A STRONG HOT SPOT THEOREM

DAVID H. BAILEY AND MICHAŁ MISIUREWICZ

Abstract. A real number α is said to be b-normal if every m-long string of digits appears in the base-b expansion of α with limiting frequency b^{-m}. We prove that α is b-normal if and only if it possesses no base-b "hot spot." In other words, α is b-normal if and only if there is no real number y such that smaller and smaller neighborhoods of y are visited by the successive shifts of the base-b expansion of α with larger and larger frequencies, relative to the lengths of these neighborhoods.

1. Introduction

We say that a real constant α is b-normal (or normal base-b) if for every m, every m-long string of base-b digits appears in the base-b expansion of α with limiting frequency b^{-m}. It is well-known that almost all real numbers are b-normal for all integers b, but there are very few results proving b-normality for specific real numbers. Many of the well-known mathematical constants, including $\pi, e, \log 2, \sqrt{2}$ and numerous others, are suspected to be b-normal for commonly used number bases, based on explicit computations, but there are no proofs as yet—not to any base for any one of these constants. The problem of normality is discussed in greater detail in [4, Chapter 4].

In [1] Bailey and Crandall established b-normality of each member of the class of real constants

$$\alpha_{b,c}(r) = \sum_{k=1}^{\infty} \frac{1}{c^k b^{c^k + r_k}},$$

where the integers $b, c > 1$ are co-prime, and where r_k is the k-th binary digit of some real number $r \in [0, 1)$. This class is uncountably infinite, even for fixed b and c, due to the fact, which can be easily shown, that if $r \neq s$, then $\alpha_{b,c}(r) \neq \alpha_{b,c}(s)$. These results extend an earlier result due to Stoneham [6]. The proof given in [1] is somewhat difficult and

Date: March 26, 2005.
2000 *Mathematics Subject Classification.* Primary 11K16, Secondary 37A30.
Key words and phrases. Normal numbers.

relies on several not-well-known results, including one by Korobov on the properties of certain pseudo-random sequences.

Recently it has been found see (see [2] and Section 4 below) that normality can be established by means of a much simpler proof, as a consequence of a result given in [5, page 77], which may be termed the "weak hot spot" theorem (here $\{\cdot\}$ denotes fractional part, and # denotes count):

Theorem 1.1. *The real constant α is b-normal if and only if there exists a constant B such that for every subinterval $[c, d)$ of the unit interval,*

$$\limsup_{n\to\infty} \frac{\#_{1\le j\le n}(\{b^j\alpha\} \in [c,d))}{n} \le B(d-c).$$

Note that Theorem 1.1 implies that if a real constant α is not b-normal, then there must exist some interval $[c_1, d_1)$ with the property that successive shifts of the base-b expansion of α visit $[c_1, d_1)$ ten times more frequently, in the limit, than its length $d_1 - c_1$; there must be another interval $[c_2, d_2)$ that is visited 100 times more often than its length; there must be a third interval $[c_3, d_3)$ that is visited 1,000 times more often than its length; etc. But we cannot infer from Theorem 1.1 that these successive intervals are nested, or that there exists a real number y (a "hot spot") such that sufficiently small neighborhoods of y are visited arbitrarily "too often" relative to the lengths of these neighborhoods.

Yet many explicit non-normal constants clearly possess hot spots. Here are two simple examples. First, consider the fraction 1/28. Obviously this is not a 10-normal number since its decimal expansion repeats. It is easy to see by examining its decimal expansion that it possesses six base-10 hot spots, namely $1/7, 2/7, \cdots, 6/7$. As a second example, consider $\sum_{n\ge1} 10^{-n^2}$, which is irrational but also clearly not 10-normal. This has zero as a base-10 hot spot.

Given the key role that Theorem 1.1 played in simplifying the proof of normality for the $\alpha_{b,c}$ constants mentioned above, it behooves researchers to seek as strong a result along this line as possible. Such a result may be useful in establishing normality for a wider class of real constants, perhaps ultimately leading to a proof of normality (in some number base, such as base 2) for constants such as π and $\log 2$.

