Arithmetic properties of automata : regular sequences

By *J. H. Loxton* and *A. J. van der Poorten* at Macquarie

Hartmanis and Stearns [7] have posed the problem of finding an algebraic irrational whose decimal expansion is real-time computable. No such number is known and indeed it seems unlikely that any such number exists. In this paper, we prove a new transcendence theorem which answers the Hartmanis-Stearns question in the negative for a restricted class of machines. More precisely, we show that the decimal expansion of an algebraic irrational cannot be generated by a finite automaton. In this sense, such expansions are relatively complicated, but we are still a long way from proving that they are random, or even normal. This connection between the Hartmanis-Stearns problem and transcendence theory was noted by Cobham [5] and our argument follows the spirit of his remarks.

In a previous paper [9], we tackled this problem by a technique which has become known as Mahler's method. The method gives transcendence results for functions of several complex variables which satisfy functional equations of a certain type. (See [8] for a general description and survey of results.) For the application to automata, one only needs functions of one complex variable evaluated at very special points, but in [9] we were not able to verify one technical hypothesis that seemed unavoidable in the transcendence proof. The present work follows a different route and replaces and completes our previous attempt in this case. The principal change is a better approach to algebraic independence, inspired by Mahler's work on the algebraic approximation of functions [11]. The auxiliary results on this subject in section 2 are completely general and of interest in their own right. The transcendence argument itself in sections 3 and 4 relies on little more than first principles and reverts to the style which made the original applications of Mahler's method so appealing.

There are a number of essentially equivalent formulations of the transcendence theorem. The application to automata following Cobham [4] is detailed in section 1. Another version suggested by Mendès France [12] yields the following result. Let $f = \sum f_n X^n$ be a formal Laurent series over a finite field $\mathbf{F}_q$ which is algebraic over the field $\mathbf{F}_q(X)$ of rational functions and suppose that the coefficients f_n take the values 0 or 1. (In fact, the last hypothesis does not lose any generality.) If f is not in $\mathbf{F}_q(X)$, then the number $\ldots f_{-1}f_0 \cdot f_1 f_2 \ldots$, regarded as expanded in any base $b \geq 2$, is transcendental.

Finite automata and Mahler functions

1. Let t be an integer greater than 1. A t-automaton consists of a finite set S of states containing a distinguished element i, the initial state, and a subset F of acceptance or final states, related by a map

$$\tau : \{0, 1, 2, \ldots, t-1\} \times S \longrightarrow S\,,$$

called the transition function. A word w in $\cup_{n=0}^{\infty}\{0, 1, 2, \ldots, t-1\}^n$ is said to be accepted by the automaton if w sends the state i to a state in F. Thus, if $i = x_0$, $w = \omega_0\omega_1 \ldots \omega_{n-1}$ with the ω_i in $\{0, 1, 2, \ldots, t-1\}$ and $x_{k+1} = \tau(\omega_k, x_k)$, then w sends x_0 to x_n, so it is accepted by the automaton if and only if x_n is in F. The language $\mathcal{L}$ of all words accepted by the automaton is said to be generated by the automaton. The words of $\mathcal{L}$ may be interpreted as natural numbers represented in base t, yielding a formal power series

$$L(X) = \sum_{n \in \mathcal{L}} X^n\,.$$

Conversely, suppose we associate a symbol with each state in S. (It is convenient below to associate the symbol j with the state x_j.) Denote by $(h)_t$ the word expressing the positive integer h in base t, read from left to right. Such an h gives rise to a symbol β_h, say, which is the symbol associated with the final state reached by the automaton after processing the word $(h)_t$. (If i is the initial state, we always set $\tau(i, 0) = i$ so that initial zeros are ignored.) Then the formal power series

$$\sum_{h \geq 0} \beta_h X^h,$$

or the sequence (β_h), is appropriately described as t-automatic. In particular, the series L arises by allotting the symbol 1 to the states of F and the symbol 0 to the states in $S \setminus F$.

These matters are of interest in apparently distant areas of mathematics. Suppose that $\sum u_h X^h$ is a power series over $\mathbf{Q}$ representing an algebraic function. By Eisenstein's Theorem the set of rational primes dividing the denominators of the coefficients is finite. Suppose p is a prime not in that set. Christol, Kamae, Mendès France and Rauzy [3] prove that the sequence $(u_h \bmod p)$ is p-automatic. More generally, Denef and Lipschitz [6], show that, for any positive integer k, the sequence $(u_h \bmod p^k)$ is p^k-automatic. Conversely [3], if (u_h) is p-automatic then the series $\sum u_h X^h$ is algebraic over $\mathbf{F}_p(X)$.

