Publ. Math. Debrecen
50 / 1-2 (1997), 107–113

An antidemocratic representation

By IMRE Z. RUZSA (Budapest)

1. Introduction

Let A be a set of positive integers and write

$$r(n) = \#\{(a, a') : a, a' \in A,\ a + a' = n\},$$
$$B = A + A = \{n : r(n) \geq 1\}.$$

P. Erdős and A. Ivič asked whether it is possible to have $d(B) = 0$ and at the same time $r(n) \to \infty$ when n is restricted to $n \in B$. We construct a class of such sets.

We write

$$A(x) = \#\big(A \cap [1, x]\big)$$

and similarly for B.

Theorem 1. *There is a set A such that $B(x) \ll x^{1-\alpha}$ while $r(n) \gg n^{\beta}$ for all $n \in B$ with certain absolute constants α, β.*

Obviously $r(n) \leq A(n) \leq B(n) + O(1)$ for all n, thus we have necessarily $\alpha + \beta \leq 1$. I cannot decide whether β can be arbitrarily near to 1.

Theorem 2. *If $\omega(x) \to \infty$ arbitrarily slowly, then there is a set A such that*

$$A(x) \ll B(x) \ll (\log x)^{\omega(x)}$$

while $r(n) \to \infty$ for $n \in B$.

I cannot decide whether $A(x) \ll (\log x)^{O(1)}$ is possible for such a set.

Supported by Hungarian National Foundation for Scientific Research, Grant No. T 017433

The construction is described in the next section. The proof that these sequences have the desired properties is given in Section 3.

2. The construction

For a positive integer n let $s(n)$ denote the number of 1's in the binary representation of n.

Let $t : \mathbb{N} \to \mathbb{N} \cup \{0\}$ be a nondecreasing function with the properties

$$t(k) \le \frac{3}{8}k, \quad t(k+1) \le 1 + t(k).$$

We consider, for each such function t, the set

$$A = A_t = \{n : 2^k \le n < 2^{k+1} \text{ and } s(n) \le t(k) \text{ for some } k\}.$$

The sets in Theorems 1 and 2 will be particular cases of this construction for suitable choices of the function t. In this section we establish a general property of these sequences.

Take an integer n, and define k by

$$2^k \le n < 2^{k+1}. \tag{2.1}$$

Develop n in base 2. Either $n = 2^k$, or we can write this development as

$$n = 2^k + 2^l + \sum_{i=0}^{l-1} \varepsilon_i 2^i \quad (k > l). \tag{2.2}$$

Throughout this section k and l will denote the numbers determined by (2.1) and (2.2). Now put

$$T(n) = \max\bigl(t(k) + t(l),\, 2t(k-1) - 1\bigr).$$

For $n = 2^k$, we define $T(n) = 2t(k-1) - 1$.

Theorem 3. *If $n \in B$, then $s(n) \le T(n)$, and*

$$r(n) \ge c \frac{2^{(2/3)t(k)}}{\sqrt{t(k)}} \tag{2.3}$$

with a positive absolute constant c and k defined by (2.1).

We shall need the following property of the function s.

Lemma 2.1. *For arbitrary positive integers m, n we have*

$$s(m+n) \le s(m) + s(n),$$

with equality if there is no i such that the i'th digits in the binary representations of m and n are both equal to 1 and strict inequality otherwise.

We leave the simple proof to the reader.

Lemma 2.2. *If $n \in B$, then*

$$s(n) \le T(n).$$

PROOF. We use representation (2.2). Assume $n = a + a'$, $a, a' \in A$, $a \ge a'$. If $a \ge 2^k$, then $a' < 2^{l+1}$, hence

$$s(a) \le t(k), \quad s(a') \le t(l),$$

and so

$$s(n) \le s(a) + s(a') \le t(k) + t(l) \le T(n).$$

If $a < 2^k$, then

$$s(a) \le t(k-1), \quad s(a') \le t(k-1),$$

and there must be a common 1 in the representations of a and a', since otherwise we could add them without carry and $n = a + a'$ would not contain 2^k. Consequently we have

$$s(n) \le s(a) + s(a') - 1 \le 2t(k-1) - 1 \le T(n).$$

If $n = 2^k$, then only the second case is possible and we get the result by the second argument. □

From now on we fix the following notation: $s(n) = s$ and

$$n = \sum_{i=1}^{s} 2^{u_i}, \quad u_1 > u_2 > \cdots > u_s, \quad u_1 = k, \ u_2 = l. \tag{2.4}$$

