ACTA ARITHMETICA
XCIII.2 (2000)

On sets of natural numbers without solution to a noninvariant linear equation

by

Tomasz Schoen (Kiel*)

Let us consider a linear equation

$$(*) \qquad a_1x_1 + \ldots + a_kx_k = b,$$

where $a_1, \ldots, a_k, b \in \mathbb{Z}$. We call the equation $(*)$ *invariant* if both $s = a_1 + \ldots + a_k = 0$ and $b = 0$, and *noninvariant* otherwise. We say that a set A is $(*)$-*free* if it contains no nontrivial solution to $(*)$ and define $r(n)$ as the size of the largest $(*)$-free set contained in $[n] = \{1, \ldots, n\}$.

The behavior of $r(n)$ has been extensively studied for many cases of invariant linear equations. The two best known examples are the equation $x + y = 2z$, when $r(n)$ is the size of the largest set without arithmetic progression of length three contained in $[n]$ (see [6]), and the equation $x_1 + x_2 = y_1 + y_2$, when $r(n)$ becomes the size of the largest Sidon subset of $[n]$ (see [3], [7], [8]).

Much less is known about the behavior of $r(n)$ for noninvariant linear equations, maybe apart from sum-free sets (see for example [1], [2], [5], [10]). The main contribution to this subject was made by Ruzsa [9] who studied properties of sets without solutions to a fixed noninvariant linear equation. Following his paper let us define

$$\begin{aligned}
\overline{\Lambda}(*) &= \sup\{\overline{\mathrm{d}}(A) : A \subseteq \mathbb{N},\ A \text{ is } (*)\text{-free}\},\\
\underline{\Lambda}(*) &= \sup\{\underline{\mathrm{d}}(A) : A \subseteq \mathbb{N},\ A \text{ is } (*)\text{-free}\},\\
\overline{\lambda}(*) &= \limsup_{n\to\infty} r(n)/n,\\
\underline{\lambda}(*) &= \liminf_{n\to\infty} r(n)/n,
\end{aligned}$$

where $\overline{\mathrm{d}}(A), \underline{\mathrm{d}}(A)$ denote the upper and lower density of the set A. Sometimes, we write just $\overline{\Lambda}, \underline{\Lambda}, \overline{\lambda}, \underline{\lambda}$ instead of $\overline{\Lambda}(*), \underline{\Lambda}(*), \overline{\lambda}(*), \underline{\lambda}(*)$.

2000 *Mathematics Subject Classification*: 11B75, 11A99.

*On leave from Adam Mickiewicz University, Poznań, Poland.

The aim of this paper is to answer the following questions posed in Ruzsa's paper [9].

1. Does there exist an absolute constant C such that for every noninvariant linear equation we have

$$C\overline{\Lambda} \geq \underline{\lambda}?$$

2. Let $\varepsilon > 0$ be an arbitrary number. Is it possible to find a noninvariant equation with $s \neq 0$ and $\overline{\lambda} < \varepsilon$?

3. Is it true that for every noninvariant linear equation we have

$$\Lambda = \overline{\Lambda} = \underline{\Lambda}?$$

4. For an integer $m > 1$, let $\varrho(m)$ denote the maximal cardinality of a $(*)$-free set $A \subseteq \mathbb{Z}_m$. Put

$$\varrho = \sup \varrho(m)/m.$$

Is it true that always

$$\overline{\lambda} = \underline{\lambda} = \max\left(\varrho, \frac{s^+ - s^-}{s^+}\right),$$

where $s^+ = \sum_{a_i>0} a_i$, $s^- = \sum_{a_i<0} a_i$ (we may assume that $s^+ > 0$ and $s^+ \geq s^-$)?

Notation. In this note $[n] = \{1, \ldots, n\}$ and $[u, w] = \{u \leq n \leq w : n \in \mathbb{N}\}$. We also set $Ak = \{ak : a \in A\}$ and $hA = \{a_1 + \ldots + a_h : a_1, \ldots, a_h \in A\}$. We use $\gcd\{A\}$ to denote the greatest common divisor of the elements of the set A, and set $s \pm A = \{s \pm a : a \in A\}$. Finally, $A(n)$ denotes the counting function of A, i.e. $A(n) = |A \cap [n]|$.

In order to deal with the first question we use the following result of Łuczak and Schoen [5].

