International Conference on "Number Theory and Related Topics" Hanoi, December 12-15, 2006

Discrete mathematics and Diophantine Problems

Michel Waldschmidt
Institut de Mathématiques de Jussieu & CIMPA

1. Borel's Conjectures, Complexity of Words
2. Transcendence, Diophantine Approximation
3. Continued Fractions

Borel's Conjectures, Complexity of Words
Transcendence, Diophantine Approximation
Continued Fractions

Abstract

One of the first goals of Diophantine Analysis is to decide whether a given number is rational, algebraic or else transcendental. Such a number may be given by its binary or decimal expansion, by its continued fraction expansion, or by other limit process (sum of a series, infinite product, integrals ...). Language theory provides sometimes convenient tools for the study of numbers given by expansions. We survey some of the main recent results on Diophantine problems related with the complexity of words.

Michel Waldschmidt http://www.math.jussieu.fr/~miw/

Happy Birthday to Professor Ha Huy Khoái

Yu. MANIN - *Cyclotomic fields and modular curves*,
Usp. Mat. Nauk **26** N°6 (1971), 7–71.
Engl. transl. : Russian Math. Surveys **26** (1971), 7–78.

First decimals of $\sqrt{2}$ http://wims.unice.fr/wims/wims.cgi

1.41421356237309504880168872420969807856967187537694807317667973
79907324784621070388503875343276415727350138462309122970249248 3
60558507372126441214970999358314132226659275055927557999505011 5
27820605714701095599716059702745345968620147285174186408891986 0
95523292304843087143214508397626036279952514079896872533965463 3
18088296406206152583523950547457502877599617298355752203375318 5
70113543746034084988471603868999706990048150305440277903164542 4
78230684929369186215805784631115966687130130156185689872372352 8
85092648612494977154218334204285686060146824720771435854874155 6
57069677653720226485447015858801620758474922657226020855844665
21458398893944370926591800311388246468157082630100594858704003 1
86480342194897278290641045072636881313739855256117322040245091 2
27700226941127573627280495738108967504018369868368450725799364 7
29060762996941380475654823728997180326802474420629269124859052 1
81004459842150591120249441341728531478105803603371077309182869 3
14710171111683916581726889419758716582152128229518488 47 . . .

First binary digits of $\sqrt{2}$ http://wims.unice.fr/wims/wims.cgi

1.0110101000001001111001100110011111110011101111001100100100000
1000101100101111101100010011011001101110101001010101111110100
1111100011101011011101100000101110101000100100111011101010000
1001100111011010001011110101100100001011000001100110011001100
1000101010100101011111100100000110000010000111010101110001010 0
0101100001110101000101100011111111001101111110111001000001111 0
1101100111001000011101110100101010000101111001000011100111000
1111011010010100111100000000010010000111001101100011111011111101
0001001110110100011010010001000000010111010000111010000101010 1
1110001111101001110010100110000010110011100011000000001000 1101
1110000110011011110111100101010110001101111001001000100010 1101
0001000010001011000101001000110000010101011110001110010001 0111
1011111000100111000110011110001101101010110101000101000111000 1
0111011011111101001110111001100101100101010011000110100001 1001
1000111110011110010000100110111111010100101111000100100000 1111
0000011011011100101100000101110111010101010010010100000 1000100
1100100000100000011001010010010101000000100111001010010 10 ...

Le fabuleux destin de $\sqrt{2}$

- *The fabulous destiny of* $\sqrt{2}$
Benoît Rittaud, Éditions *Le Pommier*, 2006.
http://www.math.univ-paris13.fr/~rittaud/RacineDeDeux

- Computation of decimals of $\sqrt{2}$:
1 542 computed by hand by Horace Uhler in 1951
14 000 decimals computed in 1967
1 000 000 decimals in 1971
$137 \cdot 10^9$ decimals computed by Yasumasa Kanada and Daisuke Takahashi in 1997 with Hitachi SR2201 in 7 hours and 31 minutes.
- Motivation : computation of π.

