János Galambos

Uniformly distributed sequences mod 1 and Cantor's series representation

Czechoslovak Mathematical Journal, Vol. 26 (1976), No. 4, 636–641

Persistent URL: http://dml.cz/dmlcz/101435

Czechoslovak Mathematical Journal, 26 (101) 1976, Praha

UNIFORMLY DISTRIBUTED SEQUENCES mod 1 AND CANTOR'S SERIES REPRESENTATION

JANOS GALAMBOS*), Philadelphia

(Received March 27, 1975)

1. SUMMARY

The aim of the present paper is two-fold. First, we show that Cantor's series representation of real numbers induces a measure on the set of sequences uniformly distributed mod 1, which measure leads to the commonly used infinite dimensional Lebesgue measure with independent components. This part therefore reverses the usual procedure of using the above measure by assumption. In other words, we "prove" that the above Lebesgue measure is the natural one to choose from among all possibilities. Secondly, we show that a set of sequences, associated with the Cantor series, and which sequences are uniformly distributed mod 1, is an "exceptional set" in the set of all sequences uniformly distributed mod 1. This therefore points to the fact that the general metric theory of uniformly distributed sequences mod 1 cannot in general be applied in connection with series representations of real numbers (for an account of such representations, see [1] and its references).

The present work is strongly related to the result of ŠALÁT [3]. However, our set-up is more general than that of Šalát and our methods of proof are completely different from his.

2. THE CANTOR SERIES

Let $\mathbf{Q} = \{q_k\}$ be a sequence of integers with $q_k \geqq 2$ for each $k \geqq 1$. Any real number $0 < x < 1$ leads to the Cantor series

$$x = \sum_{k=1}^{+\infty} \varepsilon_k(x)/q_1 q_2 \dots q_k, \tag{1}$$

*) This research was done while the author was on Research and Study Leave from Temple University and, as a Fellow of the Humboldt Foundatin, he was at the Mathematisches Seminar der Goethe-Universität, Frankfurt am Main.

where the integer coefficients $\varepsilon_k(x)$ are obtained by the following algorithm

(2) $$x = x_1\,, \quad \varepsilon_k(x) = [x_k q_k] \quad \text{and} \quad x_{k+1} = x_k q_k - \varepsilon_k(x)\,, \quad k \geqq 1\,.$$

Here, and in what follows, $[y]$ signifies the integer part of y. By the algorithm (2)

(3) $$0 \leqq \varepsilon_k(x) < q_k\,.$$

It is easily seen that any series of the form (1), satisfying (3), is the Cantor series of its sum.

Let $a_1, a_2, \ldots, a_n$ be given integers satisfying $0 \leqq a_k \leqq q_k - 1$ for $1 \leqq k \leqq n$. Then (2) immediately yields that the set

$$I_n = I_n(\mathbf{Q};\, a_1, a_2, \ldots, a_n) = \{x : \varepsilon_j(x) = a_j,\ 1 \leqq j \leqq n\}$$

is an interval of length $1/(q_1 q_2 \ldots q_n)$ and the I_n are disjoint for different sets of the a's. Therefore, denoting by $\lambda(A)$ the Lebesgue measure of the set A,

(4) $$\lambda(I_n(\mathbf{Q};\, a_1, a_2, \ldots, a_n)) = 1/q_1 q_2 \ldots q_n$$

and

(5) $$\lambda(\{x : \varepsilon_j(x) = a_j\}) = 1/q_j\,, \quad 0 \leqq a_j < q_j\,, \quad 1 \leqq j \leqq n\,.$$

We summarize the meaning of (4) and (5) as

Lemma 1. *The coefficients $\varepsilon_k(x)$, defined by* (1) *and* (2), *are stochastically independent with respect to Lebesgue measure and, for each* $k = 1, 2, \ldots, \varepsilon_k(x)$ *is uniformly distributed on the integers* $0, 1, 2, \ldots, q_k - 1$.

3. THE RESULTS

We first introduce a definition and two notations.