What we establish in this article is that, if α is not b-normal, then indeed there must exist a real number y with the property that sufficiently small neighborhoods of y are visited arbitrarily "too often" by successive shifts of the base-b expansion of α. In other words, we establish that a real constant is b-normal if and only if it has no base-b

hot spot. Two versions of this result, which in effect provide two specific notions of what is a base-b hot spot, are given in Theorems 3.4 and 3.5. As an example of an application of these results, in Section 4 we will show how Theorem 3.5 can be used to prove 2-normality for the constant

$$\alpha_{2,3}(0) = \sum_{n\geq 1} \frac{1}{3^n 2^{3^n}}, \tag{1}$$

which is one specific member of the class of alpha constants mentioned above.

2. Notation and Definitions

Whereas in the introduction we described normality in the real line, here we will deal with sequences of base-b digits, although it can be readily seen that these two notions are equivalent. In particular, Σ will denote the space of all sequences of base-b digits: $\{0, 1, \ldots, b-1\}^{\mathbb{N}}$. Consider this space with the product topology.

Let $\sigma : \Sigma \to \Sigma$ be the digit-shift transformation, that is,

$$\sigma(x_0, x_1, x_2, \ldots) = (x_1, x_2, x_3, \ldots),$$

and let λ be the measure on Σ which is the product of the measures equidistributed on $\{0, 1, \ldots, b-1\}$. Note that λ corresponds to ordinary Lebesgue measure on the real line. Then $(\Sigma, \sigma, \lambda)$ is a Bernoulli shift and the map $t \mapsto bt \pmod 1$ is its factor under the map $\varphi : (x_n)_{n=0}^{\infty} \mapsto 0.x_0x_1\ldots$ (base b). Where the context is clear, we will use x interchangeably to mean both $x \in \Sigma$ and the corresponding real number $\varphi(x)$.

A measure ν is said to be *absolutely continuous* with respect to the measure μ if $\nu(B) = 0$ whenever $\mu(B) = 0$. The map $T : U \to U$ is said to be *measure-preserving* with respect to μ if $\mu(T^{-1}B) = \mu(B)$ for every μ-measurable set B, and *ergodic* with respect to μ if $T^{-1}B = B$ implies either $\mu(B) = 0$ or $\mu(B) = 1$. See Billingsley's book or other references on ergodic theory for additional discussion of these notions [3].

If $x_0, x_1, \ldots, x_{s-1} \in \{0, 1, \ldots, b-1\}$ then we call the set

$$C_{x_0x_1\ldots x_{s-1}} = \{(y_n)_{n=0}^{\infty} \in \Sigma : y_0 = x_0, y_1 = x_1, \ldots, y_{s-1} = x_{s-1}\}$$

a *cylinder of length s*. Note that there are countably many distinct cylinders.

For sequences $x = (x_n)_{n=0}^{\infty}$, $y = (y_n)_{n=0}^{\infty} \in \Sigma$ and positive integers k, l, let $A(x, y, k, l)$ be the number of occurrences of the block

$(y_0, y_1, \ldots, y_{l-1})$ among the blocks $(x_0, x_1, \ldots, x_{l-1})$, $(x_1, x_2, \ldots, x_l)$, $\ldots$, $(x_{k-1}, x_k, \ldots, x_{k+l-2})$.

3. Results

We start with three simple lemmas. The first is a version of the Besicovitch covering lemma. The second is a well-known result in ergodic theory, included here for convenience.

Lemma 3.1. *If a set $X \subset \Sigma$ is covered by a collection D of cylinders, then there is a countable subcollection E of D that covers X and consists of pairwise disjoint cylinders.*

Proof. For each $x \in X$ we take the cylinder containing x, $C(x) \in D$, of the smallest length and set $E = \{C(x) : x \in X\}$. Clearly, E is a cover of X. If $C_1, C_2 \in E$ are distinct but not disjoint, then one of them is a subset of the other one, say, $C_1 \subset C_2$. But then $C_2 \subset C(x)$ for every $x \in C_1$, so $C_1 \neq C(x)$. Therefore $C_1 \notin E$, a contradiction. E is countable since the set of all cylinders is countable. $\square$

Lemma 3.2. *Suppose that T is ergodic with respect to μ, measure-preserving with respect to both μ and ν, and further that ν is absolutely continuous with respect to μ. Then $\mu = \nu$.*