Lemma 2.3. *If $4 \le s \le T(n)$, then*

$$r(n) \ge \binom{s-2}{[s/2]-2} \ge \frac{c2^s}{\sqrt{s}}. \tag{2.5}$$

PROOF. *First case:* $t(k) + t(l) \ge 2t(k-1)$.
Here

$$t(l) \ge 2t(k-1) - t(k) \ge t(k) - 2$$

and so

$$s \le t(k) + t(l) \le 2t(l) + 2.$$

Since $s \le t(k) + t(l)$, we can find integers q and r such that $s = q + r$, $q \le t(k)$, $r \le t(l)$. Moreover, we can achieve

$$\left[\frac{s}{2} - 1\right] \le r \le \left[\frac{s}{2}\right].$$

Indeed, if $s \ge 2t(l)$, then $r = t(l)$, $q = s - t(l)$ is such a choice, while if $s < 2t(l)$, then $r = [s/2]$, $q = s - [s/2]$ works.

Now for each decomposition

$$\{3, 4, \dots, s\} = X \cup Y$$

into disjoint sets X, Y such that $|X| = q - 1$, $|Y| = r - 1$ we can set

$$a = 2^k + \sum_{i \in X} 2^{u_i}, \quad a' = 2^l + \sum_{i \in Y} 2^{u_i}.$$

This will be a valid decomposition of n, since $s(a) = q \le t(k)$, $s(a') = r \le t(l)$. The number of such decompositions is

$$\binom{s-2}{r-1} \ge \binom{s-2}{[s/2]-2}.$$

Second case: $t(k) + t(l) \le 2t(k-1) - 1$. Then also $s \le 2t(k-1) - 1$. If we put now

$$r = \left[\frac{s-1}{2}\right], \quad q = \left[\frac{s}{2}\right],$$

then we have $r \le q \le t(k-1) - 1$ and $r + q = s - 1$.

Again we write n in the form (2.4). If we put

$$\{2, 3, \dots, s\} = X \cup Y$$

with disjoint sets X, Y such that $|X| = q$, $|Y| = r$, then the numbers

$$a = 2^{k-1} + \sum_{i \in X} 2^{u_i}, \quad a' = 2^{k-1} + \sum_{i \in Y} 2^{u_i}$$

satisfy $s(a) \le q + 1 \le t(k-1)$, $s(a') \le r + 1 \le t(k-1)$ (if $u_2 = k - 1$, then a carry may occur so strict inequality is possible), hence $a, a' \in A$ and $a + a' = n$. The number of such decompositions is

$$\binom{s-1}{r} = \binom{s-1}{[(s-1)/2]} > \binom{s-2}{[s/2]-2}. \qquad \square$$

Lemma 2.3 provides many representations for an $n \in B$ if s is not too small. For small values of s we use a different method.

Consider the representations of n in the form

$$n = \sum_{i=0}^{k} \delta_i 2^i, \tag{2.6}$$

with $\delta_i = 0, 1$ or 2.

Lemma 2.4. *Suppose $3s < k$. For every integer*

$$0 \le p \le \frac{k-3s}{4} \tag{2.7}$$

there is a representation of n in the form (2.6) *such that the number of $\delta_i = 1$ is $s + p$ and the number of $\delta_i = 2$ is at most p. Moreover, if k' is the largest subscript such that $\delta_{k'} \ne 0$, then $k' = k$ or $k - 1$, $\delta_{k'} = 1$ and $\delta_{k'-1} = 0$ or 1.*

Proof. We construct these representations recursively. The usual binary representation provides such a representation for $p = 0$.

Now we make changes to a given representation, in each step increasing the number of ones by one and the number of twos by at most one.

To achieve this, we do one of the following two operations.

If there is a substring 20 in the sequence $\delta_k, \dots, \delta_0$, we change it into 12. This increases the number of ones by 1 and does not change the number of twos.

If there is a 1000, we change it into 0112. This increases both the number of ones and the number of twos by 1.

The property that the first two nonzero δ's are 10 or 11 holds in the starting representation and it is preserved by these operations.

We repeat these steps as long as we can. Suppose that after p steps we stop. We have now $s+p$ ones, $\le p$ twos, and blocks of 0's between them (and possibly a block of zeros before them). Since none of the described operations is possible, a block of 0's (not in the initial position) can only follow a 1 and its length is at most 2, hence these give at most $2(s+p)$ zeros. Thus altogether we have at most

$$(s+p) + p + 2(s+p) = 3s + 4p$$

digits from $\delta_{k'}$ on. This means that

$$3s + 4p \ge k' + 1.$$

Since

$$n \geq 2^k > \sum_{i=0}^{k-2} 2 \cdot 2^i = 2^k - 2,$$

we have necessarily $k' \geq k-1$ and we obtain

$$3s + 4p \geq k,$$

hence the range in (2.7) is indeed covered. □

Proof of Theorem 3. From Lemma 2.2 we know that $s(n) \leq T(n)$ is necessary for $n \in B$. Now we prove the lower estimation.

