Časopis pro pěstování matematiky, roč. 102 (1977), Praha

ON A CLASS OF ARITHMETICAL SETS

H. G. MEIJER, Delft and TIBOR ŠALÁT, Bratislava

(Received August 8, 1975)

Infinite subsets of the set N of all natural numbers will be called arithmetical sets. In the paper [1] P. Erdös studied the arithmetical sets $A = \{a_1 < a_2 < \ldots\}$ with the property (P): If $i_1 < i_2 < \ldots < i_s$ is an arbitrary finite sequence of indices, then $a_{i_1} + a_{i_2} + \ldots + a_{i_s}$ does not belong to the set A. Denote by T^* the system of all arithmetical sets having the property (P).

Let k be a natural number, $k \geqq 2$. Denote by T_k the system of all arithmetical sets $A = \{a_1 < a_2 < \ldots\}$ with the following property (P_k): If $i_1 < i_2 < \ldots < i_k$ is an arbitrary sequence of indices with k terms, then the number $a_{i_1} + a_{i_2} + \ldots$ $\ldots + a_{i_k}$ does not belong to the set A. Put $T_k^* = \bigcap_{j=2}^{k} T_j$ (for $k \geqq 2$) and $T = \bigcup_{j=2}^{\infty} T_j$. We have obviously

$$T^* = \bigcap_{k=2}^{\infty} T_k^* = \bigcap_{k=2}^{\infty} T_k \quad \text{and} \quad T_2^* \supset T_3^* \supset \ldots \supset T_k^* \supset T_{k+1}^* \supset \ldots .$$

It is clear that if $A \in T_k$ or $A \in T_k^*$ and B is an arithmetical set, $B \subset A$, then $B \in T_k$ and $B \in T_k^*$, respectively. Further, it is easy to check that

$$B_1 = \{1, 3, \ldots, 2k - 1, \ldots\} \in T_2 - T_3$$

and

$$B_2 = \{1, 2, 3, 10, 10^2, \ldots, 10^n, \ldots\} \in T_3 - T_2 .$$

Hence none of the inclusions $T_2 \subset T_3$, $T_3 \subset T_2$ is valid.

If

$$A \subset N = \{1, 2, \ldots\} ,$$

then we put

$$A(n) = \sum_{a \leqq n, a \in A} 1 , \quad \delta_1(A) = \liminf_{n \to \infty} \frac{A(n)}{n} , \quad \delta_2(A) = \limsup_{n \to \infty} \frac{A(n)}{n}$$

and $\delta(A) = \lim_{n \to \infty} (A(n)/n)$ (if the limit of the right-hand side exists). It is proved in [1] that the asymptotic density $\delta(A)$ of each arithmetical set A having the property (P) is zero.

With each set $A \subset N$ we can associate a real number $\varrho(A) = \sum_{j=1}^{\infty} \varepsilon_j 2^{-j}$, where $\varepsilon_j = 1$ if $j \in A$ and $\varepsilon_j = 0$ otherwise (see [2], p. 17). The number $\varrho(A)$ will be called the dyadic value of the set A. If S is a system of sets $A \subset N$, then $\varrho(S)$ denotes the set of all numbers $\varrho(A)$, $A \in S$. Obviously we have $\varrho(S) \subset \langle 0, 1\rangle$ and $\varrho(S)$ provides a tool for measuring the size of the system S.

The purpose of this paper is to illustrate from both the metric and the topological point of view the structure of the systems T, T^*, T_k, T_k^* in terms of the just defined dyadic values of sets $A \subset N$.

1. METRIC PROPERTIES OF SETS $\varrho(T)$, $\varrho(T^*)$, $\varrho(T_k)$, $\varrho(T_k^*)$

In the following, we denote by $|M|$ and $|M|^*$ the Lebesgue measure and the outer Lebesgue measure of the set M, respectively, and by dim M the Hausdorff dimension of the set $M \subset (-\infty, +\infty)$.

We mention the following simple fact which is well-known in the theory of dyadic expansions of real numbers: If m is any natural number then the interval $(0, 1\rangle$ is a union of pairwise disjoint intervals of the form

$$I = \left(\frac{s}{2^m}, \frac{s+1}{2^m}\right\rangle \quad (0 \leqq s \leqq 2^m - 1).$$

Each interval I is associated with a sequence $\varepsilon_1^0, \varepsilon_2^0, \ldots, \varepsilon_m^0$ of numbers 0 and 1 in such a way that for the dyadic expansion $x = \sum_{k=1}^{\infty} \varepsilon_k(x) 2^{-k}$ ($\varepsilon_k(x) = 0$ or 1 and for an infinite number of k's we have $\varepsilon_k(x) = 1$)) of any number x belonging to I the equalities $\varepsilon_k(x) = \varepsilon_k^0$ $(k = 1, 2, \ldots, m)$ hold.

