Suites Automatiques

Lucas Boczkowski

1er février 2011

Résumé

Nous présentons succintement les suites automatiques ou suites engendrées par un automate, modèle le plus simple de machine. Bien que correspondant à une donnée finie - les états de l'automate - une suite automatique n'est pas nécessairement très simple ou régulière (ultimement périodique par exemple). La première partie offre une familiarisation avec la notion étudiée puis, dans les deux parties suivantes, on donne des caractérisations de l'automaticité en éclaircissant les liens entre automate et suite engendrée.

Notation. Soit une suite u, on notera parfois le n^{e} terme $u(n)$ pour une meilleure lisibilité.

1 Définition et exemples

Dans cette partie nous définissons la k-automaticité et donnons deux exemples de suites 2-automatiques. Comme on le verra, ces suites présentent des similarités. Ces ressemblances seront développées dans la partie suivante.

Les automates qu'on va considérer ont un seul état initial que l'on notera souvent q_0 et sont déterministes et complets. L'ensemble des états sera noté de manière usuelle $\mathcal{Q}$ et au lieu d'avoir certains états terminaux on dispose d'une fonction de sortie τ qui à chaque état associe une valeur "de sortie".

Définition 1.1. Soit k un entier ≥ 2. On dit qu'une suite $(u_n)_n$ est *k-automatique* s'il existe un automate $\mathcal{A} = \{\mathcal{Q}, \delta, q_o, \tau\}$ tel que :

$$\forall n, \tau(\delta(q_0, \langle n \rangle_k)) = u_n$$

où $\langle n \rangle_k$ désigne l'écriture en base k de n **à partir du chiffre de poids faible**. Autrement dit, si on lit dans l'automate $\mathcal{A}$ la suite constituée des chiffres de n en base k à partir du chiffre de poids faible, on tombe sur un état dont la valeur de sortie est u_n. Il est naturel de considérer que la fonction de sortie est à valeur dans le même ensemble que la suite et que les arêtes de l'automate sont étiquetées par $\{0, 1, \ldots, k-1\}$.

1.1 La suite de Thue-Morse

Cette suite fut en réalité d'abord étudiée par Eugène Prouhet avant d'être redécouverte par Axel Thue en 1906 comme exemple de suite sans chevauchement. Puis Marston Morse, en 1922, donna une nouvelle interprétation de la suite, dans le contexte de la géométrie différentielle.

Définition 1.2. On définit la suite de Thue-Morse, $(t_n)_n$ de la façon suivante :

$$t_0 = 0,$$
$$\forall n,\ t_{2n} = t_n,$$
$$\forall n,\ t_{2n+1} = 1 - t_n.$$

Ses premiers termes sont :

n	0	1	2	3	4	5	6	7	8	9	10	11
t_n	0	1	1	0	1	0	0	1	1	0	0	1

Grâce aux relations définissant t, on voit que t_n correspond à la somme des chiffres dans l'écriture en base 2 de n, modulo 2. La suite de Thue-Morse est 2-automatique ; elle est engendrée par l'automate suivant :

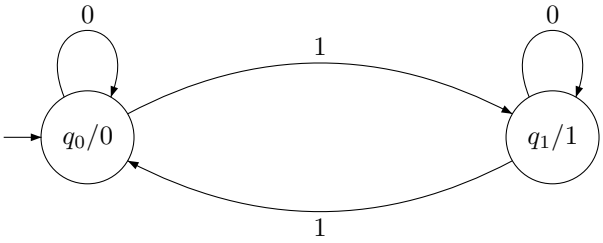

FIGURE 1 – Automate de Thue-Morse.

Remarque. q/α est une notation pour dire que la valeur de sortie à l'état q, $\tau(q)$, est α.

Le fait que l'automate comporte un nombre *fini* d'état est lié, dans un sens que l'on précisera, au fait que la suite correspond à un nombre *fini* de relations de récurrence particulières.
Donnons encore une propriété intéressante de la suite de Thue-Morse (c'est d'ailleurs souvent comme ça qu'elle est définie). On considère le morphisme 2-uniforme [1] $\sigma : \{0,1\}^* \to \{0,1\}^*$ donné par :

$$\sigma(0) = 01$$
$$\sigma(1) = 10$$

La suite $(\sigma^n(0))_n$ converge vers t (au sens de la convergence simple). Il en découle que t est point fixe de σ.