Theorem A. *If $A \subseteq \mathbb{N}$ and there is no solution to the equation $y = x_1 + \ldots + x_k$, then*

$$\overline{\mathrm{d}}(A) \leq 1/\rho(k-1),$$

where $\rho(k) = \min\{m \in \mathbb{N} : m$ does not divide $k\}$. ∎

Now we can answer the first from Ruzsa's questions in the negative.

Theorem 1. *There is no an absolute constant C such that*

$$C\overline{\Lambda} \geq \underline{\lambda}$$

for every linear equation. Moreover, for every $\varepsilon > 0$ there is an equation such that $\overline{\Lambda} < \varepsilon$ and $\underline{\lambda} > 1 - \varepsilon$.

Proof. It is enough to prove that there exists a sequence of equations $(e_1), (e_2), \ldots$ such that

$$\underline{\lambda}(e_n) \to 1 \quad \text{and} \quad \overline{\Lambda}(e_n) \to 0 \quad \text{as } n \to \infty.$$

For a natural number n set $k_n = n!+1$. Then, for every n, we have $\rho(k_n) > n$. Furthermore, denote by (e_n) the equation

$$y = x_1 + \ldots + x_{k_n}.$$

Thus, it follows from Theorem A that for every $n \in \mathbb{N}$,

$$\overline{\Lambda}(e_n) \le 1/\rho(k_n) < 1/n,$$

and so $\overline{\Lambda}(e_n) \to 0$ as $n \to \infty$.

On the other hand, for every $m \in \mathbb{N}$ the set $\{\lceil m/k_n \rceil + 1, \ldots, m\}$ contains no solutions to the equation (e_n), so

$$\underline{\lambda}(e_n) \ge (k_n - 1)/k_n.$$

Consequently, $\underline{\lambda}(e_n) \to 1$ as $n \to \infty$, which completes the proof of Theorem 1. ∎

In order to solve the second problem we make use of the following theorem of Lev [4].

THEOREM B. *Assume that* $A \subseteq [n]$ *and*

$$|A| \ge \frac{n-1}{k} + 2.$$

Then there are integers $d \le k-1$, $h \le 2k-1$ *and* m *such that*

$$\{md, (m+1)d, \ldots, (m+n-1)d\} \subseteq hA.$$

Furthermore, $d = \gcd\{A - \min A\}$ *and* h *can be chosen to be the largest multiple of* d *less than or equal to* $2k-1$. ∎

Ruzsa [9] showed that $\overline{\lambda}$ may not be bounded from below by a positive absolute constant. For every $\varepsilon > 0$ he gave an example of a noninvariant linear equation with $s = 0$ and $\overline{\lambda} < \varepsilon$ and asked: Is it possible that $s \ne 0$? We prove a more general result, which for a suitable choice of k and l provides an example of a noninvariant equation with $s \ne 0$ and arbitrarily small $\overline{\lambda}$.

THEOREM 2. *Suppose that* $k, l \in \mathbb{N}$ *and* $k > l$. *If* $A \subseteq \{1, \ldots, n\}$ *contains no solution to the equation* $x_1 + \ldots + x_k = y_1 + \ldots + y_l$, *then*

$$|A| \le \max\left(\frac{2(k-l)n}{l}, \left\lceil \frac{n}{\rho(k-l)} \right\rceil\right).$$

Proof. Suppose that the assertion does not hold, so in particular $|A| > 2(k-l)n/l$. Obviously, we can assume $2(k-l)/l < 1$. Thus, it follows from Theorem B that there exists $a \in \mathbb{N}$ such that

$$\{a, a+d, \ldots, a+(n-1)d\} \subseteq \lfloor l/(k-l) \rfloor A,$$

where $d = \gcd\{A - \min A\}$. Furthermore, for some $b \in \mathbb{N}$ we have

$$\{b, b+d, \ldots, b+(k-l)(n-1)d\} \subseteq lA.$$

Note that, since $|A| > \lceil n/\rho(k-l)\rceil$, we must have $d < \rho(k-l)$, and so $k-l \equiv 0 \pmod d$ by the definition of the function ρ.