In the following, the interval $(0, 1\rangle$ is regarded as a metric space with the Euclidean metric.

The proof of the main part of the following theorem is based on this lemma.

Lemma 1,1. *Let a be a fixed natural number. Put*

$$H(a) = \{x \in (0, 1\rangle ; \ \forall_{j>a} \varepsilon_j(x) \, \varepsilon_{j+a}(x) = 0\} .$$

Then $|H(a)| = 0$.

Proof. Let $t \geqq 1$ be an arbitrary natural number. The set $H(a)$ is contained in the union of all such intervals

$$\left(\frac{s}{2^{(2t+1)a}}, \frac{s+1}{2^{(2t+1)a}}\right\rangle \quad (0 \leqq s \leqq 2^{(2t+1)a} - 1)$$

which are associated with the sequences

(1) $$\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_{(2t+1)a}$$

of 0's and 1's having the following properties: each of the numbers $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_a$ is 0 or 1, and

(2) $$0 = \varepsilon_{a+1} \cdot \varepsilon_{2a+1} = \varepsilon_{a+2} \cdot \varepsilon_{2a+2} = \ldots = \varepsilon_{2a} \cdot \varepsilon_{3a} =$$
$$= \varepsilon_{3a+1} \cdot \varepsilon_{4a+1} = \ldots = \varepsilon_{4a} \cdot \varepsilon_{5a} = \ldots = \varepsilon_{(2t-1)a+1} \cdot \varepsilon_{2ta+1} = \ldots$$
$$\ldots = \varepsilon_{2ta} \cdot \varepsilon_{(2t+1)a} .$$

It is easy to check that the number of sequences (1) satisfying (2) is $2^a \cdot 3^{at}$. Therefore

$$|H(a)|^* \leqq \frac{2^a \cdot 3^{at}}{2^{(2t+1)a}} = \left(\frac{3^a}{4^a}\right)^t .$$

Hence we conclude $|H(a)| = 0$ since t is arbitrarily large.

Theorem 1,1. *Each of the sets $\varrho(T_k)\,(k = 2, 3, \ldots)$ is a G_δ-set (in $(0,1\rangle$) and $|\varrho(T_k)| = 0$ $(k = 2, 3, \ldots)$.*

Corollary. $|\varrho(T)| = |\varrho(T^*)| = 0$, $|\varrho(T_k^*)| = 0$ $(k = 2, 3, \ldots)$.

Proof. Let $k \geqq 2$. Denote by I_m the union of all intervals

$$\left(\frac{s}{2^m}, \frac{s+1}{2^m}\right\rangle \quad (0 \leqq s \leqq 2^m - 1),$$

which are associated with such sequences $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_m$ that if $1 = \varepsilon_{i_1} = \varepsilon_{i_2} = \ldots$ $\ldots = \varepsilon_{i_k}$, $i_1 < i_2 < \ldots < i_k$, $i_1 + i_2 + \ldots + i_k \leqq m$, then $\varepsilon_{i_1 + i_2 + \ldots + i_k} = 0$. We shall prove that

(3) $$\varrho(T_k) = \bigcap_{m=1}^{\infty} I_m .$$

If $x \in \varrho(T_k)$, then $x = \varrho(A)$, $A \in T_k$, $x = \sum_{j=1}^{\infty} \varepsilon_j(x)\, 2^{-j}$ (the dyadic expansion of x). It follows from the definition of the system T_k that $\varepsilon_{i_1 + i_2 + \ldots + i_k}(x) = 0$, $i_1 < i_2 < \ldots$ $\ldots < i_k$ if $\varepsilon_{i_l}(x) = 1$ $(l = 1, 2, \ldots, k)$. Therefore $x \in I_m$ for each $m = 1, 2, \ldots$.

