On modifying normal numbers

Christoph Aistleitner*

Abstract

A real number x is called a normal number (in a base $\beta \geq 2$) if all possible blocks of digits appear with the same asymptotic frequency in the β-ary expansion of x. This notion was introduced by E. Borel in 1909, who proved that almost all real numbers (in the sense of Lebesgue measure) are normal numbers. There exist many constructions of normal numbers, most of which are based on assembling appropriate function values to obtain the digital expansion of x. Volkmann showed that normal numbers constructed in this way can be slightly modified without losing the normality property. In the present note, we generalize the result of Volkmann.

Let $\mathcal{C}(\beta)$ denote the set of reals which have a β-ary representation with asymptotic frequency of non-zero digits equal to 0. We prove that if x is normal in base β, then $x + qy$ is also normal in the same base, for any $y \in \mathcal{C}(\beta)$ and $q \in \mathbb{Q}$.

Likewise, let $\mathcal{S}$ denote a set of positive integers of asymptotic density 0. We prove that a normal number x retains the normality property, if an arbitrary digit is inserted between positions j and $j+1$ for each $j \in \mathcal{S}$.

1 Introduction and statement of results

Let x be a real number, and let $\beta \geq 2$ be an integer. Then x can be written in the form

$$x = \lfloor x \rfloor + \sum_{j=1}^{\infty} x_j \beta^{-j}, \qquad x_j \in \{0, \ldots, \beta - 1\}.$$

Let $B_k = y_1 y_2 \ldots y_k$ be a block of k digits in base β (that is, $0 \leq y_j \leq \beta - 1,\ 1 \leq j \leq k$). Let

$$N(B_k, x, n) = \#\{j \in \mathbb{N}, 1 \leq j \leq n - k + 1 :\ x_j = y_1, \ldots, x_{j+k-1} = y_k\},$$

i.e. the function N counts the number of occurrences of the block B_k within the first n digits (after the decimal point) of the β-ary expansion of x. Then x is called *normal* in base β if for all $k \geq 1$ and all possible blocks B_k of k digits

$$\frac{N(B_k, x, n)}{n} \to \frac{1}{\beta^k} \qquad \text{as} \quad n \to \infty.$$

This notion was introduced by Borel [1] in 1909, who showed that almost all numbers (in the sense of Lebesgue measure) are normal numbers. The first example of a normal number was given by Champernowne [2], who showed that the number

$$0.1\ 2\ 3\ 4\ 5\ 6\ 7\ 8\ 9\ 10\ 11\ 12\ 13\ 14\ \ldots,$$

*Graz University of Technology, Institute of Mathematics A, Steyrergasse 30, 8010 Graz, Austria. e-mail: aistleitner@math.tugraz.at. Research supported by the Austrian Research Foundation (FWF), Project S9610-N23.

Mathematics Subject Classification: 11K16

constructed by assembling the decimal representations of the consecutive positive integers, is normal in base 10. Later, this method was generalized to prove that, for some appropriately chosen function f, the number

$$0.f(1)\ f(2)\ f(3)\ f(4)\ f(5)\dots \tag{1}$$

or

$$0.\lfloor f(1)\rfloor\ \lfloor f(2)\rfloor\ \lfloor f(3)\rfloor\ \lfloor f(4)\rfloor\ \lfloor f(5)\ \rfloor\dots \tag{2}$$

(if f is not an integer-valued function) is normal in a certain base. Various constructions of this type appear in the papers of Copeland and Erdős [4], Davenport and Erdős [5], Nakai and Shiokawa [11, 12, 13]. The results of Madritsch [8] and Madritsch, Thuswaldner and Tichy [9] are the most recent contributions in this area. For a general reference on normal numbers, see [6, 7, 16].

