT} = \varepsilon_n p_1^{\varphi_1(n)} \ldots p_l^{\varphi_l(n)}, \qquad \varepsilon_n \in \{-1, +1\}, \qquad \varphi_i(n) \in \boldsymbol{N}.$$

La série $\sum a_{r+nT} t^n$ étant rationnelle, il existe, par lemme 3.3, des entiers n_0 et N tels que l'application $n \to v_{p_1}(a_{r+n_0T+NTn})$ soit affine. Posons $v_{p_1}(a_{r+n_0T+NTn}) = \lambda n + \mu$. L'ensemble $\boldsymbol{C} = \{n: n_0 + Nn \in \boldsymbol{B}\}$ est arithmétiquement dense, et la série $b = \sum b_{r'+nT'} t^n$ où $b_{r'+nT'} p_1^{\lambda n+\mu} = a_{r'+nT'}$ $(r' = r + n_0 T, T' = NT)$ est évidemment rationnelle. Il existe donc un ensemble arithmétiquement dense $\boldsymbol{C}$ tel que $\mathscr{P}\{(b_{r'+nT'})_{n\in\boldsymbol{C}}\}$ soit fini et de cardinalité $l-1$. Par hypothèse de récurrence la série b est de Pólya donc $a_{r'+nT'}$ s'écrit sous la forme $a_{r'+nT'} = p_1^{\lambda n+\mu} cd^n$. Par le lemme 3.4, on a $a_n = c'd'^n$ pour tout $n \in \boldsymbol{N}$. Ceci achève la démonstration.

4. Factorisation des fractions rationnelles. Dans cette section, nous prouvons le théorème suivant qui est l'analogue du théorème de factorisation des polynômes.

THÉORÈME II. *Soit* $f \in \boldsymbol{Q}(t)$, *une fraction rationnelle, et* $\boldsymbol{A}$ *un ensemble arithmétiquement dense dans* $\boldsymbol{Z}$;

(1) *si* $f(\boldsymbol{A}) \subset g(\boldsymbol{Z})$ *pour une fraction rationnelle* $g \in \boldsymbol{Q}(t)$, *il existe un polynôme* $p \in \boldsymbol{Q}[t]$ *tel que* $f = g \circ p$;

(2) *si* $f(\boldsymbol{A}) \subset p(\boldsymbol{Q})$ *pour un polynôme* $p \in \boldsymbol{Q}[t]$, *il existe une fraction rationnelle* $g \in \boldsymbol{Q}(t)$ *telle que* $f = p \circ g$.

Ce résultat est une légère généralisation de [2]. Nous donnons la preuve entière par souci de complétude. Si l'on pose $p(t) = t^k$, on déduit de (2) le:

COROLLAIRE 4.1. *Soit* $f \in \boldsymbol{Q}(t)$ *et* $\boldsymbol{A}$ *un ensemble arithmétiquement dense dans* $\boldsymbol{Z}$; *si pour tout* $n \in \boldsymbol{A}$, $f(n)$ *est puissance k-ième d'un nombre rationnel, il existe une fraction rationnelle* $g \in \boldsymbol{Q}(t)$ *telle que* $f = g^k$.

A la base de la preuve est un théorème de Davenport, Lewis et Schinzel [5] que nous formulons comme lemme:

LEMME 4.2. *Soit* $F(t, t') \in \boldsymbol{Z}[t, t']$, *et* $\boldsymbol{A}$ *un ensemble arithmétiquement dense dans* $\boldsymbol{Z}$, *si pour tout* $n \in \boldsymbol{A}$, *il existe* $m \in \boldsymbol{Z}$ *tel que l'on ait* $F(n, m) = 0$, *alors il existe un polynôme* $r \in \boldsymbol{Q}[t]$ *tel que identiquement* $F(t, r(t)) = 0$.

