Survey of some recent results on the complexity of expansions of algebraic numbers

Michel Waldschmidt

Institut de Mathématiques de Jussieu + CIMPA
http://www.math.jussieu.fr/∼miw/

April 26, 2006

Diophantine Approximation and Heights
ESI — Erwin Schrödinger Institute, Wien
April 26, 2006

Émile Borel (1871–1956)

Émile Borel

- *Les probabilités dénombrables et leurs applications arithmétiques,*
  Palermo Rend. **27**, 247-271 (1909).
  Jahrbuch Database JFM 40.0283.01
  `http ://www.emis.de/MATH/JFM/JFM.html`
- *Sur les chiffres décimaux de* $\sqrt{2}$ *et divers problèmes de probabilités en chaînes,*
  C. R. Acad. Sci., Paris **230**, 591-593 (1950).
  Zbl 0035.08302

g-ary expansion of an algebraic number

Let $g \geq 2$ be an integer and x a real algebraic irrational number.

- *The* g*–ary expansion of* x *should satisfy some of the laws shared by almost all numbers (for Lebesgue's measure).*
- In particular *each digit* $0, 1, \ldots, g-1$ *should occur at least once.*
- As a consequence, *each given sequence of digits should occur infinitely often.*
- `Hint` : take a power of g.
- For instance, each of the four sequences $(0,0)$, $(0,1)$, $(1,0)$, $(1,1)$ should occur infinitely often in the binary expansion of x (take $g = 4$.)

Normal numbers

- A real number x is *normal in basis* g if its g–ary expansion has the following property :
  - each digit occurs with frequency $1/g$
  - each sequence of two digits occurs with frequency $1/g^2$
  - and so on

  A number is normal if it is normal in any basis $g \geq 2$. Borel suggested that each real irrational algebraic number should be normal.
- There is no explicitly known example of a triple (g, a, x), where $g \geq 3$ is an integer, a a digit in $\{0, \ldots, g-1\}$ and x an algebraic irrational number, for which one can claim that the digit a occurs infinitely often in the g–ary expansion of x.

Normal numbers

- Almost all numbers (for Lebesgue's measure) are normal.
- Example of a 2–normal number (Champernowne 1933, Bailey and Crandall 2001) : the *binary Champernowne number*, obtained by concatenation of the sequence of integers

  $0.1\,10\,11\,100\,101\,110\,111\,1000\,1001\,1010\,1011\,1100\,1101\,1110\ldots$

  `http://mathworld.wolfram.com/ChampernowneConstant.html`
- If a and g are coprime integers > 1, then

  $$\sum_{n \geq 0} a^{-n} g^{-a^n}$$

  is normal in basis g.

BBP numbers

- Hypothesis A of Bailey and Crandall (Experimental Math. 2001) : behaviour of orbits of the discrete dynamical system $T_g(x) = gx \pmod 1$.
- J-C. Lagarias (Experimental Math. 2001) : connection with special values of G functions
- D. Bailey, Jon Borwein, S. Plouffe (Math. Comp. 1997) : BBP numbers

$$\sum_{n\geq 1} \frac{p(n)}{q(n)} \cdot g^{-n}$$

where $g \geq 2$ is an integer, p and q relatively prime polynomials in $\mathbf{Z}[X]$ with $q(n) \neq 0$ for $n \geq 1$.

Expansion of algebraic numbers
Complexity of words
Words and transcendence
Continued fractions

BBP numbers : examples

- $\log 2$ is a BBP number in basis 2 since

$$\sum_{n\geq 1} \frac{1}{n} \cdot x^n = -\log(1-x) \quad \text{and} \quad \sum_{n\geq 1} \frac{1}{n} \cdot 2^{-n} = \log 2.$$

- $\log 2$ is a BBP number in basis $3^2 = 9$ since

$$\sum_{n\geq 1} \frac{1}{2n-1} \cdot x^{2n-1} = \log \frac{1+x}{1-x}, \quad \sum_{n\geq 1} \frac{6}{2n-1} \cdot 3^{-2n} = \log 2.$$

- π^2 is a BBP number in basis 2 and $3^4 = 81$ (D.J. Broadhurst 1999).

