Normal numbers and the Borel Hierarchy

Verónica Becher

Universidad de Buenos Aires

XXI Brazilian Logic Conference
Serra Negra, São Paulo, 12-16 May 2025

Normal numbers

A base is an integer greater than or equal to 2.

For a real number x, the expansion of x in base b is a sequence $a_1 a_2 a_3 \ldots$ of integers from $\{0, 1, \ldots, b-1\}$ such that

$$x - \lfloor x \rfloor = \sum_{k \geq 1} a_k b^{-k} = 0.a_1 a_2 a_3 \ldots$$

where infinitely many of the a_k are not equal to $b-1$.

For example, $3.14159265358979323846...$ shows the expansion of π in base 10.

Normal numbers

Definition (Borel, 1909)

A real number x is simply normal to base b if, in the expansion of x in base b, each digit occurs with limiting frequency equal to $1/b$.

A real number x is normal to base b if x is simply normal to base b^k, for every positive integer k.

A real number x is absolutely normal if x is normal to every base.

Theorem (Borel 1922, Niven and Zuckerman 1951)

A real number x is normal to base b if, for every $k \geq 1$, every block of k digits occurs in the expansion of x in base b with limiting frequency $1/b^k$.

Counterexamples of Borel normal numbers

$0.01001000100001\ldots$ is not simply normal to base 2.

$0.01010101010\ldots$ is simply normal to base 2 but not to 2^2 nor 2^3, etc.

Each number in Cantor middle-thirds set is not simply normal to base 3.

Rational numbers are not normal to any base.

Existence of normal numbers

Theorem (Borel 1909)

Almost all real numbers, with respect to Lebesgue measure, are absolutely normal.

Émile Borel, 1909:

Give an example of an absolutely normal number

Émile Borel, 1909:

Give an example of an absolutely normal number

Are the mathematical constants π, e, or $\sqrt{2}$, absolutely normal?
Or at least simply normal to some base?

Conjecture (Borel 1950)

Irrational algebraic numbers are absolutely normal.

Absolutely normal, non-effective constructions

Bulletin de la Société Mathématique de France (1917) 45:127–132; 132–144

DÉMONSTRATION ÉLÉMENTAIRE DU THÉORÈME DE M. BOREL SUR LES NOMBRES ABSOLUMENT NORMAUX ET DÉTERMINATION EFFECTIVE D'UN TEL NOMBRE;

PAR M. W. SIERPINSKI.

On appelle, d'après M. Borel, *simplement normal* par rapport à la base q (1) tout nombre réel x dont la partie fractionnaire

(1) E. BOREL, *Leçons sur la théorie des fonctions*, p. 197, Paris, 1914.

SUR CERTAINES DÉMONSTRATIONS D'EXISTENCE;

PAR M. H. LEBESGUE.

Dans une lettre, adressée à M. Borel, et qui accompagnait l'envoi de l'article précédent, M. Sierpinski se demandait si cet article devait être publié, s'il ne ferait pas double emploi avec une démonstration que j'avais indiquée à M. Borel et que celui-ci a signalée dans la deuxième édition de ses *Leçons sur la théorie des fonctions* (p. 198).

V. Becher and S. Figueira. An example of a computable absolutely normal number, 2002.

Normal to a given base

Theorem (Champernowne, 1933)

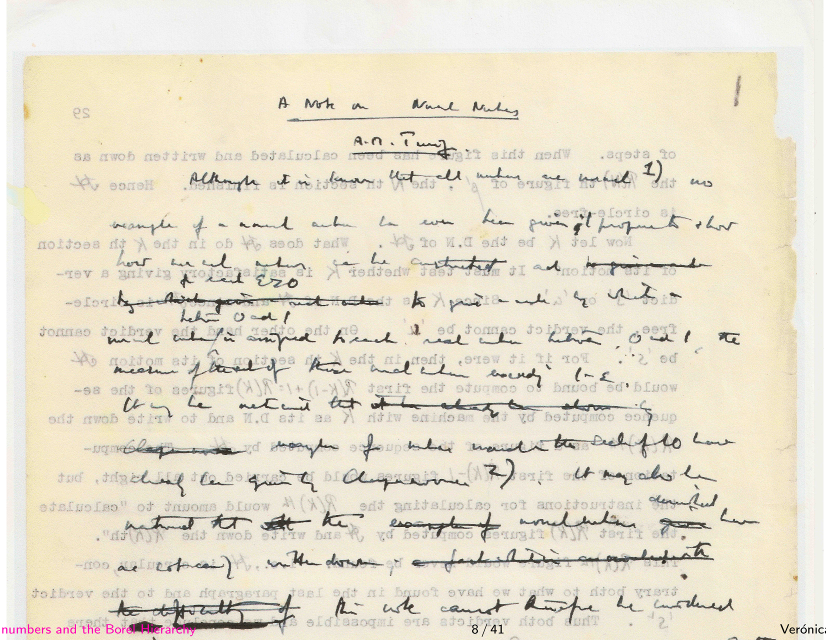
$0.123456789101112131415161718192021\ldots$ *is normal to base* 10.

