Rational Functions, Diagonals, Automata and Arithmetic

dedicated to the memory of Kurt Mahler

By *Leonard Lipshitz*[1] **at Purdue and** *Alfred J. van der Poorten*[2] **at Macquarie**

We mention some (perhaps surprising) connections among the objects of the title. Whilst the results are not new, we do give different proofs in a number of cases to emphasise the interrelationships.

1. Notions and Notation

1.1 By R we denote an integral domain; then $R[[x]]$ is the ring of formal power series over R in the variables $x = (x_1, \dots, x_n)$. We write a series $f(x) \in R[[x]]$ as $f(x) = \sum a_{i_1 \cdots i_n} x_1^{i_1} \cdots x_n^{i_n} = \sum a_\nu x^\nu$ where $\nu = (\nu_1, \dots, \nu_n)$ is a multi-index and $x^\nu = x_1^{\nu_1} \cdots x_n^{\nu_n}$. In the sequel R is usually the finite field $\mathbb{F}_p$ of p elements or the ring of p-adic integers $\mathbb{Z}_p$. One says that $f(x) \in R[[x]]$ is *algebraic* if it is algebraic over the quotient field of the polynomial ring $R[x]$. If $f(x) \in \mathbb{Z}_p[[x]]$ we say that f is *algebraic* mod p^s if there is an algebraic $g(x) \in \mathbb{Z}_p[[x]]$ with $f \equiv g \bmod p^s$.

1.2 Given $f(x, y) = \sum a_{ij} x^i y^j$ it is natural to refer to the series $I_{xy} f = \sum a_{ii} x^i$ as its *diagonal*. In general, if $f(x) = \sum a_{i_1 \cdots i_n} x_1^{i_1} \cdots x_n^{i_n}$ we define its diagonal $I_{12} f$ by $I_{12} f = \sum a_{i_1 i_1 i_3 \cdots i_n} x_1^{i_1} x_3^{i_3} \cdots x_n^{i_n}$; generally, for $k \neq l$, the other I_{kl} are defined correspondingly. By a *diagonal* we mean any composition of the I_{kl} s. The (complete) diagonal is $If = \sum a_{ii \cdots i} x_1^i$. We shall also need the *off-diagonals* $J_{kl} f = \sum_{i_k > i_l} a_{i_1 \cdots i_n} x_1^{i_1} \cdots x_n^{i_n}$.

The *Hadamard product* of series $f(x) = \sum a_\nu x^\nu$ and $g(x) = \sum b_\nu x^\nu$ is (their "child's product") $f * g(x) = \sum a_\nu b_\nu x^\nu$.

1.3 The operations of diagonal and Hadamard product are connected by:

$$f * g(x) = I_{1,n+1} \dots I_{n,2n} f(x_1, \dots, x_n) g(x_{n+1}, \dots, x_{2n}) \,; \tag{1.3.1}$$

$$I_{12} f = f * \big(\frac{1}{1 - x_1 x_2} \prod_{j=3}^{n} \frac{1}{1 - x_j} \big) \,. \tag{1.3.2}$$

Over $\mathbb{C}$, the diagonal is also given by the integral

$$\begin{aligned} I_{12} f(t, x_3 \dots, x_n) &= \frac{1}{2\pi i} \oint\limits_{x_1 x_2 = t} f(x_1, \dots, x_n) \frac{dx_1 \wedge dx_2}{dt} \\ &= \frac{1}{2\pi i} \oint\limits_{|y| = \epsilon} f(\frac{t}{y}, y, x_3, \dots, x_n) \frac{dy}{y} \,. \end{aligned} \tag{1.3.3}$$

Noting (1.3.1) and (1.3.2), one sees that if a ring of power series is closed under one of the two operations of taking diagonals or Hadamard products then it is also closed under the other operation. From (1.3.3) it follows (*cf* [19]) that the diagonal of a rational function $f(x, y)$ of two variables is an algebraic function. Indeed,

$$(If)(t) = \frac{1}{2\pi i} \oint\limits_{|y| = \epsilon} f(\frac{t}{y}, y) \frac{dy}{y} = \frac{1}{2\pi i} \oint\limits_{|y| = \epsilon} \frac{P(t, y)}{Q(t, y)} dy \,.$$

Writing $Q(t, y) = \prod \big(y - y_i(t) \big)$, where the $y_i(t)$ are algebraic, and evaluating the integral by residues verifies the claim. In fact, every algebraic power series of one variable is the diagonal of a rational power series of

[1] Supported in part by a grant from the NSF.

[2] Work supported in part by the Australian Research Council.

two variables ([19]) and, indeed, every algebraic power series in n variables is the diagonal of a rational power series in $2n$ variables ([16]). However, as we see below (§2.10), in characteristic zero diagonals of rational functions in more than two variables do not generally yield algebraic functions.

1.4 A *finite automaton* (*cf* [32]) is a 'machine' with a finite number of states $\mathcal{S}$, a finite input alphabet $\mathcal{I}$, and a finite output alphabet $\mathcal{O}$. In the immediate sequel we will take $\mathcal{I} = \{0, 1, \dots, p-1\}$ and will presume there to be n input tapes. The output $\mathcal{O}$ will consist of elements of $\mathbb{F}_p$. At each stage of the computation the automaton reads one digit from each tape and accordingly alters its internal state, reporting its new state as an element of $\mathbb{F}_p$. Thus, formally, a finite automaton is a transition function $\tau : \mathcal{I}^n \times \mathcal{S} \to \mathcal{S}$ and an output function $o : \mathcal{S} \to \mathbb{F}_p$. Some distinguished state is the initial state.