Example [10]. Consider the transition table

	0	1	2	3
x_0	x_0	x_0	x_0	x_0
x_1	x_1	x_3	x_2	x_1
x_2	x_0	x_0	x_2	x_1
x_3	x_1	x_3	x_0	x_0

If x_1 is the initial state, the generated sequence $(\beta_n)_{n \geq 0}$ is the sequence $\lim_{n\to\infty} \theta^n(1)$ generated by the uniform (regular) substitution θ

$$0 \mapsto 0000\,, \quad 1 \mapsto 1321\,, \quad 2 \mapsto 0021\,, \quad 3 \mapsto 1300\,,$$

in the following sense:

$$\theta(1) = 1321\,,\ \theta^2(1) = \theta(1321) = \theta(1)\theta(3)\theta(2)\theta(1) = 1321130000211321\,,\ \ldots\ldots$$

This sequence yields the characteristic sequence of those non-negative integers which, in base 4, can be expressed omitting the digit 2 and using only the digits 0,1, and -1; n is such an integer if and only if β_n is either 1 or 3.

Cobham [4], [5] points out that, as the example suggests, finite t-automata and uniform t-substitutions are effectively the same thing. Thus we are led to consider tag machines or substitution automata defined as follows. Let $b_0, b_1, \ldots, b_{m-1}$ be a given alphabet of letters (or symbols), and suppose we are given a substitution

$$b_0 \mapsto w_0, b_1 \mapsto w_1, \ldots, b_{m-1} \mapsto w_{m-1}$$

with words w_i of finite length. Denote by

$$\beta_0\beta_1\beta_2\ldots\ldots$$

a fixed point of the substitution. For instance, the sequence stable under the substitution of the example is

1 3 2 1 1 3 0 0 0 0 2 1 1 3 2 1 ......
1321130000 2113211321 1300000000 0000000000 0021132113 2113000021 1321......

In the present paper we are concerned with the regular case in which each word w_i is of the same length, say of $t \geq 2$ symbols. This construction then gives the same t-automatic sequences (β_h) as before.

Consider the generating function $\sum \beta_h X^h$. More conveniently, associate with each symbol b_i its characteristic function

$$g_i(X) = \sum_{h \geq 0} u_{ih} X^h\,, \qquad u_{ih} = \begin{cases} 1 & \text{if } \beta_h = b_i \\ 0 & \text{otherwise,} \end{cases}$$

so that $\sum_{h\geq 0} \beta_h X^h = \sum_{i=0}^{m-1} b_i g_i(X)$. Note that $\beta_{th}\beta_{th+1}\ldots\beta_{t(h+1)-1}$ depends only on β_h. Accordingly write

$$v_{ijk} = \begin{cases} 1 & \text{if } b_i \text{ is the } (k+1)\text{-st symbol of the word } w_j \\ 0 & \text{if not,} \end{cases}$$

so that

$$u_{i,th+k} = \sum_{j=0}^{m-1} v_{ijk} u_{jh}\,.$$

In other words,

$$\begin{aligned} g_i(X) = \sum_{s=0}^{\infty} u_{is}X^s &= \sum_{h=0}^{\infty}\sum_{k=0}^{t-1} u_{i,th+k}X^{th+k} \\ &= \sum_{j=0}^{m-1}\left(\sum_{k=0}^{t-1} v_{ijk}X^k\right)\sum_{h=0}^{\infty} u_{jh}X^{th} \\ &= \sum_{j=0}^{m-1} p_{ij}(X)g_j(X^t)\,. \end{aligned}$$

If we denote by $\mathcal{A}(X)$ the $m \times m$ matrix

$$\mathcal{A}(X) = \Big(p_{ij}(X)\Big)_{0\le i,j\le m-1}$$

and by $g(X)$ the column vector $g(X) = \big(g_0(X), g_1(X), \ldots, g_{m-1}(X)\big)'$, then we have the matrix functional equation

$$g(X) = \mathcal{A}(X)g(X^t)\,.$$

Moreover every linear combination of the $g_i(X)$ satisfies a functional equation of the shape

$$\sum_{i=0}^{m} a_i(X)f(X^{t^i}) = 0\,,$$

with polynomial coefficients $a_i(X)$.

Special cases of such functional equations were studied by Mahler in the late twenties; see [8]. It is therefore appropriate to refer to these systems of equations as Mahler systems and to their solutions as Mahler functions.

