des fonctions localement constantes sur les boules de rayon \([p^{k+1}]\), or
\(\xi_{m,n} e^{\xi_{m,n} p^{k+1}}\) est une base de \(E_{p}^{\frac{1}{p}}\), ainsi il existe \(p^{k+1}\) d'élément de \(Z_{p}\),
\((a_{n})_{n \in \mathbb{C}^{p}, n \geq p^{k+1}}\) tels que

\[ f_{1} = \frac{1}{p} \left( \sum_{n=0}^{p^{k+1}-1} a_{n} Q_{n} \right), \]

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 \([p^{k+1}]\) et le théorème se laisse alors aisément démontrer par récurrence.

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

Bibliographie


Reçu le 15. 2. 1971 (145)

ACTA ARITHMETICA
XX (1972)

The Hausdorff dimension of sets related to \(g\)-expansions

by

JÁNOS GALAMBOS (Philadelphia, Pa.)*

1. Introduction. The classical theorem of Borel states that for \(g = 2\),
the digits \(\{e_{k}\}\) in the expansion

\[ x = \sum_{k=1}^{\infty} e_{k} g^{-k}, \quad x \in (0, 1), \]

are stochastically independent with respect to Lebesgue measure, hence
the law of large numbers 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 expansion (1) if \(g\) is the (only) solution in (1, 2) of

\[ |z|^{a+1} - |z|^{a} = 1 \]

with some integer \(a \geq 1\). 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 \(e_{k}\)'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

\[ x = \frac{1}{g} = \sum_{m=1}^{\infty} \frac{1}{g^{2m}} + \sum_{k=2m+1}^{\infty} \frac{1}{g^{k}} \]

* The author was a recipient of Summer Faculty Research Award at Temple University.
for any integer \( s \geq 1 \), 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

\[
ax = a_1, \quad g^{-n_k} < a_k \leq g^{-n_{k+1}}, \quad a_{k+1} = a_k - g^{-n_k}
\]

to have the series representation

\[
a = \sum_{k=1}^{+\infty} g^{-n_k}
\]

with \( 1 \leq n_1 < n_2 < \ldots \) positive integers. (4) is evidently (1) after the zero terms having been dropped, and corresponds to the algorithm of [3].

My aim is to prove the following result.

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

\[
\varepsilon(s, M) = \sum_{k=0}^{M} g^{-nk} = 1.
\]

According to the remark in the previous paragraph, \( Z_M \) has Lebesgue measure 0. The assumption \( M, M' \geq a+1 \) is not a restriction, since the algorithm (3) implies that \( m_k = a+1 \) for all \( k \geq 2 \) (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 Luroth'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 \( |I| \leq c \) for all \( I \in D \) \( (|I| \) stands for the length of \( I \)). Let \( \mathcal{R}(U) \) be the set of all \( c \)-coverings of \( U \) and set

\[
\mathcal{H}(c, U, s) = \inf_{D \in \mathcal{R}(U)} \sum_{I \in D} |I|^s
\]

and

\[
\mathcal{H}(U, s) = \lim_{c \to 0} \mathcal{H}(c, U, s)
\]

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 \( 0 < s < s_0 \), \( \mathcal{H}(U, s) = +\infty \) and for \( s_0 < s < 1 \), \( \mathcal{H}(U, s) = 0 \). 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 \( U = \mathbb{Z}_M \), \( \mathcal{H}(M) = +\infty \) for all \( 0 < s < s_0 \), where \( s_0 \) is the root of the equation (5), i.e., \( \varepsilon(s, M) = 1 \), and for \( s_0 < s < 1 \), \( \mathcal{H}(M, s) = 0 \).

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 \( 1 \leq n_1 < n_2 < \ldots \) 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 \( j_k - j_{k-1} = a+1 \) for \( r \geq 2 \) and it is empty otherwise.

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

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

\((P \) 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 \( \varepsilon_0 \) be the solution in \( s \) of (8) and let \( s_0 < s < 1 \).

By the definition of \( Z_M \) it is evident that for any \( k \geq 1 \),

\[
Z_M \subset \bigcup_{n_1 < n_2, \ldots, n_k} \{m_1 = t_1, m_2 = t_2, \ldots, m_k = t_k\}
\]

Since

\[
\{m_1 = t_1, m_2 = t_2, \ldots, m_k = t_k\} = \{n_1 = t_1, n_2 = t_1 + t_2, \ldots, n_k = t_1 + \ldots + t_k\},
\]

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

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

which, by \( t_k \geq a+1 \), holds if \( k \) is chosen so that

\[
g^{-(a+1)k+1}(1-1/g) < c.
\]

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

\[
\mathcal{H}(c, Z_M, s) \leq \sum_{s_0 < s < 1} g^{-s(t_1 + \ldots + t_k)} \leq \left( \sum_{t=0}^{M} g^{-t} \right)^k = \varepsilon(s, M).
\]
By the definition of \( s_0 \), for all \( s_0 < s < 1 \), \( s(e, \mathcal{M}) < 1 \). Hence, since the inequality (11) holds for all \( h \) satisfying (3.0),

\[
h(e, \mathcal{M}, s) = 0 \quad \text{for} \quad s_0 < s < 1.
\]