Complexity of the g–ary expansion of an irrational algebraic real number

Let $g \geq 2$ be an integer.

- É. Borel (1909 and 1950) : *the g–ary expansion of an algebraic irrational number should satisfy some of the laws shared by almost all numbers (with respect to Lebesgue's measure).*
- In particular *each digit should occur*, hence *each given sequence of digits should occur infinitely often.*
- 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.

Conjecture 1 (Émile Borel)

- Rendiconti del Circolo matematico di Palermo, **27** (1909), 24–271.
  Comptes Rendus de l'Académie des Sciences de Paris **230** (1950), 591–593.
- **Conjecture 1.** *Let x be an irrational algebraic real number, $g \geq 3$ a positive integer and a an integer in the range $0 \leq a \leq g-1$. Then the digit a occurs at least once in the g–ary expansion of x.*
- If a real number x satisfies Conjecture 1 for all g and a, then it follows that for any g, each given sequence of digits occurs infinitely often in the g–ary expansion of x.
- This is easy to see by considering powers of g.

Borel's Conjecture 1

- For instance, Conjecture 1 with $g = 4$ implies that each of the four sequences $(0,0)$, $(0,1)$, $(1,0)$, $(1,1)$ should occur infinitely often in the binary expansion of each irrational algebraic real number x.
- K. Mahler : *For any $g \geq 2$ and any $n \geq 1$, there exist algebraic irrational numbers x such that any block of n digits occurs infinitely often in the g–ary expansion of x.*

Normal expansions

- A real number x is called *simply normal in base g* if each digit occurs with frequency $1/g$ in its g–ary expansion.
- A real number x is called *normal in base g* or *g–normal* if it is simply normal in base g^m for all $m \geq 1$.
- Hence a real number x is normal in base g if and only if, for any $m \geq 1$, each sequence of m digits occurs with frequency $1/g^m$ in its g–ary expansion.
- A real number is called *normal* if it is normal in any base $g \geq 2$.
- Hence a real number is normal if and only if it is simply normal in any base $g \geq 2$.

Borel's Conjecture 2

- **Conjecture 2.** *Let x be an irrational algebraic real number. Then x is normal.*
- Almost all real numbers (for Lebesgue's measure) are normal.
- Examples of computable normal numbers have been constructed (W. Sierpinski, H. Lebesgue, V. Becher and S. Figueira) but the known algorithms to compute such examples are fairly complicated ("ridiculously exponential", according to S. Figueira).

Example of normal numbers

An example of a 2–normal number (Champernowne 1933, Bailey and Crandall 2001) is the *binary Champernowne number*, obtained by the concatenation of the sequence of integers

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

$$= \sum_{k\geq 1} k2^{-c_k} \quad \text{with} \quad c_k = k + \sum_{j=1}^{k} [\log_2 j].$$

Further examples of normal numbers

- (Korobov, Stoneham ...) : *if a and g are coprime integers > 1, then*

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

is normal in base g.

- A.H. Copeland and P. Erdős (1946) : a normal number in base 10 is obtained by concatenation of the sequence of prime numbers

$$0.2\,3\,5\,7\,11\,13\,17\,19\,23\,29\,31\,37\,41\,43\,47\,53\,59\,61\,67\ \ldots$$

Infinite words

Let $\mathcal{A}$ be a finite alphabet with g elements.

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

Sturmian words

Assume $g = 2$, say $\mathcal{A} = \{0, 1\}$.

- 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 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).
- Examples of Sturmian words are given by 2–dimensional billiards.

Sturmian words

- Define $f_1 = 1$, $f_2 = 0$ and, for $n \geq 3$ (concatenation) : $f_n = f_{n-1} f_{n-2}$.
  The *Fibonacci word*

$$w = 0100101001001010010100100101001001\ldots$$

  is Sturmian.
- On the alphabet $\{0, 1\}$, 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 $v0$ and $v1$ are factors of w of length $m+1$*.