1.5 Such an automaton generates its characteristic function $f = \sum a_\nu x^\nu$ by the rule that a_ν is the output after the words $\nu = (\nu_1, \dots, \nu_n)$, each expressed in p-adic notation (that is, in base p), have been fully read from the input tapes. The principle of *irrelevance of symbols* is plain: the elements of the output alphabet just serve as markers—if $f = \sum a_\nu x^\nu$ is generated by a finite automaton then, for each output symbol i, so are the series $f^{(i)}(x) = \sum_{a_\nu = i} x^\nu$; and conversely. Thus our choice of output alphabet is of no matter. In §**3** we consider automata with output alphabet $\mathbb{Z}/p^s$. In the above spirit, notice that if for each $k < s$ the series $\sum a_\nu(k) x^\nu$ is generated by a finite automaton then so is the series

$$\sum \big(a_\nu(0) + a_\nu(1)p + \cdots + a_\nu(s-1)p^{s-1}\big)x^\nu ;$$

and, once again, conversely.

A basic property of finite automata is their cyclic nature: Since it has only finitely many internal states, after a sufficiently long input the automaton will be in the same state as it was after a shorter portion of that input. (Incidentally, it follows that a finite automaton can add numbers in base p, but that no finite automaton can square arbitrarily large numbers.)

1.6 We digress, both to detail the foregoing and to provide an (apparently) alternative description for finite automata. Consider, for example, the sequence (r_h)

$$0001001000\ 0111010001\ 0010111000\ 1000010010\ 0001110111\dots\dots$$

which, for the viewer's convenience and as a blow in the battle against the binary, we have split into groups of 10 symbols.

The 'pattern', which is to say a formation rule, becomes plain, by our viewing the symbols in pairs as binary integers. We then see that the resulting sequence is self reproducing under the uniform (or *regular*) 2-substitution θ on the symbols $\{0, 1, 2, 3\}$:

$$0 \mapsto 01\,, \quad 1 \mapsto 02\,, \quad 2 \mapsto 31\,, \quad 3 \mapsto 32\,.$$

Finally the output map o replacing 3 by 1, 2 by 1 and 1 by 0 yields the given sequence:

$$\begin{array}{l} 0001001000\ 0111010001\ 0010111000\ 1000010010\ 0001110111\dots\dots \\ \ 0\ 1\ 0\ 2\ 0\ \ 1\ 3\ 1\ 0\ 1\ \ 0\ 2\ 3\ 2\ 0\ \ 2\ 0\ 1\ 0\ 2\ \ 0\ 1\ 3\ 1\ 3 \\ 0102013101\ 0232020102\ 0131323101\ 3101020131\ 0102320232\dots\dots \\ 0001001000\ 0111010001\ 0010111000\ 1000010010\ 0001110111\dots\dots \end{array}$$

The 'what is going on here' is somewhat disguised by our notation. The intermediate symbols $\{0, 1, 2, 3\}$ represent the four states $\{s_0, s_1, s_2, s_3\}$ of a binary automaton ($p = 2$ in our description at **1.4**). The substitution θ provides the transition map τ as in the transition table:

	0	1
s_0	s_0	s_1
s_1	s_0	s_2
s_2	s_3	s_1
s_3	s_3	s_2

The output map o is as described above.

The sequence (r_h) is *generated* by the automaton in all of the following senses: It is (an image of) a sequence invariant under the substitution θ (that is, under the transition map). That invariant sequence is $\lim_{h\to\infty}\theta^h(0)$, where

$$\theta(0)=01\,,\ \theta^2(0)=\theta(01)=\theta(0)\theta(1)=0102\,,\ldots,\ \theta^h(0)=\theta(\theta^{h-1}(0)\theta^{h-1}(1)=\ \ldots$$

Finally, the automaton induces a map $h\mapsto r_h$ on the nonnegative integers in the following way (as is obviously equivalent to the foregoing): We have just one input tape containing the digits of h written in base 2; s_0 is the initial state. The automaton reads the digits of h successively (from left to right, disregarding irrelevant leading zeros since these leave the automaton in state s_0). The final state reached is transformed by the output map and yields r_h.

We will, in the sequel, use the sequence (r_h) as a convenient and interesting example of an *automatic sequence*. Cobham [13] shows that the interconnections displayed by the example are general.

At **1.5** we have already remarked that, more generally, a multiplicity of input tapes causes a finite automaton to induce a map $\nu\mapsto a_\nu$ on n-tuples of nonnegative integers.

1.7 The following "breaking up" procedure will be fundamental to several of our arguments: If $y(x)\in\mathbb{F}_p[[x]]$ and $S=\{0,1,\ldots,p-1\}^n$ then, for $\alpha\in S$, there are unique $y_\alpha(x)\in\mathbb{F}_p[[x]]$ such that $y(x)=\sum_{\alpha\in S}x^\alpha y_\alpha^p(x)$. In different words: $\{x^\alpha=x_1^{\alpha_1}\cdots x_n^{\alpha_n}:\alpha\in S\}$ is a basis for $\mathbb{F}_p[[x]]$ over $\big(\mathbb{F}_p[[x]]\big)^p$.

1.8 We choose to work over $\mathbb{F}_p$ for clarity and definiteness, but, with only minor modifications, the results cited remain valid if $\mathbb{F}_p$ is replaced by an arbitrary finite field $\mathbf{F}$ and $\mathbb{Z}_p$ by the complete discrete valuation ring with prime p and residue field $\mathbf{F}$.

2. Algebraic power series in characteristic p, diagonals and automata

2.1 If $y(x)\in\mathbb{F}_p[[x]]$ is algebraic then y satisfies an equation of the shape

$$\sum_{i=r}^{s}f_i(x)y^{p^i}=0\,,$$

with r, $s\in\mathbb{N}$, the $f_i\in\mathbb{F}_p[x]$ and $f_r\neq 0$. In fact, we may take $r=0$, for if $r>0$ then writing $f_i=\sum_\alpha x^\alpha f_{i\alpha}^p$ as in **1.7** we get that

$$\sum_\alpha x^\alpha\big(\sum_{i=r}^{s}f_{i\alpha}(x)y^{p^{i-1}}\big)^p=0\,.$$

Hence $\sum_{i=r-1}^{s-1}f_{i+1,\alpha}(x)y^{p^i}=0$ and some $f_{r\alpha}\neq 0$.