Normal Approximation

2. Let m be a positive integer and $f(z) = (f_1(z), f_2(z), \ldots, f_m(z))$ be a given system of m formal power series $f_j(z)$ over a field $\mathbf{F}$ which do not all have zero constant term. (More succinctly, f does not vanish at 0.) Let $\rho = (\rho_1, \rho_2, \ldots, \rho_m)$ be a system of non-negative integer parameters ρ_j with sum $\sum \rho_j = \sigma$. There is a system $(a_1(z|\rho), a_2(z|\rho), \ldots, a_m(z|\rho))$ of polynomials $a_j(z|\rho)$ over $\mathbf{F}$, not all zero, which together with the remainder function

$$r(z|\rho) = \sum_{j=1}^{m} a_j(z|\rho)f_j(z)\,,$$

satisfy the inequalities

$$\deg a_j(z|\rho) \le \rho_j - 1 \qquad (1 \le j \le m)\,,$$

$$\operatorname{ord} r(z|\rho) \geq \sigma - 1 \ .$$

(Indeed, the requirements amount to $\sigma-1$ homogeneous linear equations in the σ unknown coefficients of the polynomials.) We call such a system $(a_1(z|\rho), a_2(z|\rho), \ldots, a_m(z|\rho))$ an approximation at ρ. It is convenient to say that the system $(a_1(z|\rho), a_2(z|\rho), \ldots, a_m(z|\rho))$ is of degree less than ρ. This implies a partial order on the parameter points: ρ' lies above ρ if $\rho'_j \geq \rho_j$ for $j = 1, 2, \ldots, m$.

Suppose the functions $f_j(z)$ are linearly independent over the field $\mathbf{F}(z)$. Amongst all approximations at ρ there is, up to normalisation by an element of $\mathbf{F}^\times$, a unique approximation $(b_1(z|\rho), b_2(z|\rho), \ldots, b_m(z|\rho))$ so that ord $\sum_{j=1}^m b_j(z|\rho) f_j(z)$ is maximal; say, equal to $\sigma - 1 + \kappa(\rho)$ for some $\kappa(\rho) \geq 0$. We call this approximation the best approximation at ρ, and refer to $\kappa(\rho)$ as the excess at ρ. To see the uniqueness of the best approximation, note that if there are two distinct equally good approximations, then a suitable linear combination of them provides a better approximation. However, if the $f_j(z)$ are linearly dependent the best approximation may not be unique. In this case, we set $\kappa(\rho) = \infty$ if there is a choice of approximating polynomials making $r(z|\rho) = 0$. Finally, we say that the system $f(z) = (f_1(z), f_2(z), \ldots, f_m(z))$ is normal at ρ if $\kappa(\rho) = 0$. Normality at ρ implies that there is (up to normalisation by an element of $\mathbf{F}^\times$) a unique approximation at ρ.

Proposition A. *If the system of formal power series $f(z) = (f_1(z), f_2(z), \ldots, f_m(z))$ over $\mathbf{F}$ does not vanish at 0, and does not have a normal approximation at any parameter point ρ' lying above ρ, then there is a system of polynomials $a(z) = (a_1(z), a_2(z), \ldots, a_m(z))$, of degree less than ρ, such that*

$$\sum_{j=1}^m a_j(z) f_j(z)$$

is identically zero.

Proof. Suppose that $\kappa(\rho)$ is finite and that $\kappa(\rho')$ is positive for every ρ' lying above ρ; recall that this includes the choice $\rho' = \rho$. Choose a parameter point ρ' lying above ρ where $\kappa(\rho')$ is minimal. Set $\sum \rho'_j = \sigma'$ and consider the best approximation

$$b(z|\rho') = (b_1(z|\rho'), b_2(z|\rho'), \ldots, b_m(z|\rho'))$$

at ρ'. We claim that for some h with $h = 1, 2, \ldots,$ or m, the best approximation at

$$\rho' + \delta_h = (\rho'_1 + \delta_{h1}, \rho'_2 + \delta_{h2}, \ldots, \rho'_m + \delta_{hm}) \ ,$$

where δ_{hj} is the usual Kronecker δ, is given by $b(z|\rho')$ and thus has order $\sigma' - 1 + \kappa(\rho')$ and excess $\kappa(\rho') - 1$.

Suppose otherwise, namely that for each h the best approximation at $\rho' + \delta_h$ has order at least $\sigma' + \kappa(\rho')$. Since $\kappa(\rho')$ is finite, the best approximation at $\rho' + \delta_h$, say

$$(B_{h1}(z|\rho'), B_{h2}(z|\rho'), \ldots, B_{hm}(z|\rho')) \ ,$$

must have the property that

$$\deg B_{hh}(z|\rho') = \rho'_h \ ,$$

since otherwise we contradict the definition of $b(z|\rho')$. We may therefore suppose that the approximations have been normalised so that the leading coefficient of each $B_{hh}(z|\rho')$ is equal to 1.

