A note on a problem of Hilliker and Straus

Mirosława Jańczak

Faculty of Mathematics and CS
Adam Mickiewicz University
ul. Umultowska 87, 61-614 Poznań, Poland
mjanczak@amu.edu.pl

Submitted: May 20, 2006; Accepted: Oct 23, 2007; Published: Oct 30, 2007
Mathematics Subject Classifications: 06124, 06124

Abstract

For a prime p and a vector $\bar{\alpha} = (\alpha_1, \dots, \alpha_k) \in \mathbb{Z}_p^k$ let $f(\bar{\alpha}, p)$ be the largest n such that in each set $A \subseteq \mathbb{Z}_p$ of n elements one can find x which has a unique representation in the form $x = \alpha_1 a_1 + \dots + \alpha_k a_k,\ a_i \in A$. Hilliker and Straus [2] bounded $f(\bar{\alpha}, p)$ from below by an expression which contained the L_1-norm of $\bar{\alpha}$ and asked if there exists a positive constant $c(k)$ so that $f(\bar{\alpha}, p) > c(k) \log p$. In this note we answer their question in the affirmative and show that, for large k, one can take $c(k) = O(1/k \log(2k))$. We also give a lower bound for the size of a set $A \subseteq \mathbb{Z}_p$ such that every element of $A + A$ has at least K representations in the form $a + a'$, $a, a' \in A$.

1 Introduction

Let $f(p)$ denote the largest number n such that in any set $A = \{a_1, \dots, a_n\}$ contained in $\mathbb{Z}_p = \mathbb{Z}/p\mathbb{Z}$ at least one difference $a_i - a_j$ is incongruent to all other differences. Straus [4] estimated $f(p)$ up to a constant factor, showing that

$$\frac{1}{2}\log_2(p-1) + 1 \le f(p) < \frac{(2 + o(1))}{\log_2 3}\log_2 p$$

for all primes p. Hilliker and Straus [2] studied the following natural generalization of the problem. For a given vector $\bar{\alpha} = (\alpha_1, \dots, \alpha_k) \in \mathbb{Z}_p^k$ consider the set of all linear combinations $S = S(\bar{\alpha}, A) = \alpha_1 A + \alpha_2 A + \dots + \alpha_k A$. Let $f(\bar{\alpha}, p)$ be the largest n such that for any set $A \subseteq \mathbb{Z}_p$, $|A| = n$, one can find at least one element which has the unique representation in S. They proved that

$$f(\bar{\alpha}, p) \ge \frac{\log(p-1)}{\log(2\|\bar{\alpha}\|_1)} + 1,$$

where $\|\bar{\alpha}\|_1 = \sum_{i=1}^{k} |\alpha_i|$. They ask if the L_1-norm of a vector $\bar{\alpha}$ can be replaced by a function which depends only on k, i.e., if $f(\bar{\alpha}, p) > c(k)\log p$?

In the note we settle the above problem in the affirmative (Theorem 1 Corollary 1 below). We also show that our lower bound for $f(\bar{\alpha}, p)$ given in Theorem 1 cannot be much improved (Theorem 2). In section 3 we find a lower bound on $|A \pm A|$ for special sets A such that every element $x \in A + A$ has at least two different representations $a + a'$, $a, a' \in A$. Finally, we give a lower bound for the size of a set $A \subseteq \mathbb{Z}_p$ such that every element $t \in A + A$ has at least $K \geq 2$ representations of the form $t = a + a'$, $a, a' \in A$.

Throughout the note $\bar{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_k)$ denotes a vector with nonzero integral components, and l denote the number of different components of $\bar{\alpha}$. By $\log x$ we always mean $\log_2 x$, p is a prime, and A is a set of residues modulo p. We set $r \cdot T = \{rt : t \in T\}$ but sometimes we shall omit the dot writing for instance $\alpha_i A$ instead $\alpha_i \cdot A$. By $S = S(\bar{\alpha}, A)$ we mean the set

$$S = S(\bar{\alpha}, A) = \alpha_1 A + \alpha_2 A + \cdots + \alpha_k A,$$

and for a natural k we put

$$kA = \underbrace{A + A + \cdots + A}_{k}.$$

For $x \in \mathbb{Z}_p$ let $\nu_{\bar{\alpha}}(x) = \nu_{\bar{\alpha},A}(x)$ be the number of representation of x in $\mathbb{Z}_p$ in the form $x = \alpha_1 a_1 + \cdots + \alpha_k a_k$, where $a_1, \ldots, a_k \in A$. For $t \in \mathbb{R}$ let $\|t\|$ denotes the distance from t to the nearest integer.