2.2 Thus if $y\in\mathbb{F}_p[[x]]$ is algebraic, y satisfies an equation of the shape

$$f(x)y=\sum_{i=1}^{s}f_i(x)y^{p^i}=L\big(y^p,\ldots,y^{p^s}\big)\,,$$

where L is linear with coefficients polynomials in x. After multiplying by f^{p-1}, breaking up y and the coefficients of L as in **1.7** and then taking p-th roots, we get the equations

$$f(x)y_{\alpha_1}=L_{\alpha_1}\big(y,y^p,\ldots,y^{p^{s-1}}\big)\,. \tag{2.2.1}$$

Now multiplying by f^{p-1} and substituting for $f(x)y$ on the right hand side of (2.2.1) yields

$$f^py_{\alpha_1}=L_{\alpha_1}(f^{p-2}L(y^p,\ldots,y^{p^s}),f^{p-1}y^p,\ldots,f^{p-1}y^{p^{s-1}})\,,$$

which is linear in $y^p, \ldots, y^{p^s}$. This brings us back, more or less, to the start and shows that iterating the process described leads to equations of the shape

$$f(x)y_{\alpha_1\ldots\alpha_e} = L_{\alpha_1\ldots\alpha_e}\big(y, y^p, \ldots, y^{p^{s-1}}\big)\,.$$

If, during this procedure, we keep track of the (multi-) degree in x of $L_{\alpha_1\ldots\alpha_e}$ we see that that degree remains bounded. Hence, since $\mathbb{F}_p$ is finite, there are only a finite number of distinct $y_{\alpha_1\ldots\alpha_e}$ and we have:

2.3 Theorem([9], [16]) *If $y = \sum a_\nu x^\nu$ is algebraic then*

(F) *there is an e such that for every $(\alpha_1, \ldots, \alpha_e) \in S^e$ there is an $e' < e$ and a $(\beta_1, \ldots, \beta_{e'}) \in S^{e'}$ such that*

$$y_{\alpha_1\ldots\alpha_e} = y_{\beta_1\ldots\beta_{e'}}\ ;$$

equivalently,

(A) *there is an e such that for all $j = (j_1, \ldots, j_n)$ with the $j_i < p^e$ there is an $e' < e$ and a $j' = (j'_1, \ldots, j'_n)$ with the $j'_i < p^{e'}$ such that*

$$a_{p^e\nu+j} = a_{p^{e'}\nu+j'} \text{ for all } \nu\,.$$

2.4 Conversely, if y satisfies (F) then taking $y_{\beta_1\ldots\beta_{e'}}$ and breaking it up $e-e'$ times, we see that the $y_{\alpha_1\ldots\alpha_e}$ satisfy a system of the form

$$y_{\alpha_1\ldots\alpha_e} = \sum x^\gamma y^{p^{e-e'}}_{\beta_1\ldots\beta_e}\ . \tag{2.4.1}$$

Viewing this system as a system of equations in the $y_{\delta_1\ldots\delta_e}$ we see (because, happily, the derivative of a p-th power vanishes) that its Jacobian is 1. By the Lemma below it follows that the $y_{\alpha_1\ldots\alpha_e}$, and hence y, must be algebraic.

2.5 Lemma([25], p.286) *If $\mathbb{K} \subset \mathbb{L}$ are fields and $y_1, \ldots, y_N \in \mathbb{L}$ satisfy a system of N polynomial equations $F_i(Y_1, \ldots, Y_N) = 0$ over $\mathbb{K}$ with $J(y_1, \ldots, y_N) \neq 0$, where $J = \det\big(\partial F_i/\partial Y_j\big)$ is the Jacobian of the system, then the y_i are all algebraic over $\mathbb{K}$.*

To see this, observe that if the conclusion were false there would be a nontrivial derivation D on $\mathbb{K}(y_1, \ldots, y_N)$ which is trivial on $\mathbb{K}$. But by applying D to the equations $F_i(y_1, \ldots, y_N) = 0$ we obtain

$$\big(\frac{\partial F_i}{\partial Y_j}(y_1, \ldots, y_N)\big)(Dy_i) = 0\,,$$

which forces $Dy_i = 0$ for all i.

2.6 If we interpret the equations (2.4.1), in the unknowns $y_{\delta_1\ldots\delta_e}$, over $\mathbb{Z}_p$ instead of over $\mathbb{F}_p$, they have Jacobian $\equiv 1 \bmod p$. Hence, by the Implicit Function Theorem for $\mathbb{Z}_p[[x]]$ (which is just Hensel's Lemma; an elementary proof is given at [16], p.50) these equations have a unique solution $\tilde{y}_{\alpha_1\ldots\alpha_e} \in \mathbb{Z}_p[[x]]$ with $\overline{\tilde{y}}_{\alpha_1\ldots\alpha_e} = y_{\alpha_1\ldots\alpha_e}$, where the $\bar{}$ denotes reduction mod p. By Lemma 2.5 the $\tilde{y}_{\alpha_1\ldots\alpha_e}$ are also algebraic. Thus we see that every algebraic power series in $\mathbb{F}_p[[x]]$ is the reduction of an algebraic power series in $\mathbb{Z}_p[[x]]$. Below, at **4.3** we mention a further, rather extraordinary, "lifting theorem" whereby every algebraic power series in $\mathbb{F}_p[[x]]$ is shown to lift to a series in $\mathbb{Z}[[x]]$ which is a solution of a system of functional equations.

2.7 Suppose that the series $\sum a_\nu x^\nu$ is generated by a finite automaton $\mathcal{M}$, as explained at **1.4-6**. Choose e so large that every state that $\mathcal{M}$ enters in the course of any computation is entered in a computation of length less than e. Then the a_ν satisfy version (A) of Theorem **2.3**.