Now consider the matrix

$$\mathcal{B}(z|\rho') = \Big(B_{hj}(z|\rho')\Big)_{1\le h,j\le m} \ .$$

We refer to $\mathcal{B}(z|\rho')$ as the approximation matrix at ρ'. On the one hand we have

$$\det \mathcal{B}(z|\rho') = z^{\sigma'} + \text{terms of lower degree}\,.$$

This is plain since the diagonal supplies the unique term of degree σ'.

On the other hand, since f does not vanish at 0, we can choose an index i so that $f_i(0) \neq 0$ and we can use column operations to replace the j–th column of $\det\big(\mathcal{B}(z|\rho')\big)$ by the entries

$$f_i(z)^{-1}\sum_{j=1}^{m} B_{hj}(z|\rho')f_j(z) \qquad (h = 1, 2, \ldots, m\,).$$

Hence

$$\operatorname{ord}\big(\det \mathcal{B}(z|\rho')\big) \geq \sigma' + \kappa(\rho')\,.$$

These two conclusions entail $\kappa(\rho') = 0$, contrary to our hypothesis that $\kappa(\rho') > 0$ for every ρ' lying above ρ.

Thus, for some h, the excess at $\rho' + \delta_h$ is at most $\kappa(\rho') - 1$ contrary to the minimality in our choice of ρ'. It follows finally that $\kappa(\rho')$ must be infinite for all ρ' lying above ρ proving the proposition.

We restate the proposition in the form in which we use it later and draw some further conclusions from the proof.

Theorem 1 (Normality zig-zag theorem for linearly independent series)**.** *Let $f(z) = (f_1(z), f_2(z), \ldots, f_m(z))$ be a system of m formal power series over a field $\mathbf{F}$ which does not vanish at 0 and which has entries linearly independent over the ring of polynomials $\mathbf{F}[z]$. Then $f(z)$ is normal at an infinite sequence of parameter points $\rho(0) = 0, \rho(1), \rho(2), \ldots\ldots$ with $\rho(n+1)$ lying above $\rho(n)$ and $\sum_{j=1}^{m}\rho_j(n) = n$ for all n and $\min_{1\le j\le m}\rho_j(n) \to \infty$ as $n \to \infty$. At each point ρ of this normality zig-zag the approximation matrix*

$$\mathcal{B}(z|\rho) = \Big(B_{hj}(z|\rho)\Big)_{1\le h,j\le m}$$

of polynomials $B_{hj}(z|\rho)$ *satisfying* $\deg B_{hj}(z|\rho) \le \rho_j - 1 + \delta_{hj} \quad (h, j = 1, 2, \ldots, m)$, *and* $\operatorname{ord} \sum_{j=1}^{m} B_{hj}(z|\rho) f_j(z) \ge \sigma \quad (h = 1, 2, \ldots, m)$, *remains nonsingular when* z *is specialised to any element of* $\mathbf{F}^{\times}$.

Proof. The existence of the required normality zig-zag follows from the construction in the proof of the proposition. Also, from the proof we see that normality at ρ entails

$$\det \mathcal{B}(z|\rho) = z^{\sigma}$$

for some non-zero constant c and this gives the last assertion of the theorem.

The notions employed above arise from the manuscript of Mahler on ‘Perfect Systems’ [11] as expanded upon by Coates [2]. Our theorem is an extension of the normality zig-zag theorem of Coates.

Linear independence

3. Let m be a positive integer and $f(z) = (f_1(z), f_2(z), \ldots, f_m(z))$ be a given system of m power series $f_j(z)$ over an algebraic number field $\mathbf{K}$, which are linearly independent over $\mathbf{K}[z]$ and analytic at the origin. Suppose $f(z)$ does not vanish at 0. Let ρ be a parameter point at which $f(z)$ is normal and at which $\rho_{\min} = \min_{1 \le j \le m} \rho_j$ is as large as will be required in the sequel. The existence of ρ is guaranteed by theorem 1. Let

$$\mathcal{B}(z|\rho) = \Big(B_{hj}(z|\rho) \Big)_{1 \le h, j \le m}$$

be the approximation matrix at ρ.