Hypothesis A of Bailey and Crandall

Hypothesis A of Bailey and Crandall : *Let*

$$\theta := \sum_{n\geq 1} \frac{p(n)}{q(n)} \cdot g^{-n}$$

where $g \geq 2$ *is a positive integer,* $R = p/q \in \mathbf{Q}(X)$ *a rational function with* $q(n) \neq 0$ *for* $n \geq 1$ *and* $\deg p < \deg q$. *Set* $y_0 = 0$ *and*

$$y_{n+1} = gy_n + \frac{p(n)}{q(n)} \pmod 1$$

Then the sequence $(y_n)_{n\geq 1}$ *either has finitely many limit points or is uniformly distributed modulo* 1.

Number of 1's in the binary expansion of an algebraic number

D. Bailey, J. Borwein, R. Crandall and C. Pomerance.
On the Binary Expansions of Algebraic Numbers,
Journal de Théorie des Nombres de Bordeaux, vol. **16** (2004), pp. 487-518. MR214495

If x *is a real algebraic number of degree* $d \geq 2$*, then the number of* 1*'s among the first* N *digits in the binary expansion of* x *is at least* $CN^{1/d}$*, where* C *is a positive number which depends only on* x.

Number of 1's in the binary expansion of an algebraic number

- For any integer $d \geq 2$, the number
$$\sum_{n\geq 0} 2^{-d^n}$$
is transcendental (result due to K. Mahler, 1929). *Fredholm number :* $\sum_{n\geq 0} 2^{-2^n}$. *A. J. Kempner (1916)*
- The number
$$\sum_{n\geq 0} 2^{-F_n},$$
having 1 at the Fibonacci numbers positions $1, 2, 3, 5, 8\dots$ is transcendental. *(also follows from Mahler's method).*

Mahler's method

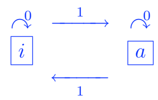
- Mahler (1930, 1969) : Let $d \geq 2$; the function $f(z) = \sum_{n\geq 0} z^{-d^n}$ satisfies $f(z^d) + z = f(z)$ for $|z| < 1$.
- Claim by J.H. Loxton and A.J. van der Poorten (1982–1988) using Mahler's method : *automatic irrational numbers are transcendental.*
- P.G. Becker (1994) : *for any given non–eventually periodic automatic sequence* $\mathbf{u} = (u_1, u_2, \dots)$, *the real number*
$$\sum_{k\geq 1} u_k g^{-k}$$
is transcendental, provided that the integer g *is sufficiently large (in terms of* $\mathbf{u}$*).*

Transcendence of automatic numbers

- **Theorem** (B. Adamczewski, Y. Bugeaud, F. Luca, 2004 – conjecture of A. Cobham, 1968) : *The sequence of digits of a real irrational algebraic number is not automatic.*
- In other terms if the sequence of g-ary digits of a real number x is given by a finite automaton, then x is transcendental.
- Tool : Schmidt's subspace Theorem.

Finite automata

- Automaton : States i, a, b... Transitions : 0 or 1.
- Example : the automaton



with $f(i) = 0$, $f(a) = 1$

- produces the sequence $a_0a_1a_2\dots$ where, for instance, a_9 is $f(i) = 0$ since $1001[i] = 100[a] = 10[a] = 1[a] = i$.
- This is the Thue-Morse sequence, where the $n+1$-th term a_n is 1 if the number of 1's in the binary expansion of n is odd, 0 if it is even.
  The Thue-Morse number is $\sum_{n\geq 0} a_n 2^{-n}$.

The Thue-Morse sequence 01101001100101101 . . .