Conversely, if the a_ν satisfy (A) one can construct a finite automaton $\mathcal{M}$ that generates $\sum a_\nu x^\nu$. $\mathcal{M}$ will be equipped with a table detailing the identifications (A) and an output list of the values of the a_j for $j = (j_1, \ldots, j_n)$ with the $j_i < p^e$. $\mathcal{M}$ computes as follows: It reads e digits from each tape. Then it uses the

table (A) to replace those e digits by e' digits. It reads a further $e-e'$ digits and iterates. At each stage it outputs the appropriate value from its output list. So we have:

2.8 Theorem([9], [16]) *$\sum a_\nu x^\nu \in \mathbb{F}_p[[x]]$ is algebraic if and only if it is generated by a finite automaton.*

This has the immediate
2.9 Corollary([19], [15], [16]) *Let f, $g \in \mathbb{F}_p[[x]]$ be algebraic. Then*
(i) *Every diagonal of f is algebraic, as is every off-diagonal.*
(ii) *The Hadamard product $f * g$ is algebraic.*
(iii) *Each characteristic series $f^{(i)} = \sum_{a_\nu = i} x^\nu$ is algebraic.*

2.10 We emphasise that the situation in characteristic 0 is very different. Neither diagonals nor (equivalently, see **1.3**) Hadamard products preserve algebraicity. For the latter, the standard example is

$$(1-4x_1)^{-1/2} = \sum \binom{2h}{h} x_1^h\,, \text{ but } \sum \binom{2h}{h}^2 x_1^h$$

is not algebraic. We note that the first remark is just the useful identity

$$\binom{2h}{h} = (-4)^h \binom{-\frac{1}{2}}{h}\,.$$

Congenial facts such as this are of interest in constructions useful to logicians; see, for example [23]. Then, with a little work and some first-year introductory calculus* the latter series is given by the integral

$$\frac{2}{\pi}\int_0^{\pi/2} \frac{dt}{\sqrt{(1-16x_1\sin^2 t)}}\,.$$

This is a complete elliptic integral well known not to represent an algebraic function.

3. Algebraic power series mod p^s

3.1 Theorem *If $f \in \mathbb{Z}_p[[x]]$ is algebraic and I is any diagonal then If is algebraic* mod p^s *for all $s \in \mathbb{N}$.*
Proof. It is enough to show the Theorem for $I = I_{12}$. As above, $\bar{}$ denotes reduction mod p. Then $\overline{f} \in \mathbb{F}_p[[x]]$ is algebraic and hence so is $I_{12}\overline{f}$. By **2.6** there is an algebraic $g \in \mathbb{Z}_p[[x_1, x_3, \ldots, x_n]]$ such that $\overline{g} = I_{12}\overline{f}$. Similarly, there are algebraic h_1, $h_2 \in \mathbb{Z}_p[[x]]$ such that $\overline{h}_1 = J_{12}\overline{f}$ and $\overline{h}_2 = J_{21}\overline{f}$, whilst $I_{12}h_1 = I_{21}h_2 = 0$. (Before lifting to characteristic 0 write $J_{12}\overline{f} = x_1 k(x_1, x_1x_2, x_3, \ldots, x_n)$ and lift $k(x_1, z, x_3, \ldots, x_n)$.) Set $\tilde{g} = g(x_1x_2, x_3, \ldots, x_n)$ and note that $I_{12}\tilde{g} = g$. Now, by induction on s, $I_{12}\frac{1}{p}(f - \tilde{g} - h_1 - h_2)$ is algebraic mod p^{s-1},

3.2 Corollary *If f, $g \in \mathbb{Z}_p[[x]]$ are algebraic then their Hadamard product $f * g$ is algebraic* mod p^s, *for all s.*

Next we generalise Theorem **2.3**:
3.3 Theorem([15], [16]) *If $f(x) = \sum a_\nu x^\nu \in \mathbb{Z}_p[[x]]$ is algebraic then for every s there is an e such that for every $j = (j_1, \ldots, j_n)$ with the $j_i < p^e$ there is an $e' < e$ and a $j' = (j_1', \ldots, j_n')$ with the $j_i' < p^{e'}$ such that*

$$a_{\nu p^e + j} \equiv a_{\nu p^{e'} + j'} \mod p^s \text{ for all } \nu$$

(and hence the a_ν mod p^s can be generated by a finite automaton).

* Nowadays, probably at best second-year, if one is lucky.

Proof. For $i \in \{1, \dots, p-1\}$ each characteristic series $\overline{f}^{(i)}$ is algebraic. Lift $\overline{f}^{(i)}$ to $f^{(i)} \in \mathbb{Z}_p[[x]]$ with $f^{(i)}$ algebraic and $\overline{f^{(i)}} = \overline{f}^{(i)}$. Then $g_i = *^{p^s}\overline{f}^{(i)}$, the p^s-fold Hadamard product of $\overline{f}^{(i)}$ is algebraic mod p^s. Of course each coefficient $a_{\nu i}$ of $g_i = \sum a_{\nu i}x^\nu$ is congruent to either 0 or 1 mod p^s. Hence $g = \frac{1}{p}(f - \sum i g_i)$ is algebraic mod p^{s-1} and the result follows by induction.

4. Automata, functional equations and arithmetic

4.1 We return to the example 2-automatic sequence (r_h) of **1.6** to notice the pattern:

$$0001001000\ 0111010001\ 0010111000\ 1000010010\ 0001110111\dots\dots$$
$$0001001000\ 0111010001\ 0010111000\ 1000010010\ 0001110111\dots\dots$$

The second row is (r_{2h}). Remarkably, it coincides with the original sequence, illustrating that $r_h = r_{2h}$. The third row is (r_{2h+1}). With careful attention, we see that $r_{4h} = r_{4h+1}$ but $r_{4h+2} \neq r_{4h+3}$. Setting $P(X) = \sum(-1)^{r_h}X^h$, these observations amount to the functional equation

$$P(X) = P(X^2) + XP(-X^2)\,. \tag{4.1.1}$$

Noticing that for an arbitrary sequence (i_h), with $i_h \in \{0, 1\}$, one has

$$2\sum i_h X^h = (1-X)^{-1} - \sum(-1)^{i_h}X^h\,,$$

we find that the function $R(X) = \sum r_h X^h$ satisfies

$$2X(1-X^4)R(X^4) + (1-X^4)(1-X)R(X^2) - (1-X^4)R(X) + X^3 = 0\,. \tag{4.1.2}$$

To discover this functional equation directly we recall the generated sequence

$$0102013101\ 0232020102\ 0131323101\ 3101020131\ 0102320232\dots\dots$$

and the defining substitutions

$$0 \mapsto 01\,,\quad 1 \mapsto 02\,,\quad 2 \mapsto 31\,,\quad 3 \mapsto 32\,.$$

