Publ. Math. Debrecen
54 ***Supplement*** (1999), 619–623

Most numbers obey no probability laws

By CRISTIAN CALUDE (Auckland) and TUDOR ZAMFIRESCU (Dortmund)

Abstract. Therefore these numbers have now been punished: they have to appear in the authors' paper. The parallelism between category and measure fails to hold true for random reals (i.e. reals having the sequence of digits in some base CHAITIN-MARTIN-LÖF random, [2], [3]), even for simple normal numbers: the corresponding sets have measure one, but are of first category [7]. In this note we shall prove that a class of pseudo-random reals is residual. This result will be then used to directly derive the fact that, in the sense of category, most reals do not obey any probability laws. Moreover, we show that in a strong sense, nearly all numbers contain all possible words in all possible languages.

A real number in [0, 1] written in base b is called *disjunctive* [5] in case the sequence of its digits contains all possible words over that base. Disjunctivity is a "qualitative" analogue of normality [1], a weaker form pseudo-randomness. Let us denote by $\mathcal{D}$ the set of all numbers which are disjunctive in any base. We call these numbers *absolutely disjunctive*. Let x be a real number in [0, 1] written in base b, and $x_{(n)} \in \mathbf{Q}$ have precisely n digits, namely the first n digits of x (completed with zeros if necessary). If the word a appears (without overlaps) exactly k times in $x_{(n)}$, put

$$p_{a,n}(x) = \frac{kL(a)}{n},$$

where $L(a)$ means the length of (number of digits in) a. Let

$$p_a^-(x) = \liminf_{n\to\infty} p_{a,n}(x), \quad p_a^+(x) = \limsup_{n\to\infty} p_{a,n}(x).$$

Usually, we say that a appears with probability p in x if

$$p_a^-(x) = p_a^+(x) = p.$$

According to the law of large numbers, in almost every real number from $[0,1]$, every word appears with its "natural" probability, so for example the ones appear with probability 1/2 if base 2 is used. This happens for almost all, but not exactly all of them. In the sense of Baire categories, how do most numbers behave? Here "most" means "those in a residual set", i.e. "all, except those in a set of first category".

Theorem 1. *For most numbers $x \in [0,1]$, using any base and choosing any word a written in the same base,*

$$p_a^-(x) = 0 \quad \textit{and} \quad p_a^+(x) = 1.$$

PROOF. Choose arbitrarily the base b and the word a written in base b. Let, for some $\alpha \in (0,1) \cap \mathbf{Q}$,

$$\mathcal{R}_{\alpha,n}^+ = \{x \in [0,1] : \exists\, m \geq n \quad \text{s.t. } p_{a,m}(x) \geq \alpha\}.$$

We claim that $\mathcal{R}_{\alpha,n}^+$ contains an open set dense in $[0,1]$. Indeed, choose $y \in [0,1)$ and $\varepsilon > 0$ arbitrarily. Let $q \geq n$ satisfy $b^{-q} < \varepsilon$. To the digits of $y_{(q)}$ we add the word a as many times as needed in order to get a number $z = z_{(m)}$ (with $z_{(n)} = y_{(n)}$) satisfying $p_{a,m}(z) \geq \alpha$. Then the whole interval $(z, z + b^{-m})$ lies in $\mathcal{R}_{\alpha,n}^+$ and each of its points has distance at most $b^{-m} < b^{-q} < \varepsilon$ from y. The claim is proven. Analogously,

$$\mathcal{R}_{\alpha,n}^- = \{x \in [0,1] : \exists\, m \geq n \quad \text{s.t. } \; p_{a,m}(x) \leq \alpha\}$$

contains an open dense set. Of course, $\mathcal{R}_{\alpha,n}^-$ and $\mathcal{R}_{\alpha,n}^+$ depend on b and a. Thus,

$$\mathcal{R}_{a,b} = \bigcap_{\alpha,n} (\mathcal{R}_{\alpha,n}^- \cap \mathcal{R}_{\alpha,n}^+)$$

is residual in $[0,1]$. Notice that

$$\bigcap_n \mathcal{R}_{\alpha,n}^- = \{x \in [0,1] : p_a^-(x) \leq \alpha\},$$

$$\bigcap_{\alpha,n} \mathcal{R}_{\alpha,n}^- = \{x \in [0,1] : p_a^-(x) 0\},$$

and similarly for $+$ instead of $-$. Therefore $\cap_{a,b} \mathcal{R}_{a,b}$ is exactly the set of real numbers the theorem speaks about, and it is residual in $[0,1]$.

A consequence of this result is that most numbers in $[0,1]$ are absolutely disjunctive. As another immediate consequence we get a result due to OXTOBY and ULAM ([7], p. 877):

Corollary 1. *The law of large numbers is false in the sense of category.*

PROOF. Indeed, the set of all numbers $x \in [0,1]$ such that in their dyadic development the digits 0 and 1 appear with probability one-half lies in the complement of the residual set from Theorem 1.

Random numbers are transcendental, as they are non-computable [2]. Liouville numbers, in which arbitrarily sparce ones occur, are transcendental numbers which seem to be "typically" non-random. However, the contrary is true, as shown by the following corollary, which strongly improves a result due to JÜRGENSEN and THIERRIN [5], saying that, for one arbitrary base, uncountably many Liouville numbers are disjunctive.