Write $t = t(k)$. We may suppose that $t \geq 9$, otherwise the theorem is obvious. If $s > (2/3)t - 2$, we apply Lemma 2.3. Assume now $s \leq (2/3)t - 2$. We use Lemma 2.4 with

$$p = \left[\frac{2}{3}t\right] - s - 2.$$

The condition $p \leq (k-3s)/4$ follows from the assumption $t \leq (3/8)k$.

Take representation (2.6), where there are $s + p = [(2/3)t] - 2$ ones and at most p twos. Let k' be the largest subscript with $\delta_i \neq 0$; we know that $k' = k$ or $k-1$. Let U be the set of those i's that satisfy $i \leq k'-2$ and $\delta_i = 1$; we have $|U| \leq s + p - 1$ and

$$|U| \geq s + p - 2 = \left[\frac{2}{3}t\right] - 4.$$

Now let $U = X \cup Y$ be a decomposition of U into sets of size

$$|X| = \left[\frac{|U|}{2}\right], \quad |Y| = \left[\frac{|U|+1}{2}\right].$$

The number of such decompositions is

$$\binom{|U|}{[|U|/2]} \gg \frac{2^{|U|}}{\sqrt{|U|}} \gg \frac{2^{(2/3)t}}{\sqrt{t}}.$$

Each of these decompositions induces a representation $n = a + a'$, namely

$$a = 2^m + \sum_{i \in X} 2^i + \sum_{\delta_i = 2} 2^i,$$
$$a' = 2^{m'} + \sum_{i \in Y} 2^i + \sum_{\delta_i = 2} 2^i,$$

where m and m' are defined as follows. If $\delta_{k'-1} = 0$, then $m = m' = k' - 1$; if $\delta_{k'-1} = 1$, then $m = k'$, $m' = k' - 1$. We have to show that $a, a' \in A$. To this end observe that

$$s(a) \leq s(a') \leq 1 + p + |Y| \leq 1 + p + \frac{s+p}{2} \leq t - s - 1.$$

Moreover $a, a' \geq 2^{k'-1} \geq 2^{k-2}$ and $t(k-2) \geq t - 2 \geq t - s - 1$ which shows that $a, a' \in A$. □

3. Proof of Theorems 1 and 2

Proof of Theorem 1. We take a $0 < \gamma < 1/4$ and put $t(k) = [\gamma k]$. Then for $n \leq x$ and $n \in B$ we have $s(n) \leq 2\gamma k$ with $k = [\log_2 x]$. Thus we have

$$B(x) \leq \sum_{i \leq 2\gamma k} \binom{k+1}{i} \ll 2^{(1-\alpha)k} \ll x^{1-\alpha}$$

for any

$$\alpha < 1 + 2\gamma \log_2 2\gamma + (1 - 2\gamma) \log_2 (1 - 2\gamma)$$

(this follows by Stirling's formula and a routine calculation). Moreover for an $n \in B$ we know

$$r(n) \gg 2^{(2\gamma/3)k} / \sqrt{k} \gg n^{\beta}$$

with any $\beta < 2\gamma/3$. □

With this method we can achieve $\beta = 1/6 - \varepsilon$. If γ passes $1/4$, then the "typical" numbers with approximately $k/2$ digits will be in B, thus we have $d(B) = 1$.

Proof of Theorem 2. If $t(k) \to \infty$, then $r(n) \to \infty$ for $n \in B$ by Theorem 3. On the other hand, if $n \leq x$ and $n \in B$, then n is the sum of at most $2t(k)$ powers of 2 with $k = [\log_2 x]$. Hence

$$B(x) \leq \sum_{i=0}^{2t(k)} (k+1)^i \ll t(k)(k+1)^{2t(k)},$$

and this will be $O\big((\log x)^{\omega(x)}\big)$ if $t(k)$ tends to infinity sufficiently slowly. □

IMRE Z. RUZSA
MATHEMATICAL INSTITUTE OF THE HUNGARIAN ACADEMY OF SCIENCES
BUDAPEST, P.O.BOX 127.
H–1364 HUNGARY
E-mail: ruzsa@math-inst.hu

(Received February 9, 1996)