Proof. Given a measurable set B, apply the ergodic theorem to $f(t) = I_B(t)$, the indicator function of B:

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} f(T^k x) = \int f(t)\, d\mu = \mu(B) \tag{2}$$

for μ-a.e. x. Since ν is absolutely continuous with respect to μ, (2) holds for ν-a.e. x as well. Since T preserves ν, we can write

$$\begin{aligned} \nu(B) &= \int f(x)\, d\nu = \frac{1}{n} \sum_{i=0}^{n-1} \int f(T^i x)\, d\nu \\ &= \int \frac{1}{n} \sum_{i=0}^{n-1} f(T^i x)\, d\nu \to \int \mu(B) d\nu = \mu(B), \end{aligned}$$

by the dominated convergence theorem. $\square$

In the following, given $x \in \Sigma$, $C(m, x)$ will denote the cylinder of length m containing x.

Lemma 3.3. *Suppose ν is a measure on Σ with the property that for ν-almost every x,*

$$\liminf_{m\to\infty} b^m \nu(C(m, x)) < \infty. \tag{3}$$

Then ν is absolutely continuous with respect to λ.

Proof. Let B be any set with $\lambda(B) = 0$, and let $\varepsilon > 0$ be given. Then there exists a set Q with $\nu(Q) < \varepsilon$ and $M > 0$ such that the left-hand side of (3), as a function of x, is smaller than M except on Q. Since B has λ measure zero, there is a set $U \supset B$, which is a union of cylinders, such that $\lambda(U) < \varepsilon/M$. Then for every $x \in B$ there is $n(x)$ such that $C(n(x), x) \subset U$. By the choice of Q, if $x \in B \setminus Q$ then there exists $m(x) \geq n(x)$ such that $b^{m(x)}\nu(C(m(x), x)) \leq M$, or in other words that $\nu(C(m(x), x)) \leq M\lambda(C(m(x), x))$. Since $m(x) \geq n(x)$, we have $C(m(x), x) \subset U$.

The collection of cylinders $C(m(x), x)$ covers $B \setminus Q$. By Lemma 3.1 there is a countable subcollection (C_k) of pairwise disjoint cylinders from this collection that also covers $B \setminus Q$. Its union is contained in U, so

$$\nu(B \setminus Q) \leq \sum_k \nu(C_k) \leq M \sum_k \lambda(C_k) \leq M\lambda(U) < M(\varepsilon/M) = \varepsilon.$$

Thus

$$\nu(B) = \nu(B \setminus Q) + \nu(B \cap Q) < \varepsilon + \varepsilon = 2\varepsilon,$$

which implies that $\nu(B) = 0$. Thus ν is absolutely continuous with respect to λ. □

Theorem 3.4. *Assume that $x \in \Sigma$ and x is not b-normal. Then there exist $y \in \Sigma$ and an increasing sequence of positive integers $(k_n)_{n=1}^{\infty}$ such that for every m the limit*

$$a_m = \lim_{n\to\infty} \frac{A(x, y, k_n, m)}{k_n} \tag{4}$$

exists and

$$\lim_{m\to\infty} b^m a_m = \infty. \tag{5}$$

Proof. Let M be the space of all probability measures on Σ with the weak-* topology. It is well known that this space is compact.

Let $x = (x_n)_{n=0}^{\infty} \in \Sigma$. Then b-normality of x is equivalent to

$$\lim_{k\to\infty} \frac{A(x, y, k, m)}{k} = \frac{1}{b^m} \tag{6}$$

for every y, m. We have

$$A(x, y, k, m) = \sum_{j=0}^{k-1} \delta_{\sigma^j(x)}(C_{y_0 y_1 \cdots y_{m-1}}),$$

where δ_z is the probability measure concentrated at z (the Dirac's delta at z). Therefore (6) is equivalent to

$$\lim_{k\to\infty} \frac{1}{k} \sum_{j=0}^{k-1} \delta_{\sigma^j(x)}(C_{y_0 y_1 \cdots y_{m-1}}) = \lambda(C_{y_0 y_1 \cdots y_{m-1}}).$$

