Algebraic Automatic Continued Fractions in Characteristic 2

Yining Hu
HUST

One World Numeration Seminar
February 14, 2023

Automatic Sequences

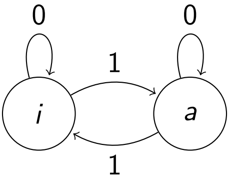
A sequence is said to be d-**automatic** if it can be generated by a d-DFAO (*deterministic finite automaton with output*). For an integer $d \geq 2$, a d-DFAO is defined to be a 6-tuple

$$M = (Q, \Sigma, \delta, q_0, \Delta, \tau)$$

where Q is the set of states with $q_0 \in Q$ being the initial state, $\Sigma = \{0, 1, \ldots, d-1\}$ the input alphabet, $\delta : Q \times \Sigma \to Q$ the transition function, Δ the output alphabet, and $\tau : Q \to \Delta$ the output function.

Automatic Sequences

The d-DFAO M generates a sequence $(c_n)_{n\geq 0}$ in the following way: for each non-negative integer n, the base-d expansion of n is read by M from right to left starting from the initial state q_0, and the automaton moves from state to state according to its transition funciton δ. When the end of the string is reached, the automaton halts in a state q, and the automaton outputs the symbol $c_n = \tau(q)$.



$$\tau(i) = 0, \ \tau(a) = 1.$$

The sequence $(t(n))_n$ generated by this automaton is the *Thue-Morse* sequence whose first terms are $0110100110010110\cdots$.

Automaticity, Algebricity and Transcendence

Theorem 1 (Christol, Kamae, Mendès France, Rauzy 1979)

The formal power series $\sum_{n=0}^{\infty} u_n x^n$ *in* $\mathbb{F}_p[x]$ *is algebraic over* $\mathbb{F}_p(x)$ *if and only if* $(u_n)_n$ *is p-automatic.*

Theorem 2 (Adamczewski, Bugeaud 2007)

An automatic real number is either rational or transcendental.

Theorem 3 (Bugeaud 2013)

The continued fraction expansion of an algebraic number of degree at least three cannot be generated by a finite automaton.

Continued Fractions for Power Series

Let K be a field. Let

$$f(z) = c_n z^n + c_{n-1} z^{n-1} + \cdots + c_0 + c_{-1} z^{-1} + \cdots \tag{1}$$

be an arbitrary element of $K((1/z))$. Define the integer part of $f(z)$ as

$$[f(z)] = c_n z^n + c_{n-1} z^{n-1} + \cdots + c_0. \tag{2}$$

Set $f_0 = f$, $a_0 = [f_0]$, $f_0 = a_0 + 1/f_1$, $a_1 = [f_1]$, $f_1 = a_1 + 1/f_2$, $a_2 = [f_1]$, ...
Then $a_0 \in K[z]$ and $a_j \in K[z] \backslash K$ for $j \geq 1$, and $f(z)$ admits the following continued fraction expansion

$$f(z) = a_0(z) + \cfrac{1}{a_1(z) + \cfrac{1}{a_2(z) + \cfrac{1}{a_3(z) + \cfrac{1}{\ddots}}}}. \tag{3}$$

Previous Work

Hu and Han studied continued fractions defined by the Thue-Morse sequence and the period-doubling sequence, and proved the algebraicity for some cases. Bugeaud and Han proved the algebraicity of Thue-Morse continued fractions. Hu and Lasjaunias proved the algebraicity of period-doubling continued fractions.

Recall that the Thue-Morse sequence can be defined by the substitution $0 \to 01$, $1 \to 10$, And the period-doubling sequence can be defined by the substitution $0 \to 01$, $1 \to 00$.

More sequences

We notice that the period-doubling sequence $\mathbf{p} = 01000101\ldots$ is the limit of the following sequence of finite words:

$$
\begin{aligned}
&0\\
&010\\
&0100010\\
&010001010100010\\
&\ldots
\end{aligned}
$$

More precisely, let W_0 be the empty word, let $\boldsymbol{\varepsilon} = (01)^\infty$, and $W_{n+1} = W_n, \varepsilon_n, W_n$ for $n \geq 0$. Then $\mathbf{p} = \lim W_n$. Therefore it is natural to consider sequences defined in the same way but with W_0 any finite word and ϵ any periodic sequence. We let $\mathcal{P}(W_0, \epsilon)$ denote the sequence defined this way and let $\mathcal{P}$ denote the family of such sequences.

We notice that the Thue-Morse sequence $\mathbf{t} = 01101001\ldots$ is the sum of the period-doubling sequence $\mathbf{p} = 10111010\ldots$. More precisely, if we define σ to be the operator that maps a binary sequence $(u_n)_n$ to $(\sum_{j=0}^{n-1} u_j \pmod 2)_n$, then $\mathbf{t} = \sigma(\mathbf{p})$. It is therefore natural to consider continued fractions defined by $\sigma^n(\mathbf{p})$ for $n \geq 2$.