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-Schmidt Theorem due to Ridout (1957).
Recall the lectures by Min Ru and Paul Vojta.
• **Reference :** Yuri Bilu's Lecture in the Bourbaki Seminar, November 2006 :
The many faces of the Subspace Theorem [after Adamczewski, Bugeaud, Corvaja, Zannier...]
http://www.math.u-bordeaux.fr/~yuri/publ/subspace.pdf

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.*

Automata

A *finite automaton* consists of
• the *input alphabet* $\mathcal{A}$, usually the set of digits $\{1, 2, \ldots, g-1\}$;
• the set $\mathcal{Q}$ of states, a finite set of 2 or more elements, with one element called the *initial state* i singled out;
• The *transition map* $\mathcal{Q} \times \mathcal{A} \to \mathcal{Q}$, which associates to every state a new state depending on the current input;
• the *output alphabet* $\mathcal{B}$, together with the *output map* $f : \mathcal{Q} \to \mathcal{B}$.

Example : powers of 2

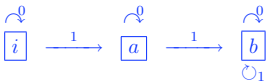
The sequence of binary digits of the number

$$\sum_{n\geq 0} 2^{-2^n} = 0.11010001000000001000\cdots = 0.a_1a_2a_3\ \ldots$$

with

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

is automatic : $\mathcal{A} = \mathcal{B} = \{0, 1\}$, $\mathcal{Q} = \{i, a, b\}$, $f(i) = 0$, $f(a) = 1$, $f(b) = 0$,



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 $p(m) = O(m)$.*

Powers of 2 *(continued)*

The complexity $p(m)$ of the automatic sequence of binary digits of the number

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

is at most $2m$:

$$\begin{array}{rccccccc} m = & 1 & 2 & 3 & 4 & 5 & 6 & \cdots \\ p(m) = & 2 & 4 & 6 & 7 & 9 & 11 & \cdots \end{array}$$

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).

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.

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

$$\mathbf{u} = 0121221222122221222221222222221222\ldots$$

generated by $0 \mapsto 012$, $1 \mapsto 12$, $2 \mapsto 2$, the number $\eta = \sum_{k\geq 1} u_k 3^{-k}$ 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*.
- For instance for the Thue-Morse-Mahler numbers

$$\xi_g = \sum_{n\geq 0} \frac{a_n}{g^n}$$

(where $a_n = 0$ if the sum of the binary digits in the expansion of n is even, $a_n = 1$ if this sum is odd) the exponent of irrationality is ≤ 5.

Liouville numbers and exponent of irrationality

• An *exponent of irrationality* for $\xi \in \mathbf{R}$ is a number $\kappa \geq 2$ such that there exists $C > 0$ with

$$\left|\xi - \frac{p}{q}\right| \geq \frac{C}{q^\kappa} \quad \text{for all} \quad \frac{p}{q} \in \mathbf{Q}.$$

• A *Liouville number* is a real number with no finite exponent of irrationality.

• **Liouville's Theorem.** *Any Liouville number is transcendental.*

• In the theory of *dynamical systems*, a *Diophantine number* (or *a number satisfying a Diophantine condition*) is a real number which is not Liouville. *References :* M. Herman, J.C. Yoccoz.

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 ?*

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)

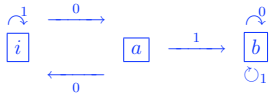
- J.H. Evertse, 1996.
- M. Queffélec, 1998 : *transcendence of the Thue–Morse continued fraction*.
- P. Liardet and P. Stambul, 2000.
- J-P. Allouche, J.L. Davison, M Queffélec and L.Q. Zamboni, 2001 : *transcendence of Sturmian or morphic continued fractions*.
- C. Baxa, 2004.
- B. Adamczewski, Y. Bugeaud,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

The Baum-Sweet sequence

- 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 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 automatic, associated with the automaton



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

International Conference on "Number Theory and Related Topics" Hanoi, December 12-15, 2006

Discrete mathematics and Diophantine Problems

Michel Waldschmidt
Institut de Mathématiques de Jussieu & CIMPA

1. Borel's Conjectures, Complexity of Words
2. Transcendence, Diophantine Approximation
3. Continued Fractions