The proof is by direct counting.
It is unknown if it is normal to bases that are not powers of 10.

Generalizations:

squares Besicovitch 1935,
primes Copeland and Ërdos 1946,
de Bruijn words Ugalde, 2000.

Absolutely normal, effective-construction



Alan Turing, Collected Works, Pure Mathematics, J.L. Britton editor, 1992.

A Note on Normal Numbers

Although it is known that almost all numbers are normal [1)] no example of a normal number has ever been given . I propose to shew how normal numbers may be constructed and to prove that almost all numbers are normal constructively

Consider the R -figure integers in the scale of t ($t \geq 2$). If γ is any sequence of figures in that scale we denote by $N(t,\gamma,n,R)$ the number of thesein which γ occurs exactly n times. Then it can be proved without difficulty that

$$\frac{\sum_{n=1}^{R} n\, N(t,\gamma,n,R)}{\sum_{n=1}^{R} N(t,\gamma,n,R)} = \frac{R-r+1}{R}\, t^{-r}$$

where $l(\gamma) = r$ is the lenght of the sequence γ : it is also possible[2)] to prove that

$$\sum_{|n-Rt^{-r}|>k} N(t,\gamma,n,R) < 2t^{R} e^{-k^2 t^r/4R} \qquad (1)$$

provided $\frac{kt^r}{R} < \cdot 3$

Let α be a real number and $S(\alpha,t,\gamma,R)$ the number of occurrences

ON COMPUTABLE NUMBERS, WITH AN APPLICATION TO THE ENTSCHEIDUNGSPROBLEM

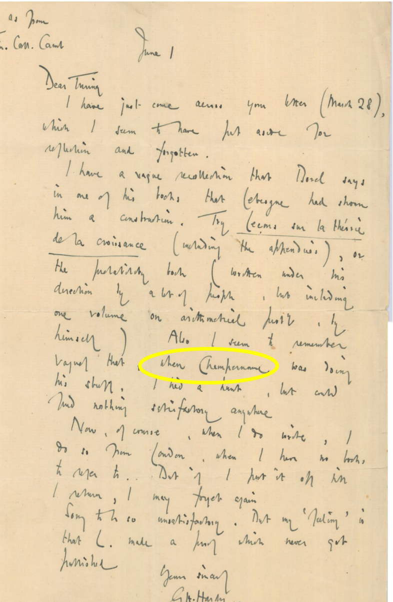
By A. M. Turing.

[Received 28 May, 1936.—Read 12 November, 1936.]

The "computable" numbers may be described briefly as the real numbers whose expressions as a decimal are calculable by finite means. Although the subject of this paper is ostensibly the computable *numbers*, it is almost equally easy to define and investigate computable functions of an integral variable or a real or computable variable, computable predicates, and so forth. The fundamental problems involved are, however, the same in each case, and I have chosen the computable numbers for explicit treatment as involving the least cumbrous technique. I hope shortly to give an account of the relations of the computable numbers, functions, and so forth to one another. This will include a development of the theory of functions of a real variable expressed in terms of computable numbers. According to my definition, a number is computable if its decimal can be written down by a machine.

In §§ 9, 10 I give some arguments with the intention of showing that the computable numbers include all numbers which could naturally be regarded as computable. In particular, I show that certain large classes of numbers are computable. They include, for instance, the real parts of all algebraic numbers, the real parts of the zeros of the Bessel functions, the numbers π, e, etc. The computable numbers do not, however, include all definable numbers, and an example is given of a definable number which is not computable.

Letter exchange between Turing and Hardy 1937 ? (AMT/D/5)



June 1

Dear Turing,
I have just came across your letter (March 28) which I seem to have put aside for reflection and forgotten.

