Sums and $k$-sums in abelian groups of order $k$

Weidong Gao$^{a,1}$, Imre Leader$^{b,*}$

$^a$ Department of Computer Science and Technology, University of Petroleum, Beijing 102200, China
$^b$ Department of Pure Mathematics and Mathematical Statistics, Centre for Mathematical Sciences, Wilberforce road, Cambridge CB3 0WB, UK

Received 26 March 2002; revised 25 May 2005
Available online 28 December 2005
Communicated by David Goss

Abstract

Let $G$ be an abelian group of order $k$. How is the problem of minimizing the number of sums from a sequence of given length in $G$ related to the problem of minimizing the number of $k$-sums? In this paper we show that the minimum number of $k$-sums for a sequence $a_1, \ldots, a_r$ that does not have 0 as a $k$-sum is attained at the sequence $b_1, \ldots, b_{r-k+1}, 0, \ldots, 0$, where $b_1, \ldots, b_{r-k+1}$ is chosen to minimise the number of sums without 0 being a sum. Equivalently, to minimise the number of $k$-sums one should repeat some value $k-1$ times. This proves a conjecture of Bollobás and Leader, and extends results of Gao and of Bollobás and Leader.

© 2005 Elsevier Inc. All rights reserved.

0. Introduction

Given a sequence $a_1, \ldots, a_r$ in $\mathbb{Z}_k$, the integers modulo $k$, a $k$-sum is a sum of the form $a_{i_1} + \cdots + a_{i_k}$, where $i_1 < \cdots < i_k$. How large can $r$ be without 0 being a $k$-sum? It is clear that we may have $r = 2k - 2$, by taking $a_1 = \cdots = a_{k-1} = 0$ and $a_k = \cdots = a_{2k-2} = 1$. Erdős, Ginzburg and Ziv [5] showed that this is best possible. In other words, they showed that if we have $a_1, \ldots, a_{2k-1}$ in $\mathbb{Z}_k$ then some $k$-sum is 0. Since then, numerous other proofs of this result have been found—see Alon and Dubiner [1] for a general survey.

* Corresponding author.
E-mail address: i.leader@dpmms.cam.ac.uk (I. Leader).

$^1$ The research partially supported by the National Natural Science Foundation of China, grant number 19971058, and the Education Committee Foundation of China.
We are interested here in two extensions of the Erdös–Ginzburg–Ziv theorem. Bollabás and Leader [2] gave a ‘quantitative’ version, showing that, given \(a_1, \ldots, a_n \in \mathbb{Z}_k\), where \(k \leq r \leq 2k - 1\), if 0 is not a k-sum then there are at least \(r - k + 1\) k-sums. This clearly implies the Erdös–Ginzburg–Ziv theorem, by putting \(r = 2k - 1\). Note that the restriction to 0 not being a k-sum is necessary, as otherwise we could make all the \(a_i\) equal, and note also that the result is best possible, as may be seen by taking \(a_1 = \cdots = a_k - 1 = 0\) and \(a_{k+1} = \cdots = a_r = 1\).

In a different direction, Gao [7] related sums to k-sums in general abelian groups, as follows. For \(G\) a finite abelian group, the Davenport constant \(s(G)\) of \(G\) is the minimal \(n\) such that, whenever \(a_1, \ldots, a_n \in G\), some (non-empty) sum of the \(a_i\) is 0. For example, the Davenport constant of \(\mathbb{Z}_k\) is easily seen to be \(k\). It is believed that \(s(\mathbb{Z}_k^n) = (n - 1)(k - 1) + 1\)—this has been proved by Olson when \(k\) is a prime or prime-power [12] and when \(n = 2\) [13]. The determination of the Davenport constant is one of the most fascinating unsolved problems concerning finite abelian groups: see Geroldinger and Schneider [10] for some results and counterexamples.

Gao [7] proved that, if we write \(s'(G)\) for the minimal \(n\) such that, whenever \(a_1, \ldots, a_n \in G\), some \(k\)-sum of the \(a_i\) is 0, then \(s'(G) = s(G) + k - 1\). Note that in one direction this is obvious: if \(a_1, \ldots, a_r\) has no non-empty sum being 0, then certainly 0 is not a \(k\)-sum of \(a_1, \ldots, a_{r+k-1}\), where \(a_{r+1} = \cdots = a_{r+k-1} = 0\). This result instantly implies the Erdös–Ginzburg–Ziv theorem, as \(s(\mathbb{Z}_k) = k\).

Let us remark in passing that the family of \(k\)-sums from a sequence has been studied by several authors. Olson [14] gave a sufficient condition for the family of \(k\)-sums from a sequence \(a_1, \ldots, a_{2k-1}\) in an abelian group \(G\) of order \(k\) to be the entire group \(G\); this result was extended by Gao [6] to deal with sequences \(a_1, \ldots, a_r\), for general \(r\). Hamidoune, Ordaz and Ortuño [11] gave a sufficient condition for 0 to be a \(k\)-sum from a sequence \(a_1, \ldots, a_r\), in terms of the number of \(a_i\) that are allowed to assume the same value.