Pour $F(t, t') = f(t) - t'^k$, ce résultat donne le corollaire suivant, déjà mentionné dans l'introduction et dont on peut d'ailleurs donner une preuve élémentaire (cf. [3]):

COROLLAIRE 4.3. *Soit* $f \in \boldsymbol{Z}[t]$, *et* $\boldsymbol{A} \subset \boldsymbol{N}$ *arithmétiquement dense*; *si* $f(n)$ *est puissance k-ième d'un entier pour tout* $n \in \boldsymbol{A}$, *il existe un polynôme* $g \in \boldsymbol{Z}[t]$ *tel que* $f = g^k$.

Nous utiliserons dans la preuve du théorème II le lemme suivant, dont la première partie est classique et la deuxième est immédiate.

LEMME 4.4. *Deux polynômes* $a, b \in \boldsymbol{Z}[t]$ *sont sans facteur commun sauf peut-être une constante, si et seulement s'il existe un entier* N *tel que* $(a(n), b(n))$ *divise* N *pour tout* $n \in \boldsymbol{Z}$; *de plus, si* $n \equiv n' (\mathrm{mod} N)$, *alors*

$$(a(n), b(n)) = (a(n'), b(n')).$$

Preuve du théorème II. a). La première partie du théorème découle du lemme 4.2 comme suit: Posons $f = a/b$, $g = a'/b'$, avec

$a, b, a', b' \in \boldsymbol{Z}[t]$, et soit $F \in \boldsymbol{Z}[t, t']$ défini par:

$$F(t, t') = a(t)b'(t') - a'(t')b(t).$$

Par hypothèse, le lemme 4.2 peut être appliqué, et il existe un polynôme $p \in \boldsymbol{Q}[t]$ tel que $F\big(t, p(t)\big) = 0$ identiquement; autrement dit, on a $f = g \circ p$.

b). Pour prouver la deuxième partie, posons $f = a/b$, avec $a, b \in \boldsymbol{Z}[t]$ sans facteur commun, le coefficient directeur de b étant positif. Supposons d'abord le polynôme p unitaire et à coefficients entiers:

$$p(t) = t^k + p_1 t^{k-1} + \ldots + p_k, \quad p_1, \ldots, p_k \in \boldsymbol{Z}.$$

Pour tout $n \in \boldsymbol{A}$, il existe deux entiers y_n, z_n premiers entre eux, avec $z_n > 0$, tels que $f(n) = p(y_n/z_n)$, d'où:

$$\frac{a(n)}{b(n)} = \frac{y_n^k + p_1 y_n^{k-1} z_n + \ldots + p_k z_n^k}{z_n^k},$$

et comme cette dernière fraction ne peut être réduite, on a, pour tout $n \in \boldsymbol{A}$:

$$b(n) = \varepsilon_n\big(a(n), b(n)\big) z_n^k \quad \text{avec} \quad \varepsilon_n \in \{-1, 1\}. \tag{4.1}$$

c). Soit maintenant N l'entier dont l'existence est assurée par le lemme 4.4, et posons $d = \big(a(0), b(0)\big)$. Les polynômes $\bar{a}$ et $\bar{b}$ définis par:

$$\bar{a}(t) = \frac{a(tN)}{d}, \quad \bar{b}(t) = \frac{b(tN)}{d}$$

sont à coefficients entiers par construction; l'ensemble $\boldsymbol{B} = \{n \in \boldsymbol{Z}: nN \in \boldsymbol{A}\}$ est arithmétiquement dense dans $\boldsymbol{Z}$ par 2.1 (ii), et il résulte de (4.1) que:

$$\bar{b}(n) = \varepsilon_{nN}(z_{nN})^k \quad \text{pour tout } n \in \boldsymbol{B}. \tag{4.2}$$

Le polynôme $b' = \bar{b}^2$ vérifie les hypothèses du corollaire 4.3, et il existe un polynôme $u \in \boldsymbol{Z}[t]$ tel que $b' = u^{2k}$. On peut supposer le coefficient directeur de u positif comme l'est celui de $\bar{b}$; on en tire $\bar{b} = u^k$.