1. Un morphisme est dit k-uniforme si l'image de toute lettre est un mot de taille k

1.2 La suite de Rudin-Shapiro

Définition 1.3. On définit la suite de Rudin-Shapiro, $(r_n)_n$ de la façon suivante :

$$\begin{aligned} r_0 &= 0, \\ \forall n,\ r_{2n} &= r_n, \\ \forall n,\ r_{4n+1} &= r_n, \\ \forall n,\ r_{4n+3} &= 1 - r_{2n+1}. \end{aligned}$$

Ses premiers termes sont :

n	0	1	2	3	4	5	6	7	8	9	10	11
r_n	0	0	0	1	0	0	1	0	0	0	0	1

Il s'agit encore d'une suite automatique. Elle est engendrée par l'automate suivant :

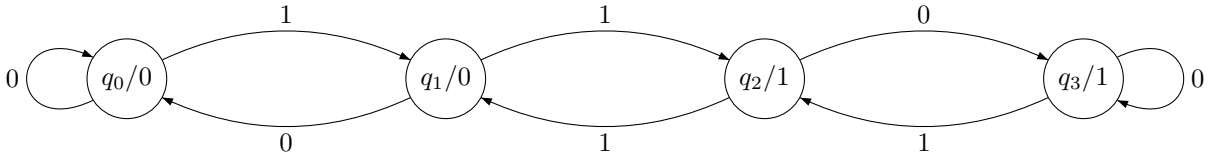

FIGURE 2 – L'automate de Rudin-Shapiro.

Cette suite donne le nombre d'occurrences de "11" dans l'écriture de n en base 2, le tout modulo 2. On ne peut pas trouver comme pour la suite de Thue-Morse un morphisme 2-uniforme dont la suite soit point fixe. Il faut d'abord "agrandir" l'alphabet. Plaçons nous sur $\{A, B, C, D\}$ et définissons σ sur cet alphabet par :

$$\begin{aligned} \sigma(A) &= AB \\ \sigma(B) &= AC \\ \sigma(C) &= DB \\ \sigma(D) &= DC \end{aligned}$$

Le mot $\sigma^\infty(A)$ commence par :

$$ABACABDBABACDCAC...$$

La suite de Rudin-Shapiro s'obtient en prenant l'image de ce point fixe pour σ par l'application de projection $\pi : \{A, B, C, D\} \to \{0, 1\}$ définie par :

$$\begin{aligned} \pi(A) &= 0 = \pi(B) \\ \pi(C) &= 1 = \pi(D) \end{aligned}$$

2 Les liens

Au vu des ressemblances entre les deux suites, il est pertinent d'introduire certaines définitions.

Définition 2.1. Pour une suite u donnée et un entier k, le *k-noyau* que l'on notera $K_u^{(k)}$ ou plus simplement K_u lorsqu'il n'y aura pas d'ambiguïté est :

$$K_u^{(k)} \stackrel{def}{=} \{(u(k^\alpha n + r))_n \ / \alpha \in \mathbb{N},\ 0 \leq r < \alpha\}$$

Dans les deux cas, le noyau était fini (cela suit pratiquement de la définition que l'on a donnée des suites t et r). Donnons à présent une définition pertinente pour le lien avec les morphismes k-uniformes.

Définition 2.2. Une suite u à valeur dans un alphabet A qu'elle est dite *l'image d'un point fixe d'un morphisme k-uniforme* $\sigma : B \to B$ s'il existe $w \in B^{\mathbb{N}}$ point fixe de σ et π une application de projection de B dans A tel que $\forall n, u_n = \pi(w_n)$.

En écrivant le fait d'être un point fixe lettre à lettre on obtient la caractérisation suivante qui sera souvent utile.