- For $n \geq 0$ define $a_n = 0$ if the sum of the binary digits in the expansion of n is even, $a_n = 1$ if this sum is odd : the *Thue-Morse sequence* $(a_n)_{n\geq 0}$ starts with

$$0\,1\,1\,0\,1\,0\,0\,1\,1\,0\,0\,1\,0\,1\,1\,0\,1\ldots$$

- No sequence of three consecutive identical blocks :

$$\begin{array}{c} 0\;0\;0 \\ 1\;1\;1 \\ 01\;01\;01 \\ 10\;10\;10 \\ 001\;001\;001 \\ \ldots \end{array}$$

Powers of 2

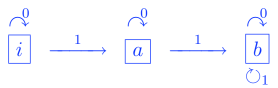
The binary number

$$\sum_{n\geq 0} 2^{-2^n} = 0.1101000100000001000\cdots = 0.a_1a_2a_3\ldots$$

with

$$a_n = \begin{cases} 1 & \text{if } n \text{ is a power of } 2, \\ 0 & \text{otherwise} \end{cases}$$

is produced by the automaton



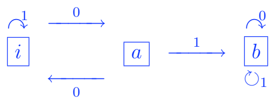
with $f(i) = 0$, $f(a) = 1$, $f(b) = 0$.

The Baum-Sweet sequence

- For $n \geq 0$ define $a_n = 1$ if the binary expansion of n contains no block of consecutive 0's of odd length, $a_n = 0$ otherwise : the Baum-Sweet sequence $(a_n)_{n \geq 0}$ starts with

$$1\,1\,0\,1\,1\,0\,0\,1\,0\,1\,0\,0\,1\,0\,0\,1\,1\,0\,0\,1\,0\ldots$$

- This sequence is produced by the automaton



with $f(i) = 1$, $f(a) = 0$, $f(b) = 0$.

Words

- We consider an alphabet A with g letters. The free monoid A^* on A is the set of *finite words* $a_1 \ldots a_n$ where $n \geq 0$ and $a_i \in A$ for $1 \leq i \leq n$. The law on A^* is called *concatenation*.
- The number of letters of a finite word is its *length* : the length of $a_1 \ldots a_n$ is n.
- The number of words of length n is g^n for $n \geq 0$. The single word of length 0 is the empty word e with no letter. It is the neutral element for the concatenation.

Infinite words

- We shall consider *infinite words* $w = a_1 \dots a_n \dots$
  A *factor of length* m of such a w is a word of the form $a_k a_{k+1} \dots a_{k+m-1}$ for some $k \geq 1$.
- The *complexity* of an infinite word w is the function $p(m)$ which counts, for each $m \geq 1$, the number of distinct factors of w of length m.
- Hence for an alphabet A with g elements we have $1 \leq p(m) \leq g^m$ and the function $m \mapsto p(m)$ is non–decreasing.
- According to Borel's suggestion, the complexity of the sequence of digits in basis g of an irrational algebraic number should be $p(m) = g^m$.

Automatic sequences

- Let $g \geq 2$ be an integer. An infinite sequence $(a_n)_{n\geq 0}$ is said to be g–automatic if a_n is a finite-state function of the base–g representation of n : this means that there exists a finite automaton starting with the g–ary expansion of n as input and producing the term a_n as output.
- A. Cobham, 1972 : *Automatic sequences have a complexity* $p(m) = O(m)$.

Morphisms

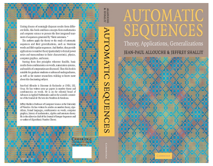
- Let A and B be two finite sets. A map from A to B^* can be uniquely extended to a homomorphism between the free monoids A^* and B^*. We call *morphism from A to B* such a homomorphism.
- A morphism ϕ from A into itself is said to be *prolongable* if there exists a letter a such that $\phi(a) = au$, where u is a non–empty word such that $\phi^k(u) \neq e$ for every $k \geq 0$. In that case, the sequence of finite words $(\phi^k(a))_{k\geq 1}$ converges in $A^{\mathbf{N}}$ (endowed with the product topology of the discrete topology on each copy of A) to an infinite word $w = au\phi(u)\phi^2(u)\phi^3(u)\dots$. This infinite word is clearly a fixed point for ϕ and we say that w *is generated by the morphism* ϕ.

Recurrent morphisms, binary morphisms, morphic sequences

- If, moreover, every letter occurring in w occurs at least twice, then we say that w is generated by a *recurrent morphism.*
- If the alphabet A has two letters, then we say that w is generated by a *binary morphism.*
- More generally, an infinite sequence w in $A^{\mathbf{N}}$ is said to be *morphic* if there exist a sequence u generated by a morphism defined over an alphabet B and a morphism ϕ from B to A such that $w = \phi(u)$.