d). Soit maintenant $\boldsymbol{B}' = \{n \in \boldsymbol{B}: |n| \geqslant N_0\}$, où N_0 est assez grand pour que $|n| \geqslant N_0$ implique $u(n) > 0$ si le degré l de u est pair, $nu(n) > 0$ si l est impair. L'ensemble $\boldsymbol{B}'$ est arithmétiquement dense dans $\boldsymbol{Z}$ par 2.1 (iii). De (4.2), on déduit que pour tout $n \in \boldsymbol{B}'$, on a $u(n) = z_{nN}$ si k ou l est pair, et $u(n) = -z_{nN}$ si k et l sont impairs; donc l'équation

$$\frac{\bar{a}(n)}{\bar{b}(n)} = p\left(\frac{m}{u(n)}\right)$$

a, pour tout $n \in \boldsymbol{B}'$, la solution $m = y_{nN}$ dans le premier cas et $m = -y_{nN}$ dans le deuxième cas. On peut donc appliquer le lemme 4.2 au polynôme

$$F(t, t') = \bar{a}(t) - \big(t'^k + p_1 t'^{k-1} u(t) + \ldots + p_k\big(u(t)\big)^k\big),$$

d'où l'on tire l'existence d'un polynôme $v \in \boldsymbol{Z}[t]$ tel que $\bar{a}/\bar{b} = p \circ (v/u)$, d'où finalement $f = p \circ g$ avec $g(t) = \dfrac{v(t/N)}{u(t/N)}$.

e). Supposons maintenant que le polynôme p soit à coefficients entiers:

$$p(t) = p_0 t^k + p_1 t^{k-1} + \ldots + p_k$$

et soit $p' \in \boldsymbol{Z}[t]$ le polynôme unitaire:

$$p'(t) = t^k + \sum_{i=1}^{k} p_i' t^{k-i}$$

où $p_i' = p_0^{i-1} p_i$ pour $i = 1, \ldots, k$. Clairement, on a $p'(p_0 t) = p_0^{k-1} p(t)$, et si par conséquent $f(\boldsymbol{A}) \subset p(\boldsymbol{Q})$, on a $f'(\boldsymbol{A}) \subset p'(\boldsymbol{Q})$, avec $f' = p_0^{k-1} f$. En vertu de ce qui précède, il existe une fraction rationnelle $g' \in \boldsymbol{Q}(t)$ telle que $f' = p' \circ g'$, d'où l'on déduit que $p_0 f' = p_0^k f = (p_0 p') \circ g'$, et comme $p_0 p'(t) = p(t/p_0)$, également $f = p \circ g$, avec $g = g'/p_0$.

Soit enfin $p \in \boldsymbol{Q}[t]$ quelconque, et m le pgcd des dénominateurs des coefficients de p. Le polynôme mp est donc à coefficients entiers, et comme $f(\boldsymbol{A}) \subset p(\boldsymbol{Q})$ implique que $mf(\boldsymbol{A}) \subset mp(\boldsymbol{Q})$, il existe d'après ce qui précède une fraction rationnelle $g \in \boldsymbol{Q}(t)$ telle que l'on ait $mf = mp \circ g$, d'où l'on tire que $f = p \circ g$. Le théorème est complètement démontré.

5. Factorisation de suites récurrentes. Soit a une fonction adrationnelle à valeurs rationnelles, et $\mathscr{A} = \sum_{n=0}^{\infty} a(n) t^n$ la série rationnelle associée. Nous donnons ici des réponses partielles au problème de factorisation dans les cas $a(\boldsymbol{N}) \subset p(\boldsymbol{Z})$, où p est un polynôme, et $a(\boldsymbol{N}) \subset b(\boldsymbol{N})$, où b est une autre fonction adrationnelle.