There exist only few results on modifications of normal numbers. Recently, Pellegrino [15] showed that if the number

$$0.x_1x_2x_3\dots$$

is normal, then the number

$$0.x_1(x_1x_2)(x_1x_2x_3)\dots$$

is normal as well.
Another result, due to Volkmann [17], is tailor-made for normal numbers which are constructed using the technique described in equations (1), (2). Let x be a normal number, whose β-ary expansion is given by assembling positive integers $p_1, p_2, p_3, \dots$ in their β-ary representation

$$x = 0.p_1\ p_2\ p_3\dots,$$

where $p_j \to \infty$ as $j \to \infty$. Let $q_1, q_2, q_3, \dots$ be positive integers. Then the normality of x also implies the normality of the number

$$\hat{x} = 0.(p_1+q_1)\ (p_2+q_2)\ (p_3+q_3)\dots,$$

provided that $\log q_j = o(\log p_j)$ as $j \to \infty$. In other words, Volkmann showed that if assembling a certain sequence of blocks of digits produces a normal number, then assembling a slightly modified version of this sequence produces a normal number as well.

Let $\mathcal{C}(\beta)$ denote the set of all reals for which the set of indices of nonzero digits (in their expansion in base β) has asymptotic density zero. More precisely, $\mathcal{C}(\beta)$ contains all real numbers $x = \lfloor x\rfloor.x_1x_2x_3\dots$ for which

$$\lim_{n\to\infty} \frac{\#\{1 \le j \le n:\ x_j \ne 0\}}{n} = 0.$$

Moreover, let $\mathcal{C}(\beta)_{\mathbb{Q}}$ denote the set of all reals which are a rational multiple of a number in $\mathcal{C}(\beta)$, i.e.

$$\mathcal{C}(\beta)_{\mathbb{Q}} := \{x \in \mathbb{R}:\ x = yz \quad \text{for some} \quad y \in \mathbb{Q},\ z \in \mathcal{C}(\beta)\}.$$

We mention that in particular $\mathbb{Q} \subset \mathcal{C}(\beta)_{\mathbb{Q}}$, and that $x+y \in \mathcal{C}(\beta)_{\mathbb{Q}}$ for all $x, y \in \mathcal{C}(\beta)_{\mathbb{Q}}$.

We will prove the following two theorems:

Theorem 1 *Let $\beta \geq 2$, and let $y \in \mathcal{C}(\beta)_{\mathbb{Q}}$. If x is normal in base β, then $x+y$ is also normal in base β.*

Theorem 2 *Let $\beta \geq 2$, and assume that x is normal in base β. Let $\mathcal{S}$ be a set of indices which is of asymptotic density zero, and assume that y results from adding an arbitrary digit in the base-β-expansion of x between positions j and $j+1$ for each $j \in \mathcal{S}$. Then y is also normal in base β.*

Remark 1: The assumptions of Theorem 1 imply not only the normality of $x+y$, but also the normality of $x-y$. In fact, the normality of x also implies the normality of $-x$ (see [10]), and thus by Theorem 1 the normality of x implies the normality of $-x+y$ and $-(-x+y) = x-y$.

Remark 2: By applying Theorem 2 repeatedly, we can insert a block of at most K digits (for some fixed K) in the β-ary representation of x between positions j and $j+1$ for each $j \in \mathcal{S}$ without loosing the normality property, provided $\mathcal{S}$ is of asymptotic density zero.

Remark 3: Theorem 1 and Theorem 2 together imply the aforementioned result of Volkmann. We will a give a short explanation of this assertion at the end of the present paper.

In connection with the two theorems above we pose the following two open problems:

Open problem 1: Let $\beta \geq 2$. Write $\mathcal{D}(\beta)$ for the set of reals which preserve normality if added to a normal number. More precisely,

$$\mathcal{D}(\beta) := \{y \in \mathbb{R} : \; x+y \text{ is normal in base } \beta \text{ for all } x \text{ which are normal in base } \beta\}.$$

Then by Theorem 1 we have $\mathcal{C}_{\mathbb{Q}}(\beta) \subset \mathcal{D}(\beta)$. Find a complete characterization of $\mathcal{D}(\beta)$.

Open problem 2: Characterize the set $\mathcal{D}^*$ of all real numbers, which preserve *absolute normality* if added to an absolutely normal number.

Concerning Problem 1, we remark that the set $\mathcal{D}(\beta)$ is necessarily of Lebesgue measure zero: normality of x implies normality of $-x$. Therefore $\mathcal{D}(\beta)$ can not contain any normal number x, since $(-x)+x = 0$ and 0 is not normal; by the result of Borel [1] the set of normal numbers has full measure, so necessarily $\lambda(\mathcal{D}(\beta)) = 0$. Concerning Problem 2, we note that $\mathbb{Q} \subset \mathcal{D}^*$, and that $\mathcal{D}^*$ of course also has measure zero.