Let $x \in (0, 1\rangle$, $x = \sum_{j=1}^{\infty} \varepsilon_j(x)\, 2^{-j}$, $x \notin \varrho(T_k)$. Denote U the system of all arithmetical sets. Then $\varrho(U) = (0, 1\rangle$ and $\varrho : U \to (0, 1\rangle$ is a one-to-one mapping (cf. [2], p. 18). Hence $(0, 1\rangle = \varrho(T_k) \cup \varrho(U - T_k)$, the sets on the right-hand side being disjoint. Hence $x \in \varrho(U - T_k)$, $x = \varrho(A)$, $A \in U - T_k$. Since $A \notin T_k$, there exists such a sequence $i_1 < i_2 < \ldots < i_k$ of natural numbers that $\varepsilon_{i_l}(x) = 1$ $(l = 1, 2, \ldots, k)$ and $\varepsilon_{i_1 + i_2 + \ldots + i_k}(x) = 1$. Hence $x \notin I_p$, where $p = i_1 + i_2 + \ldots + i_k$, therefore $x \notin \bigcap_{m=1}^{\infty} I_m$. The equality (3) is proved.

From (3) it follows immediately that $\varrho(T_k)$ is a G_δ-set in $(0, 1\rangle$.

Let $k \geqq 2$, let

$$a_1^0 < a_2^0 < \ldots < a_{k-1}^0 \tag{4}$$

be a sequence of natural numbers. Denote by $T_k(a_1^0, \ldots, a_{k-1}^0)$ the system of all sets $A \in T_k$ of the form

$$A = \{a_1^0 < a_2^0 < \ldots < a_{k-1}^0 < a_k < a_{k+1} < \ldots\}.$$

Then

$$\varrho(T_k) = \bigcap \varrho(T_k(a_1^0, \ldots, a_{k-1}^0)),$$

the union on the right-hand side being taken over all finite sequences of the form (4). Hence it suffices to prove that

$$|\varrho(T_k(a_1^0, \ldots, a_{k-1}^0))| = 0 \tag{5}$$

for each sequence (4).

In the notation used in Lemma 1,1 we have obviously

$$\varrho(T_k(a_1^0, \ldots, a_{k-1}^0)) \subset H(a),$$

where $a = a_1^0 + \ldots + a_{k-1}^0$. Hence (5) follows from Lemma 1,1. The proof of Theorem 1,1 is complete.

The proof of the following lemma is based on an idea from [1]. The lemma will be useful in the proof of Theorem 1,2.

Lemma 1,2. *If $A \in T_m^*$ $(m \geqq 2)$, then $\delta_2(A) \leqq 1/m$. Moreover, there is an $A \in T_m^*$ such that $\delta(A) = 1/m$.*

Proof. Let $A = \{a_1 < a_2 < \ldots\} \in T_m^*$. Since $A \in T_m^*$ $(m \geqq 2)$, the elements of the sets $P_1, P_2, \ldots, P_m$ do not belong to the set A, where

$$\begin{aligned}
P_1 &= \{a_1 + a_2,\ a_1 + a_3,\ \ldots,\ a_1 + a_j,\ \ldots\},\\
P_2 &= \{(a_1 + a_2) + a_3,\ (a_1 + a_2) + a_4,\ \ldots,\ (a_1 + a_2) + a_j,\ \ldots\},\\
&\vdots\\
P_m &= \{(a_1 + \ldots + a_m) + a_{m+1},\ (a_1 + \ldots + a_m) + a_{m+2},\ \ldots\\
&\quad \ldots,\ (a_1 + \ldots + a_m) + a_{m+j},\ \ldots\}.
\end{aligned}$$

The sets $P_1, P_2, \ldots, P_m$ are pairwise disjoint. Indeed, if $P_i \cap P_l \neq \emptyset$ for $i \neq l$, $i, l \leqq m$, then there exist such numbers s, d, $s \geqq i + 1$, $d \geqq l + 1$ that

$$(a_1 + \ldots + a_i) + a_s = (a_1 + \ldots + a_l) + a_d.$$

Let $i < l$. Then

$$a_s = a_{i+1} + \ldots + a_l + a_d \tag{6}$$

and the number of summands on the right-hand side of (6) is equal to $l - i + 1 \leqq m$. Hence (6) contradicts the assumption $A \in T_m^*$.