4.2 It is quite as convenient to deal with the general case (as in [11]): Accordingly, let $a_0, a_1, \dots, a_{m-1}$ be a given alphabet and suppose we are given a substitution

$$a_0 \mapsto w_0,\quad a_1 \mapsto w_1,\quad \dots,\quad a_{m-1} \mapsto w_{m-1}$$

with words w_i, in the a_j, each of length $t \geq 2$. Denote by $s_0 s_1 s_2 \dots\dots$ a sequence fixed by the substitution and consider its generating function $\sum s_h X^h$. Denote the characteristic function of each symbol a_i by $f_i(X) = \sum_{s_h = a_i} X^h = \sum_{h\geq 0} u_{ih}X^h$, so that $\sum_{h\geq 0} s_h X^h = \sum_{i=0}^{m-1} a_i f_i(X)$. Note that $s_{th}s_{th+1}\dots s_{t(h+1)-1}$ depends only on s_h. Accordingly, set $v_{ijk} = 1$ if a_i is the $(k+1)$-st symbol of the word w_j (and 0 otherwise), so that $u_{i,th+k} = \sum_{j=0}^{m-1} v_{ijk}u_{jh}$. In other words,

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

If we denote by $A(X)$ the $m \times m$ matrix $A(X) = \big(p_{ij}(X)\big)$ and by $f(X)$ the column vector $f(X) = \big(f_0(X), f_1(X), \dots, f_{m-1}(X)\big)'$, then we have the matrix functional equation

$$f(X) = A(X)f(X^t)\,.$$

Moreover, it is plain that every linear combination of the $f_i(X)$ over the field of rational functions satisfies an equation of the shape $\sum_{i=0}^{m} c_i(X)g(X^{t^i}) = 0$, with polynomial coefficients $c_i(X)$.

Special cases of such functional equations were studied by Kurt Mahler in the late twenties; see [28]. It is therefore fitting to refer to these systems of equations as Mahler systems and to their solutions as Mahler functions.

4.3 Let $f(x_1) \in \mathbb{Q}[[x_1]]$ be algebraic. Because of Eisenstein's Theorem f has a reduction $\overline{f}$ mod p for almost all primes p, and that reduction is, of course, an algebraic element of $\mathbb{F}_p[[x_1]]$. By **2.6** such an algebraic power series lifts to an algebraic power series in $\mathbb{Z}_p[[x_1]]$ and its reductions mod p^s have coefficients generated by an automaton which reads its input in base p. Treating its coefficients, which are of the shape $a_h(0) + a_h(1)p + \cdots + a_h(s-1)p^{s-1}$, as elements of $\mathbb{Z}$ the new series satisfies a Mahler p-functional equation by the argument of **4.2** above. The theorem of Pólya-Carlson (the most convenient reference is [33], part VIII, Chap. 3, §5) tells us that the new series is either rational, in which case its sequence of coefficients is (eventually) periodic, or it represents a transcendental function with the unit circle as natural boundary. (In the light of the remark at **1.8** we also refer the reader to [3], Chap. 5.)

Thus the reductions of algebraic power series are either rational or, when viewed in characteristic zero, are transcendental functions satisfying Mahler functional equations.

Our example $R(X) = \sum r_h X^h$, when viewed as an element of $\mathbb{F}_2[[X]]$, satisfies the algebraic equation

$$(1+X)^5R^2 + (1+X)^4R + X^3 = 0\,.$$

When viewed as an element of $\mathbb{C}[[X]]$ the series is a transcendental function and the algebraic equation lifts to the Mahler 2-functional equation (4.1.2)

$$2X(1-X^4)R(X^4) + (1-X^4)(1-X)R(X^2) - (1-X^4)R(X) + X^3 = 0\,.$$

In much this spirit, there is a theorem of Cobham:

4.4 Theorem([12]) *If $f(x_1) = \sum_{i\in\mathcal{S}} x_1^i$ is generated by both an s- and a t-automaton and s and t are multiplicatively independent integers (equivalently, $\log s/\log t$ is irrational), then the set $\mathcal{S}$ is a finite union of arithmetic progressions (equivalently, the given function is rational).*

Recall that by "irrelevance of symbols" (*cf* **1.5**) there is no loss of generality in supposing given algebraic power series over finite fields to have just coefficients 0 and 1. We have that, if $\gcd(p,q) = 1$, a power series algebraic as an element of both $\mathbb{F}_p[[x]]$ and $\mathbb{F}_q[[x]]$, is rational. Incorporating the remarks of **4.3** we see that a power series is either rational, or is algebraic in at most one characteristic and transcendental in all others.

The argument in [12] seems easy locally, but is difficult globally. It remains of interest to find a more direct proof (*cf* [38]).

4.5 The definition of finite b-automaton entails that a sequence of integers generated by the automaton exhibits some 'digit pattern' in the base b representation of the integers. For example, the sequence (r_h) counts the number of occurrences (mod 2) of the pair 11 in the binary expansion of h:

0	1	10	11	100	101	110	111	1000	1001	1010	1011	1100	1101	1110	1111	10000	10001	10010	10011	...
0	0	0	1	0	0	1	0	0	0	0	1	1	1	0	1	0	0	0	1	...

In these terms, Cobham's Theorem at **4.4** states that a sequence of integers has 'digit pattern' in multiplicatively independent bases if and only if it is (eventually) periodic.

In characteristic zero it makes sense to speak of values taken by a series. In an attempt to decrease the awe of the transcendental we sketch the proof of the theorem:

4.6 Theorem([29]) *Let $t \geq 2$ be an integer. If $s_0s_1s_2\ldots\ldots$ is a t-automatic sequence and $f(X) = \sum s_hX^h$ is a transcendental Mahler t-function then, for integers $b \geq 2$, the numbers $f(b^{-1})$ are transcendental.*

To understand transcendence theory one need only know the fundamental lemma of the subject:

If $n \in \mathbb{Z}$ and $|n| < 1$ then $n = 0$.

