des fonctions localement constantes sur les boules de rayon $|p^{u+1}|$, or $(Q_n)_{0\leqslant n<p^{u+1}}$ est une base de E_{u+1} [1], ainsi il existe p^{u+1} élément de Z_p, $(a_n)_{0\leqslant n<p^{u+1}}$ tels que

$$f_1 = \frac{1}{p}\Big(f - \sum_{n=0}^{p^{u+1}-1} a_n Q_n\Big),$$

soit une fonction de $C(Z_p, Z_p)$. D'autre part, l'hypothèse (ii) du théorème et le lemme 4 montrent que f_1 est une fonction localement constante sur les boules de rayon $|p^{u+2}|$ et le théorème se laisse alors aisément démontrer par récurrence.

Notons que ce théorème admet une généralisation et une réciproque [7].

Bibliographie

[1] Yvette Amice, *Interpolation p-adique,* Bull. Soc. Math. France 92 (1964), p. 117–160.

[2] N. C. Ankeny, M. Artin, and S. Chowla, *On the class number of a quadratric number field,* Annals of Math. 56 (1952), p. 479–493.

[3] H. Bauer, *Numerische Bestimmung von Klassenzahlen reeller zyklischer Zahlkörper,* Journal of Number Theory 1 (1969), p. 161–162.

[4] Jean Fresnel, *Nombres de Bernoulli et fonctions L p-adiques,* Annales de l'Institut Fourier 17 (2) (1967), p. 281–333.

[5] — *Applications arithmétiques de la formule p-adique des résidus,* Sém. Delange Pisot Poitou (1966–1967), 8ème année N° 18.

[6] Laurent Gruson, *Algèbres de Banach ultramétriques,* Journées Poitou-Aquitaine de la S. M. F., Poitiers 1967.

[7] Eve Helsmoortel, *Comportement local des fonctions continues sur un compact régulier d'un corps local,* C. R. Acad. Sc. Paris, 271 (12), p. 546–548.

[8] Kenkichi Iwasawa, *On p-adic L-functions,* Annals of Math. 89 (1969), p. 198–205.

[9] Marc Krasner, *Prolongement analytique uniforme et multiforme dans le corps valués complets,* Les tendances géométriques en algèbre et théorie des nombres, N° 143, Colloques internationaux du C.N.R.S.

[10] T. Kubota and H. W. Leopoldt, *Eine p-adische Theorie der Zetawerte. I. Einführung der p-adischen Dirichletschen L-Funktionen,* J. Reine Angew. Math. 2 14/215 (1964), p. 328–333.

[11] H. W. Leopoldt, *Eine Verallgemeinerung der Bernoullishen Zahlen,* Abh. Math. Sem. Univ. Hamburg 22 (1958), p. 131–140.

[12] — *Zur Arithmetik in abelschen Zahlkörpern,* J. Reine Angew. Math. 209 (1962), p. 54–71.

[13] — *Über Fermatquotienten von Kreiseinheiten und Klassenzahlformeln modulo p,* Rend. Circ. Mat. Palermo 9 (1960), p. 39–50.

[14] I. S. Slavutski, *On Mordell's theorem,* Acta Arith. 11 (1965), p. 57–66.

Reçu le 15. 2. 1971 (145)

ACTA ARITHMETICA
XX (1972)

The Hausdorff dimension of sets related to g-expansions

by

JÀNOS GALAMBOS (Philadelphia, Pa.)*

1. Introduction. The classical theorem of Borel states that for $g = 2$, the digits $\{e_k\}$ in the expansion

$$x = \sum_{k=1}^{+\infty} e_k g^{-k}, \quad x \in (0,1), \tag{1}$$

are stochastically independent with respect to Lebesgue measure, hence the law of large numbres implies that the relative frequency of 0's among $e_1, e_2, \ldots, e_N$ tends to $p = \frac{1}{2}$. By this same theorem we also have that, with any prescribed m, for almost all $x \in (0,1)$, there are infinitely many k's such that $e_k = e_{k+1} = \ldots = e_{k+m} = 0$. These problems become very difficult if we take g in (1) to be $1 < g < 2$. The e_k's are no more independent, though the results quoted above for $g = 2$ remain to hold (except that $p = \frac{1}{2}$; it will be an expression in terms of g), see [8], [3] and [4]. Recently I came across a problem in mathematical statistics [5] where I needed the distribution of the length between two consecutive one's in the expression (1) if g is the (only) solution in $(1, 2)$ of