I have a vague recollection that Borel says in one of his books that Lebesgue had shown him a construction. Try Leçons sur la théorie de la croissance (including the appendices), or the productivity book (written under his direction by a lot of people, but including one volume on arithmetical prosy, by himself).

Also I seem to remember vaguely that when Champernowne was doing his stuff I had a hunt, but could not find nothing satisfactory anywhere.

Now, of course, when I do write, I do so from London, where I have no books to refer to. But if I put it off till my return, I may forget again.
Sorry to be so unsatisfactory. But my 'feeling' is that Lebesgue made a proof which never got published.
Yours sincerely,

G.H. Hardy

Turing's algorithm for computing normal numbers

Corrected and completed Becher, Figueira, Picchi, 2007.

Theorem (Turing 1937?)

A constructive version of Borel's theorem showing the set of absolutely normal numbers in $[0,1]$ has measure Lebesgue measure 1.

Turing's proof

For each k, n,

- $E_{k,n}$ is a finite union of open intervals with rational endpoints.
- measure($E_{k,n}$)$= 1 - \dfrac{1}{k} + \dfrac{1}{k+n}$.
- $E_{k,n+1} \subset E_{k,n}$.

$E_{k,n+1}$ results from removing from $E_{k,n}$ the points that are not candidates to be normal, according to the inspection of an initial segment of their expansions.

For each k, the set $\bigcap_n E_{k,n}$ has measure $1 - \dfrac{1}{k}$ and consists entirely of absolutely normal numbers.

Turing's algorithm for computing normal numbers

Theorem (Turing 1937?)

There is an algorithm that, given an integer k and an infinite sequence ν of zeros and ones, produces an absolutely normal number $\alpha(k, \nu)$ in the unit interval, expressed in base two.

Turing's algorithm for computing normal numbers

At each step n, divide the current interval I_n in two halves,
Choose the half that includes normal numbers in large-enough measure.
The output $\alpha(k, \nu)$ is the trace of the left/right selection at each step.

Start $I_0 = (0, 1)$, $(I_n)_{n \geq 0}$ nested and decrasing, $|I_n| = 2^{-n}$.
The real number produced by the algorithm is $\alpha(k, \nu) = \bigcap_{n \geq 1} I_n$.

Computation of the n-th digit of $\alpha(k, \nu)$ in base 2 requires doubly exponential in n elementary operations.

Other exponential algorithms: Schmidt 1961/1962; Levin 1971 proved in Alvarez and Becher 2015; the effectivisation of Sierpińskis' construction, Becher, Figueira 2002 ....

Turing's Note on Normal Numbers

- Proved the existence of computable normal numbers.
- Gave a better answer to Borel's question: an algorithm!
- Started effective mathematics: concepts specified by finitely definable approximations could be made computational.

Algorithm in polynomial time

Theorem (Becher, Heiber and Slaman, 2013)

There is an algorithm that computes an absolutely normal number with just above quadratic time-complexity.

The algorithm is based on Turing's.
Speed is gained by :

testing the extension instead of the whole initial segment.

slowing convergence to normality.

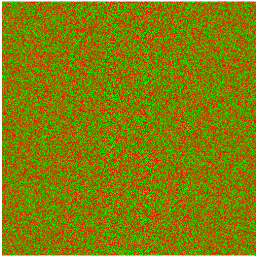
Lutz and Mayordomo (2013) and Figueira and Nies (2013) algorithm based on martingales.

Lutz and Mayordomo obtained in nearly linear time (2021)

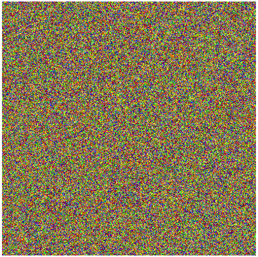
Algorithm in polynomial time

Output of algorithm Becher, Heiber and Slaman, 2013 programmed by Martin Epszteyn.

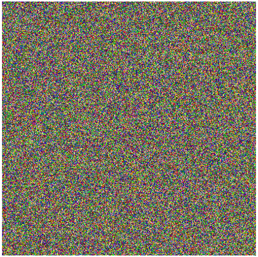
$0.4031290542003809132371428380827059102765116777624189775110896366\ldots$



base 2



base 6



base10

Plots of the first 250000 digits of the output of our algorithm.