We shall operate with the valuations on $\mathbf{K}$ and we employ the usual notation. In particular, $|\cdot|_v$ denotes a normalised valuation on $\mathbf{K}$ and $\mathbf{K}_v$ denotes the corresponding v-adic completion. The essence in normalising the valuations is to do so coherently, in order to obtain the product formula. A useful normalisation, is to start from the naïve valuation

$$\begin{cases} \|p\|_v = p^{-1} & \text{if } v \mid p \text{ and } p \text{ is a rational prime,} \\ \|n\|_v = n & \text{if } v \mid \infty \text{ and } n \text{ is a non-negative integer,} \end{cases}$$

and then to set

$$|x|_v = \|x\|^{[\mathbf{K}_v : \mathbf{Q}_v]/[\mathbf{K}:\mathbf{Q}]}.$$

Let V_- be a finite set of valuations of $\mathbf{K}$. For each v in V_-, choose δ_v with $0 < \delta_v < 1$ such that each of the series $f_j(z)$ converges for z in $\mathbf{K}_v$ with $|z|_v \le \delta_v$ and such that the $f_j(z)$ have no simultaneous zeros on this disc. Let V_+ be the set of the remaining valuations on $\mathbf{K}$.

Let β be an element of $\mathbf{K}^\times$ satisfying $|\beta|_v \le \delta_v$ for v in V_- and $|\beta|_v \ge 1$ for v in V_+. For each v in V_-, we can assign a value in $\mathbf{K}_v$ to $f_j(\beta)$. We suppose there are s linearly independent linear relations

$$\sum_{j=1}^{m} c_{ij} f_j(\beta) = 0 \qquad (i = 1, 2, \ldots, s\,),$$

in $\mathbf{K}_v$ for each v in V_- and with coefficients c_{ij} from $\mathbf{K}$.

By Theorem 1, the matrix

$$\mathcal{B}(\beta|\rho) = \Big(B_{hj}(\beta|\rho) \Big)_{1\le h,j\le m} \,.$$

is nonsingular. Set $B_{hj}(\beta|\rho) = b_{hj} \quad (h, j = 1, 2, \ldots, m\,)$. The s vectors $(c_{i1}, c_{i2}, \ldots, c_{im})$ and $m - s$ of the m vectors $(b_{h1}, b_{h2}, \ldots, b_{hm})$ are linearly independent, so there is no loss of generality in supposing, for convenience, that the determinant

$$\Delta = \begin{vmatrix} c_{11} & c_{12} & \ldots & c_{1m} \\ \vdots & \vdots & \ddots & \vdots \\ c_{s1} & c_{s2} & \ldots & c_{sm} \\ b_{s+1,1} & b_{s+1,2} & \ldots & b_{s+1,m} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m1} & b_{m2} & \ldots & b_{mm} \end{vmatrix}$$

does not vanish.

We proceed to estimate Δ at each place v of $\mathbf{K}$. As usual, set $\log^+ x = \max(0, \log x)$. For an archimedean valuation v, let $|b|_v$ be an upper bound for the v-adic length of the polynomials $B_{hj}(z|\rho)$ and for a non-archimedean v let $|b|_v$ denote the v-adic maximum of the coefficients of these polynomials. Also, let $|c|_v$ denote an upper bound for the v-adic values of the c_{ij}. Let V denote the set of all the valuations of $\mathbf{K}$. The absolute logarithmic height $\mathrm{h}(x)$ for $x \in \mathbf{K}$ is given by

$$\mathrm{h}(x) = \sum_{v\in V} \log^+ |x|_v \,,$$

and it is useful to write

$$\mathrm{h}(b) = \sum_{v\in V} \log^+ |b|_v \text{ and } \mathrm{h}(c) = \sum_{v\in V} \log^+ |c|_v \,.$$

Take $v \in V_+$ so that $|\beta|_v \ge 1$. We have the crude upper bound

$$\log|\Delta|_v \le \log^+ |m!|_v + s\log^+ |c|_v + (m-s)\log^+ |b|_v + (\sigma - s\rho_{\min})\log|\beta|_v \ .$$

Next, take v in V_- so that $|\beta|_v < 1$. Choose an index h so that $f_h(\beta) \neq 0$ in $\mathbf{K}_v$. By operating on the h-th column of Δ, we can write

$$\Delta = (f_h(\beta))^{-1} \begin{vmatrix} c_{11} & \dots & 0 & \dots & c_{1m} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ c_{s1} & \dots & 0 & \dots & c_{sm} \\ b_{s+1,1} & \dots & r_{s+1} & \dots & b_{s+1,m} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ b_{m1} & \dots & r_m & \dots & b_{mm} \end{vmatrix},$$

where we have set $r_i = r_i(\beta) = \sum b_{ij} f_j(\beta)$ for $i = 1, 2, \dots, m$, and have used the fact that $\sum c_{ij} f_j(\beta) = 0$ in K_v for $i = 1, 2, \dots, s$. For $v \in V_-$, we may set $|(f_h(\beta))^{-1} r_i|_v \leq |r|_v |\beta|_v^\sigma \quad (i = 1, 2, \dots, m)$ where $|r|_v$ is independent of β. Of course, we lose no generality if we choose $|b|_v$ so that $|r|_v \leq |b|_v$. We then have, on expanding by the h-th column,

$$\log|\Delta|_v \leq \log^+ |m!|_v + s\log^+ |c|_v + (m-s-1)\log^+ |b|_v + \log^+ |r|_v + \sigma \log|\beta|_v \, .$$