$$z^{a+1} - z^a = 1 \tag{2}$$

with some integer $a \geqslant 1$. Note that if $a = 0$, then $g = 2$, hence the number theoretical problems related to (1) with this g are natural generalizations of Borel's investigations. In the present note I shall evaluate the Hausdorff dimension of the set of those x's for which the distance between any consecutive one's is bounded by M, a given number. Before giving the precise statement, however, I should specify the algorithm for (1). Without an algorithm, the digits e_k are not defined, as the following example shows this. Let $g = \frac{1}{2}(\sqrt{5}+1)$, the solution of (2) with $a = 1$. Then for $x = \frac{1}{2}(\sqrt{5}-1)$ we have

$$x = \frac{1}{g} = \sum_{m=1}^{s} \frac{1}{g^{2m}} + \sum_{k=2s+3}^{+\infty} \frac{1}{g^k}$$

* The author was a recipient of Summer Faculty Research Award at Temple University.

for any integer $s \geqslant 1$, i. e., there are infinite many ways to have the series (1). I use the following algorithm, being the natural extension of decimal expansions: define the sequence $n_1, n_2, \ldots$ of integers by

$$x = x_1, \quad g^{-n_k} < x_k \leqslant g^{-n_k+1}, \quad x_{k+1} = x_k - g^{-n_k} \tag{3}$$

to have the series representation

$$x = \sum_{k=1}^{+\infty} g^{-n_k} \tag{4}$$

with $1 \leqslant n_1 < n_2 < \ldots$ positive integers. (4) is evidently (1) after the zero terms having been dropped, and corresponds to the algorithm of [8].

My aim is to prove the following result.

THEOREM. *Let g be the (only) solution of (2) satisfying $1 < g < 2$. Let $M \geqslant a+1$ be an integer and let Z_M denote the set of those $x \in (0,1)$ for which $m_k = n_k - n_{k-1}$ are bounded by M for all $k \geqslant 1$, where $n_0 = 0$. The Hausdorff dimension $H(M)$ of Z_M is the solution in s of the equation*

$$w(s, M) = \sum_{k=a+1}^{M} g^{-ks} = 1. \tag{5}$$

According to the remark in the previous paragraph, Z_M has Lebesgue measure 0. The assumption $M \geqslant a+1$ is not a restriction, since the algorithm (3) implies that $m_k \geqslant a+1$ for all $k \geqslant 2$ (see Lemma 1 in the next section). I shall prove Theorem in the next section. In addition to my results [3] and [4], I shall make use of the technique developed by Salát [9] to investigate the Hausdorff dimension of certain sets related to Lüroth's expansion.

2. The proof of Theorem. Let us state the definition of the Hausdorff dimension used in this paper.

DEFINITION. Let $c > 0$ be a given real number and let $U \subset (0,1)$. A collection D of a denumerable number of intervals I is called a *c-covering of U* if $U \subset \bigcup_{I \in D} I$ and if $|I| \leqslant c$ for all I in D ($|I|$ stands for the length of I). Let $\mathscr{A}(c, U)$ be the set of all c-coverings of U and set

$$h(c, U, s) = \inf_{D \in \mathscr{A}(c,U)} \sum_{I \in D} |I|^s \tag{6}$$

and

$$h(U, s) = \lim h(c, U, s) \quad \text{as} \quad c \to 0+. \tag{7}$$

It is known (see [1] and [10]) that the limit (7) always exists and also that there is one and only one s_0 such that for all $0 < s < s_0$, $h(U, s) = +\infty$ and for $s_0 < s < 1$, $h(U, s) = 0$. This unique s_0 is called the Hausdorff dimension of U.