Lemme 2.3. *Si σ est un morphisme k-uniforme d'un alphabet B dans lui même alors w est point fixe si et seulement si :*

$$\forall r < k, \sigma(w_n)_r = w_{kn+r}$$

Les suites t et r sont image de point fixe d'un morphisme 2-uniforme. On est maintenant en mesure de donner une généralisation :

Proposition 2.4 (Cobham, 1972)**.** *Soit $k \in \mathbb{N}^*$ et soit $(u_n)_n$ une suite à valeurs dans $\{0, \ldots, k-1\}$, alors les trois conditions suivantes sont équivalentes :*

i) la suite $(u_n)_n$ est k-automatique

ii) le noyau de u, K_u, est fini

iii) la suite $(u_n)_n$ est image d'un point fixe d'un morphisme k-uniforme

Démonstration. $i) \Rightarrow ii)$. Soit $\mathcal{A}$ un automate engendrant u. Prenons $v \in K_u$, on peut trouver $\alpha \in \mathbb{N}$ et $r \in [0; k-1]$ tels que $v = (u(k^\alpha n + r))_n$. La notation $\widehat{\langle r \rangle_k}$ désigne ici l'écriture de r en base k à laquelle on ajoute des 0 à gauche pour obtenir un mot de taille α. On construit l'automate $\mathcal{A}'$ copie de $\mathcal{A}$ où l'on a remplacé l'état initial q_0 par l'état $q(v) := \delta(q_0, \widehat{\langle r \rangle_k})$. Comme lire n dans $\mathcal{A}'$ revient à se placer sur l'état $q(v)$ dans $\mathcal{A}$ puis faire les calculs dans $\mathcal{A}$, on voit que $\mathcal{A}'$ engendre v [2]. Finalement, $v \mapsto q(v)$ est une injection. K_u a au plus autant d'éléments qu'il y a d'états dans $\mathcal{A}$: il est bien fini.

$ii) \Rightarrow i)$. Comme automate engendrant u, il suffit de prendre $\mathcal{A} = \{\mathcal{Q}, \delta, q_o, \tau\}$ défini de la façon suivante [3] :

2. multiplier par k^α en base k correspond à ajouter α 0 à la droite du mot écrit en base k

3. d'après ce qui précède, on sait qu'il nous faudra au moins autant d'états que d'éléments de K_u

– l'ensemble des états $\mathcal{Q}$ est exactement K_u
– $q_0 = u$
– la fonction de transition δ est définie par

$$\delta(v, r) = v' \Leftrightarrow \forall n \in \mathbb{N}, v'(n) = v(k.n + r)$$

– la fonction de sortie τ est donnée par :

$$\tau(v) = v_0$$

Par récurrence sur la longueur de l'écriture en base k de n notée l_n, on montrerait : $\delta(q_0, \langle n \rangle_k) = (n' \mapsto u(k^{l_n}.n' + n))$, et donc $\tau(\delta(q_0, \langle n \rangle_k)) = u(k^{l_n}.0 + n) = u(n)$ ce que l'on voulait.

ii) $\Rightarrow$ *iii)*. K_u est fini, on peut écrire :

$$K_u = \{v_1, v_2, \ldots, v_d\} \text{ avec } v_1 = u$$

Posons $X = [\,0; k-1]$. L'alphabet sur lequel on va travailler est X^d. La suite qui sera point fixe est donnée, pour $n \in \mathbb{N}$, par $V_n := (v_1(n), v_2(n), \ldots, v_d(n))$ Pour chaque $r \in [\,0; k-1]$ on définit l'application :

$$s_r : X^d \to X^d \text{ telle que } \forall n \in \mathbb{N},\ V_{kn+r} = s_r(V_n)$$

On considère le morphisme $\sigma : (X^d)^* \to (X^d)^*$ k-uniforme donné par :

$$\forall W \in X^d,\ \sigma(W) = s_0(W)s_1(W)\ldots s_{k-1}(W)$$

où le produit est à prendre au sens de la concaténation. Par le critère 2.3, la suite $(V_n)_n := (v_1(n), v_2(n), \ldots, v_d(n))_n$ est bien point fixe de σ. $(u_n)_n$ est l'image de $(V_n)_n$ par la projection $\pi_1 : (\alpha_1, \ldots, \alpha_n) \in X^d \mapsto \alpha_1 \in X$

iii) $\Rightarrow$ *ii)*. Donnons nous B,σ,w et π comme dans la définition 2.2. Il suffit de montrer que :

$$\mathrm{Card}(K_w) < \infty$$

En écrivant $\sigma = s_0 s_1 \ldots s_{k-1}$ avec les s_i des applications de B dans lui même, le critère 2.3 assure que w doit vérifier les relations :

$$\forall n \in \mathbb{N}, \forall r \in [\,0; k-1]\ \ s_r(w(n)) = w(kn + r)$$

Il en découle que si $r = \sum_{j=0}^{\alpha-1} r_j k^j$, alors :

$$\forall n \in \mathbb{N},\ w(k^\alpha n + r) = s_{r_0} s_{r_1} \ldots s_{r_{\alpha-1}}(w(n))$$