2 Proofs

If x is normal in base β, then qx is also normal in base β for any $q \in \mathbb{Q}\backslash\{0\}$, see [3, 10]. Every $y \in \mathcal{C}_{\mathbb{Q}}(\beta)$ can be written in the form $q\hat{y}$, where $q \in \mathbb{Q}$ and $\hat{y} \in \mathcal{C}(\beta)$. Then $x+y$ is normal if and only if $x/q + \hat{y}$ is normal. Thus it is sufficient to prove Theorem 1 for $y \in \mathcal{C}(\beta)$ instead of $y \in \mathcal{C}_{\mathbb{Q}}(\beta)$.

A number x is normal in base β if and only if the sequence $(\langle \beta^k x \rangle)_{k \geq 1}$, where $\langle \cdot \rangle$ denotes the fractional part of a real number, is uniformly distributed modulo one. This was first observed by Wall [18] (see also [14, p. 110]). By Weyl's criterion (see [6] or [7]) a sequence $(z_k)_{k\geq 1}$ of reals from the unit interval is uniformly distributed modulo one (u.d. mod 1) if and only if

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} e^{2\pi i h z_k} = 0 \qquad \text{for all} \qquad h \in \mathbb{Z},\ h \neq 0. \tag{3}$$

Lemma 1 *Let $y \in \mathcal{C}(\beta)$. Then*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \langle \beta^k y \rangle = 0.$$

Proof: Let $y = \lfloor y \rfloor + \langle y \rangle = \lfloor y \rfloor + 0.y_1 y_2 y_3 \ldots$. Then, by assumption,

$$\lim_{n\to\infty} \frac{\#\{1 \leq j \leq n:\ y_j \neq 0\}}{n} = 0. \tag{4}$$

It is easily seen that

$$\langle \beta^k y \rangle = 0.y_{k+1} y_{k+2} y_{k+3} \cdots \tag{5}$$

By (4), for any fixed $K \geq 1$,

$$\lim_{n\to\infty} \frac{\#\{1 \leq j \leq n:\ (y_j, \ldots, y_{j+K-1}) \neq (0, \ldots, 0)\}}{n} = 0.$$

By (5), this implies

$$\begin{aligned} & \frac{1}{n} \sum_{k=1}^{n} \langle \beta^k y \rangle \\ \leq \quad & \left(\frac{1}{n} \sum_{k=1}^{n} \beta^{-K+1} \right) + \frac{\#\{1 \leq j \leq n:\ (y_j, \ldots, y_{j+K-1}) \neq (0, \ldots, 0)\}}{n} \\ \leq \quad & \beta^{-K+1} + o(1). \end{aligned}$$

Since K was arbitrary, this proves the lemma.

Proof of Theorem 1: Assume that x is normal in base β. By Weyl's criterion this implies

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} e^{2\pi i h \beta^k x} = 0 \qquad \text{for all} \qquad h \in \mathbb{Z},\ h \neq 0. \tag{6}$$

As mentioned before we may assume w.l.o.g. that $y \in \mathcal{C}(\beta)$. Then for all $h \in \mathbb{Z},\ h \neq 0$ we have

$$\begin{aligned} \left| \frac{1}{n} \sum_{k=1}^{n} e^{2\pi i h \beta^k (x+y)} \right| &= \left| \frac{1}{n} \sum_{k=1}^{n} e^{2\pi i h \beta^k x} \left(e^{2\pi i h \beta^k y} - 1 \right) + \frac{1}{n} \sum_{k=1}^{n} e^{2\pi i h \beta^k x} \right| \\ &\leq \frac{1}{n} \sum_{k=1}^{n} \left| e^{2\pi i h \beta^k y} - 1 \right| + \left| \frac{1}{n} \sum_{k=1}^{n} e^{2\pi i h \beta^k x} \right| \\ &\leq \frac{1}{n} \sum_{k=1}^{n} 2\pi |h| \langle \beta^k y \rangle + \left| \frac{1}{n} \sum_{k=1}^{n} e^{2\pi i h \beta^k x} \right|. \end{aligned} \tag{7}$$

The first sum in (7) tends to 0 as $n \to \infty$ by Lemma 1. The second sum in (7) also tends to 0 by (6). This proves Theorem 1.