Hence in order to prove Theorem we have to show that, in case of $U = Z_M$, $h(Z_M, s) = +\infty$ for all $0 < s < s_0$, where s_0 is the root of the equation (5), i. e., $w(s_0, M) = 1$, and for $s_0 < s < 1$, $h(Z_M, s) = 0$.

Before turning to the details, let us quote some results of [3] and [4], which are related to the structure of Z_M and which will be needed in the sequel.

LEMMA 1. *Let $1 \leqslant n_1 < n_2 < \ldots$ be defined by (3). Then the set $\{n_1 = j_1, n_2 = j_2, \ldots, n_k = j_k\}$ is an interval of length*

$$g^{-j_k+1}(1-1/g)$$

if $j_r - j_{r-1} \geqslant a+1$ for $r \geqslant 2$ and it is empty otherwise.

LEMMA 2. *The functions $m_k = n_k - n_{k-1}$, $k \geqslant 1$, where $n_0 = 0$, are stochastically independent with respect to Lebesgue measure and for $k \geqslant 2$, their distribution is given by*

$$P(m_k = t) = \begin{cases} g^{-t} & \text{for } t \geqslant a+1, \\ 0 & \text{otherwise} \end{cases}$$

(P is Lebesgue measure).

Lemma 2 is quoted to justify the claim that $P(Z_M) = 0$, as indeed it follows from Lemma 2 and from the law of large numbers.

Let us now give the details of the proof. Let s_0 be the solution in s of (5) and let $s_0 < s < 1$.

By the definition of Z_M it is evident that for any $k \geqslant 1$,

$$Z_M \subset \bigcup_{\substack{a+1 \leqslant t_i \leqslant M \\ 1 \leqslant i \leqslant k}} \{m_1 = t_1, m_2 = t_2, \ldots, m_k = t_k\}. \tag{8}$$

Since

$$\{m_1 = t_1, m_2 = t_2, \ldots, m_k = t_k\} = \{n_1 = t_1, n_2 = t_1 + t_2, \ldots, n_k = t_1 + \ldots + t_k\}, \tag{9}$$

in view of Lemma 1, (8) provides a c-covering of Z_M, if

$$g^{-t_1-t_2-\ldots-t_k+1}(1-1/g) < c$$

which, by $t_i \geqslant a+1$, holds if k is chosen so that

$$g^{-(a+1)k+1}(1-1/g) < c. \tag{10}$$

Fix k to satisfy (10). Then by (6) and (8) and again by Lemma 1,

$$h(c, Z_M, s) \leqslant \sum_{a+1 \leqslant t_i \leqslant M} g^{-s(t_1+\ldots+t_k)} \leqslant \Big(\sum_{t=a+1}^{M} g^{-st}\Big)^k = w^k(s, M). \tag{11}$$

By the definition of s_0, for all $s_0 < s < 1$, $w(s, M) < 1$. Hence, since the inequality (11) holds for all k satisfying (10),

(12) $$h(c, Z_M, s) = 0 \quad \text{for} \quad s_0 < s < 1.$$

Clearly, $h(Z_M, s) = 0$ for $s_0 < s < 1$, and the definition thus implies that $H(M) \leqslant s_0$. Note that if $M = a+1$, $s_0 = 0$, therefore $H(a+1) = 0$. In the sequel therefore we assume that $M > a+1$, which implies that $s_0 > 0$.

Let $0 < s < s_0$. By Definition, it is sufficient to prove that $h(Z_M, s) > 0$. To show this, we shall give in several steps a lower estimate for $h(c, Z_M, s)$ which is positive and which at the final stage will not depend on c. The details are as follows.