Notice, firstly, that to show that a number α is transcendental amounts to demonstrating that no nontrivial expression $a_0 + a_1\alpha + \cdots + a_m\alpha^{m-1}$, with the a_i in $\mathbb{Z}$, can vanish. Thus one seeks to establish the linear independence of numbers $1, \alpha, \ldots, \alpha^{m-1}$ over $\mathbb{Z}$ for all m. Similarly, to prove that numbers $\alpha_1, \ldots, \alpha_n$ are algebraically independent one considers all monomials $\alpha^\mu = \alpha_1^{\mu_1} \cdots \alpha_n^{\mu_n}$, and proves the linear independence over $\mathbb{Z}$ of arbitrary finite collections of the α^μ.

We have a single transcendental function $f(z)$ satisfying a functional equation of the shape

$$c_n(z)f(z^{t^n}) + \cdots + c_1(z)f(z^t) + c_0(z)f(z) = c(z)\,,$$

with $c(z)$ and the $c_i(z) \in \mathbb{Z}[z]$ and wish to show that $\alpha = f(b^{-1})$ is transcendental. Set $f_i(z) = f(z^{t^i})$. We take a finite set of monomials $f^\mu = f_0^{\mu_0} \cdots f_n^{\mu_n}$, linearly independent and stable under the transformation $z \mapsto z^t$ over $\mathbb{Q}(z)$. For the present application the column vector g, say, of these monomials may be supposed to satisfy a matrix functional equation of the shape $g(z) = A(z)g(z^t)$, with an $m \times m$ matrix A of polynomials with integer coefficients.

We apply this as follows: Suppose (contrary to what we want to show) that there is, over $\mathbb{Z}$, a nontrivial linear relation

$$a_1g_1(b^{-1}) + \cdots + a_mg_m(b^{-1}) := a \cdot g(b^{-1}) = 0\,.$$

Then, by the iterated form $g(z) = A^{(k)}(z)g(z^{t^k})$ of the functional equation, we have for $k = 0, 1, \ldots\ldots$

$$0 = a \cdot A^{(k)}(b^{-1})g(b^{-t^k}) := a(k) \cdot g(b^{-t^k})\,.$$

This allows us to study the functions g_i at points conveniently close to the origin.

To prove the transcendence result, we construct a nonzero integer which, given the relation just mentioned, has absolute value less than 1. Our construction requires some notions from the theory of simultaneous polynomial approximation of functions that derive from work of Mahler ([30]).

In the sequel $\rho = (\rho_1, \ldots, \rho_m)$ denotes an m-tuple of integers with sum σ. By ord we denote the order of vanishing at $z = 0$. We say that the vector of polynomials $(p_1, \ldots, p_m)$ with $\deg p_j \le \rho_j - 1$, approximates the vector g at ρ if

$$\operatorname{ord}\bigl(r(z) = \sum p_j(z)g_j(z)\bigr) \ge \sigma - 1\,.$$

If no approximations at ρ have ord $r(z) > \sigma - 1$ then the vector g is said to be *normal* at ρ. We claim that if g is normal at ρ then there exists a (unique) $m \times m$ matrix $P_\rho = \bigl(p_{ij}\bigr)$ of polynomials with the following properties:

(i) The p_{ii} are monic and $\deg p_{ii} = \rho_i$; whilst, off the diagonal, $\deg p_{ij} \le \rho_i - 1$.
(ii) $\operatorname{ord}\bigl(r_i(z) = \sum p_{ij}(z)g_j(z)\bigr) \ge \sigma$.
(iii) $\det P_\rho = z^\sigma$.

This is not hard to see. The polynomials p_{ii} must have exact degree ρ_i, since otherwise we contradict normality at ρ. Plainly $\det P_\rho = z^\sigma +$ lower order terms, whilst, by multiplying by a g_i, and applying (ii), we see that $\operatorname{ord}(\det P_\rho) \ge \sigma$.

Thus $\det P_\rho$ vanishes only at $z = 0$. Set $B = b^{t^k}$. There is then no loss of generality in supposing that the vector $a(k)$ and the last $m - 1$ rows of $P_\rho(B^{-1})$ are linearly independent over $\mathbb{C}$. Now consider the $m \times m$ determinant Δ whose rows are these m vectors, noting that Δ is a nonzero rational number. There is some integer Q_ρ, depending only on ρ and not on k, so that all the $Q_\rho p_{ij}$ have integer coefficients and there is a constant $C > 0$, independent of ρ and k, so that $B^C a(k)$ is an integer vector. Set $\rho_{\min} = \min \rho_i$. Then, on expanding the determinant Δ by its first row, we see that

$$B^C Q_\rho^{m-1} B^{\sigma - \rho_{\min}} \Delta \tag{4.6.1}$$

is an integer.

Now multiply Δ by $g_1(B^{-1})$, say. After simple manipulation, $g_1(B^{-1})\Delta$ assumes the shape

$$\begin{vmatrix} 0 & a_2(k) & \ldots & a_m(k) \\ r_2(B^{-1}) & p_{22}(B^{-1}) & \ldots & p_{2m}(B^{-1}) \\ \vdots & \vdots & \ddots & \vdots \\ r_m(B^{-1}) & p_{m2}(B^{-1}) & \ldots & p_{mm}(B^{-1}) \end{vmatrix}\,.$$

Expanding by the first column shows that Δ is surprisingly small; only the factor $B^{-\sigma}$ depends jointly on the parameters k and ρ.

By Theorem 1 of [29] we may suppose that the vector g is normal at a sequence of parameter points ρ with $\rho_{\min} \to \infty$. Thence, choose ρ so that $\rho_{\min}$ is much larger than C and subsequently select k so that B dominates the remaining quantities (which depend only on ρ and not on k). It follows that the nonzero integer in (4.6.1) has absolute value less than 1. This contradiction proves the transcendence result.