Let $x \in A$ be an arbitrary number with $x < n$. Then

$$x(k-l) \le (k-l)(n-1)d.$$

Hence

$$b + x(k-l) \in \{b, b+d, \ldots, b+(k-l)(n-1)d\} \subseteq lA.$$

Thus, there exist $x_1, \ldots, x_l$, $y_1, \ldots, y_l \in A$ such that

$$b = x_1 + \ldots + x_l \quad \text{and} \quad b + x(k-l) = y_1 + \ldots + y_l.$$

Hence, we arrive at

$$x_1 + \ldots + x_l + x(k-l) = y_1 + \ldots + y_l,$$

which is a contradiction. ■

For any fixed $t \in \mathbb{N}$, set $k = (2t+3)t!$ and $l = (2t+2)t!$, which implies $\rho(k-l) > t$. Thus, Theorem 2 gives $\overline{\lambda} < 1/t$ for the equation $x_1 + \ldots + x_k = y_1 + \ldots + y_l$.

Finally, we show that for the equation $x_1 + x_2 = ky$, where $k \ge 10$, neither $\Lambda = \overline{\Lambda} = \underline{\Lambda}$, nor $\underline{\lambda} = \max\left(\varrho, \frac{s^+ - s^-}{s^+}\right)$, which answers the third and the fourth question of Ruzsa. As a matter of fact, we prove that one can have $\overline{\Lambda} < \underline{\Lambda} < \underline{\lambda}$.

Let us make first the following elementary observation.

Fact. *Let A be a set of positive integers with no solution to the equation $x_1 + x_2 = ky$, where k is fixed positive integer. Then $\underline{\Lambda} \le 1/2$.*

Proof. Every set $A \in \mathbb{N}$ with $\underline{\mathrm{d}}(A) > 1/2$ contains in its sum-set $A+A$ each natural number from some point on. Thus, the sets $A+A$ and Ak may not be disjoint. ■

Example 1. For a given $k > 2$ define

$$S = \left(\bigcup_{n=0}^{\infty}\left[\frac{k^{2n}}{2^n} + 1, \frac{k^{2n+1}}{2^{n+1}}\right]\right) \cap \mathbb{N}.$$

It is clear that there is no solution to the equation $x_1 + x_2 = ky$ in the set S and $\overline{\mathrm{d}}(S) = k(k-2)/(k^2-2)$, so $\overline{\Lambda} \ge k(k-2)/(k^2-2)$. The next theorem shows that, in fact, $\overline{\Lambda} = k(k-2)/(k^2-2)$.

Theorem 3. *If $A \subseteq \mathbb{N}$ contains no solutions to the equation $x_1 + x_2 = ky$, where $k \ge 10$, then*

$$\overline{\mathrm{d}}(A) \le \frac{k(k-2)}{k^2-2}.$$

Proof. Assume $\overline{\mathrm{d}} = \overline{\mathrm{d}}(A) \ge k(k-2)/(k^2-2)$. For a given ε with $1/k^3 > \varepsilon > 0$ choose n_ε so that $A(i) < (\overline{\mathrm{d}}+\varepsilon)i$ for every $i > n_\varepsilon$. Let n be such that $n > kn_\varepsilon$ and $(\overline{\mathrm{d}}-\varepsilon)n < A(n)$. Furthermore set $m = \min A$.

First, assume

$$A \cap \left[\frac{4n}{k^2-2}, \frac{2(k^2-2k-2)n}{k(k^2-2)}\right] \neq \emptyset.$$

For each $y_0 \in A \cap [4n/(k^2-2), n/k]$ and $x < ky_0$, either $x \notin A$ or $ky_0 - x \notin A$, so

$$A(n) \le \frac{ky_0}{2} + (n - ky_0) \le \frac{k^2-2k-2}{k^2-2}n < (\overline{\mathrm{d}}-\varepsilon)n,$$

which contradicts the choice of n. The case

$$A \cap \left[\frac{n}{k}, \frac{2(k^2-2k-2)n}{k(k^2-2)}\right] \neq \emptyset$$

can be dealt with in a similar way.

Now suppose

$$A \cap \left[\frac{4n}{k^2-2}, \frac{2(k^2-2k-2)n}{k(k^2-2)}\right] = \emptyset.$$

Set

$$\begin{aligned} A_1 &= A \cap \left[\frac{2n}{k^2} + \frac{m}{k}, \frac{4n}{k(k^2-2)}\right), \\ A_2 &= A \cap \left(\frac{2(k^2-2k-2)n}{k(k^2-2)}, \frac{2n}{k}\right], \end{aligned}$$

and assume that neither of these sets is empty, otherwise the proof follows the same lines. Observe $(A_1k - m) \cap A = \emptyset$ and $(A_1k - m) \subseteq [2n/k, n]$. Since A has no solutions to the equation $x_1 + x_2 = ky$ we get

$$|A \cap [ks - n, n]| \le n - ks/2,$$

where $s = \min A_2$. Moreover, since $k \ge 10$, we have $k \max A_1 \le ks - n$. These yield

$$|A \cap [2n/k, n]| \le n - 2n/k - |A_1| - n + ks/2,$$

so that

$$\begin{aligned} A(n) &\le (\overline{\mathrm{d}}+\varepsilon)2n/k^2 + |A_1| + |A_2| + ks/2 - 2n/k - |A_1| + O(1) \\ &\le (\overline{\mathrm{d}}+\varepsilon)2n/k^2 + n - 2n/k + O(1). \end{aligned}$$