La première inclusion se ramène, dans le cas particulier où la série $\mathscr{A}$ n'a que le pôle 1, à l'inclusion des ensembles de valeurs prises par deux polynômes. De même, le deuxième problème se ramène au premier si la série $\mathscr{B} = \sum_{n=0}^{\infty} b(n) t^n$ n'a que le pôle 1. Nous examinons le cas particulier du premier problème où $\mathscr{A}$ possède un pôle unique et simple sur son cercle de convergence. Nous n'avons pu déterminer les conséquences de la deuxième relation d'inclusion que dans le cas où la série $\mathscr{B}$ est une série de Pólya, mais pensons que la même conclusion est valable sans cette restriction.

PROPOSITION 5.1. *Soient*

$$\mathscr{A} = \sum_{n=0}^{\infty} a(n) t^n \quad \text{et} \quad \mathscr{B} = \sum_{n=0}^{\infty} b(n) t^n$$

deux séries rationnelles normalisées à coefficients rationnels, et soit $\boldsymbol{A}$ un ensemble arithmétiquement dense dans $\boldsymbol{N}$; si $a(\boldsymbol{A}) \subset b(\boldsymbol{N})$ et si $\mathscr{B}$ est une série de Pólya, alors $\mathscr{A}$ est une série de Pólya, et de plus, il existe un entier $M \geqslant 1$, et des entiers r_m, $M_m \geqslant 0$ pour $m = 0, \ldots, M-1$, tels que l'on ait:

$$a(m+nM) = b(r_m + nM_m) \quad \text{pour tout } n \in \boldsymbol{N}.$$

Preuve. La condition $a(\boldsymbol{A}) \subset b(\boldsymbol{N})$ exprime le fait que la suite $(a(n))_{n\in\boldsymbol{N}}$ est une suite de Pólya généralisée. En vertu du théorème I, la série $\mathscr{A}$ est une série de Pólya. Il suffit donc de prouver la deuxième assertion pour une série $\mathscr{A}$ telle que $a(n)$ soit de la forme $a(n) = ac^n$ pour tout $n \in \boldsymbol{N}$, avec $a, c \in \boldsymbol{Q}$, et on peut supposer que $c \notin \{0, 1, -1\}$. Posons alors

$$\mathscr{B} = \sum_{m=0}^{M-1} \frac{b(m)t^m}{1 - d_m t^M},$$

et pour tout $k = 0, \ldots, M-1$, soit

$$A_k = \{n \in \boldsymbol{N}, \exists s \in \boldsymbol{N}\colon\ a(n) = b(k)d_k^s\}.$$

Nous allons prouver que si $A_k \neq \emptyset$, alors A_k est de la forme $A_k = r_k + M_k\boldsymbol{N}$, avec $r_k, M_k \in \boldsymbol{N}$. Pour cela, soit r_k le plus petit élément de A_k supposé non vide. L'ensemble

$$I_k = \{n \in \boldsymbol{Z},\ \exists s \in \boldsymbol{Z}\colon\ c^n = d_k^s\}$$

est un idéal, donc de la forme $I_k = M_k\boldsymbol{Z}$, et $n \in I_k \cap \boldsymbol{N}$ si $r_k + n \in A_k$. L'ensemble A_k est donc bien de la forme annoncée. Il en découle que $B = \bigcup\{A_k\colon\ k = 0, \ldots, M-1\}$ est, à un nombre fini d'éléments près, une union de progressions arithmétiques de même raison M, et comme l'ensemble $\boldsymbol{A}$ est arithmétiquement dense, on a $B = \boldsymbol{N} \setminus S$ où S est un ensemble fini. Puisque $\mathscr{B}$ est une série rationnelle normalisée, on a $S = \emptyset$ et la proposition est établie.