Let $n > a_1 + \ldots + a_m + m$. The number of elements of the set P_1 lying in the interval $\langle 1, n \rangle$ is obviously equal to $A(n - a_1) - 1$, similarly the number of elements of the set P_2 lying in that interval is equal to $A(n - (a_1 + a_2)) - 2$, etc. Since the sets P_j ($j = 1, 2, \ldots, m$) are pairwise disjoint, we obtain

$$(7) \qquad (A(n - a_1) - 1) + (A(n - (a_1 + a_2)) - 2) + \ldots + (A(n - (a_1 + \ldots + a_m)) - m) \leqq n .$$

A simple estimation yields

$$(8) \qquad \begin{aligned} A(n - a_1) &\geqq A(n) - a_1 , \\ A(n - (a_1 + a_2)) &\geqq A(n) - (a_1 + a_2) , \\ &\vdots \\ A(n - (a_1 + \ldots + a_m)) &\geqq A(n) - (a_1 + \ldots + a_m) . \end{aligned}$$

From (7), (8) we get

$$\frac{A(n)}{n} \leqq \frac{b_m + c_m}{nm} + \frac{1}{m},$$

where

$$b_m = \frac{m(m + 1)}{2}, \quad c_m = a_1 + (a_1 + a_2) + \ldots + (a_1 + \ldots + a_m) .$$

The inequality $\delta_2(A) \leqq 1/m$ follows now immediately.

Further, the set $A = \{1, m + 1, 2m + 1, \ldots, jm + 1, \ldots\}$ belongs to the system T_m^* and $\delta(A) = 1/m$. The proof is complete.

Since $|\varrho(T^*)| = 0$, $|\varrho(T_k^*)| = 0$ $(k = 2, 3, \ldots)$ the question of the Hausdorff dimension of the sets $\varrho(T^*)$, $\varrho(T_k^*)$ $(k \geqq 2)$ arises. In what follows we give upper and lower estimates for $\dim \varrho(T_k^*)$ and the precise value of $\dim \varrho(T^*)$.

Denote by d the function defined on the interval $\langle 0, 1 \rangle$ in the following way: $d(0) = d(1) = 0$ and

$$d(\zeta) = \frac{\zeta \log \zeta + (1 - \zeta) \log (1 - \zeta)}{\log \frac{1}{2}}$$

for $\zeta \in (0, 1)$.

It is easy to see that

$$(9) \qquad \lim_{\zeta \to 0+} d(\zeta) = 0 .$$

Theorem 1,2. (i) For each $k \geqq 2$, the inequality $\dim \varrho(T_k^*) \geqq 1/k$ holds.

(ii) For each $k \geqq 2$, the inequality $\dim \varrho(T_k^*) \leqq d(1/k)$ holds.

(iii) $\dim \varrho(T^*) = 0$.

Remark. The estimate for $k = 2$ in (ii) is trivial since $d(\frac{1}{2}) = 1$.

Proof. (i) Put (for $k \geqq 2$)

$$C_k = \{1,\ k + 1,\ 2 . k + 1,\ \ldots,\ lk + 1,\ \ldots\} .$$

Evidently $C_k \in T_k^*$. Denote by S_k the system of all arithmetical sets which are subsets of the set C_k. Then $S_k \subset T_k^*$ and so

$$\varrho(S_k) \subset \varrho(T_k^*)\,. \tag{10}$$

Denote by 2^{C_k} the system of all subsets of the set C_k. Then it is easy to see that the set $\varrho(2^{C_k}) - \varrho(S_k)$ is countable, hence

$$\dim \varrho(2^{C_k}) = \dim \varrho(S_k)\,. \tag{11}$$

But $\varrho(2^{C_k})$ is equal to the set of all such real numbers $x = \sum_{j=1}^{\infty} \varepsilon_j \,.\, 2^{-j}$ that $\varepsilon_j = 0$ for $j \neq lk + 1$ $(l = 0, 1, \ldots)$ and $\varepsilon_{lk+1} = 0$ or 1 $(l = 0, 1, \ldots)$.

The Hausdorff dimension of the set $\varrho(2^{C_k})$ can be established by virtue of Theorem 2,7 from [3]. The following special result is a consequence of this theorem:

Let P be a set of natural numbers, let $\{\varepsilon_j^0\}$, $j \in P$ be a sequence of numbers 0 and 1. Denote by

$$Z = Z(P; \{\varepsilon_j^0\}, j \in P)$$

the set of all such $x = \sum_{j=1}^{\infty} \varepsilon_j \,.\, 2^{-j}$ that $\varepsilon_j = \varepsilon_j^0$ for $j \in P$ and $\varepsilon_j = 0$ or 1 for $j \in N - P$. Then

$$\dim Z = \liminf_{n\to\infty} \frac{\log \prod_{j \leqq n, j \in N-P} 2}{n \log 2}\,.$$

Put $P = N - C_k$, $\varepsilon_j^0 = 0$ for $j \in P$. Then we get

$$\dim \varrho(2^{C_k}) = \liminf_{n\to\infty} \frac{\log \prod_{j \leqq n, j \in C_k} 2}{n \log 2} = \tag{12}$$

$$= \liminf_{n\to\infty} \frac{\log 2^{[(n-1)/k]}}{n \log 2} = \liminf_{n\to\infty} \frac{[(n-1)/k]}{n} = \frac{1}{k}\,.$$

($[u]$ denotes the integer which satisfies $[u] \leqq u < [u] + 1$). From (10), (11) and (12) we obtain $\dim \varrho(T_k^*) \geqq 1/k$.