Automatic sequences and morphic sequences

- **Theorem (A. Cobham)** : *automatic sequences are the same as uniform morphic sequences.*
- Jean-Paul Allouche and Jeffrey Shallit
  Automatic Sequences : Theory, Applications, Generalizations,
  Cambridge University Press, 2003.



`http ://www.cs.uwaterloo.ca/∼shallit/asas.html`

Example 1 : the Fibonacci word

Take $A=\{a,b\}$.

- Start with $f_1=b$, $f_2=a$ and define (concatenation) : $f_n=f_{n-1}f_{n-2}$.
- Hence $f_3=ab$ $\quad f_4=aba$ $\quad f_5=abaab$
  $f_6=abaababa$ $\quad f_7=abaababaabaab$
  $f_8=abaababaabaababaababa$
- The *Fibonacci word*

  $$w=abaababaabaababaababaabaababaabaab\dots$$

  is generated by a binary recurrent morphism : it is the fixed point of the morphism $a\mapsto ab$, $b\mapsto a$;
  under this morphism, the image of f_n is f_{n+1}.

Example 2 : the Thue-Morse word *abbabaabbaababbab* ...

- In the Thue-Morse sequence $0110100110010110 1\dots$ replace 0 by a and 1 by b.
  The *Thue-Morse word*
  $$w = abbabaabbaababbab\dots$$
  is generated by a binary recurrent morphism : it is the fixed point of the morphism $a \mapsto ab$, $b \mapsto ba$.

The Thue-Morse-Mahler number

- The *Thue-Morse-Mahler number in basis* $g \geq 2$ is the number
  $$\xi_g = \sum_{n \geq 0} \frac{a_n}{g^n}$$
  where $(a_n)_{n \geq 0}$ is the Thue-Morse sequence. The g–ary expansion of ξ_g starts with
  $$0.1101001100101101\dots$$
- These numbers were considered by K. Mahler who proved in 1929 that they are transcendental.

Example 3 : the Rudin–Shapiro sequence

- The Rudin–Shapiro word $aaabaabaaaabbbab\ldots$. For $n \geq 0$ define $r_n \in \{a,\ b\}$ as being equal to a (respectively b) if the number of occurrences of the pattern 11 in the binary representation of n is even (respectively odd).
- Let σ be the morphism defined from the monoid B^* on the alphabet $B = \{1, 2, 3, 4\}$ into B^* by : $\sigma(1) = 12$, $\sigma(2) = 13$, $\sigma(3) = 42$ and $\sigma(4) = 43$. Let

$$\mathbf{u} = 121312421213\ldots$$

be the fixed point of σ begining with 1 and let φ be the morphism defined from B^* to $\{a, b\}^*$ by : $\varphi(1) = aa$, $\varphi(2) = ab$ and $\varphi(3) = ba$, $\varphi(4) = bb$. Then the Rudin-Shapiro word is $\varphi(\mathbf{u})$, hence it is morphic.

Example 4 : powers of 2

The binary automatic number

$$\sum_{n\geq 0} 2^{-2^n} = 0.1101000100000001000\cdots$$

yields the word

$$\mathbf{v} = v_1 v_2 \cdots v_n \cdots = bbabaaabaaaaaaabaaa\cdots$$

where

$$v_n = \begin{cases} b & \text{if } n \text{ is a power of } 2, \\ a & \text{otherwise.} \end{cases}$$

The complexity $p(m)$ of $\mathbf{v}$ is bounded by $2m$:

$m =$	1	2	3	4	5	6	$\cdots$
$p(m) =$	2	4	6	7	9	11	$\cdots$

Sturmian words

Assume $g = 2$, say $A = \{a, b\}$.

- A word is periodic if and only if its complexity is bounded.
- If the complexity $p(m)$ a word w satisfies $p(m) = p(m+1)$ for one value of m, then $p(m+k) = p(m)$ for all $k \geq 0$, hence the word is periodic. It follows that *a non–periodic word* w *has a complexity* $p(m) \geq m + 1$.
- An infinite word of minimal complexity $p(m) = m + 1$ is called *Sturmian* (Morse and Hedlund, 1938).
- Two dimensional billiards produce Sturmian words.