Proof of Theorem 2: Let $x = \lfloor x \rfloor + \langle x \rangle = \lfloor x \rfloor + 0.x_1x_2x_3\dots$, and assume that x is normal in base β. Write $y = 0.y_1y_2y_3\dots$ for the number which results from adding an arbitrary digit in the base-β-expansion of x between positions j and $j+1$, for each $j \in \mathcal{S}$. To be able to distinguish the digits of x and y we introduce the functions u, w, S and the set $\mathcal{T}$.

By assumption,

$$\frac{\#\{1 \le j \le n : \ j \in \mathcal{S}\}}{n} \to 0 \quad \text{as} \quad n \to \infty. \tag{8}$$

For $m \ge 1$, define

$$S(m) = \#\{1 \le k \le m-1 : \ k \in \mathcal{S}\}.$$

The function $S(m)$ counts how many additional digits are inserted into the expansion of x prior to position m. Then by assumption

$$\frac{S(m)}{m} \to 0 \quad \text{as} \quad m \to \infty.$$

Write $u(j)$ for the position of the digit x_j in the expansion of y, i.e.

$$u(j) = j + S(j).$$

Define

$$\mathcal{T} := \{j \ge 1 : \ \text{there exists an } k \ge 1 \text{ such that } j = k + S(k)\}.$$

Then

$$\lim_{n\to\infty} \sum_{j=1}^{n} \frac{\mathbb{1}(j \notin \mathcal{T})}{n} = 0. \tag{9}$$

We define $w(j)$ as the index of the digit y_j in the expansion of x, provided it exists. More precisely, $w(j)$ is the index k for which

$$k + S(k) = j.$$

This number $w(j)$ exists if $j \in \mathcal{T}$.

Fix $K \ge 1$, and set

$$\mathcal{S}_K := \{j \ge 1 : \ (\mathbb{1}_{j\in\mathcal{S}} + \dots + \mathbb{1}_{j+K-1\in\mathcal{S}}) \ge 1\}. \tag{10}$$

Then, by (8),

$$\lim_{n\to\infty} \sum_{j=1}^{n} \frac{\mathbb{1}(j \in \mathcal{S}_K)}{n} = 0. \tag{11}$$

For $w(j) \notin \mathcal{S}_K$ we have

$$(x_{w(j)}, \dots, x_{w(j)+K-1}) = (y_j, \dots, y_{j+K-1}),$$

which implies

$$\left| \langle \beta^{w(j)} x \rangle - \langle \beta^j y \rangle \right| \le \beta^{-K+1},$$

and for any $h \in \mathbb{Z},\ h \neq 0$,

$$\left| e^{2\pi i h \beta^{w(j)} x} - e^{2\pi i h \beta^j y} \right| \leq 2\pi |h| \beta^{-K+1}. \tag{12}$$

Let $n \geq 1$. We may assume that $n \in \mathcal{T}$, since the difference between two consecutive elements of $\mathcal{T}$ is at most 2. Then there exists an $m = m(n)$ such that

$$n = m + S(m),$$

and

$$\{w(j) \leq n,\ j \in \mathcal{T}\} = \{1, \dots, m\}.$$

Using (12), we get

$$\begin{aligned}
& \left| \frac{1}{n} \sum_{j=1}^{n} e^{2\pi i h \beta^j y} \right| \\
= & \left| \frac{1}{n} \sum_{\substack{1 \leq j \leq n,\\ j \in \mathcal{T}}} e^{2\pi i h \beta^j y} + \frac{1}{n} \sum_{\substack{1 \leq j \leq n,\\ j \notin \mathcal{T}}} e^{2\pi i h \beta^j y} \right| \\
\leq & \left| \frac{1}{n} \sum_{\substack{1 \leq j \leq n,\\ j \in \mathcal{T}, w(j) \in \mathcal{S}_K}} e^{2\pi i h \beta^j y} \right| + \left| \frac{1}{n} \sum_{\substack{1 \leq j \leq n,\\ j \in \mathcal{T}, w(j) \notin \mathcal{S}_K}} e^{2\pi i h \beta^j y} \right| + \frac{\#\{1 \leq j \leq n :\ j \notin \mathcal{T}\}}{n} \\
\leq & \left| \frac{1}{n} \sum_{\substack{1 \leq j \leq n,\\ j \in \mathcal{T}, w(j) \notin \mathcal{S}_K}} e^{2\pi i h \beta^{w(j)} x} \right| + \frac{2\pi |h|}{\beta^{K-1}} + \frac{\#\{1 \leq j \leq n :\ j \notin \mathcal{T} \text{ or } w(j) \in \mathcal{S}_K\}}{n}.
\end{aligned} \tag{13}$$