The product formula for a non-zero element x of $\mathbf{K}$ states that

$$\sum_{v \in V} \log|x|_v = 0 \, .$$

By adding the estimates for $\log|\Delta|_v$ over all v in V we obtain

$$\begin{aligned} 0 &= \sum_{v \in V} \log|\Delta|_v \\ &\leq \sum_{v \in V} \log^+ |m!|_v + s \sum_{v \in V} \log^+ |c|_v + (m-s) \sum_{v \in V} \log^+ |b|_v - s\rho_{\min} \sum_{v \in V_+} \log|\beta|_v \ , \end{aligned}$$

that is,

$$s\rho_{\min}\mathrm{h}(\beta) - (m-s)\mathrm{h}(b) - s\mathrm{h}(c) - \mathrm{h}(m!) \leq 0 \, .$$

Let $f(z) = (f_1(z), f_2(z), \dots, f_m(z))'$ be a vector whose components are formal power series over the field $\overline{\mathbf{Q}}$ of all algebraic numbers. Suppose that the vector $f(z)$ satisfies the system of functional equations

$$f(z) = \mathcal{A}(z) f(z^t) \, ,$$

where $\mathcal{A}(z)$ is an $m \times m$ matrix of rational functions over $\overline{\mathbf{Q}}$, and t is is a rational integer greater than 1. The coefficients of the series $f_j(z)$ lie in an algebraic number field, $\mathbf{K}$ say, of finite degree because a finite number of coefficients of the series and the coefficients of the entries of $\mathcal{A}(z)$ determine the series completely. We therefore lose no generality in the proposition below in assuming everything is defined over an algebraic number field.

Proposition B. *Let $f(z)=(f_1(z),f_2(z),\dots,f_m(z))'$ be a vector whose components are formal power series over an algebraic number field* $\mathbf{K}$ *and linearly independent over the ring of polynomials* $\mathbf{K}[z]$ *and suppose that the vector $f(z)$ satisfies the system of functional equations*

$$f(z)=\mathcal{A}(z)f(z^t)\,,$$

where $\mathcal{A}(z)$ is a non-singular $m\times m$ matrix of rational functions over $\mathbf{K}$, *and t is a rational integer greater than 1. Let α be a nonzero number in* $\mathbf{K}$ *and let V_- denote the set of valuations of* $\mathbf{K}$ *where $|\alpha|_v<1$. Suppose that the series $f_j(\alpha)$ converge v-adically for each j and each v in V_- and that the matrices $\mathcal{A}(\alpha^{t^k})\quad(t=1,2,\dots)$ are defined and nonsingular. Let*

$$a(z)=a_1f_1(z)+a_2f_2(z)+\cdots+a_mf_m(z)$$

be a linear form with coefficients a_j in $\mathbf{K}$ *and not all zero. Then $a(\alpha)$ cannot vanish in $\mathbf{K}_v$ for every v in V_-.*

Proof. The proof is by contradiction. That is, we assume that $a(\alpha)=0$ in $\mathbf{K}_v$ for each v in V_-.

We may suppose that the vector $f(z)$ does not vanish at 0, by multiplying $f(z)$ and $\mathcal{A}(z)$ by appropriate powers of z, if necessary. The functional equations ensure that the series $f_j(z)$ are v-adically analytic at the origin for all places v of $\mathbf{K}$. Set

$$\beta=\alpha^{t^k}.$$

If k is sufficiently large, then β satisfies the requirements of the previous work in this section, that is $|\beta|_v\le\delta_v$ for each v in V_- and $|\beta|_v\ge1$ for the remaining valuations on $\mathbf{K}$. (To verify that $f_j(\beta)\ne0$ for some j, note that the analyticity at the origin entails that none of the f_j can vanish at more than finitely many of the points α^{t^k}.)