4.7 We have shown that if $s_0 s_1 s_2 \ldots\ldots$ is a sequence generated by a finite automaton then a number $\sum s_i b^{-i} = s_0.s_1 s_2 \ldots\ldots$ (with the "decimal" point indicating expansion in base b) is either rational or transcendental. In different words: *The digits of an irrational algebraic number cannot be generated by a finite automaton.*

4.8 There is transcendence theory in characteristic $p > 0$. For example ([31]), consider, over $\mathbb{F}_p$, the formal power series

$$\sum_{h\geq 0} \binom{\lambda}{h} X^h = (1+X)^\lambda = (1+X)^{\sum_{i\geq 0} \lambda_i p^i} = \prod_{i\geq 0} (1+X^{p^i})^{\lambda_i} = \sum_{h\geq 0} \prod_{i\geq 0} \binom{\lambda_i}{h_i} X^h ;$$

with $\lambda = \sum \lambda_i p^i \in \mathbb{Z}_p$ and $h = \sum h_i p^i$ expanded in base p. It is natural to guess that $(1+X)^\lambda$ yields a formal power series algebraic over $\mathbb{F}_p(X)$ if and only if λ is rational. Indeed, if the series is algebraic then the sequence $\left(\binom{\lambda}{h}\right)$ is p-automatic, and that entails that the sequence $\left(\binom{\lambda}{p^h}\right)$ is periodic. But, as we saw above, $\binom{\lambda}{p^h} = \lambda_h$, so the sequence (λ_h) is periodic, which is to say that λ is rational as we had predicted. More generally, similar, though more sophisticated arguments [2] show that if $f \in \mathbb{F}_p[[X]]$, and is algebraic with constant coefficient 1, then the formal power series $f^{\lambda_1}, \ldots, f^{\lambda_s}$ are algebraically independent over $\mathbb{F}_p(X)$ if and only if the p-adic integers $1, \lambda_1, \ldots, \lambda_s$ are linearly independent over $\mathbb{Z}$.

The classical theory of diophantine approximation in fields of positive characteristic is described by Geijsel [20].

4.9 The growing literature on the subject of finite automata is readily reached by iteration of references. Our example automatic sequence (r_h), the Rudin-Shapiro sequence, is a hero of the story "FOLDS!"[14]. A more recent survey is that of Allouche [1].

5. Diagonals of rational functions

5.1 In this section we consider the class of power series in $\mathbb{Q}[[x]]$ which occur as the diagonals of rational functions $f(x) = P(x)/Q(x)$ with P, Q polynomials and $Q(0, \ldots, 0) \neq 0$.

Many combinatorial generating functions are diagonals of rational functions (*cf* [34]). Moreover, functions of number-theoretic interest arise in this way, as, for example, with Apéry's function ([37])

$$A(x_1) = \sum_{h=0}^{\infty} \sum_{k=0}^{\infty} \binom{h}{k}^2 \binom{h+k}{k}^2 x_1^h ,$$

which occurs in the proof of the irrationaliy of $\zeta(3)$, and which is the diagonal

$$A(x_1) = I\{(1-x_1)[(1-x_2)(1-x_3)(1-x_4)(1-x_5) - x_1 x_2 x_3]\}^{-1} .$$

At **1.3** we mentioned that every algebraic power series is the diagonal of a rational power series, whilst we saw at **3.1** that every such diagonal is algebraic mod p^s for all s and almost all p (since if f(x) is a rational power series over $\mathbb{Q}$ then $f(x) \in \mathbb{Z}_p[[x]]$ for almost all p).

The complete diagonals of rational power series have many other interesting properties:

5.2 Theorem(*cf* [8]) *If $f(x_1)$ is the diagonal of a rational power series over $\mathbb{Q}$ then*

(i) f has positive radius of convergence r_p at every place p of $\mathbb{Q}$ and $r_p = 1$ for almost all p.

(ii) for almost all p the function f is bounded on the disc $D_p(1^-) = \{t \in \mathbb{C}_p : |t| < 1\}$, where $\mathbb{C}_p$ is the completion of the algebraic closure of the p-adic rationals $\mathbb{Q}_p$ and, for almost all p, $\sup\{|f(t)| : t \in D_p(1^-)\} = 1$.

(iii) f satisfies a linear differential equation over $\mathbb{Q}[x_1]$; and

(iv) this equation is a Picard-Fuchs equation.

5.3 Whilst (i) and (ii) are reasonably straightforward, (iii) and (iv) are more difficult and are proved by Deligne using resolution of singularities (see [6], the footnote on p.5). Dwork [18] has given a proof which avoids resolution. An elementary proof of (iii) and its generalisations appears in [26].

Picard-Fuchs equations are equations that "come from geometry"—they are satisfied by those analytic functions (of the parameter) which are the periods of differential forms along cycles in pencils of curves. (A number of concrete examples are worked out in [36].) The singular points of these equations are regular with rational exponents.

One readily notices that $f(x_1) = \sum a_i x_1^i$ satisfying a linear differential equation with coefficients from $\mathbb{Q}[x_1]$ is equivalent to the Taylor coefficients a_i satisfying a linear recurrence relation

$$p_0(h)a_{h+k} = p_1(h)a_{h+k-1} + \cdots + p_k(h)a_h \ ,$$

where the p_j are polynomials; see [35], and [27] for generalisations to several variables. Algebraic functions over $\mathbb{Q}$, of course, satisfy differential equations of the sort just mentioned. The linear recurrence can be useful in computing their Taylor coefficients (*cf* [10]).

5.4 There is a great deal of folklore on the subject matter of Theorem **5.2**. There are conjectures of Bombieri and Dwork in [17] to the effect that if the solutions of a linear differential equation, over $\mathbb{Q}[x_1]$, have "large" radii of convergence in $\mathbb{C}_p$ for almost all p, then they are functions that "come from geometry". This has been verified for the Apéry function ([17]). Christol [8] conjectures that if $f(x_1)$ satisfies the first three properties cited in **5.2** then f is the diagonal of a rational function; a special case is proved in [8].