Let $\varepsilon > 0$ be an arbitrary real number. In view of (6), there is a c-covering D of Z_M such that

(13) $$h(c, Z_M, s) + \varepsilon \geqslant \sum_{I \in D} |I|^s.$$

By the definition of Z_M,

(14) $$Z_M = \bigcap_{k=1}^{+\infty} \bigcup_{\substack{t_i \in (a+1, M) \\ 1 \leqslant i \leqslant k}} \{m_1 = t_1, m_2 = t_2, \ldots, m_k = t_k\}.$$

Define

(15) $$I(t_1, t_2, \ldots, t_k) = \{m_1 = t_1, m_2 = t_2, \ldots, m_k = t_k\}.$$

Lemma 1 states that the sets $I(t_1, t_2, \ldots, t_k)$ are intervals. We call these sets fundamental intervals. Take the closures of the fundamental intervals occuring in (14). This results in adding a denumerable set to Z_M, let this be denoted by Z_M^*. Evidently, the Hausdorff dimension of Z_M and of Z_M^* coincide, hence the covering D in (13) can be taken to cover Z_M^*. It is also known that the covering intervals can be taken to be open. Now, since an open system D covers a closed set Z_M^*, the Heine–Borel theorem, [6], p. 72, yields that a finite subset of D already covers Z_M^*. Evidently, it can be assumed that the endpoints of the intervals left from D to cover Z_M^* are elements of Z_M^*. As a matter of fact, an interval I can be taken to be the closed interval $[u, v]$ where $u = \inf\{I \cap Z_M^*\}$ and $v = \sup\{I \cap Z_M^*\}$. Let this new system of intervals be denoted by D_1, i. e., D_1 is a finite collection of intervals which is a c-covering of Z_M^* and such that the end-points of the intervals belonging to D_1 are elements of Z_M^*. Since D_1 was obtained from D by dropping some of its elements and by possibly reducing the length of those intervals left, we have from (13) that

(16) $$h(c, Z_M, s) + \varepsilon \geqslant \sum_{I \in D_1} |I|^S.$$

We shall now give a lower estimate for the terms on the right hand side of (16) in terms of fundamental intervals which estimate will be independent of c. Since, by Lemma 1, the length of a fundamental interval (15) tends to 0 with k, the description of the structure of D_1 above clearly shows the following fact. Starting with the interval $(0, 1)$ and constructing the fundamental intervals with $k = 1$, then with $k = 2$, and so on, there will be a first step when, for $I \in D_1$, the fundamental interval $I(t_1, t_2, \ldots, t_k)$ contains I but the endpoints of I belong to $I(t_1, \ldots, t_k, d_1)$ and $I(t_1, \ldots, t_k, d_2)$, respectively, with $d_1 \neq d_2$. Since, by construction, the fundamental intervals $I(t_1, \ldots, t_k, d_i)$, $i = 1, 2$, occur in the representation (14), we have that all the parameters $t_1, t_2, \ldots, t_k, d_1$ and d_2 are between $a+1$ and M and also for any y for which $I(t_1, \ldots, t_k, d_i, y)$ can cover the endpoints of I, we have that $a+1 \leqslant y \leqslant M$ (the endpoints of I are elements of Z_M^*, and any point of Z_M^* belongs to a fundamental interval with its parameters being between $a+1$ and M). This however implies that the length of I can not be arbitrarily small in terms of $I(t_1, \ldots, t_k)$, namely, the least value of $|I|$ is obtained if $d_1 = M$ and $d_2 = M-1$, and then necessarily I contains all fundamental intervals $I(t_1, \ldots, t_k, M-1, y)$ with $y > M$. These intervals are disjoint and their length, by Lemma 1, is

$$\sum_{y=M+1}^{+\infty} g^{-t_1-\cdots-t_k-(M-1)-y+1}(1-1/g) = \{g^{-2M+1}/(g-1)\}|I(t_1, \ldots, t_k)|.$$

Considering the interval $(0, 1)$ as a fundamental interval (15) with $k = 0$, the above argument applies to this case, too, hence we have got that for any $I \in D_1$, there is a well defined fundamental interval containing I and the argument above yields that

(17) $$\sum_{I \in D_1} |I|^s \geqslant \{g^{-(2M+1)s}/(g-1)^s\} \sum_{I \in D_2} |I|^s$$

where D_2 is the collection of the fundamental intervals obtained by the construction above. It is evident that two fundamental intervals are either disjoint or one is contained in the other one. Drop those elements of D_2 which are contained in another element of D_2 and let the remaining set be denoted by D_3. This results in an additional decrease in (17), hence (16) yields the inequality