(ii) Denote by Z_k the system of all arithmetical sets A with $\delta_2(A) \leqq 1/k$. Then on account of Lemma 1,2 we have $T_k^* \subset Z_k$. It is well-known that $\dim \varrho(Z_k) = d(1/k)$ (cf. [2], p. 195 or [5], Theorem 51). From these facts we get $\dim \varrho(T_k^*) \leqq d(1/k)$.

(iii) We shall give two proofs for (iii).

Proof I. Since $T^* \subset T_k^*$ $(k = 2, 3, \ldots)$ according to (ii) we have

$$\dim \varrho(T^*) \leqq d\left(\frac{1}{k}\right) \quad (k = 2, 3, \ldots)$$

and so (see (8))

$$\dim \varrho(T^*) \leqq \lim_{k\to\infty} d\left(\frac{1}{k}\right) = 0\,.$$

Proof II. Denote by W_0 the system of all arithmetical sets A with $\delta_1(A) = 0$. Then

$$\dim \varrho(W_0) = 0 \tag{13}$$

(see [2], p. 195). We have mentioned already that if $A \in T^*$, then $\delta(A) = 0$ (cf. [1]). Hence

$$T^* \subset W_0 . \tag{14}$$

From (13), (14) we get $\dim \varrho(T^*) = 0$. The proof is complete.

2. TOPOLOGICAL PROPERTIES OF SETS $\varrho(T_k)$, $\varrho(T_k^*)$, $\varrho(T)$, $\varrho(T^*)$

In this part of the paper we shall complete the first part by proving some further properties of the sets $\varrho(T_k)$, $\varrho(T_k^*)$, $\varrho(T)$, $\varrho(T^*)$. These sets are viewed as subsets of the metric space $(0, 1\rangle$ with the usual Euclidean metric.

It was already proved in the first part of the paper that the sets $\varrho(T_k)$ $(k \geqq 2)$ are G_δ-sets. This fact implies easily

Theorem 2,1. *The sets $\varrho(T^*)$, $\varrho(T_k^*)$ $(k \geqq 2)$ are G_δ-sets, $\varrho(T)$ is a $G_{\delta\sigma}$-set in $(0, 1\rangle$.*

Proof. Theorem 2,1 follows at once from Theorem 1,1 and from the equalities

$$\varrho(T_k^*) = \bigcap_{j=2}^{k} \varrho(T_j)\,, \quad \varrho(T^*) = \bigcap_{k=2}^{\infty} \varrho(T_k^*)\,, \quad \varrho(T) = \bigcup_{k=2}^{\infty} \varrho(T_k)\,. \tag{15}$$

Finally, we shall show that the sets studied in this part of the paper are poor from the topological point of view.

Theorem 2,2. (i) *The sets $\varrho(T^*)$, $\varrho(T_k^*)$, $\varrho(T_k)$ $(k \geqq 2)$ are nowhere-dense sets in $(0, 1\rangle$.*

(ii) *The set $\varrho(T)$ is a set of the first Baire category in $(0, 1\rangle$.*

Proof. Part (ii) follows from (i) in virtue of (15). Further,

$$\varrho(T^*) \subset \varrho(T_k^*) \subset \varrho(T_k) \quad (k = 2, 3, \ldots)\,,$$

hence it suffices to prove that $\varrho(T_k)$ $(k \geqq 2)$ is a nowhere-dense set in $(0, 1\rangle$.

Let $k \geqq 2$. On account of the well-known criterion of the nowhere-density of sets in metric spaces it is sufficient to prove that each open interval $I \subset (0, 1\rangle$ contains an interval J which is disjoint with the set $\varrho(T_k)$ (cf. [4], p. 74).

Let $I \subset (0, 1\rangle$ be an open interval. Choose a natural number m such that for a suitable s, $0 \leqq s \leqq 2^m - 1$, we have

$$I_1 = \left(\frac{s}{2^m}, \frac{s+1}{2^m}\right) \subset I\,.$$