Finally, let us mention a simple but important observation that for every $x, d_1, d_2 \in \mathbb{Z}_p$, $d_1 \neq 0$,

$$\nu_{\bar{\alpha},A}(x) = \nu_{\bar{\alpha},d_1 A + d_2}(d_1 x + d_2 \sum_{i=1}^{k} \alpha_i). \qquad (1)$$

2 A lower bound for $f(\bar{\alpha}, p)$

First we present a simple argument which shows that in the inequality $f(\bar{\alpha}, p) \geq \frac{\log(p-1)}{\log(2\|\bar{\alpha}\|_1)} + 1$, proved by Hilliker and Straus [2], one can replace the factor $(\log(\|\bar{\alpha}\|_1))^{-1}$ by a constant depending only on k.

Theorem 1. *For every $\bar{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_k)$ we have*

$$f(\bar{\alpha}, p) \geq \frac{\log p}{l \log 2k}.$$

Proof. Let $A = \{a_1, \ldots, a_n\}$ be a set such that for every element $x \in S$ we have $\nu_{\bar{\alpha}}(x) \geq 2$ and $|A| = f(\bar{\alpha}, p) + 1$. Let $T = \alpha_1 A \cup \cdots \cup \alpha_k A \subseteq \mathbb{Z}_p$. Because of (1) we can and shall assume that $a_1 = 0$.

Dirichlet approximation theorem implies that there exists r, $0 < r < p$, such that for every $x \in T$ we have

$$\left\| \frac{rx}{p} \right\| \leq p^{-\frac{1}{|T|-1}}.$$

Hence, for all $\alpha_1 a_1 + \cdots + \alpha_k a_k \in S$ we have

$$\left\| \frac{r(\alpha_1 a_1 + \cdots + \alpha_k a_k)}{p} \right\| \leq \left\| \frac{r\alpha_1 a_1}{p} \right\| + \cdots + \left\| \frac{r\alpha_k a_k}{p} \right\| \leq kp^{-\frac{1}{|T|-1}}.$$

We shall show that

$$p^{-\frac{1}{|T|-1}} \geq \frac{1}{2k}. \tag{2}$$

Indeed, suppose that the above inequality does not hold and $p^{-\frac{1}{|T|-1}} < \frac{1}{2k}$, so that $r \cdot T \subseteq (-\frac{p}{2k}, \frac{p}{2k})$. Let $x_i \in \alpha_i r \cdot A$ $(i = 1, \ldots, k)$. Observe, that for every $x_1 + \cdots + x_k \in r \cdot S$ we have

$$\|x_1 + \cdots + x_k\| < \|x_1\| + \cdots + \|x_k\| < \frac{1}{2}.$$

Hence, if m_i $(i = 1, \ldots, k)$ is the largest element in $\alpha_i r \cdot A$ considered as a subset of $(-\frac{p}{2k}, \frac{p}{2k})$, then, clearly, $m_1 + m_2 + \cdots + m_k$ has exactly one representation in S, because the effect modulo is not possible. Therefore

$$p^{-\frac{1}{|T|-1}} \geq \frac{1}{2k}.$$

Hence

$$|T| \geq \frac{\log p}{\log 2k} + 1,$$

and, since the cardinality of T is at most $l(|A| - 1) + 1$,

$$f(\bar{\alpha}, p) + 1 = |A| \geq \frac{\log p}{l \log 2k} + 1,$$

completing the proof of Theorem 1. □

Since $l \leq k$ as an immediate consequence of Theorem 1 we get the following result.