5.5 The results of §**3** constitute a vast range of congruences for the Taylor coefficients of diagonals of rational functions. Such, and other, congruences have been obtained for the coefficients of the Apéry function (for example [21], by elementary arguments, and [4], [36], [5]). It would be interesting to see if these congruences can be obtained by the methods of §**3**, which are, of course, effective.

5.6 We conclude by mentioning Grothendieck's Conjecture: *If a linear homogeneous differential equation with coefficients from $\mathbb{Q}[x_1]$, and of order n, has, for almost all p, n independent solutions in $\mathbb{F}_p[[x_1]]$ then all its solutions are algebraic.* This has been proved in a number of special cases (for example, for Picard-Fuchs equations) by Katz (see [24]). Some results have also been obtained by elementary methods ([22]).

References

[1] *Jean-Paul Allouche*, Automates finis en théorie des nombres, Expositiones Math., **5** (1987), 239–266

[2] *J.-P. Allouche, M. Mendès France and A. J. van der Poorten*, Indépendance algébrique de certaines séries formelles, Bull. Soc. Math. France **116** (1988)

[3] *Yvette Amice*, Les nombres p-adiques, Presses Universitaires de France (1975)

[4] *F. Beukers*, Some congruences for the Apéry numbers, J. Number Theory **21** (1985), 144–155

[5] *F. Beukers*, Another congruence for the Apéry numbers, J. Number Theory **25** (1987), 201–210

[6] *Gilles Christol*, Diagonales de fractions rationnelles et équations différentielles, Groupe d'étude d'analyse ultramétrique (Amice, Christol, Robba), Inst. Henri Poincaré, Paris, 1982/83 n°18

[7] *Gilles Christol*, Diagonales de fractions rationnelles et équations de Picard-Fuchs, Groupe d'étude d'analyse ultramétrique (Amice, Christol, Robba), Inst. Henri Poincaré, Paris, 1984/85 n°13

[8] *Gilles Christol*, Fonctions hypergéométriques bornées, Groupe d'étude d'analyse ultramétrique (Amice, Christol, Robba), Inst. Henri Poincaré, Paris, 1986/87 n°10

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

[10] *D. V. Chudnovsky and G. V. Chudnovsky*, On expansion of algebraic functions in power and Puiseux series, IBM Research Report **RC 11365**, IBM Research Centre, Yorktown Heights, New York,(1985)

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

[12] *Alan Cobham*, On the base dependence of sets of numbers recognisable by finite automata, Math. Systems Theory, **3** (1969), 186–192

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

[14] *Michel Dekking, Michel Mendès France and Alf van der Poorten*, FOLDS! The Mathematical intelligencer, **4** (1982), 130–138; II: Symmetry disturbed, *ibid.* 173–181; III: More morphisms, *ibid.* 190–195

[15] *P. Deligne*, Intégration sur un cycle évanescent, Invent. Math. **76** (1984), 129–143

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

[17] *Bernard Dwork*, On Apéry's differential operator, Groupe d'étude d'analyse ultramétrique (Amice, Christol, Robba), Inst. Henri Poincaré, Paris, 1979/81 n°25

[18] *B. Dwork*, Differential systems associated with families of singular hypersurfaces (preprint)

[19] *H. Furstenberg*, Algebraic functions over finite fields, J. Alg. **7** (1967), 271–277

[20] *J. M. Geijsel*, Transcendence in fields of positive characteristic, PhD Thesis, Amsterdam (1978) = Mathematical Centrum Tracts (Amsterdam) **91** (1979)

[21] *I. Gessel*, Some congruences for the Apéry numbers, J. Number Theory **14** (1982), 362–368

[22] *Taira Honda*, Algebraic differential equations, Symposia Mathematica **XXIV** Academic Press (1981), 169–204

[23] *J. P. Jones and Y. V. Matijasevič*, Exponential diophantine representation of recursively enumerable sets, Proceedings of the Herbrand Symposium (Logic Colloquium '81, European Meeting of the Association of Symbolic Logic, July 16–24, 1981, Marseille, France), Studies in Logic **107**, North-Holland Publishers, Amsterdam (1982), 159–177

[24] *N. Katz*, Algebraic solutions of differential equations, Invent. Math. **18** (1972), 1-118

[25] *Serge Lang*, Algebra, Addison-Wesley (1965)

[26] *L. Lipshitz*, The diagonal of a D-finite power series is D-finite, J. Alg. **113** (1988), 373–378

[27] *L. Lipshitz*, D-finite power series, J. Alg. (to appear)

[28] *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

[29] *J. H. Loxton and A. J. van der Poorten*, Arithmetic properties of automata: regular sequences, J. für Math. (to appear)

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

[31] *M. Mendès France and A. J. van der Poorten*, Automata and the arithmetic of formal power series, Acta Arith. **46** (1986), 211–214

[32] *Marvin L. Minsky*, Computation: finite and infinite machines, Prentice-Hall (1967)

[33] *G. Pólya and G. Szegő*, Problems and theorems in analysis, Springer-Verlag (translation of 4th edition, 1976)
[34] *Richard P. Stanley*, Generating functions, in *Gian-Carlo Rota*, ed., Studies in Combinatorics, MAA Studies in Mathematics, **17** (1978), 100–141
[35] *Richard P. Stanley*, Differentiably finite power series, European J. Comb. **1** (1980), 175–188
[36] *Jan Stienstra and Frits Beukers*, On the Picard-Fuchs equation and the formal Brauer group of certain elliptic $K3$-surfaces, Math. Ann. **271** (1985), 269–304
[37] *Alfred van der Poorten*, A proof that Euler missed... Apéry's proof of the irrationality of $\zeta(3)$; An informal report, The Mathematical Intelligencer, **1** (1979), 195–203
[38] *A. J. van der Poorten*, Remarks on automata, functional equations and transcendence, Sém. théorie des nombres de Bordeaux, année 1986-1987, exp. n°27, 11pp

Department of Mathematics, Purdue University, West Lafayette, IN 47907, USA
School of Mathematics, Physics, Computing and Electronics, Macquarie University, NSW 2109, Australia