Théorème III. *Soit $\mathscr{A} = \sum_{n=0}^{\infty} a(n)t^n$ une série rationnelle à coefficients entiers ayant un pôle unique simple sur son cercle de convergence, et p un polynôme à coefficients complexes; si $a(\boldsymbol{N}) \subset p(\boldsymbol{Z})$, il existe une série rationnelle $\mathscr{B} = \sum_{n=0}^{\infty} b(n)t^n$ à coefficients entiers telle que l'on ait $a = p \circ b$.*

Nous ne savons pas remplacer l'inclusion $a(\boldsymbol{N}) \subset p(\boldsymbol{Z})$ par $a(\boldsymbol{A}) \subset p(\boldsymbol{Z})$, où $\boldsymbol{A}$ est un ensemble arithmétiquement dense.

Pour $p(t) = t^k$, on obtient le résultat suivant dû à Ch. Pisot (cf. [1]):

Corollaire 5.2. *Soit $\mathscr{A}$ comme dans l'énoncé précédent; si $a(n)$ est puissance k-ième d'un entier pour tout $n \in \boldsymbol{N}$, il existe une série rationnelle $\mathscr{B} = \sum_{n=0}^{\infty} b(n)t^n$ à coefficients entiers telle que $a = b^k$.*

Soit d le degré du polynôme p. Si d est impair, l'équation $a(n) = p(y)$ a au plus une solution si $u(n)$ est assez grand. Si par contre, d est pair, cette équation peut avoir une ou deux solutions. Le caractère rationnel de la série $\sum b(n)t^n$ dépend bien entendu de la solution $b(n) = y$ retenue. En fait, le lemme suivant dit essentiellement que si d est pair, l'équation $a(n) = p(y)$ a soit toujours une, soit toujours deux solutions, et donc qu'un choix convenable de $b(n)$ est possible.

Lemme 5.3. *Soit $p \in \boldsymbol{C}[t]$ un polynôme; si l'ensemble*

$$K = \{x \in \boldsymbol{N}, \exists y \in -\boldsymbol{N}\colon\ p(x) = p(y)\}$$

est infini, il existe $s \in \boldsymbol{Q}$ tel que

$$p(s+t) = p(-s-t).$$

Preuve. Posons $p(t) = p_0t^d + p_1t^{d-1} + \ldots + p_d$, avec d pair. Pour tout $x \in K$ assez grand, il existe un entier $y = y(x) \in -\boldsymbol{N}$ unique tel que $p(x) = p\big(y(x)\big)$; posons:

$$y(x) = -x - \zeta(x) = -x\big(1 + \varepsilon(x)\big).$$

Comme $\left(\frac{y(x)}{x}\right)^d \to \frac{p\big(y(x)\big)}{p(x)} = 1$, lorsque $x \to \infty$ dans K, on a $\varepsilon(x) \to 0$ lorsque $x \to \infty$ dans K.

Pour tout $x \in K$, l'équation $p(x) = p\big(-x\big(1+\varepsilon(x)\big)\big)$ donne:

$$p_0\big(1+\varepsilon(x)\big)^d = p_0 + \frac{p_k}{x^k}\big(1 - (-1)^k\big(1+\varepsilon(x)\big)^{d-k}\big) + O(x^{-1-k}),$$

où k est le plus petit entier tel que $p_k \neq 0$ (et $k = d+1$ si $p_1 = \ldots = p_d = 0$). En développant, on obtient $\varepsilon(x) = O(x^{-1})$, d'où $\zeta(x) = O(1)$.

Comme $\zeta(x)$ est entier pour tout $x \in K$, il existe un entier $z \in \boldsymbol{Z}$, et un sous-ensemble infini L de K tel que l'on ait $\zeta(x) = z$ pour tout $x \in L$, donc $p(-x-z) = p(x)$ pour tout $x \in L$, donc pour tout $x \in \boldsymbol{Z}$. Posant $s = -z/2$, on en déduit l'assertion.

Preuve du théorème III. Par le lemme, l'équation $a(n) = p(y)$ possède, pour n assez grand, ou bien une solution $b(n) = y$, ou bien deux solutions $b(n) > 0$ et y_n' liées par la relation $b(n) + y_n' = h$, où h est un entier fixe.