Corollary 1. *For any* $\bar{\alpha}$

$$f(\bar{\alpha}, p) \geq \frac{\log p}{k \log 2k}. \quad \square$$

From Theorem 1 it follows that, in particular, for $\bar{\alpha}^{(k)} = (1, 1, \ldots, 1)$ we have

$$f(\bar{\alpha}^{(k)}, p) \geq \frac{\log p}{\log 2k}.$$

Our next result shows that in general this bound cannot be much improved.

Theorem 2. *For every* $\varepsilon > 0$, $k \geq 2$ *and every prime* $p > p_\varepsilon$ *we have*

$$f(\bar{\alpha}^{(k)}, p) < \left(\frac{2 + 3\varepsilon}{\log(2k - 1)} \right) \log p + 3.$$

Proof. Our construction of a set A is a straightforward generalization of the one presented in [2]. Put

$$R = \{0, \pm 1, \pm 2, \dots, \pm z_k\},$$

where

$$z_k = \left\lceil \frac{k(2k-1)^m - 1}{k-1} \right\rceil$$

and $\frac{2k}{\varepsilon} < m < \log_{2k-1}\left(\frac{\varepsilon}{k}\log_{2k-1} p\right)$. Thus, the set $(k-1)R$ consists of all residues modulo $k(2k-1)^m$. We recursively define a descending sequence $a_1, a_2, \dots, a_l$ setting

$$a_1 = (p-r)/k,\ p \equiv r \mod k(2k-1)^m,\ r \in (k-1)R,$$

$$a_{i+1} = \begin{cases} a_i/(2k-1) & \text{if } a_i \equiv 0 \mod (2k-1) \\ (a_i - r_i)/k & \text{if } a_i \not\equiv 0 \mod (2k-1), \end{cases} \tag{3}$$

where $r_i \equiv a_i \mod k(2k-1)^m$. The last element a_l of this sequence satisfies

$$a_l \geq z_k + 1,\ a_{l+1} \in R. \tag{4}$$

Define

$$A = R \cup \{\pm a_1, \dots, \pm a_l\}.$$

We need to show that every element $x \in S$ has at least two different representations. It is clear that if $z = a_1 + \cdots + a_i + \cdots + a_j + \cdots + a_k$ with $a_i \neq a_j$, then $z = a_1 + \cdots + a_j + \cdots + a_i + \cdots + a_k$ is another representation of z. It remains to show that each element ka, where $a \in A$, has at least two representations in S. If $a = 0$ then it is indeed the case, since

$$ka = 0 + \cdots + 0 = 1 + (-1) + 0 + \cdots + 0.$$

For $0 < a < z_k$ we have

$$ka = (a-1) + (a+1) + \underbrace{a + \cdots + a}_{k-2}.$$

Finally, if $a = z_k$, then by (3) and (4)

$$z_k + 1 \leq a_l \leq (2k-1)z_k.$$

Hence

$$(k-1)z_k - 1 \geq ka - a_l \geq -(k-1)z_k.$$

Observe that $ka - a_l \in (k-1)R$. So, there exist $b_1, \dots, b_{k-1} \in R$ such that $ka = a_l + b_1 + \cdots + b_{k-1}$.

Now we show that every element ka_j has at least two representations in S. If $j \geq 2$, then by construction of the sequence we have either $a_j = a_{j-1}/(2k-1)$, or $a_j = (a_{j-1} - r_{j-1})/k$. If $a_j = a_{j-1}/(2k-1)$, then $(2k-1)a_j = a_{j-1}$ and

$$ka_j = a_{j-1} - (k-1)a_j = a_{j-1} + (k-1)(-a_j).$$

If $a_j = (a_{j-1} - r_{j-1})/k$, then

$$ka_j = a_{j-1} + \underbrace{(-r_{j-1})}_{\in (k-1)R}.$$

If $j = 1$ then ka_1 has the following two representations in S:

$$a_1 = (p - r)/k, \text{ where } r \equiv p \pmod{k(2k-1)^m}, \text{ and } r \in (k-1)R,$$

$$ka_1 = p - r \equiv -r \pmod{p}.$$

It means that

$$ka_1 = \underbrace{a_1 + \cdots + a_1}_{k} = 0 + \underbrace{(-r)}_{\in (k-1)R}.$$