The third term in (13) tends to zero as $n \to \infty$ by (9) and (11). The second term in (13) can be made arbitrarily small since K was arbitrary. Finally, the first term in (13) is bounded by

$$\begin{aligned}
& \frac{\#\{1 \leq j \leq n, j \in \mathcal{T}, w(j) \in \mathcal{S}_K\}}{n} + \frac{1}{n} \sum_{\substack{1 \leq j \leq n,\\ j \in \mathcal{T}}} e^{2\pi i h \beta^{w(j)} x} \\
= & \frac{\#\{1 \leq j \leq m,\ j \in \mathcal{S}_K\}}{n} + \frac{1}{n} \sum_{j=1}^{m} e^{2\pi i h \beta^j x} \\
\leq & \frac{\#\{1 \leq j \leq m,\ j \in \mathcal{S}_K\}}{m} + \frac{1}{m} \sum_{j=1}^{m} e^{2\pi i h \beta^j x},
\end{aligned}$$

which also tends to zero by (11) and since x is normal in base β. This proves Theorem 2.

In conclusion we give a short proof of Remark 3. Let $p_1, p_2, p_3, \dots$ be positive integers (in a fixed base β) and assume that $p_j \to \infty$ as $j \to \infty$. Let $q_1, q_2, q_3, \dots$ also be positive integers, and assume that

$$\log q_j = o(\log p_j) \qquad \text{as} \qquad j \to \infty. \tag{14}$$

We want to show that the normality of the number

$$x = 0.p_1\ p_2\ p_3\ \dots \tag{15}$$

implies the normality of

$$\hat{x} = 0.(p_1 + q_1)\ (p_2 + q_2)\ (p_3 + q_3)\ \dots .$$

Since finitely many digits do not affect the normality property, we can assume that $q_j \leq p_j$ for $j \geq 1$. This means that the number of digits of $(p_j + q_j)$ exceeds the number of digits of p_j at most by 1. Whenever the number of digits of $(p_j + q_j)$ is larger than the number of digits of p_j (that is, whenever a carry occurs) we replace p_j (understood as a block of digits) by $\hat{p_j} =$ "$0p_j$". In other words, the block $\hat{p_j}$ consists of the same digits as p_j, but with an additional leading digit "0". Write $d(m)$ for the number digits (in base β) of an integer m. Then for any $M \geq 1$ the number of digits of the combination of the blocks $p_1\ \dots\ p_M$ is $\sum_{j=1}^M d(p_j)$. Consider the expansion of x, truncated somewhere between the end of the block p_M and the end of the block p_{M+1}. The number of positions, where additional digits are inserted, is at most $M+1$, while the number of digits of the truncated expansion is at least $\sum_{j=1}^M d(p_j)$. By assumption $\lim_{j\to\infty} p_j = \infty$, which implies $\lim_{j\to\infty} d(p_j) = \infty$ and

$$\lim_{M\to\infty} \frac{M+1}{\sum_{j=1}^M d(p_j)} = 0.$$

Thus the set of positions in the β-ary expansion of x, where additional digits are inserted, is of asymptotic density zero, and by Theorem 2 the normality of the number (15) implies the normality of

$$\widetilde{x} = 0.\hat{p_1}\ \hat{p_2}\ \hat{p_3}\ \dots .$$

Now replace the blocks $q_1, q_2, q_3, \dots$ by blocks $\hat{q_1}, \hat{q_2}, \hat{q_3}, \dots$, which are padded with so many additional leading digits "0" that the number of digits of $\hat{q_j}$ equals the number of digits of $\hat{p_j}$, for all $j \geq 1$. Set