Available from `http://www.dc.uba.ar/people/profesores/becher/software/ann.zip`

Our main tool

Turing's algorithm: $I_0 = (0,1)$. For each $n \geq 1$,

$$I_n = \left(\frac{a}{2^n}, \frac{a+1}{2^n}\right) \text{ for } a \in \{0,1\}, \text{ and } I_n \supset I_{n+1}.$$

Instead, for each n, we consider a t-sequence of intervals

$$\left(I_n^{(2)}, I_n^{(3)}, \ldots, I_n^{(t_n)}\right)$$

such that

$$I_{n-1}^{(t_{n-1})} \supset I_n^{(2)},$$
$$I_n^{(2)} \supset I_n^{(3)} \supset \ldots \supset I_n^{(t_n)},$$

and for each $b = 2, 3, \ldots, t_n$,

$$I_n^{(b)} = \left(\frac{a}{b^p}, \frac{a+1}{b^p}\right), \text{ for some } a \text{ and } p, \text{ and } I_n^{(b)} \supset I_{n+1}^{(b)}.$$

Alexander Kechris, before 1994:

What is the descriptive complexity of the set of absolutely normal numbers?

Only after having the construction of absolutely normal numbers in polynomial time were able to answer this question.

Borel hierarchy for subsets of the real numbers

The Borel hierarchy for subsets of the real numbers is the stratification of the σ-algebra generated by the open sets with the usual interval topology.

A set is $\mathbf{\Sigma}^0_1$ if it is open.

A set is $\mathbf{\Pi}^0_1$ if it is closed.

A set is $\mathbf{\Sigma}^0_{n+1}$ if it is countable union of $\mathbf{\Pi}^0_n$ sets.

A set is $\mathbf{\Pi}^0_{n+1}$ if it is a countable intersection of $\mathbf{\Sigma}^0_n$ sets.

Examples of subsets in $\mathbb{R}$.

$\mathbf{\Sigma}^0_1$ (open) : $(a,b) \cup (b,c)$, endpoints in $\mathbb{Q}$

$\mathbf{\Pi}^0_1$ (closed) : $\{0\} = \mathbb{R} \setminus \left(\bigcup_{n\geq 1}(-n,0) \cup \bigcup_{n\geq 1}(0,n) \right)$

$\mathbf{\Sigma}^0_2$ (F_σ) : $\mathbb{Q} = \bigcup_{q\in\mathbb{Q}}\{q\}$. Also any countable set is $\mathbf{\Sigma}^0_2$.

$\mathbf{\Pi}^0_2$ (G_δ) : Irrationals$= \bigcap_{q\in\mathbb{Q}} \mathbb{R} \setminus \{q\}$.

Borel hierarchy for subsets of the real numbers

A set $A \subseteq \mathbb{R}$ is hard for a class $\mathcal{C}$ if for every $C \in \mathcal{C}$ there is continuous $f : \mathbb{R} \to \mathbb{R}$ such that $C = f^{-1}(A)$.

A set is complete for a class if it is hard for this class and belongs to the class.

By Wadge theorem, a $\mathbf{\Sigma}^0_n$ subset of $\mathbb{R}$ is $\mathbf{\Sigma}^0_n$-complete iff it is not $\mathbf{\Pi}^0_n$.

Effective Borel Hierarchy for subsets of the real numbers

When we restrict to intervals with rational endpoints and computable countable unions and intersections, we obtain the effective Borel hierarchy.

One way to present the finite levels of the effective Borel hierarchy is by means of the arithmetical hierarchy of formulas in the language of second-order arithmetic, with quantification just on integers.

Effective Borel Hierarchy for subsets of the real numbers

Atomic formulas assert algebraic identities between integers or membership of real numbers in intervals with rational endpoints.

A formula is Π^0_0 and Σ^0_0 if all its quantifiers are bounded.
A formula is Σ^0_{n+1} if it has the form $\exists x\, \theta$ where θ is Π^0_n,
A formula is Π^0_{n+1} if it has the form $\forall x\, \theta$ where θ is Σ^0_n.

A set of real numbers is Σ^0_n (respectively Π^0_n) iff membership in that set is definable by a formula which is Σ^0_n (respectively Π^0_n).

$A \subseteq \mathbb{R}$ is hard for a class $\mathcal{C}$ if for every $C \in \mathcal{C}$ there is computable $f : \mathbb{R} \to \mathbb{R}$ such that $C = f^{-1}(A)$.

In the effective case Wadge theorem does not hold, so the oly way to show that a set is complete for a class is to show that it is hard for this class and belongs to the class.