In fact, we also want to consider sequences of the form $\sigma^n(\mathbf{u})$ where $n \geq 1$ and $\mathbf{u} \in \mathcal{P}$. It turns out more convenient to consider yet a larger family $\mathcal{G}$ defined as follows: Let $\boldsymbol{\Upsilon}$ be a periodic $(0,1)$-sequence. Let u_0 and v_0 be two finite words. For $n \geq 0$, define

$$u_{n+1} = u_n, u_n$$
$$v_{n+1} = v_n, v_n$$

if $\Upsilon_n = 0$, and

$$u_{n+1} = u_n, v_n$$
$$v_{n+1} = v_n, u_n$$

otherwise. We let $\mathcal{G}(u_0, v_0, \boldsymbol{\Upsilon})$ denote $\lim_n u_n$ and let $\mathcal{G}$ denote the family of such sequences.

Continued Fractions in Characteristic 2

Let $A = \{a_0, \ldots, a_k\}$ be a finite alphabet. We treat a_j as formal variables and define

$$\mathbb{F}_2[A] := \mathbb{F}_2[a_0, \ldots, a_k]$$
$$\mathbb{F}_2(A) := \mathbb{F}_2(a_0, \ldots, a_k)$$
$$\mathbb{F}_2((A)) := \mathbb{F}_2((\frac{1}{a_0}, \ldots, \frac{1}{a_k})).$$

Here $\mathbb{F}_2((\frac{1}{a_0}, \ldots, \frac{1}{a_k}))$ denotes the ring of power series of the form

$$\varphi = \sum_{n_0, \ldots, n_k \geq N} c_{n_0, n_1, \ldots, n_k} a_0^{-n_0} \cdots a_k{}^{-n_k}, \tag{4}$$

where N is an integer and $c_{n_0, n_1, \ldots, n_k} \in \mathbb{F}_2$.
Let $(u_n)_{n \geq 0}$ be a sequence taking values in A. It defines a formal power series $\sum_{n \geq 0} u_n z^n$ in $\mathbb{F}_2[A][[z]]$.
We define a norm on $\mathbb{F}_2((A))$ by assigning a series of the form (4) the number 2^{-m}, where $m = \min\{n_0 + n_1 \cdots + n_k \mid \; c_{n_0, n_1, \ldots, n_k} \neq 0\}$ (with the convention that $\min \emptyset = \infty$) is the *valuation* of the series φ.

Continued Fractions in Characteristic 2

The continued fraction $\mathrm{CF}(\mathbf{u}) = [u_0, u_1, \ldots]$ is defined as the limit of the sequence $([u_0, u_1, \ldots, u_n])_n$:

$$[u_0] = u_0,$$
$$[u_0, u_1, \ldots, u_n] = u_0 + \frac{1}{[u_1, \ldots, u_n]} \in \mathbb{F}_2((A)),$$

for $n \geq 1$. For example,

$$\begin{aligned}[u_0, u_1, u_2] &= u_0 + \frac{1}{u_1 + \frac{1}{u_2}} \\ &= u_0 + \frac{u_1^{-1}}{1 + (u_1 u_2)^{-1}} \\ &= u_0 + u_1^{-1} + u_1^{-2} u_2^{-1} + u_1^{-3} u_2^{-2} + \cdots\end{aligned}$$

Continued Fractions in Characteristic 2

Define

$$M_n = \begin{pmatrix} 1 & \frac{1}{u_n} \\ \frac{1}{u_n} & 0 \end{pmatrix} \begin{pmatrix} 1 & \frac{1}{u_{n-1}} \\ \frac{1}{u_{n-1}} & 0 \end{pmatrix} \cdots \begin{pmatrix} 1 & \frac{1}{u_0} \\ \frac{1}{u_0} & 0 \end{pmatrix} \tag{5}$$

then

$$[u_0, u_1, \ldots, u_n] = \frac{M_{n,0,1}}{M_{n,0,0}}. \tag{6}$$

In general, we do not have the convergence of $(M_{n,0,1})_n$ and $(M_{n,0,0})_n$, but we do have the convergence of $\left(\frac{M_{n,0,1}}{M_{n,0,0}}\right)_n$, which is proved in the same way as in the case of classical contineud fraction for real numbers.

Main result

Theorem 4 (H.)

If $\mathbf{u} = \mathcal{P}(W_0, \varepsilon)$ is a sequence in $\mathcal{P}$, where ε has period n, then $\mathrm{CF}(\mathbf{u})$ is algebraic over $\mathbb{F}_2(A)$ of degree at most 2^n.

Theorem 5 (H.)

If $\mathbf{u} = \mathcal{G}(u_0, v_0, \boldsymbol{\Upsilon})$ is a sequence in $\mathcal{G}$, where $\boldsymbol{\Upsilon}$ has period k, and contains an even number of 1's in one period, then $\mathrm{CF}(\mathbf{u})$ is algebraic over $\mathbb{F}_2(A)$ of degree at most 2^k.

Remark For a sequence $\mathbf{u} \in \mathcal{P} \cup \mathcal{G}$, $\sigma(\mathbf{u}) \in \mathcal{G}$.