Bollabás and Leader [2] conjectured the following extension of their result and the result of Gao: the minimum number of \(k\)-sums for a sequence \(a_1, \ldots, a_r\) from \(G\) that does not have 0 as a \(k\)-sum is attained at the sequence \(b_1, \ldots, b_{r-k+1}, 0, \ldots, 0\), where \(b_1, \ldots, b_{r-k+1}\) is chosen to minimise the number of sums without 0 being a sum. Our main aim in this paper is to prove this conjecture. This is a common generalisation of the above two results: one could view it as a quantitative version of the result of Gao, and as ‘explaining’ the result of Bollobás and Leader (as the problem of minimizing the number of sums in \(\mathbb{Z}_k\) without 0 being a sum is easily seen to be solved by taking all \(a_i = 1\)).

The plan of the paper is as follows. In Section 1 we prove this result, and make some related remarks and conjectures. In Section 2 we obtain some bounds on the number of sums for subsequences of a given sequence.

1. The minimum number of \(k\)-sums

Let us start with some notation. Let \(G\) be a finite abelian group, and let \(S = (a_1, \ldots, a_n)\) be a sequence of elements in \(G\). By \(\sigma(S)\) we denote the sum \(\sum_{i=1}^{n} a_i\). By \(\sum(S)\) we denote the set that consists of all elements of \(G\) that can be expressed as the sum of a non-empty subsequence of \(S\):

\[
\sum(S) = \{a_{i_1} + \cdots + a_{i_l} : 1 \leq i_1 < \cdots < i_l \leq n\}.
\]
For every $1 \leq m \leq n$, we denote by $\sum_{\leq m}(S)$ the set consisting of all elements in $G$ which can be expressed as the sum of a subsequence $T$ of $S$ with $1 \leq |T| \leq m$:

$$\sum_{\leq m}(S) = \{a_{i_1} + \cdots + a_{i_l} : 1 \leq l \leq m \text{ and } 1 \leq i_1 < \cdots < i_l \leq n\}.$$  

By $\sum_m(S)$ we denote the set of all elements in $G$ that can be expressed as the sum of a subsequence $T$ of $S$ with $|T| = m$:

$$\sum_m(S) = \{a_{i_1} + \cdots + a_{i_m} : 1 \leq i_1 < \cdots < i_m \leq n\}.$$  

If $U$ is a subsequence of $S$, we write $S \setminus U$ for the subsequence obtained by deleting the terms of $U$ from $S$; if $U$ and $V$ are disjoint subsequences of $S$, we write $UV$ for the subsequence obtained by adjoining the terms of $U$ to $V$.

Our aim is to prove the following result.

**Theorem 1.** Let $G$ be an abelian group of order $k$, and let $r \geq k$. Let $S = (a_1, \ldots, a_r)$ be a sequence of $r$ elements in $G$. Suppose that $0 \not\in \sum_k(S)$. Then, there is a sequence $T$ of $r - k + 1$ elements in $G$ such that $|\sum_k(S)| \geq |\sum(T)|$ and $0 \not\in \sum(T)$.

Our main tool will be the following lemma from [7].

**Lemma 2.** Let $G$ be an abelian group of order $k$, and let $S = (a_1, \ldots, a_k)$ be a sequence of $k$ elements in $G$. Let $h$ be the maximal number $t$ such that there is an element of $G$ which occurs $t$ times in $S$. Then $0 \in \sum_{\leq h}(S)$.

**Proof of Theorem 1.** Let $h$ be the maximal number $t$ such that there is an element $x$ (say) in $G$ which occurs $t$ times in $S$. Without loss of generality we may assume that $x = 0$ (otherwise, we consider the sequence $(-x + a_1, \ldots, -x + a_r)$ instead of $S$). By rearranging the subscripts we may assume that

$$S = (a_1, \ldots, a_{r-h}, 0, \ldots, 0).$$  

Let $W$ be a maximal subsequence (in length) of $(a_1, \ldots, a_{r-h})$ such that $\sigma(W) = 0$. We will show that $|W| \leq k - h - 1$.  

If $k - h \leq |W| \leq k$, then $W(0, \ldots, 0)$ is a $k$-subsequence with sum zero, contradicting $0 \not\in \sum_k(S)$.

If $|W| > k$, then apply Lemma 2 repeatedly: we obtain disjoint subsequences $W_1, \ldots, W_t$ of $W$ such that $\sigma(W_i) = 0$, $1 \leq |W_i| \leq h$, $|W \setminus W_1 \cdots \setminus W_t| \leq k$ and $|W \setminus W_1 \cdots \setminus W_{t-1}| > k$ for $i = 1, \ldots, t$. Now we have $k - h < |W \setminus W_1 \cdots \setminus W_t| \leq k$ and $\sigma(W \setminus W_1 \cdots \setminus W_t) = 0$, and therefore

$$W_1 \cdots \setminus W_t)(\underbrace{0, \ldots, 0}_{k-|W \setminus W_1 \cdots \setminus W_t|})$$

is a $k$-subsequence with sum zero, contradicting $0 \not\in \sum_k(S)$. This proves the assertion (1).
By rearranging the subscripts we may assume that
\[ S = (a_1, \ldots, a_{r-k+1+l}, b_1, \ldots, b_w, 0, \ldots, 0) \]
with \( W = (b_1, \ldots, b_w) \) and \( l = k - w - h - 1 \).