Effective case implies the general case

Every Σ^0_n set is $\boldsymbol{\Sigma}^0_n$ and every Π^0_n set is $\boldsymbol{\Pi}^0_n$.

For every $\boldsymbol{\Sigma}^0_n$ set A there is a Σ^0_n formula and a real parameter such that membership in A is defined by that Σ^0_n formula relative to that real parameter.

Since computable maps are continuous, proofs of hardness in the effective hierarchy yield proofs of hardness in general by relativization.

The set of normal numbers to a given base is Π^0_3

A real $0.a_1a_2a_3\ldots$ is simply normal to base b if for all d in $\{0,\ldots,b-1\}$,

$$\lim_{n\to\infty} \frac{\#\{j : 1 \le j \le n, a_j = d\}}{n} = \frac{1}{b}.$$

That is, if for all digits d in base b, for all rational $\varepsilon > 0$, there exists $n_0 \in \mathbb{N}$ such that for all $n \ge n_0$,

$$\left| \frac{\#\{j : 1 \le j \le n, a_j = d\}}{n} - \frac{1}{b} \right| < \varepsilon.$$

The formula is

$$\forall d \forall \epsilon \exists n_0 \ \forall n \ \varphi(x, d, b, n, \varepsilon)$$

where φ has one free real variable x, one free integer variable b and 3 quantifiers only on integers.

A real x is normal to base b if x is simply normal to all integer powers of b. Then, the defining formula is

$$\forall p \forall d \forall \epsilon \exists n_0 \ \forall n \ \varphi(x, d, b^p, n, \varepsilon)$$

This is $\forall\exists\forall$, matching the 3 in the subscript of Π^0_3.

Normal Numbers in the Borel Hierarchy

Theorem (Ki and Linton 1994)

The set of real numbers that are normal to a fixed base is Π^0_3*-complete.*

Theorem (Becher, Heiber, Slaman 2014)

The set of real numbers that are absolutely normal is Π^0_3*-complete.*

Achim Ditzen conjectured in 1994, we confirmed it:

Theorem (Becher and Slaman 2014)

The set of real numbers normal to some base is Σ^0_4-complete.

Normal numbers in multiplicatively dependent bases

Two integers x and y are multiplicatively dependent if there are positive integers r and s such that $x^r = y^s$.

Example: 2 and 8 are multiplicatively dependent but 2 and 6 are multiplicatively independent.

Integers that are not perfect powers are pairwise multiplicatively independent.

Theorem (Maxfield 1953)

Let b and b' multiplicatively dependent. For any real number x, x is normal to base b if and only if x is normal to base b'.

Two theorems

Theorem (Becher and Slaman 2014)

Fix a base s. There is a computable $f : \mathbb{N} \to \mathbb{Q}$ monotonically decreasing to 0 such that for any $g : \mathbb{N} \to \mathbb{Q}$ monotonically decreasing to 0, there is an absolutely normal number x whose discrepancy for base s eventually dominates g, and whose discrepancy for each base multiplicatively independent to s is eventually dominated by f. Furthermore, x is computable from g.

Theorem (Becher and Slaman 2014)

it answers Brown, Moran and Pearce 1985
For any given set of bases closed under multiplicative dependence, there are real numbers that are normal to each base in the given set, but not simply normal to any base in its complement.

A fixed point

Theorem (Becher and Slaman 2014)

For any Π^0_3 formula φ in second order arithmetic there is a computable real number x such that, for any non-perfect power b, x is normal to base b iff $\varphi(x, b)$ is true.

Tools

In addition to discrete tools we use analytic tools.

Weyl's criterion: A sequence $(x_n)_{n\geq 1}$ of real numbers is uniformly distributed modulo 1 (u.d. mod 1) for Lebesgue measure iff for every non-zero integer t, $\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} e^{2\pi i t x_k} = 0.$

A number x normal to base b if and only if $(b^k x)_{k\geq 0}$ is u.d. mod 1 for Lebesgue measure. That is, iff, for every non-zero integer t,

$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1} e^{2\pi i t b^k x} = 0.$$

Tools

We are interested in denying simple normality to a given base while having normality to all other bases that are multiplicatively independent to it. This is an almost-everywhere porperty, albeit not in the sense of Lebesgue.

Cassels 1959 proved that for the the Cantor middle-thirds set with respect to the uniform measure, almost all numbers are normal to every base that is not a power of 3. Schmidt 1961/1962 generalized this to other Cantor sets.