$$q = 0.\hat{q_1}\ \hat{q_2}\ \hat{q_3}\ \dots .$$

Then by (14) the asymptotic frequency of non-zero digits in the β-ary representation of q is zero. More precisely, the number of nonzero digits in the expansion of q, truncated at the end of the M-th block $\hat{q_M}$, is at most $A(M) := \sum_{j=1}^M d(q_j)$, while the total number of digits in the truncated expansion is at least $B(M) := \sum_{j=1}^M d(p_j)$. By (14) we have $\lim_{j\to\infty} d(q_j)/d(p_j) = 0$, and thus the ratio of $A(M)/B(M) \to 0$ as $M \to \infty$ by the Stolz-Cesàro theorem. If one truncates inside $\hat{q_M}$ before the block q_M, then the ratio is bounded by

$$\frac{A(M-1)}{B(M-1)+r} < \frac{A(M-1)}{B(M-1)}$$

(where r denotes the number of padding zeros at the end of the truncated expansion). If one truncates at the block q_M, then the ratio is bounded by

$$\frac{A(M)-r}{B(M)-r} < \frac{A(M)}{B(M)}$$

(where r is the number of missing digits from the end of the truncated expansion to the end of the block q_M). Thus the asymptotic frequency of nonzero digits in the expansion of q is equal to 0. Therefore $q \in \mathcal{C}(\beta)$, and by Theorem 1 the number $\widetilde{x} + q$ is normal. Finally, it is easily seen that the numbers $\widetilde{x}$ and q are constructed in such a way that $\widetilde{x} + q = \hat{x}$. Thus $\hat{x}$ is also normal, which proves Remark 3.

References

[1] E. Borel. Les probabilités dénombrables et leurs applications arithmétiques. *Rend. Circ. Palermo* **27**:247-271, 1909.

[2] D.G. Champernowne. The construction of decimals in the scale of ten. *J. London Math. Soc.* **8**:254-260, 1933.

[3] K.T. Chang. A note on normal numbers. *Nanta Math.* **9**(1):70-72, 1976.

[4] A.H. Copeland and P. Erdős. Note on normal numbers. *Bull. Amer. Math. Soc.* **52**:857-860, 1946.

[5] H. Davenport and P. Erdős. Note on normal decimals. *Canadian J. Math.* **4**:58-63, 1952.

[6] M. Drmota and R. Tichy. *Sequences, discrepancies and applications.* Lecture Notes in Mathematics 1651, Springer-Verlag, 1997.

[7] L. Kuipers and H. Niederreiter. *Uniform distribution of sequences.* Pure and Applied Mathematics. Wiley-Interscience, 1974.

[8] M. Madritsch. Generating normal numbers over Gaussian integers. *Acta Arith.* **135**:63-90, 2008.

[9] M. Madritsch, J. Thuswaldner and R. Tichy. Normality of numbers generated by the values of entire functions. *J. Number Theory* **128**(5):1127-1145, 2008.

[10] J.E. Maxfield. Normal k-tuples. *Pacific J. Math.* **3**(1):189-196, 1953.

[11] Y. Nakai and I. Shiokawa. A class of normal numbers. *Japan J. Math.* **16**:17-29, 1990.

[12] Y. Nakai and I. Shiokawa. A class of normal numbers II. *Number Theory and Cryptography. London Math. Soc. Lecture Notes Ser. A.* **154**, Cambridge Univ. Press:204-210, 1990.

[13] Y. Nakai and I. Shiokawa. Discrepancy estimates for a class of normal numbers. *Acta Arith.* **62**:271-284, 1994.

[14] I. Niven. *Irrational numbers.* Carus Math. Monographs, John Wiley and Sons, 1956.

[15] D. Pellegrino. On normal numbers. *Proyecciones* **25**(1):19-30, 2006.

[16] O. Strauch and Š. Porubský. *Distribution of sequences: a sampler.* Series of the Slovak Academy of Sciences 1, Peter Lang, 2005.

[17] B. Volkmann. On modifying constructed normal numbers. *Ann. Fac. Sci. Toulouse Math. (5)* **1**(3):269-285, 1979.

[18] D.D. Wall. *Normal numbers.* Ph.D. thesis, University of California, Berkeley, 1949.