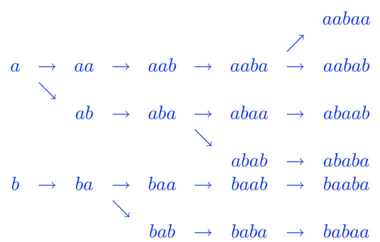
The Fibonacci word is Sturmian

- The Fibonacci word

  $$w = abaababaabaababaababaabaababaabaab\dots$$

  is Sturmian.
- On the alphabet $\{a,\ b\}$, a Sturmian word w is characterized by the property that *for each* $m \geq 1$, *there is exactly one factor* v *of* w *of length* m *such that both* va *and* vb *are factors of* w *of length* $m + 1$.

The Fibonacci word $abaababaabaababaababaabaababaabaab\ldots$ is Sturmian



Transcendence and Sturmian words

- S. Ferenczi, C. Mauduit, 1997 : A number whose sequence of digits is Sturmian is transcendental.
  Combinatorial criterion : *the complexity of the g–ary expansion of every irrational algebraic number satisfies*

$$\liminf_{m\to\infty}(p(m)-m)=+\infty.$$

- *Tool :* a p–adic version of the Thue–Siegel–Roth Theorem due to Ridout (1957).

Further transcendence results on g-ary expansions of real numbers

- J-P. Allouche and L.Q. Zamboni(1998).
- R.N. Risley and L.Q. Zamboni(2000).
- B. Adamczewski and J. Cassaigne (2003).

Complexity of the g-ary expansion of an algebraic number

- **Theorem** (B. Adamczewski, Y. Bugeaud, F. Luca 2004). *The binary complexity p of a real irrational algebraic number x satisfies*

$$\liminf_{m\to\infty} \frac{p(m)}{m} = +\infty.$$

- Corollary (conjecture of A. Cobham (1968)) : *If the sequence of digits of an irrational real number x is automatic, then x is transcendental.*

Irrationality measures for automatic numbers

- Further progress by B. Adamczewski and J. Cassaigne (2006) – solution to a Conjecture of J. Shallit (1999) : *A Liouville number cannot be generated by a finite automaton.*
- The irrationality measure of the automatic number associated with $\sigma(0) = 0^n1$ and $\sigma(1) = 1^n0$ is at least n.
- For the Thue-Morse-Mahler numbers for instance the exponent of irrationality is ≤ 5.

Christol, Kamae, Mendes-France, Rauzy

The result of B. Adamczewski, Y. Bugeaud and F. Luca implies the following statement related to the work of G. Christol, T. Kamae, M. Mendès-France and G. Rauzy (1980) :

Corollary. *Let $g \geq 2$ be an integer, p be a prime number and $(u_k)_{k\geq 1}$ a sequence of integers in the range $\{0, \ldots, p-1\}$. The formal power series*

$$\sum_{k\geq 1} u_k X^k$$

and the real number

$$\sum_{k\geq 1} u_k g^{-k}$$

are both algebraic (over $\mathbf{F}_p(X)$ and over $\mathbf{Q}$, respectively) if and only if they are rational.

Schmidt's subspace Theorem

For $\mathbf{x} = (x_0, \ldots, x_{m-1}) \in \mathbf{Z}^m$, define $|\mathbf{x}| = \max\{|x_0|, \ldots, |x_{m-1}|\}$.

W.M. Schmidt (1970) : *Let $m \geq 2$ be a positive integer, S a finite set of places of $\mathbf{Q}$ containing the infinite place. For each $v \in S$ let $L_{0,v}, \ldots, L_{m-1,v}$ be m independent linear forms in m variables with algebraic coefficients in the completion of $\mathbf{Q}$ at v. Let $\epsilon > 0$. Then the set of $\mathbf{x} = (x_0, \ldots, x_{m-1}) \in \mathbf{Z}^m$ such that*

$$\prod_{v \in S} |L_{0,v}(\mathbf{x}) \cdots L_{m-1,v}(\mathbf{x})|_v \leq |\mathbf{x}|^{-\epsilon}$$

is contained in the union of finitely many proper subspaces of $\mathbf{Q}^m$.