Definition. A sequence $z_1, z_2, \ldots$ of real numbers is said to be *uniformly distributed* mod 1 if the sequence $\vartheta_k = z_k - [z_k]$ satisfies the limit relation below. For any given α, $0 \leqq \alpha \leqq 1$, let $N_n(\alpha)$ be the number of those $k \leqq n$, for which $\vartheta_k < \alpha$. Then, as $n \to +\infty$,

(6) $$\lim N_n(\alpha)/n = \alpha\,.$$

Since, in the above definition, an arbitrary sequence $\{z_k\}$ is transformed into a sequence $\{\vartheta_k\}$ with $0 \leqq \vartheta_k < 1$ for each k, in the sequel we restrict ourselves to sequences with components from the interval $[0, 1]$. We shall denote by $\mathcal{U}$ the set of all sequences ϑ_k, $k = 1, 2, \ldots$, with $0 \leqq \vartheta_k < 1$ and for which (6) holds. In addition, in all statements in the sequel, $\mathbf{Q}$ stands for a sequence $\{q_k\}$ of integers with $q_k \geqq 2$ for $k = 1, 2, \ldots$ and we put

$$S_n = \frac{1}{n} \sum_{k=1}^{n} \frac{1}{q_k}. \tag{7}$$

Before we formulate our results, we quote a theorem of Šalát [3] as

Lemma 2. *For a given* $\mathbf{Q}$, *let* $T_{\mathbf{Q}}$, *a map of* $\mathscr{U}$ *into the interval* $[0, 1]$, *be defined by*

$$T_{\mathbf{Q}}\{\vartheta_k\} = \sum_{k=1}^{+\infty} [\vartheta_k q_k] / q_1 q_2 \ldots q_k . \tag{8}$$

Then, if $S_n \to 0$ *as* $n \to +\infty$,

$$\lambda(T_{\mathbf{Q}}\mathscr{U}) = 1 .$$

We now state our results.

Theorem 1. *Let* X_k, $k = 1, 2, \ldots$, *be a sequence of independent random variables and let* X_k *be uniformly distributed on the integers* $0, 1, 2, \ldots, q_k - 1$, *where* $q_k \geqq 2$ *for each* k. *Then the sequence* $\vartheta_k = X_k / q_k$, $k = 1, 2, \ldots$, *belongs to* $\mathscr{U}$ *almost surely if, and only if, as* $n \to +\infty$,

$$\lim S_n = 0 . \tag{9}$$

Theorem 2. *For a given* $\mathbf{Q}$, *we introduce the measure*

$$\lambda_{\mathbf{Q}}(A) = \lambda(T_{\mathbf{Q}} A)$$

on subsets A *of* $\mathscr{U}$ *for which* $T_{\mathbf{Q}} A$ *is Borel measurable* (see (8)). *Then the only measure on subsets of* $\mathscr{U}$ *which is consistent with each* $\lambda_{\mathbf{Q}}$ *is the inflnitive dimensional Lebesgue measure* λ_∞ *with independent components.*

Theorem 3. *Let* $\mathbf{Q}$ *be such that* $nS_n \to +\infty$ *with* n. *Let* $\varepsilon_k(x)$, $k = 1, 2. \ldots$, *be defined by* (2) *and put* $\vartheta_k = \vartheta_k(x) = \varepsilon_k(x)/q_k$. *Then, for almost all* x *in* $[0, 1)$, *the discrepancy*

$$D_n = \sup_{0 \leqq \alpha \leqq 1} \left| \frac{N_n(\alpha)}{n} - \alpha \right| \geqq \tfrac{1}{2} S_n ,$$

for sufficiently large n.

Some comments are in order. First of all, we should add that, in view of Lemma 1, Theorem 1 is essentially due to Šalát [3]. However, since his proof is very dependent on the algorithm (2), it is not clear from his argument that the conclusion remains to hold in the generality of Theorem 1. Naturally, our proof is completely different from his. Secondly, we wish to emphasize that Theorems 2 and 3 indeed contain our claim expressed in the Summary. As a matter of fact, Theorem 2 implies that a departure from λ_∞ on subsets of $\mathscr{U}$ means a departure from Lebesgue measure on $[0, 1]$ for expansions. While this is, of course, permitted in principle, any measure, different from that of Lebesgue, puts emphasis on certain subintervals of $[0, 1]$ rather than giving general information on $[0, 1]$. On the other hand, Theorem 3 shows that those elements of $\mathscr{U}$ which are "produced" by the Cantor algorithm (2), are exceptional ones for many choices of **Q** by violating the rule valid for λ_∞-almost all elements of $\mathscr{U}$. Namely, by the iterated logarithm theorem,