Comme il y a au plus $\mathrm{Card}(B)^{\mathrm{Card}(B)}$ applications de B dans lui même, il y a un nombre fini d'applications de la forme $s_{r_0} s_{r_1} \ldots s_{r_{\alpha-1}}$ ce qui achève la preuve. □

3 Le théorème de Christol

Nous présentons dans cette partie un résultat dû à Gilles Christol. Ici, les suites considérées sont q-automatiques, avec $q = p^n$ et p premier. Le théorème donne une caractérisation surprenante du fait d'être automatique en terme d'algébricité de série sur le corps $\mathbb{F}_q(X)$.

Notation. $\mathbb{F}_q[[X]]$ désigne l'ensemble des séries formelles à coefficients dans $\mathbb{F}_q$. Soit v une suite à valeur dans $\mathbb{F}_q$, on notera $F(v) = F_v := \sum_n v_n X^n$.

Pour nous donner une idée du résultat, voyons ce qui se passe avec nos deux exemples de suites automatiques :

Exemple 3.1. On se place dans $\mathbb{F}_2$ [4]. En utilisant les relations définissant t, il vient :

$$\begin{aligned} F(X) := \sum_n t_n X^n &= \sum_n t_{2n} X^{2n} + \sum_n t_{2n+1} X^{2n+1} \\ &= \sum_n t_n X^{2n} + \sum_n X^{2n+1} + \sum_n t_n X^{2n+1} \\ &= F(X^2) + \frac{X}{1+X^2} + XF(X^2) \\ &= (1+X)F(X)^2 + \frac{X}{1+X^2} \end{aligned}$$

F est donc algébrique sur le corps $\mathbb{F}_2(X)$ des fractions rationnelles modulo 2. Elle est racine du polynome P à coefficients dans $\mathbb{F}_2(X)$:

$$P(Y) := (1+X)Y^2 + Y + \frac{X}{1+X^2}$$

Exemple 3.2. Par des calculs analogues s'obtient la relation suivante pour la suite de Rudin-Shapiro (r_n) :

$$(1+X)^5 F^2(X) + (1+X)^4 F(X) + X^3 = 0$$

Avant de poursuivre, il est utile d'introduire une famille d'opérateurs sur les séries formelles.

Définition 3.3. Soit $r \in \{0, \dots, q-1\}$. On définit l'opérateur Λ_r :

$$\Lambda_r : \sum_n a_n X^n \in \mathbb{F}_q[[X]] \mapsto \sum_n a_{qn+r} X^n$$

Le lemme suivant se vérifie par un calcul simple.

Lemme 3.4. *Soit $r \in \{0, \dots, q-1\}$ et $F, G \in \mathbb{F}_q[[X]]$. Alors :*

$$\Lambda_r(FG^q) = \Lambda_r(F)G$$

4. on confond donc + et -, et de plus, pour une série formelle F, $F(X^2) = F(X)^2$

Précisons maintenant ce que veut dire être algébrique sur $\mathbb{F}_q$ pour une série formelle F.

Lemme 3.5. *Soit $F = \sum_n a_n X^n$ une série formelle à coefficients dans $\mathbb{F}_q$. Alors F est algébrique si et seulement si on peut trouver des polynômes $B_0(X)$, $B_1(X), \ldots, B_t(X)$, non tous nuls tels que :*

$$B_0 F + B_1 F^q + \cdots + B_t F^{q^t} = 0$$

On peut supposer en outre que $B_0 \neq 0$.