(12)

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

Let \( 0 < s < s_0 \). By definition, it is sufficient to prove that \( h(Z_{\mathcal{M}}, s) > 0 \). To this end, we shall show in several steps a lower estimate for \( h(e, Z_{\mathcal{M}}, s) \) 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 \( Z_{\mathcal{M}} \) such that

\[
h(e, Z_{\mathcal{M}}, s) + \varepsilon \geq \sum_{I \in D} |I|^a.
\]

(13)

By the definition of \( Z_{\mathcal{M}} \),

\[
Z_{\mathcal{M}} = \bigcap_{h=1}^{\infty} \bigcup_{1 \leq i < k < b} \{e_1 = t_1, e_2 = t_2, \ldots, e_k = t_k\}.
\]

(14)

Define

\[
I(t_1, t_2, \ldots, t_k) = \{e_1 = t_1, e_2 = t_2, \ldots, e_k = t_k\}.
\]

(15)

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 occurring in (14). This results in adding a denumerably infinite set to \( Z_{\mathcal{M}} \), let this be denoted by \( Z_{\mathcal{M}}^* \). Evidently, the Hausdorff dimension of \( Z_{\mathcal{M}} \) and of \( Z_{\mathcal{M}}^* \) coincide, hence the covering \( D \) in (13) can be taken to cover \( Z_{\mathcal{M}}^* \). It is also known that the covering intervals can be taken to be open. Now, since an open system \( D \) covers a closed set \( Z_{\mathcal{M}}^* \), the Heine-Borel theorem, [8], p. 72, yields that a finite subset of \( D \) already covers \( Z_{\mathcal{M}}^* \). Evidently, it can be assumed that the endpoints of the intervals left from \( D \) to cover \( Z_{\mathcal{M}}^* \) are elements of \( Z_{\mathcal{M}}^* \). As a matter of fact, an interval \( I \) can be taken to be the closed interval \( [u, v] \) where \( u = \inf \{I \cap Z_{\mathcal{M}}^*\} \) and \( v = \sup \{I \cap Z_{\mathcal{M}}^*\} \). Let this new system of intervals be denoted by \( D_I \). The fundamental intervals are elements of \( Z_{\mathcal{M}}^* \). Since \( D_I \) 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

\[
h(e, Z_{\mathcal{M}}, s) + \varepsilon \geq \sum_{I \in D_I} |I|^a.
\]

(16)

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 \( D_I \) 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 \( I \in D_I \), 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 \( I(t_2, \ldots, t_k, d_2) \), 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 \( t_i, t_2, \ldots, t_k, d_1 \) 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 \( a+1 < y < M \). The length of \( I \) is estimated by \( L(t_1, \ldots, t_k, y) \), and any point of \( Z_{\mathcal{M}}^* \) belongs to a fundamental interval with its parameters being between \( a+1 \) and \( M \). This particular implies that the length of \( I \) is not be arbitrarily small in terms of \( L(t_1, \ldots, t_k, d_i) \), namely, the least value of \( |I| \) is obtained if \( d_1 = M \) and \( d_2 = M-1 \), and we necessarily \( I \) contains all fundamental intervals \( I(t_1, \ldots, t_k, M-1, y) \). These intervals are disjoint and their length, by Lemma 1, is

\[
\sum_{k=0}^{\infty} g^{-k-M+1} (1-1/g) = \{g^{-(M+1)}/(g-1)\} [I(t_1, \ldots, t_k)].
\]

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 \( I \in D_I \), there is a well defined fundamental interval containing \( I \) and the argument above yields that

\[
\sum_{I \in D_I} |I| \geq \{g^{-(M+1)}/(g-1)\} \sum_{I \in D_I^*} |I|^a.
\]

(17)

where \( D_I^* \) 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 \( D_I \) which are contained in another element of \( D_I \) and let the remaining set be denoted by \( D_I^* \). This results in an additional decrease in (17), hence (16) yields the inequality

\[
h(e, Z_{\mathcal{M}}, s) + \varepsilon \geq \{g^{-(M+1)}/(g-1)\} \sum_{I \in D_I^*} |I|^a.
\]

(18)

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

\[
h(e, Z_{\mathcal{M}}, s) + \varepsilon \geq g^{-(M+1)}/(g-1).
\]

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

$$h(Z^*_M, s) \geq g^{-\frac{(k+1)s}{2}}(g-1)^s.$$  

(20)

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

$$\sum_{y=a+1}^{M} |I(t_1, \ldots, t_k, y)|^s = |I(t_1, \ldots, t_k)|^s \sum_{u=a+1}^{M} g^{-su}$$

(21) and since for $0 < s < s_0$,

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

(21) yields that if we replace in $D_k$ all of its elements of order $k$ by the single element $I(t_1, \ldots, t_k)$, the right hand side of (18) is further decreased and by this step the maximal order of $D_k$ 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