Denote by a the vector $a=(a_1,a_2,\dots,a_m)'$ so that the given linear form is $a(z)=a\cdot f(z)$. Iteration of the functional equation for $f(z)$ leads to

$$f(z)=\mathcal{A}^{(k)}(z)f(z^{t^k})\quad\text{with}\quad\mathcal{A}^{(k)}(z)=\mathcal{A}(z)\mathcal{A}(z^t)\mathcal{A}(z^{t^2})\cdots\mathcal{A}(z^{t^{k-1}})\,.$$

We can substitute for $f(z)$ in the linear form to obtain

$$a(z)=a\cdot f(z)=a\cdot\mathcal{A}^{(k)}(z)f(z^{t^k})=(\mathcal{A}^{(k)}(z))'a\cdot f(z^{t^k})\,.$$

We therefore set $c=(\mathcal{A}^{(k)}(\alpha))'\,a$ and $c(z)=c\cdot f(z)$ so that $c(\beta)=0$ in $\mathbf{K}_v$ for each v in V_-. The nonsingularity of the $\mathcal{A}^{(k)}(\alpha)$ guarantees that the linear form $c(z)$ does not vanish identically.

Now apply the fundamental inequality developed earlier in this section. We take $\beta=\alpha^{t^k}$ and $c=(\mathcal{A}^{(k)}(\alpha))'a$, so that $\mathrm{h}(\beta)=t^k\mathrm{h}(\alpha)$ and $\mathrm{h}(c)\le Ct^k$, with a positive constant C

independent of k. The remaining terms are independent of k. Choose the parameter point ρ so that $\rho_{\min} > 2C/\mathrm{h}(\alpha)$. Then choose k so that $t^k > (\mathrm{h}(m!) + (m-s)\mathrm{h}(b))/s\rho_{min}\mathrm{h}(\alpha)$. These choices contradict the fundamental inequality. So we do indeed have $a(\alpha) \neq 0$ in $\mathbf{K}_v$ for some v in V_-, as required.

The proposition yields a very general result on the nonvanishing of certain linear forms in the values of our functions at certain algebraic points. This is remarkable, because the result is obtained without any explicit data on the normality zig-zags possessed by the given system $f(z) = (f_1(z), f_2(z), \ldots, f_m(z))$, and without any information on the size of the coefficients of the polynomials comprising the approximation matrix at ρ. There is a price to be paid, however, because the contradiction in the proof depends on the assumption of a "global" relation valid in all the relevant valuations. We note that just such a condition arises in Bombieri's work on G–functions [1].

Algebraic independence

4. The main theorem on algebraic independence is a straightforward corollary of the last proposition. We restrict the statement to the situation of direct interest for the analysis of automata. There are some obvious extensions to number fields, but these consequences of the proposition are not particularly satisfying.

Theorem 2 (Algebraic independence of automatic numbers)**.** *Let $f(z) = (f_1(z), f_2(z), \ldots$ $\ldots, f_m(z))'$ be a vector whose components are formal power series over the field of rational numbers. Suppose, further, that the vector $f(z)$ satisfies the system of functional equations*

$$f(z) = \mathcal{A}(z)f(z^t)\,,$$

where $\mathcal{A}(z)$ is a non-singular $m \times m$ matrix of rational functions over $\mathbf{Q}$ and t is is a rational integer greater than 1. Set $\alpha = 1/n$, where n is an integer greater than 1, and suppose α is in the domain of convergence of each of the m series $f_j(z)$ and that the matrices

$$\mathcal{A}(\alpha^{t^k}) \quad (k = 1, 2, \ldots\ldots)$$

are defined and nonsingular. Then

$$\text{transc. deg.}_{\mathbf{Q}}\,(f_1(\alpha), f_2(\alpha), \ldots, f_m(\alpha)) = \text{transc. deg.}_{\mathbf{C}(z)}\,(f_1(z), f_2(z), \ldots, f_m(z))\ .$$

Proof. Suppose that there is an algebraic relation between the numbers $f_1(\alpha), \ldots, f_m(\alpha)$ with rational coefficients which is not obtained by specialisation of an algebraic relation between the functions $f_1(z), \ldots, f_m(z)$ with rational function coefficients. Then there is a set of monomials $f^\mu = f_1^{\mu_1} \cdots f_m^{\mu_m}$ with $0 \leq \mu_j \leq M$, say, such that the $f^\mu(\alpha)$ are linearly dependent over $\mathbf{Q}$ but the $f^{\mu}(z)$ are linearly independent over $\mathbf{C}(z)$. Let f^* be a vector whose components are these monomials supplemented by further monomials to yield a basis for the set of all monomials of degree not exceeding M. Because the components of f^* are a basis of monomials, the functional equation $f(z) = \mathcal{A}(z)f(z^t)$ determines a unique transformation $\mathcal{A}^*(z) = \Omega(\mathcal{A}(z))$ such that $f^*(z) = \mathcal{A}^*(z))f^*(z^t)$. Moreover, since $\mathcal{A}(z)$ is non-singular, we can reverse the argument and see that $\Omega(\mathcal{A}(z))^{-1} = \Omega(\mathcal{A}(z)^{-1})$. In particular, $\mathcal{A}^*(\alpha^{t^k})$ is non-singular for every positive integer k. In the terms of Proposition B, everything is now defined over $\mathbf{Q}$, and V_- consists of just the ordinary valuation on $\mathbf{Q}$. By the proposition, the components of $f^*(\alpha)$ are linearly independent over $\mathbf{Q}$, contrary to our hypothesis. Thus