(18) $$h(c, Z_M, s) + \varepsilon \geqslant \{g^{-(2M+1)s}/(g-1)^s\} \sum_{I \in D_3} |I|^s.$$

If D_2 contained the fundamental interval $(0, 1)$, then D_3 contains a single element, $(0, 1)$ only, hence for this case we have from (18) that

(19) $$h(c, Z_M, s) + \varepsilon \geqslant g^{-(2M+1)s}/(g-1)^s$$

where $\varepsilon > 0$ is arbitrary. Since the right hand side of (19) does not depend on c, it gives that, if $D_3 = \{(0, 1)\}$,

$$h(Z_M, s) \geqslant g^{-(2M+1)s}/(g-1)^s. \tag{20}$$

Let us turn to the case when $(0, 1) \notin D_3$. Let k in (15) be called the order of the fundamental interval occuring in (15). Let $I(t_1, \ldots, t_k)$ be an element of D_3 with largest order in D_3. Such an element exists since D_3 is finite and by assumption, $k > 0$. Since D_3 is a covering of Z_M^* and each element of D_3 does cover points of Z_M^*, we have that all t_i, $1 \leqslant i \leqslant k$, are between $a+1$ and M, namely, a fundamental interval $I(u_1, \ldots, u_v)$ with at least one $u_i > M$ contains no points of Z_M^*. By the maximal property of k and by the elements of D_3 being disjoint, we therefore have that the points of Z_M^* which belong to the fundamental intervals $I(t_1, \ldots, t_{k-1}, y)$ with $a+1 \leqslant y \leqslant M$, can be covered only if D_3 contains all of these fundamental intervals. By Lemma 1, however,

$$\sum_{y=a+1}^{M} |I(t_1, \ldots, t_{k-1}, y)|^s = |I(t_1, \ldots, t_{k-1})|^s \sum_{u=a+1}^{M} g^{-us} \tag{21}$$

and since for $0 < s < s_0$,

$$\sum_{u=a+1}^{M} g^{-us} > 1,$$

(21) yields that if we replace in D_3 all of its elements of order k by the single element $I(t_1, \ldots, t_{k-1})$, the right hand side of (18) is further decreased and by this step the maximal order of D_3 is reduced. Repeating this step k times, we arrive at (19), hence we proved (20) for the general case. As Definition implies, the proof is complete.

Finally, I wish to draw the attention of the reader to the works [2] and [7], where further properties of g-expansions are investigated; see also the relevant references to [7].

References

[1] P. Billingsley, *Ergodic Theory and Information*, New York 1965.

[2] J. Cigler, *Ziffernverteilung in Θ-adischen Brüchen*, Math. Zeitshr. 75 (1961), pp. 8–13.

[3] J. Galambos, *Probabilistic theorems concerning expansions of real numbers*, Periodica Math. Hungarica (to appear in the A. Rényi Memorial Volume).

[4] — *A generalization of a theorem of Borel concerning the distribution of digits in dyadic expansions*, Amer. Math. Monthly 78 (1971), pp. 774–779

[5] — *On a model for a fair distribution of gifts*, J. Appl. Probability 8 (1971), pp. 681–690.

[6] R. C. James, *Advanced Calculus*, Belmont, Cal., 1966.

[7] W. Philipp, *Mixing sequences of random variables and probabilistic number theory*, Mem. Amer. Math. Soc. 114 (1971).

[8] A. Rényi, *Representations for real numbers and their ergodic properties*, Acta Math. Acad. Sci. Hung. 8 (1957), pp. 477–493.

[9] T. Šalát, *Zur metrischen Theorie der Lürothschen Entwicklungen der reellen Zahlen*, Czechosl. Math. J. 18 (1968), pp. 489–522.

[10] — *О мере Хаусдорфа линейных множеств*, Czechosl. Math. J. 11 (1961), pp. 24–56.

DEPARTMENT OF MATHEMATICS
TEMPLE UNIVERSITY
Philadelphia, Pennsylvania

Received on 15. 5. 1971 (167)