This has to happen for every cylinder $C_{y_0 y_1 \cdots y_{m-1}}$. Therefore b-normality of x is equivalent to

$$\lim_{k\to\infty} \frac{1}{k} \sum_{j=0}^{k-1} \delta_{\sigma^j(x)} = \lambda. \tag{7}$$

Since x is not b-normal, (7) does not hold. Therefore, by compactness of M, we get

$$\lim_{n\to\infty} \frac{1}{k_n} \sum_{j=0}^{k_n-1} \delta_{\sigma^j(x)} = \mu \tag{8}$$

for some increasing sequence $(k_n)_{n=1}^{\infty}$ and some measure $\mu \in M$, $\mu \neq \lambda$. By Lemma 3.2, the measure μ cannot be absolutely continuous with respect to λ, since λ is ergodic and μ is preserved by the shift (as the limit of (8)). Therefore, by Lemma 3.3 there exists a point $y \in \Sigma$ such that

$$\lim_{m\to\infty} b^m \mu(C(m,y)) = \infty. \tag{9}$$

With the notation of that lemma, we have

$$A(x,y,k,m) = \sum_{j=0}^{k-1} \delta_{\sigma^j(x)} C(m,y).$$

Therefore, by (8), the limit in (4) exists and is equal to $\mu(C(m,y))$. This means that $a_m = \mu(C(m,y))$, so by (9) we get (5). $\square$

The following is an immediate consequence of Theorem 3.4:

Theorem 3.5. *If $x \in \Sigma$ and x is not b-normal, then there is some $y \in \Sigma$ such that*

$$\liminf_{m\to\infty} \limsup_{n\to\infty} \frac{b^m A(x,y,n,m)}{n} = \infty$$

Conversely, if for all $y \in \Sigma$,

$$\liminf_{m\to\infty} \limsup_{n\to\infty} \frac{b^m A(x,y,n,m)}{n} < \infty$$

then x is b-normal.

4. Application

As an example of an application of this result, here we give a proof of the fact, established first by Stoneham [6] and more recently in [1], that

$$\alpha = \sum_{m=1}^{\infty} \frac{1}{3^m 2^{3^m}} \tag{10}$$

is 2-normal.

Theorem 4.1. *The number α is 2-normal.*

Proof. As in the Introduction, we use the notation $\{\cdot\}$ to mean fractional part. First we note that the successive shifted binary fractions of α can be written as

$$\{2^n\alpha\} = \left\{ \sum_{m=1}^{\lfloor \log_3 n \rfloor} \frac{2^{n-3^m} \bmod 3^m}{3^m} \right\} + \sum_{m=\lfloor \log_3 n \rfloor + 1}^{\infty} \frac{2^{n-3^m}}{3^m}. \tag{11}$$

As it turns out, the first term of this expression can be generated by means of the recursion $z_0 = 0$ and, for $n \geq 1$, $z_n = \{2z_{n-1} + r_n\}$, where $r_n = 1/n$ if $n = 3^k$ for some integer k, and zero otherwise. The first few members of the z sequence are given as follows:

$0, 0, 0,$
$\frac{1}{3}, \frac{2}{3}, \frac{1}{3}, \frac{2}{3}, \frac{1}{3}, \frac{2}{3},$
$\frac{4}{9}, \frac{8}{9}, \frac{7}{9}, \frac{5}{9}, \frac{1}{9}, \frac{2}{9},$ (repeated 3 times),
$\frac{13}{27}, \frac{26}{27}, \frac{25}{27}, \frac{23}{27}, \frac{19}{27}, \frac{11}{27}, \frac{22}{27}, \frac{17}{27}, \frac{7}{27}, \frac{14}{27}, \frac{1}{27}, \frac{2}{27}, \frac{4}{27}, \frac{8}{27}, \frac{16}{27}, \frac{5}{27},$
$\frac{10}{27}, \frac{20}{27},$ (repeated 3 times), etc.

It is proven in [1] that indeed this sequence has the pattern evident here: it is a concatenation of triply repeated segments, where each individual segment consists of fractions with numerators, at stage m, that range over all integers relatively prime to the denominator 3^m. We omit this proof here. From this pattern it follows that if $n < 3^{p+1}$ then z_n is a multiple of $1/3^p$.