Corollary 2. *Most Liouville numbers are absolutely disjunctive.*

PROOF. Since the residual set from Theorem 1 is a subset of $\mathcal{D}$, most numbers from $[0,1]$ are in $\mathcal{D}$. But most reals are Liouville numbers [6], [8].

So, most numbers from $[0,1]$ lie in $\mathcal{D}$. Simultaneously, $\mathcal{D}$ contains all elements of $[0,1]$ obeying the law of large numbers and has therefore measure 1. This suggests that $\mathcal{D}$ may contain nearly all elements of $[0,1]$. But what means "nearly all"? Already Denjoy used a notion of index which is essentially what DOLZHENKO [4] later called porosity. A set $M \subset [0,1]$ is said to be *porous at* $x \in [0,1]$ if there is a number $\beta > 0$ and a sequence of points $\{x_n\}_{n=1}^{\infty}$ converging to x such that

$$(x_n - \beta|x - x_n|, \quad x_n + \beta|x - x_n|) \cap M \emptyset.$$

Further, M is called *porous* if it is porous at each of its points, and it is called *σ-porous* if it is a countable union of porous sets. We say that *nearly all* points of $[0,1]$ enjoy property P if the set of points not enjoying P is σ-porous [10]. By Lebesgue's density theorem, every porous set has measure zero and therefore a set containing nearly all elements is large from both the measure and the category points of view. But the complement of a null set of first category may well not contain nearly all elements, as ZAJÍČEK [9] proved.

Theorem 2. *Nearly all numbers in* $[0,1]$ *are absolutely disjunctive.*

PROOF. Let a be a word written in base b. Let $\mathcal{P}_{a,b}$ be the set of all numbers in $[0,1]$ which, written in base b, do not contain a. We show that $\mathcal{P}_{a,b} \setminus \{1\}$ is porous. Let $y \in \mathcal{P}_{a,b} \setminus \{1\}$. Consider the arbitrary natural number n. Using the notation from the preceding proof, $y - y_{(n)} \leq b^{-n}$. Add a to the digits of $y_{(n)}$ and get another number $z = z_{(m)}$ with $m = n + L(a)$. Clearly,

$$(z, z + b^{-m}) \cap \mathcal{P}_{a,b} = \emptyset.$$

For the midpoint z_m of $(z, z + b^{-m})$, we have $|z_m - y| < b^{-n}$ and

$$b^{-m}/2 = 2^{-1}b^{-L(a)}b^{-n} > 2^{-1}b^{-L(a)}|z_m - y|,$$

whence

$$(z_m - 2^{-1}b^{-L(a)}|y - z_m|,\ \ z_m + 2^{-1}b^{-L(a)}|y - z_m|) \cap \mathcal{P}_{a,b} = \emptyset.$$

while $z_m \to y$ when $n \to \infty$.

Hence $\mathcal{P}_{a,b} \setminus \{1\}$ is porous and $\cup_{a,b}\mathcal{P}_{a,b}$, the set in the statement, is σ-porous.

References

[1] É. Borel, Les probabilités dénombrables et leurs applications arithmétiques, *Rend. Circ. Mat. Palermo* **27** (1909), 247-271.

[2] C. Calude, Information and randomness: An algorithmic perspective, *Springer-Verlag, Berlin*, 1994.

[3] G. J. Chaitin, Information, randomness and incompleteness, papers on algorithmic information theory, *World Scientific, Singapore, 2nd ed.*, 1990.

[4] E. Dolzhenko, Boundary properties of arbitrary functions,, *Izv. Akad. Nauk SSSR Ser. Mat.* **31** (1967), 3–14.

[5] H. Jürgensen and G. Thierrin, Some structural properties of ω-languages, *13th Nat. School with Internat. Participation "Applications of Mathematics in Technology"*, (1988), 56–63, *Sofia.*

[6] J. C. Oxtoby, Measure and category, *Springer-Verlag, Berlin*, 1971.

[7] J. C. Oxtoby and S. M. Ulam, Measure-preserving homeomorphisms and metrical transitivity, *Ann. Math.* **42** (1941), 874–925.

[8] W. M. Priestley, Sets thick and thin, *Amer. Math. Monthly* **83** (1976), 648-650.

[9] L. Zajíček, Porosity and σ-porosity, *Real Analysis Exchange* **13** (1987/88), 314–350.

[10] T. Zamfirescu, Porosity in convexity, *Real Analysis Exchange* **15** (1989/90), 424–436.

CRISTIAN CALUDE
DEPARTMENT OF COMPUTER SCIENCE
THE UNIVERSITY OF AUCKLAND
PRIVATE BAG 92109, AUCKLAND
NEW ZEALAND

E-mail: cristian@cs.auckland.ac.nz

TUDOR ZAMFIRESCU
FACHBEREICH MATHEMATIK
UNIVERSITY OF DORTMUND
46 DORTMUND
GERMANY

E-mail: zamfi@steinitz.mathematik.uni-dortmund.de