$$\lambda_\infty(\{\vartheta_k\} : \limsup_{n=+\infty} (\sqrt{n}) D_n(n \log\log n)^{1/2} \leqq c) = 1, \tag{10}$$

while, by Theorem 3, for the sequence $\vartheta_k = \varepsilon_k(x)/q_k$,

$$\limsup (D_n/S_n) \geqq \tfrac{1}{2}. \tag{11}$$

Obviously, (10) and (11) can coincide only in exceptional cases. As an Example, take $q_k = \max(2, [\log\log k])$. Then (11) yields $D_n \log\log n \geqq \frac{1}{2}$ for large n.

4. PROOFS

Proof of Theorem 1. Put

$$Y_k(\alpha) = \begin{cases} 1 & \text{if } \vartheta_k < \alpha \\ 0 & \text{otherwise}. \end{cases}$$

Then, by assumption, the variables $Y_k(\alpha)$, $k \geqq 1$, are independent and

$$P(Y_k(\alpha) = 1) = \alpha + O(1/q_k),$$

where the constant in the error term $O(\cdot)$ is uniformly bounded in k. Therefore, by the strong law of large numbers and by (9), as $n \to +\infty$,

$$\frac{N_n(\alpha)}{n} = \frac{1}{n} \sum_{k=1}^{n} Y_k(\alpha) \to \alpha$$

almost surely. Conversely, assume that, as $n \to +\infty$,

$$\limsup S_n = s > 0.$$

Then there is a subsequence $n(t)$, $t = 1, 2, \ldots$, of integers such that, as $t \to +\infty$

$$\lim S_{n(t)} = s. \tag{12}$$

Let Z_k be the indicator of the event $\{X_k = 0\}$. That is, $Z_k = 1$ if $X_k = 0$ and $Z_k = 0$ otherwise. Then evidently $Z_k \leqq Y_k(\alpha)$ for each k and for any $0 < \alpha < 1$. Therefore, for any $0 < \alpha < 1$,

$$(13) \qquad \frac{N_n(\alpha)}{n} \geqq \frac{1}{n} \sum_{k=1}^{n} Z_k .$$

Since $P(Z_k = 1) = 1/q_k$, another appeal to the strong law of large numbers yields, in view of (12) and (13), that almost surely, as $n \to +\infty$,

$$(14) \qquad \limsup N_n(\alpha)/n \geqq s > 0 , \quad 0 < \alpha < 1$$

(14), however, contradicts (6) for all $0 < \alpha < s$, that is, almost surely, the sequence ϑ_k' $k \geqq 1$, is not an element of $\mathcal{U}$. Theorem 1 is thus established.

Proof of Theorem 2. Let μ be a measure on subsets of $\mathcal{U}$ and let us evaluate the distribution $F_k(z) = \mu(\vartheta_k < z)$ of ϑ_k for a fixed k under our assumption (we can restrict ourselves to $0 < z < 1$). Let first z be rational. Writing $z = t/q$, where t and q are positive integers with $t \leqq q$ and $q \geqq 2$, we choose a sequence $\mathbf{Q}$, the k-th component q_k of which is q and for which $S_n \to 0$ as $n \to +\infty$. Then by Lemma 2 and by the definition of $\lambda_{\mathbf{Q}}$, in evaluating $\lambda_{\mathbf{Q}}(\vartheta_k < z)$ we can replace $[\vartheta_k q_k]$ by $\varepsilon_k(x)$ of (1) and (2). Thus

$$\lambda_{\mathbf{Q}}(\vartheta_k < z) = \lambda_{\mathbf{Q}}(\vartheta_k q_k < t) = \lambda_{\mathbf{Q}}([\vartheta_k q_k] \leqq t - 1) =$$
$$= \lambda(\varepsilon_k(x) \leqq t - 1) = t/q = z .$$