Démonstration. Si F est algébrique sur $\mathbb{F}_q(X)$, la famille $\left(F^{q^t}\right)_{t\geq 0}$ est liée, d'où une relation de liaison non triviale de la forme de l'énoncé. Réciproquement une telle relation assure l'algébricité de F.
Prenons $t \in \mathbb{N}$ minimal tel que l'on ait une relation de liaison comme ci-dessus et montrons qu'alors $B_0 \neq 0$. Sinon :

$$B_1 F^q + \cdots + B_t F^{q^t} = 0$$

Donc, en utilisant le lemme 3.4 :

$$\Lambda_r\left(B_1 F^q + \cdots + B_t F^{q^t}\right) \overset{3.4}{=} \Lambda_r\big(B_1\big)F + \cdots + \Lambda_r\big(B_t\big)F^{q^{t-1}} = 0$$

Ce qui contredit la minimalité de t et achève la preuve. □

Théorème 3.6 (Christol, 1979)**.**[5] *Soit p un nombre premier ≥ 2 et q une puissance de p. Une suite $(u_n)_n$ à valeurs dans $\mathbb{F}_q$ est q-automatique si et seulement si la série formelle $F(u) = \sum_n u_n X^n$ est algébrique sur $\mathbb{F}_q(X)$.*

Démonstration. Sens direct. Soit $(u_n)_n$ une suite q-automatique. On se sert de l'équivalence prouvée en 2.4 : le q-noyau K_u est fini. Posons comme avant $d = \mathrm{Card}(K_u)$.

$$F(u) = \sum_{r=0}^{q-1} X^r \sum_n u_{qn+r} X^{nq}$$

Et comme dans $\mathbb{F}_q$, $G(X^q) = G(X)^q$

$$F(u) = \sum_{r=0}^{q-1} X^r \left(\sum_n u_{qn+r} X^n\right)^q$$

Ce qui montre que

$$F(u) \in \mathrm{Vect}_{v\in K_u}\langle F(v)^q\rangle$$

Puis, de la même manière :

$$\forall k \leq d : F(u)^{q^k} \in \mathrm{Vect}_{v\in K_u}\langle F(v)^{q^{k+1}}\rangle$$

5. en fait la version donnée ici est due à Christol, Kamae, Mendès France et Rauzy. Dans sa version initiale, le théorème ne portait que sur des suites à valeurs dans $\{0, 1\}$

Et donc par récurrence :

$$\forall k \leq d : F(u)^{q^k} \in \text{Vect}_{v \in K_u} \langle F(v)^{q^{d+1}} \rangle$$

Or,

$$\dim \text{Vect}_{v \in K_u} \langle F(v)^{q^{d+1}} \rangle \leq \text{Card}(K_u)$$

Ainsi la famille $\{F(u), F(u)^q, \ldots, F(u)^{q^d}\}$ est liée.

Sens indirect. Supposons la série $F(u)$ algébrique sur $\mathbb{F}_q(X)$. D'après le lemme 3.5, on peut trouver une relation de liaison de la forme :

$$\sum_{i=0}^{t} B_i(X) F_u(X)^{q^i} = 0$$

Où les B_i sont des polynômes $\in \mathbb{F}_q[X]$ avec $B_0 \neq 0$. Posons $G_u = F_u / B_0$. Il vient $G_u = \sum_{i=0}^{t} C_i G_u^{q^i}$ avec $C_i = -B_i B_0^{q^i - 2}$. Soit $N := max\{d^\circ B_0, d^\circ C_1, \ldots, d^\circ C_t\}$ et soit $\mathcal{H}$ l'ensemble suivant :

$$\mathcal{H} := \left\{ H \in \mathbb{F}_q[[X]] \Big/ H = \sum_{i=0}^{t} D_i G_u^{q^i} \text{ et } D_i \in \mathbb{F}_q[X], d^\circ D_i \leq N \right\}$$

$\mathcal{H}$ est un ensemble fini contenant $F_u = B_0 G_u$. On introduit comme dans 3.3, pour chaque $r < q$, l'application Λ_r. Pour montrer que K_u est fini, il suffit de voir que $\mathcal{H}$ est stable par chacune des Λ_r. Soit $H \in \mathcal{H}$. Ecrivons $H = \sum_{i=0}^{t} D_i G_u^{q^i}$ avec les D_i polynômes de degré $\leq N$.

$$\begin{aligned} \Lambda_r(H) &= \Lambda_r\left(D_0 G_u + \sum_{i=1}^{t} D_i G_u^{q^i} \right) = \Lambda_r\left(\sum_{i=1}^{t} (D_0 C_i + D_i) G_u^{q^i} \right) \\ &\overset{3.4}{=} \sum_{i=1}^{t} \Lambda_r\big(D_0 C_i + D_i\big) G_u^{q^{i-1}} \end{aligned}$$

Comme $d^\circ \Lambda_r\big(D_0 C_i + D_i\big) \leq \frac{d^\circ\left(D_0 C_i + D_i\right)}{q} \leq \frac{2N}{q} \leq N$, on peut conclure que $\Lambda_r(H) \in \mathcal{H}$. □

4 Divers

Dans cette section, on présente quelques autres résultats importants sur les suites automatiques.