Thus,

$$(\overline{\mathrm{d}}-\varepsilon)n \le A(n) \le (\overline{\mathrm{d}}+\varepsilon)2n/k^2 + n - 2n/k + O(1),$$

which gives

$$\overline{\mathrm{d}} \le \frac{k(k-2)}{k^2-2}. \ \blacksquare$$

Example 2. Let $n \in \mathbb{N}$ and set

$$T = \left(\left[\frac{8n}{k(k^4-2k^2-4)} + 1, \frac{4n}{k^4-2k^2-4} \right] \cup \left[\frac{4(k^2-2)n}{k(k^4-2k^2-4)} + 1, \frac{2(k^2-2)n}{k^4-2k^2-4} \right] \cup \left[\frac{2n}{k} + 1, n \right] \right) \cap \mathbb{N}.$$

It is not difficult to see that $x_1 + x_2 = ky$ with $x_1, x_2, y \in T$ is not possible. Moreover

$$|T| = \left(\frac{k(k-2)}{k^2-2} + \frac{8(k-2)}{k(k^2-2)(k^4-2k^2-4)} \right) n + O(1),$$

so

$$\underline{\lambda} \geq \frac{k(k-2)}{k^2-2} + \frac{8(k-2)}{k(k^2-2)(k^4-2k^2-4)}.$$

(In fact, it is shown in [11] that the lower bound above is the actual value of $\underline{\lambda}$ for the equation $x_1 + x_2 = ky$.)

Since $s^+ = k$ and $s^- = 2$ we have $(s^+ - s^-)/s^+ = 1 - 2/k$. On the other hand, using the same argument as in the proof of the Fact one can show that for every set $A \subseteq \mathbb{Z}_m$ with no solutions to the equation $x_1 + x_2 = ky$, we have $|A| \leq m/2$, thus $\varrho \leq 1/2$. Finally, we obtain

$$\underline{\lambda} \geq \frac{k(k-2)}{k^2-2} + \frac{8(k-2)}{k(k^2-2)(k^4-2k^2-4)} > 1 - \frac{2}{k} = \max\left(\varrho, \frac{s^+ - s^-}{s^+} \right).$$

References

[1] N. Alon, *Independent sets in regular graphs and sum-free sets of finite groups*, Israel J. Math. 73 (1991), 247–256.

[2] J.-M. Deshoulliers, G. Freimen, V. Sós and M. Temkin, *On the structure of sum-free sets*, *2*, Astérisque 258 (1999), 149–161.

[3] P. Erdős, V. Sós and A. Sárközy, *On sum sets of Sidon sets*, J. Number Theory 47 (1993), 329–347.

[4] V. Lev, *Optimal representation by sumsets and subset sums*, ibid. 62 (1997), 127–143.

[5] T. Łuczak and T. Schoen, *On the infinite sum-free sets of natural numbers*, ibid. 66 (1997), 211–224.

[6] K. F. Roth, *On certain sets of integers*, J. London Math. Soc. 28 (1953), 104–109.

[7] I. Z. Ruzsa, *On infinite Sidon sequences*, J. Number Theory 68 (1998), 63–71.

[8] —, *Solving a linear equation in a set of integers I*, Acta Arith. 65 (1993), 259–282.

[9] —, *Solving a linear equation in a set of integers II*, ibid. 72 (1995), 385–397.

[10] T. Schoen, *On the density of universal sum-free sets*, Combin. Probab. Comput. 8 (1999), 277–280.
[11] —, *Subsets of* $\{1, \ldots, n\}$ *with no solutions to the equation* $x+y = kz$, in preparation.

Mathematisches Seminar
Universität zu Kiel
Ludewig-Meyn-Str. 4
24098 Kiel, Germany
E-mail: tos@numerik.uni-kiel.de

Received on 18.6.1999
and in revised form on 6.12.1999 (3630)