Finally, we estimate the cardinality of A. Note that

$$|A| = 2l + 2z_k + 1 = 2l + 2\left\lceil \frac{k(2k-1)^m - 1}{k-1} \right\rceil + 1 < 2l + 2\frac{k(2k-1)^m}{k-1} + 3.$$

Observe that $a_{i+1} < a_i$ for all i and $a_{i+1} = a_i/(2k-1)$ for all except at most one out of every $m+1$ consecutive terms $a_j, a_{j+1}, \ldots, a_{j+m}$. We have also $a_{j+1} \le a_j/k$ if $a_{j+1} = (a_j - r_j)/k$, where $r_j \equiv a_j \pmod{k(2k-1)^m}, r_j \in (k-1)R$. Thus

$$a_{j+m+1} < k^{-1} a_j (2k-1)^{-m}$$

and

$$\frac{k(2k-1)^m}{k-1} \le a_l < pk^{\frac{-l-1}{m+1}}(2k-1)^{1-\frac{lm}{m+1}}.$$

Hence

$$l < \frac{1}{m}\left(1 - m^2 + (m+1)\frac{\log p}{\log(2k-1)}\right) < (1 + 1/m)\frac{\log p}{\log(2k-1)}.$$

Consequently,

$$\begin{aligned}
|A| &< 2\big(1 + 1/m\big)\frac{\log p}{\log(2k-1)} + 2\frac{k(2k-1)^m}{k-1} + 3 \\
&< \; 2(1 + \varepsilon/(2k))\frac{\log p}{\log(2k-1)} + 2\varepsilon/(k-1)\frac{\log p}{\log(2k-1)} + 3 \\
&= \; \left(2 + \frac{3k-1}{k(k-1)}\varepsilon\right)\frac{\log p}{\log(2k-1)} + 3 \\
&\le \; \big(2 + 3\varepsilon\big)\frac{\log p}{\log(2k-1)} + 3
\end{aligned}$$

for $\frac{2k}{\varepsilon} < m < \log_{2k-1}\left(\frac{\varepsilon}{k}\log_{2k-1} p\right)$ and $k \ge 2$. $\square$

Next result shows that for each α the order of magnitude of $f(\bar{\alpha}, p)$ is at most $\log^2 p$. This improves the upper bound for $f(\bar{\alpha}, p)$ in [2].

Theorem 3. *For every $\bar{\alpha} = (\alpha_1, \ldots, \alpha_k)$ we have*

$$f(\bar{\alpha}, p) \leq 4 \log^2 p.$$

Proof. Observe that if $\bar{\alpha} = (1, \alpha_2)$ and $\bar{\alpha}' = (1, \alpha_2, \alpha_3, \ldots, \alpha_k)$, then $f(\bar{\alpha}, p) \geq f(\bar{\alpha}', p)$. Let S be a set such that for every element $x \in S + S$ we have $\nu_{(1,1)} \geq 2$ and $|S| \leq 2 \log p$. Let $a_1, a_2 \in A = S + \alpha_2 S$. Then

$$\begin{aligned} a_1 + \alpha_2 a_2 &= (s_1 + \alpha_2 s_2) + \alpha_2(s_3 + \alpha_2 s_4) \\ &= s_1 + \alpha_2(s_2 + s_3) + \alpha_2^2 s_4 \\ &= s_1 + \alpha_2(s_2' + s_3') + \alpha_2^2 s_4 \\ &= (s_1 + \alpha_2 s_2') + \alpha_2(s_3' + \alpha_2 s_4) \\ &= a_1' + \alpha_2 a_2' \end{aligned}$$

for some $a_1, a_2, a_1', a_2' \in A$ and $s_1, s_2, s_3, s_4, s_2', s_3' \in S$. Thus

$$f(\bar{\alpha}, p) \leq |A| \leq |S|^2 \leq 4 \log^2 p.$$

□

3 The cardinality of sumsets

In this section we estimate the cardinality of $A - B$, where A is such that every element of $A + A$ has at least two representations, and B is an arbitrary subset of $\mathbb{Z}_p$. The main result of this section can be stated as follows.

Theorem 4. *If $A \subseteq \mathbb{Z}_p$ and for any element $x \in A + A$ we have $\nu_{(1,1)}(x) \geq 2$, then for any $B \subseteq \mathbb{Z}_p$*

$$|A - B| \geq |B| \left(\frac{\log p}{\log 12} - |B| \right).$$

Proof. Our argument is based on the following result of Ruzsa [3].