Tools

To show that for the middle set Cantor set with the uniform measure μ, almost all numbers are normal to all bases that multiplicatively independent to 3, one has to prove that that, for almost all x with respect to μ for each base b multiplicatively independent to 3, for every non-zero integer t, $\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} e^{2\pi i t b^k x} = 0$.

It suffice to prove the following.

Lemma (direct application of Davenport, Erdős, LeVeque's Theorem 1963)

Let μ be a measure, I an interval and b a base. If there is constant c such that for every non-zero integer t,

$$\int_I \left| \sum_{k=0}^{n-1} e^{2\pi i t b^k x} \right|^2 d\mu(x) \ < n^{2-c}$$

then for μ-almost all x in interval I are normal to base b.

Deterministic numbers

In 1976 Gérard Rauzy defined the deterministic numbers for base b as those real numbers that, when added to a base-b normal number, preserve normality to base b.

They include the rational numbers, numbers with Sturmian expansion.

Theorem (Airey, Jackson, Mance 2022)

The set of deterministic numbers for any given base b is Π^0_3-complete.

Difference Borel Hierarchy

The difference hierarchy over a pointclass is generated by taking differences of sets.

Here we are interested in the class D_2-$\mathbf{\Pi}^0_3$ which consists of all the sets that are difference between two sets in $\mathbf{\Pi}^0_3$.

The class D_2-Π^0_3 is the effective counterpart.

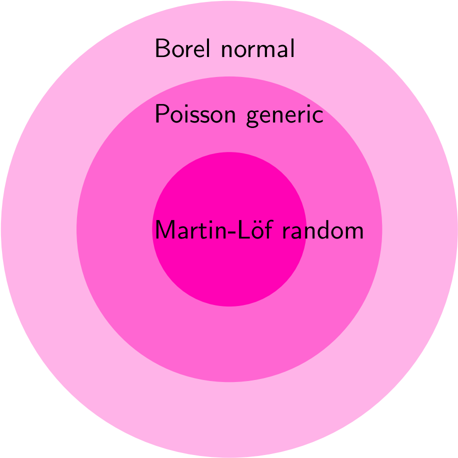
Some complexity results for difference sets

Theorem (Jackson, Mance, Vandehey 2021)

The base-2 normal but not base-3 normal is $D_2(\Pi^0_3)$-complete.

Any countable set can be written as a countable union of singleton sets and is thus in Σ^0_2. Hence, the theorem given another proof that this difference set is uncountable.

Borel normal numbers and other properties of full measure



Random numbers: Martin-Löf 1966

Poisson generic: Zeev Rudnick; Peres and Weiss (2020) ; Alvarez, Becher and Mereb (2023),

Alvarez, Becher, Cesaratto, Mereb, Peres, Weiss (2025)

Poisson generic numbers

Years ago Zeev Rudnick defined the Poisson generic numbers as those where the distribution of long blocks of digits in the initial segments of the fractional expansions is the Poisson distribution.

Definition (Zeev Rudnick)

Let λ be a positive real number. A real x is λ-Poisson generic in base b if for every non-negative integer i,

$$\lim_{k\to\infty} Z^{\lambda}_{i,k}(x) = e^{-\lambda}\frac{\lambda^i}{i!},$$

where $Z^{\lambda}_{i,k}(x)$ is the proportion of blocks of length k that occur exactly i times in the first $\lfloor \lambda b^k \rfloor$ digits of the expansion of x in base b.
A sequence is Poisson generic if it is λ-Poisson generic, for all positive λ.

Theorem (Becher, Jackson, Kwietniak, Mance 2023)

The set of Poisson generic numbers Π^0_3-complete.
The set of b-normal that are not Poisson generic in base b is $D_2(\Pi^0_3)$-complete.

Some complexity results for other numeration systems

Continued fraction, Cantor series expansions, β-expansions, generic-point in subshifts with specification all $\mathbf{\Pi}^0_3$-complete.

Beros 2017; Airey, Jackson, Kwietniak, Mance 2020; Airey, Jackson, Mance 2022

Theorem (Jackson, Mance, Vandehey 2021,2025)

Moreover, the set of numbers that are continued fraction normal but not normal to any base-b is $D_2(\Pi^0_3)$-hard.

This is the only known proof that this difference set is uncountable.

The End

Conjecture (Borel 1950)

Irrational algebraic numbers are absolutely normal.