$$\text{transc. deg.}_{\mathbf{Q}}\,(f_1(\alpha), f_2(\alpha), \ldots, f_m(\alpha)) \geq \text{transc. deg.}_{\mathbf{C}(z)}\,(f_1(z), f_2(z), \ldots, f_m(z))\ .$$

Since the opposite inequality is trivial, we have the theorem.

Concluding remarks

5. As explained in §1, certain Mahler functions correspond to finite automata. These functions have power series expansions involving only finitely many distinct coefficients and so their radius of convergence, at every valuation, is 1. Let $f(X) = \sum \beta_h X^h$ be a t-automatic power series. The series represents a rational function if and only if the sequence (β_h) is (eventually) periodic. By the Pólya-Carlson Theorem, the series either represents a rational function or a function with the unit circle as natural boundary. In the latter case f is transcendental over $\mathbf{C}(z)$. Then the transcendence arguments apply to f and, in particular, for every integer $b \geq 2$ the number $f(1/b)$ is transcendental. This is to say that the t-automatic number

$$\beta_0{\cdot}\beta_1\beta_2\cdots\cdots\,,$$

presented in base b is either rational or transcendental. In the contrapositive, as stated in the introduction,

The sequence of digits (f_h) of an irrational algebraic number $\ldots f_{-1}f_0{\cdot}f_1f_2\cdots\cdots$, presented in any base $b \geq 2$, cannot be generated by a t-automaton.

Our second introductory remark is explained as follows. If $f = \sum f_h X^h$ is a formal Laurent series over a finite field $\mathbf{F}_q$ and algebraic over the field $\mathbf{F}_q(X)$ of rational functions, then, by [3] the sequence (f_h) is p-automatic (with p the characteristic of $\mathbf{F}_q$). Since the expansion $\ldots f_{-1}f_0{\cdot}f_1f_2\cdots\cdots$, cannot represent a rational number, it must be transcendental.

References

[1] *E. Bombieri*, On G-functions, in Recent progress in analytic number theory, H. halberstam and C. Hooley, eds. (Academic Press, 1981), Chapter 24, Vol. 2, 1–67

[2] *J.H. Coates*, On the algebraic approximation of functions I, II, III, Proc. Kon. Nederl. Akad. v. Wetenschappen Ser. A **69** (1966), 421–434, 435–448, 449–461= Indag. Math. **28** (1966), 421–434, 435–448, 449–461

[3] *G. Christol, T. Kamae, M. Mendès France and G. Rauzy*, Suites algébriques, automates et substitutions, Bull. Soc. Math. France, **108** (1980), 401–419

[4] *A. Cobham*, On the Hartmanis-Stearns problem for a class of tag machines, Technical report **RC 2178**, IBM Research Centre, Yorktown Heights, New York, 1968

[5] *A. Cobham*, Uniform tag sequences, Math. Systems Theory **6** (1972), 164–192

[6] *J. Denef and L. Lipschitz*, Algebraic power series and diagonals, J. Number Theory, **26** (1987), 46–67

[7] *J. Hartmanis and R.E. Stearns*, On the computational complexity of algorithms, Trans. Amer. Math. Soc. **117** (1965), 285–306

[8] *J.H. Loxton and A.J. van der Poorten*, Transcendence and algebraic independence by a method of Mahler, in Transcendence theory – advances and applications, A. Baker and D.W. Masser, eds. (Academic Press, London and New York, 1977), Chapter 15, 211–226

[9] *J.H. Loxton and A.J. van der Poorten*, Arithmetic properties of the solutions of a class of functional equations, J. für Math. **330** (1982), 159–172

[10] *J.H. Loxton and A.J. van der Poorten*, An awful problem about integers in base four (to Paul Erdős on his 75th birthday), Acta Arith.

[11] *K. Mahler*, Perfect systems, Compositio Math. **19** (1968), 95–166

[12] *M. Mendès France*, Algebraic numbers and automata theory, Queen's papers in pure and applied mathematics **54** (1980), 79–89

School of Mathematics, Physics, Computing and Electronics, Macquarie University, Australia 2109