Par hypothèse, $a(n)$ est, pour n assez grand, de la forme

$$a(n) = b_0\alpha^n + \sum_{i=1}^{r} P_i(n)\alpha_i^n,$$

où α est réel, $|\alpha| > |\alpha_i|$ et $P_i \in \boldsymbol{C}[t]$ pour $i = 1, \ldots, r$. Posant

$$\gamma_n = \frac{1}{b_0\alpha^n}\Big(\sum_{i=1}^{r} P_i(n)\alpha_i^n\Big),$$

il existe des nombres $C > 0$, et $0 < \varrho < 1$ tels que $|\gamma_n| \leqslant C\varrho^n$ pour n assez grand.

Considérons alors le développement de Puiseux de $x = p(y)$:

$$y = u_0 x^{1/d} + u_1 + x^{-1/d}O(1),$$

et posons d'autre part:

$$a(n)^{1/d} = \big(b_0\alpha^n(1+\gamma_n)\big)^{1/d} = (b_0\alpha^n)^{1/d}\Big(1 + \sum_{j=1}^{l-1} d_j\gamma_n^j + \gamma_n^l O(1)\Big),$$

où l est choisi suffisamment grand pour que $|\alpha^{1/d}\varrho^l| < 1$. D'où:

$$b(n) = u_0 b_0^{1/d}\alpha^{n/d}\Big(1 + \sum_{j=1}^{l-1} d_j\gamma_n^j\Big) + u_1 + u_0 b_0^{1/d}\alpha^{n/d}\gamma_n^l O(1) + \alpha^{-n/d}O(1),$$

ce qui montre qu'en posant:

$$w(n) = u_0 b_0^{1/d}\alpha^{n/d}\Big(1 + \sum_{j=1}^{l-1} d_j\gamma_n^j\Big) + u_1,$$

$b(n)$ s'écrit sous la forme $b(n) = w(n) + \varepsilon(n)$, où $\sum_{n=0}^{\infty} w(n)t^n$ est une série rationnelle, et où la série $\sum_{n=0}^{\infty} \varepsilon(n)t^n$ a un rayon de convergence > 1. Par le théorème de Pólya–Carlson [7], la série $\sum_{n=0}^{\infty} b(n)t^n$ est rationnelle et le théorème est établi.

Je remercie J. F. Perrot pour les discussions que j'ai pu avoir avec lui pendant la rédaction de cet article.

Références

[1] B. Benzaghou, *Algèbres de Hadamard*, Bull. Soc. Math. France 98 (1970), p. 209–252.

[2] J. Berstel, *Sur des fractions rationnelles particulières*, C. R. Acad. Sc. Paris 270 (1970), p. 304–306.

[3] — *Contribution à l'étude des propriétés arithmétiques des langages formels*, Thèse, Paris 1972.

[4] R. D. Carmichael, *On sequences of integers defined by recurrence relations*, Quart. J. Math. 48 (1920), p. 343–372.

[5] H. Davenport, D. J. Lewis and A. Schinzel, *Polynomials of certain special types*, Acta Arith. 9 (1964), p. 107–116.

[6] G. Pólya, *Arithmetische Eigenschaften der Reihenentwicklung rationaler Funktionen*, J. Reine Angew. Math. 151 (1921), p. 1–31.

[7] — *Über gewisse notwendige Determinantenkriterien für die Fortsetzbarkeit einer Potenzreihe*, Math. Ann. 99 (1928), p. 687–706.

[8] G. Pólya und G. Szegö, *Aufgaben und Lehrsätze aus der Analysis*, Springer, 3. ed., 1964.

[9] G. Rauzy, *Suites partiellement récurrentes*, Ann. Inst. Fourier 16 (1966), fasc. 1, p. 159–234.

[10] — *Ensembles arithmétiquement denses*, C. R. Acad. Sc. Paris 265 (1967), p. 37–38.

UNIVERSITÉ PIERRE ET MARIE CURIE
Paris, France

Reçu le 7. 3. 1974
et dans la forme modifiée le 2. 12. 1974 (543)