These fractions constitute an accurate set of approximations to the sequence $(\{2^n\alpha\})$ of shifted fractions of α. In fact, by examining (11) it can be readily seen that

$$|\{2^n\alpha\} - z_n| < \frac{1}{2n}. \tag{12}$$

Suppose we are given some binary sequence y. As before, let $C(m, y)$ be the cylinder of length m containing y. This cylinder, translated to a subset of the real unit interval, is $[c, d)$, where $c = 0.y_1y_2y_3 \ldots y_m$, and d is the next largest binary fraction of length m, so that $d - c = 2^{-m}$.

We seek an estimated upper bound for $A(\alpha, y, n, m)$. Observe that $A(\alpha, y, n, m)$ is equal to the number of those j between 0 and $n - 1$ for which $\{2^j\alpha\} \in [c, d)$. Also observe, in view of (12), that if $\{2^j\alpha\} \in [c, d)$, then $z_j \in [c - 1/(2j), d + 1/(2j))$.

Let n be any integer greater than 2^{2m}, and let 3^p denote the largest power of 3 less than or equal to n, so that $3^p \le n < 3^{p+1}$. Now note that for $j \ge 2^m$, we have $[c - 1/(2j), d + 1/(2j)) \subset [c - 2^{-m-1}, d + 2^{-m-1})$. Since the length of this latter interval is 2^{-m+1}, the number of multiples of $1/3^p$ that it contains is either $\lfloor 3^p 2^{-m+1} \rfloor$ or $\lfloor 3^p 2^{-m+1} \rfloor + 1$. Thus there can be at most three times this many j's less than n for which $z_j \in [c - 2^{-m-1}, d + 2^{-m-1})$. Therefore we can write

$$\begin{aligned}
\frac{2^m A(\alpha, y, n, m)}{n} &= \frac{2^m \#_{0 \le j < n}\big(\{2^j\alpha\} \in [c, d)\big)}{n} \\
&\le \frac{2^m \big[2^m + \#_{2^m \le j < n}\big(z_j \in [c - 2^{-m-1}, d + 2^{-m-1})\big)\big]}{n} \\
&\le \frac{2^m \big[2^m + 3(3^p 2^{-m+1} + 1)\big]}{n} < 8.
\end{aligned}$$

We have shown that for all $y \in \Sigma$ and all $m > 0$,

$$\limsup_{n\to\infty} \frac{2^m A(x, y, n, m)}{n} \le 8,$$

so by Theorem 3.5, α is 2-normal. $\square$

5. Acknowledgments

The first author wishes to acknowledge Daniel Rudolph of the University of Maryland and Jacob Feldman of the University of California, Berkeley, who helped sketch out some of the key ideas in this paper. The first author's work was supported by the Director, Office of Computational and Technology Research, Division of Mathematical, Information, and Computational Sciences of the U.S. Department of Energy, under contract number DE-AC03-76SF00098. The second author's work was supported by NSF grant DMS 0139916.

References

[1] David H. Bailey and Richard E. Crandall, "Random Generators and Normal Numbers," *Experimental Mathematics* **11** (2004), no. 4, 527-546; available at `http://crd.lbl.gov/~dhbailey/dhbpapers/normal-em.pdf`

[2] David H. Bailey, "A Hot Spot Proof of Normality for the Alpha Constants," available at http://crd.lbl.gov/~dhbailey/dhbpapers/alpha-normal.pdf

[3] Patrick Billingsley, *Ergodic Theory and Information*, John Wiley, New York, 1965.

[4] Jonathan M. Borwein and David H. Bailey, *Mathematics by Experiment: Plausible Reasoning in the 21st Century*, A K Peters, Natick, MA, 2004.

[5] L. Kuipers and H. Niederreiter, *Uniform Distribution of Sequences*, Wiley-Interscience, New York, 1974.

[6] R. Stoneham, "On Absolute (j, ε)-Normality in the Rational Fractions with Applications to Normal Numbers," *Acta Arithmetica* **22** (1973), 277-286.

Lawrence Berkeley Laboratory, 1 Cyclotron Road, Berkeley, CA 94720

E-mail address: dhbailey@lbl.gov

Department of Mathematical Sciences, IUPUI, 402 N. Blackford Street, Indianapolis, IN 46202-3216

E-mail address: mmisiure@math.iupui.edu