Since μ is consistent with each $\lambda_{\mathbf{Q}}$, we have got that, for any rational z,

$$(15) \qquad F_k(z) = \mu(\vartheta_k < z) = z .$$

Let now z be irrational. Then there is an infinite sequence of rational numbers t/q such that $q \to +\infty$ and

$$\frac{t-1}{q} < z \leqq \frac{t}{q} .$$

By the monotonity of $F_k(z)$ and by (15) for rational values,

$$\frac{t-1}{q} \leqq F_k(z) \leqq \frac{t}{q} ,$$

that is, $|F_k(z) - z| < 1/q$ for infinitely many q's with $q \to +\infty$. Therefore (15) holds for all $0 < z \leqq 1$. In evaluating the multivariate distribution of $(\vartheta_{k_1}, \vartheta_{k_2}, \ldots$ $\ldots, \vartheta_{k_m})$, we can progress as in the one dimensional case above. As above, for rational $z_j = t_j/r_j$, $1 \leqq j \leqq m$, the distribution

$$\lambda_{\mathbf{Q}}(\vartheta_{k_1} < z_1, \vartheta_{k_2} < z_2, \ldots, \vartheta_{k_m} < z_m) =$$
$$= \lambda(\varepsilon_{k_1}(x) < z_1, \varepsilon_{k_2}(x) < z_2, \ldots, \varepsilon_{k_m}(x) < z_m)$$

where **Q** is again a sequence $\{q_k\}$ with $S_n \to 0$ as $n \to +\infty$ and for which $q_{k_j} = r_j$, $1 \leqq j \leqq m$. Thus by Lemma 1 and by the consistency assumption, for any $1 \leqq \leqq k_1 < k_2 < \ldots < k_m$ and for rational z's,

$$\mu(\vartheta_{k_1} < z_1, \vartheta_{k_2} < z_2, \ldots, \vartheta_{k_m} < z_m) = z_1 z_2 \ldots z_m . \tag{16}$$

We can now show componentwise as in the one dimensional case that (16) remains to hold for abitrary $0 < z_j \leqq 1$, $1 \leqq j \leqq m$. The proof of Theorem 2 is thus complete.

Proof of Theorem 3. For the proof we need the following formula of NIEDERREITER [2]:

$$D_n = \frac{1}{2n} + \max_{1 \leqq i \leqq n} \left| \vartheta_i^* - \frac{2i - 1}{2n} \right| , \tag{17}$$

where ϑ_i^*, $1 \leqq i \leqq n$, is the ϑ_i, $1 \leqq i \leqq n$, rearranged in non-decreasing order. For proving Theorem 3, we simply show that "many" $\varepsilon_k(x)$, and thus $\vartheta_k(x)$, vanish and we then apply (17). For showing this, let us put

$$Z_k = Z_k(x) = \begin{cases} 1 & \text{if } \varepsilon_k(x) = 0 \\ 0 & \text{otherwise} . \end{cases}$$

Then $\lambda(Z_k(x) = 1) = 1/q_k$ and the strong law of large numbers yields that, for almost all x in $[0, 1]$, as $n \to +\infty$

$$T_n = \sum_{k=1}^{n} Z_k \sim nS_n . \tag{18}$$

On the other hand, (17) implies that

$$D_n \geqq \frac{1}{2n} + \frac{2T_n - 1}{2n} = T_n/n$$

and (18) therefore terminates the proof.

References

[1] *Galambos, J.:* On infinite series representations of real numbers. Compositio Math. *27*, 197—204 (1973).

[2] *Niederreiter, H.:* Discrepancy and convex programming. Ann. Math. Pura Appl. *93*, 89—97 (1972).

[3] *Šalát, T.:* Zu einigen Fragen der Gleichverteilung (mod 1). Czechosl. Math. J. *18*, 476—488 (1968).

Author's addresses: Department of Mathematics, Temple University, Philadelphia, Pa. 19121, U.S.A. and Mathematisches Seminar der Goethe Universität, 6 Frankfurt/Main, Robert-Mayer-Str. 10, BRD.