Ridout's Theorem

- Ridout's Theorem : *for any real algebraic number α, for any $\epsilon > 0$, the set of $p/q \in \mathbf{Q}$ with $q = 2^k$ and $|\alpha - p/q| < q^{-1-\epsilon}$ is finite.*
- In Schmidt's Theorem take $m = 2$, $S = \{\infty,\ 2\}$,

  $L_{0,\infty}(x_0, x_1) = L_{0,2}(x_0, x_1) = x_0$,
  $L_{1,\infty}(x_0, x_1) = \alpha x_0 - x_1, \qquad L_{1,2}(x_0, x_1) = x_1$.

  For $(x_0, x_1) = (q, p)$ with $q = 2^k$, we have

  $|L_{0,\infty}(x_0, x_1)|_\infty = q, \qquad |L_{1,\infty}(x_0, x_1)|_\infty = |q\alpha - p|$,
  $|L_{0,2}(x_0, x_1)|_2 = q^{-1}, \qquad |L_{1,2}(x_0, x_1)|_2 = |p|_2 \leq 1$.

Further transcendence results

Consequences of Nesterenko 1996 result on the transcendence of values of theta series at rational points.

- The number $\sum_{n\geq 0} 2^{-n^2}$ is transcendental (D. Bertrand 1997 ; D. Duverney, K. Nishioka, K. Nishioka and I. Shiokawa 1998).
- For the word

$$\mathbf{u} = 012122122212222122222122222212222\ldots$$

generated by the non–recurrent morphism $0 \mapsto 012$, $1 \mapsto 12$, $2 \mapsto 2$, the number $\eta = \sum_{k\geq 1} u_k 3^{-k}$ is transcendental.

Complexity of the continued fraction expansion of an algebraic number

- Similar questions arise by considering the continued fraction expansion of a real number instead of its g–ary expansion.
- Open question – A.Ya. Khintchine (1949) : *are the partial quotients of the continued fraction expansion of a non–quadratic irrational algebraic real number bounded ?*
- No known example so far !

Transcendence of continued fractions

- J. Liouville, 1844
- É. Maillet, 1906, O. Perron, 1929
- H. Davenport and K.F. Roth, 1955
- A. Baker, 1962
- J.L. Davison, 1989

Transcendence of continued fractions (continued)

- J.H. Evertse, 1996.
- M. Queffélec, 1998 : *transcendence of the Thue–Morse continued fraction.*
- P. Liardet and P. Stambul, 2000.
- J-P. Allouche, J.L. Davison, M Queffélec and L.Q. Zamboni, 2001 : *transcendence of Sturmian or morphic continued fractions.*
- C. Baxa, 2004.
- B. Adamczewski, Y. Bugeaud, J.L. Davison, 2005 : *transcendence of the Rudin-Shapiro and of the Baum–Sweet continued fractions.*

Transcendence of continued fractions

- Open question : *Do there exist algebraic numbers of degree at least three whose continued fraction expansion is generated by a morphism ?*
- B. Adamczewski, Y. Bugeaud (2004) : *The continued fraction expansion of an algebraic number of degree at least three cannot be generated by a binary morphism.*

Further open problems

- Provide an explicit example of an automatic real number $x > 0$ such that $1/x$ is not automatic.
- Show that
$$\log 2 = \sum_{n\geq 1} \frac{1}{n} 2^{-n}$$
is not 2-automatic.
- Show that
$$\pi = \sum_{n\geq 0} \left(\frac{4}{8n+1} - \frac{2}{8n+4} - \frac{1}{8n+5} - \frac{1}{8n+6} \right) 2^{-4n}$$
is not 2-automatic.

Further open problems

Let $(e_n)_{n\geq 1}$ be an infinite sequence over $\{0, 1\}$ that is not ultimately periodic. Prove or disprove : *at least one of the two numbers*

$$\sum_{n\geq 1} e_n 2^{-n}, \qquad \sum_{n\geq 1} e_n 3^{-n}$$

is transcendental.

Diophantine Approximation and Heights
ESI — Erwin Schrödinger Institute, Wien
April 26, 2006