Example

Let u_0 and v_0 be single letters a and b. Let $\boldsymbol{\Upsilon} = 1^\infty$. Then $\mathcal{G}(u_0, v_0, \boldsymbol{\Upsilon})$ is the Thue-Morse sequence $\mathbf{t} = abbabaab\ldots$ Define

$$m_0 = \begin{pmatrix} 1 & 1/a \\ 1/a & 0 \end{pmatrix}, \quad w_0 = \begin{pmatrix} 1 & 1/b \\ 1/b & 0 \end{pmatrix}$$

and for $k \geq 0$,

$$m_{k+1} = w_k m_k \quad w_{k+1} = m_k w_k.$$

By (5) and (6),

$$\mathrm{CF}(\mathbf{t}) = \lim_k \frac{m_{2k+1,0,1}}{m_{2k+1,0,0}}.$$

We will show that every entry of $(m_{2k+1})_k$ converges, and the limit is algebraic. The key to the proof is the relation

$$m_3 = (m_1 + d/co + d^2/r/co^2) \cdot I, \tag{$\star$}$$

where

$$d = \det(m_1),\ r = \mathrm{tr}(m_1),\ co = m_1 + w_1 + r,\ I = r \cdot co^2.$$

Example

In fact, the relation $(\star)$ holds regardless of the particular form of m_0 and w_0; it remains true if we had taken m_0 and w_0 to be any 2×2 matrix in characteristic 2. Therefore, it actually gives us the relation between m_{2k+1} and m_{2k+3} for all $k \geq 0$. For example, if we define $m_0' = m_2$, and $w_0' = w_2$, and define m_k', w_k', d', r', co', l' accordingly, then by $(\star)$,

$$m_3' = (m_1' + d'/co' + d'^2/r'/co'^2) \cdot l'.$$

We verify directly

$$d' = d^4,\ r' = r \cdot l,\ co' = co \cdot l,\ l' = l^4.$$

Therefore

$$m_5 = m_3' = (m_3 + d^4/co/l + d^8/r/co^2/l^3) \cdot l^4$$
$$= l^{1+4} \cdot (m_1 + (d + d^4/l^2)/co + (d^2 + d^8/l^4)/r/co^2)$$

Example

We continue in this way to find

$$m_7 = l^{1+4+16} \cdot (m_1 + (d + d^4/l^2 + d^4/l^{2+8})/co + (d^2 + d^8/l^4 + d^{32}/l^{4+16})/r/co^2)$$

etc. Define

$$f = l^{1+4+16+64+\cdots}$$

$$H = d + d^4/l^2 + d^{16}/l^{2+8} + d^{64}/l^{2+8+32} + \cdots .$$

It is easy to prove by induction that

$$\lim_k m_{2k+1} = f \cdot (m_1 + H/co + H^2/r/co^2).$$

Both f and H are algebraic:

$$f^4 = f/l$$

$$H^4/l^2 = H + d.$$

Therefore $\mathrm{CF}(\mathbf{t}) \in \mathbb{F}_2(a,b)[H]$, and is algebraic of degree at most 4.

Idea of the Proof

Let m_0, w_0 be two 2×2 matrices with entries in a field of characteristic 2. Let $m_1 = w_0 m_0$, $w_1 = m_0 w_0$. We define m_{1s} and w_{1s} for all binary word s inductively as follows:

$$
\begin{aligned}
m_{1s0} &= m_{1s}^2 \\
m_{1s1} &= w_{1s} m_{1s} \\
w_{1s0} &= w_{1s}^2 \\
w_{1s1} &= m_{1s} w_{1s}.
\end{aligned}
$$

Define $d = \det(m_1)$, $r = \mathrm{tr}(m_1)$ (the trace of m_1), and $co = m_1 + w_1 + r$. For all binary word s, define $t(s)$ as the sum of digits of s modulo 2.
Let s be an arbitrary non-empty binary word. For $j < |s|$, define $s(j)$ to be $s_0 \ldots s_{j-1}$, the prefix of length j of s. Define $e(s)$ inductively as follows: If s is the empty word, then $e(s) = 0$; define $e(s0) = 2e(s) + t(s0)$, $e(s1) = 2e(s) + t(s1)$. Define and $c_j(s) = d^{2^{j-1}} / r^{2^j - 1 - e(s(j))} / co^{e(s(j))}$.
We write c_j for short when there is no ambiguity.

Idea of the Proof

Proposition 6

Let s be a binary word of length $k \geq 1$. *If* $t(s) = 0$, *then*

$$m_{1s} = (m_1 + c_1 + \cdots + c_k) \cdot d^{2^{k-1}} / c_k, \tag{7}$$

$$w_{1s} = (w_1 + c_1 + \cdots + c_k) \cdot d^{2^{k-1}} / c_k; \tag{8}$$

if $t(s) = 1$, *then*

$$m_{1s} = (w_1 + c_1 + \cdots + c_k) \cdot d^{2^{k-1}} / c_k, \tag{9}$$

$$w_{1s} = (m_1 + c_1 + \cdots + c_k) \cdot d^{2^{k-1}} / c_k. \tag{10}$$

Thank You!