4.1 Conséquences du théorème de Christol

Cobham a établi le théorème qui suit, dont on peut trouver une démonstration dans [1] :

Théorème 4.1. *Soit k et l deux entiers multiplicativement indépendants*[6] *et soit u une suite k **et** l-automatique. Alors u est ultimement périodique.*

Ceci joint au théorème de Christol permet d'affirmer :

Corollaire 4.2. *Soit q_1 et q_2 multiplicativement indépendants et u telle que F_u soit à la fois algébrique sur $\mathbb{F}_{q_1}$ et $\mathbb{F}_{q_2}$. Alors u est ultimement périodique.*

Remarque. L'énoncé précédant est ambigu : on n'a défini F_u dans $\mathbb{F}_q[[X]]$ que si u est à valeurs dans $\mathbb{F}_q$. Il suffit pour parler en toute généralité, si u est à valeur dans un ensemble fini Δ de cardinal $\leq q$, de choisir une injection $\phi : \Delta \to \mathbb{F}_q$ et de considérer $\sum_n \phi(u_n)X^n$.

En particulier, par contraposée, les suites de Thue-Morse ou de Rudin-Shapiro ne sont pas 3-automatiques. Il est intéressant de comparer 4.2 à la conjecture :

Conjecture. *Soit $(u_n)_{n\geq 0} \in \{0,1\}^{\mathbb{N}}$ telle que les deux nombres réels $\sum_{n\geq 0} u_n 2^{-n}$ et $\sum_{n\geq 0} u_n 3^{-n}$ sont algébriques sur $\mathbb{Q}$ alors ces deux nombres sont rationnels.*

Par ailleurs, on montre en construisant un automate que si deux suites u et v à valeurs dans $\{0,\ldots,k-1\}$ sont k-automatiques, leur produit uv l'est encore. En invoquant le théorème de Christol, on obtient le

Corollaire 4.3. *Soit u et v à valeurs $\{0,\ldots,q-1\}$ telles que F_u et F_v sont algébriques sur $\mathbb{F}_q(X)$. Alors le produit de Hadamard de F_u et F_v égal à F_{uv} est aussi algébrique.*

4.2 Complexité

Prenons $k \in \mathbb{N}$ et considérons les suites à valeurs dans $\{0,\ldots,k-1\}$. Pour presque toute (au sens de la mesure de Lebesgue) suite et tout mot w arbitraire de $\{0,\ldots,k-1\}$, w apparaît (une infinité de fois) dans l'écriture en base 2 de x. Ceci n'est pas vrai pour les suites automatiques. Plus précisément, on dispose du résultat suivant :

Proposition 4.4. *Soit $u \in \{0,\ldots,k-1\}^{\mathbb{N}}$ k-automatique. Notons $\rho_u(n)$ le nombre de mots de taille n qui apparaissent dans u. Alors :*

$$\rho_u(n) = \begin{cases} O(1), & \text{si } u \text{ est ultimement périodique} \\ \Theta(n), & \text{sinon} \end{cases}$$

Remarque 1. Le premier cas se produit comme suggéré dans l'introduction. On peut montrer qu'une suite ultimement périodique est 1-automatique et réciproquement.

Remarque 2. Ceci suggère que la notion de suite automatique est, en un certain sens, orthogonale à celle de suite "aléatoire". Sur un alphabet de taille k : pour une telle suite on a dit que $\rho_u(n) = k^n$.

6. i.e. $k^\alpha = l^\beta \Rightarrow \alpha = \beta = 0$

Remerciements

Je tiens à remercier Jean-Paul Allouche pour m'avoir reçu si rapidement et conseillé si efficacement.

Références

[1] Jean-Paul Allouche et Jeffrey Shallit. *Automatic Sequences*, Cambridge University Press (2003).

[2] Jean-Paul Allouche. *Automates finis en théorie des nombres*, Expo. Math 5 (1987).

[3] Michel Dekking, Michel Mendès France et Alf van der Poorten. *Folds !*, Math. Intelligencer 5 (1982).