Lemma 1. *Let $A, B \subseteq G$ be finite sets and G be an abelian group. Then there exists a set $X \subseteq G$ such that $B \subseteq X + A - A$ and $|X| \leq \frac{|B-A|}{|A|}$.*

Let X be a set whose existence is guaranteed by Lemma 1, i.e.,

$$|X| \leq \frac{|A - B|}{|B|} \quad \text{and} \quad A \subseteq X + B - B. \tag{5}$$

By Dirichlet's theorem applied to the set $X \cup B$ there is an integer $0 < r < p$ such that for any element $z \in X \cup B$

$$\left\| \frac{rz}{p} \right\| \leq p^{-\frac{1}{|X|+|B|}}.$$

For every $a \in A$ there exist $b_1, b_2 \in B$ and $x \in X$ such that $a = x + b_1 - b_2$. Hence

$$\left\| \frac{ra}{p} \right\| \leq \left\| \frac{rx}{p} \right\| + \left\| \frac{rb_1}{p} \right\| + \left\| \frac{rb_2}{p} \right\| \leq 3p^{-\frac{1}{|X|+|B|}}.$$

Moreover, arguing as in the proof of Theorem 1 (cf. (2)), we get

$$3p^{-\frac{1}{|X|+|B|}} \geq \frac{1}{4}.$$

Thus

$$|X| \geq \frac{\log p}{\log 12} - |B|,$$

and, from (5),

$$|A - B| \geq |B||X| \geq |B| \left(\frac{\log p}{\log 12} - |B| \right).$$

□

Corollary 2. *If $A \subseteq \mathbb{Z}_p$ and for any element $x \in A + A$ we have $\nu_{(1,1)}(x) \geq 2$, then*

$$|A \pm A| \geq \left\lfloor \frac{\log p}{2 \log 12} \right\rfloor^2.$$

Proof. Pick any set $B \subseteq \pm A$ with $|B| = \left\lfloor \frac{\log p}{2 \log 12} \right\rfloor$ and apply Theorem 4 for the sets A and B. □

Let $f_K(p)$ be the largest n such that for any set $A \subseteq \mathbb{Z}_p$ with at most $f_K(p)$ elements there exists at least one element in $A+A$ with less then K representations. As a corollary from Theorem 4 we obtain the following lower bound for $f_K(p)$.

Corollary 3. *For every $K \geq 2$ we have*

$$f_K(p) \geq \sqrt{K} \left\lfloor \frac{\log p}{2 \log 12} \right\rfloor - 1.$$

Proof. Let us assume that $A \subseteq \mathbb{Z}_p$, for each element $x \in A+A$ we have $\nu_{(1,1)}(x) \geq K \geq 2$, and $|A| = f_K(p) + 1$. By Corollary 2 we get

$$|A + A| > \left\lfloor \frac{\log p}{2 \log 12} \right\rfloor^2. \tag{6}$$

Since

$$K|A + A| \leq \sum_{t \in A+A} \nu_{(1,1)}(t) = |A|^2,$$

it follows that

$$\frac{|A|^2}{K} \geq |A + A|. \tag{7}$$

From (6) and (7), we get

$$f_K(p) + 1 = |A| \geq \sqrt{K} \left\lfloor \frac{\log p}{2 \log 12} \right\rfloor ,$$

and so

$$f_K(p) \geq \sqrt{K} \left\lfloor \frac{\log p}{2 \log 12} \right\rfloor - 1 .$$

□

References

[1] J. Browkin, B. Diviš, A. Schinzel, *Addition of sequences in general fields,* Monatshefte für Mathematik **82** (1976), 261–268.

[2] D. L. Hilliker, E. G. Straus, *Uniqueness of linear combinations (mod p),* Journal of Number Theory **24** (1986), 1–6.

[3] I. Z. Ruzsa, *An analog of Frieman's theorem in groups,* Asterisque **258** (1999), 323–326.

[4] E. G. Straus, *Differences of residues* (mod p), Journal of Number Theory **8** (1976), 40–42.