Set \( U = (a_1, \ldots, a_{r-k+1+l}) \). By the maximality of \( W \) we have \( 0 \notin \sum(U) \). Set \( T = (a_1, \ldots, a_{r-k+1}) \), and set \( b = \sum_{i=r-k+1}^{r-k+1+l} a_i \) (if \( l = 0 \) we set \( b = 0 \)). For every \( x \in \sum(T) \), we have \( b + x \in \sum(U) \) as a sum over a subsequence of size \( \geq l + 1 \). Therefore, \( b + x \in \sum(UW) \) as a sum over a subsequence of size \( \geq l + 1 + w \geq k - h \). Similarly to the proof of assertion (1) one can prove that \( b + x \in \sum_k(S) \). This gives that \( |\sum_k(S)| \geq |\sum(T)| \). \( \square \)

Note that of course Theorem 1 is best possible: if \( T = (b_1, \ldots, b_{r-k+1}) \) and \( S' = (b_1, \ldots, b_{r-k+1}, 0, \ldots, 0) \) then \( |\sum_k(S)| \geq |\sum(T)| = |\sum_k(S')| \).

Recall that the exponent of an abelian group is the greatest order of any of its elements. From the proof of Theorem 1 we see the following:

**Corollary 3.** Let \( G \) be an abelian group of order \( k \), let \( m \) be the exponent of \( G \), and let \( l, r \) be two integers with \( l \leq k \), \( m \mid l \) and \( r \geq 1 \). Let \( S = (a_1, \ldots, a_r) \) be a sequence of \( r \) elements in \( G \). Suppose that \( 0 \notin \sum_l(S) \). Then there is a sequence \( T \) of \( r - l + 1 \) elements in \( G \) such that \( |\sum_l(S)| \geq |\sum(T)| \) and \( 0 \notin \sum(T) \).

Let \( G \) be a finite abelian group of order \( k \) and exponent \( m \). Relating to the Davenport constant of \( G \), for any positive integer \( q \) we write \( s_{qm}(G) \) for the smallest integer \( t \) such that every sequence \( S \) of \( t \) elements in \( G \) satisfies \( 0 \notin \sum_{qm}(S) \). It is easy to see that \( s_{qm}(G) \geq qm + s(G) - 1 \), with equality holding for \( q \geq k/m \) (see [7]). Let \( l(G) \) be the smallest integer \( w \) such that \( s_{qm}(G) = qm + s(G) - 1 \) holds for every \( q \geq w \). It was shown in [8] that \( s(G)/m \leq l(G) \leq k/m \).

So far, very little seems to be known about \( l(G) \).

**Conjecture 4.** Let \( G \) be an abelian group of order \( k \), let \( m \) be the exponent of \( G \), and let \( l, r \) be two integers with \( l \leq ml(G) \), \( m \mid l \) and \( r \geq 1 \). Let \( S = (a_1, \ldots, a_r) \) be a sequence of \( r \) elements in \( G \). Suppose that \( 0 \notin \sum_l(S) \). Then there is a sequence \( T \) of \( r - l + 1 \) elements in \( G \) such that \( |\sum_l(S)| \geq |\sum(T)| \) and \( 0 \notin \sum(T) \).

2. Zero-sum-free subsequences

Let \( G \) be a finite abelian group, and \( S = (a_1, \ldots, a_l) \) a sequence of elements in \( G \). We say \( S \) is a zero-sum sequence if \( \sigma(S) = 0 \); and we say \( S \) is zero-sum-free if \( S \) contains no nonempty zero-sum subsequence, or equivalently if \( 0 \notin \sum(S) \).

For every positive integer \( r \) in the interval \( \{1, \ldots, s(G) - 1\} \), let
\[ f_G(r) = \min_{S, |S|=r} \left| \sum(S) \right|, \]
where \( S \) runs over all zero-sum-free sequences of \( r \) elements in \( G \). How does the function \( f_G \) behave?
Theorem 5. Let $G$ be a finite abelian group of exponent $m$. Then

(i) If $1 \leq r \leq m - 1$ then $f_G(r) = r$.

(ii) If $(6, m) = 1$ and $G$ is not cyclic then $f_G(m) = 2m - 1$.

In proving Theorem 5 we will make use of the following results, due to Bovey, Erdős and Niven [3] and Eggleton and Erdős [4], respectively (see also [9]). We write $f(S)$ for $|\sum(S)|$.

Lemma 6. Let $G$ be an abelian group, and let $S$ be a zero-sum-free sequence of elements in $G$. Let $S_1, \ldots, S_t$ be disjoint nonempty subsequences of $S$. Then, $f(S) \geq \sum_{i=1}^t f(S_i)$.

Lemma 7. Let $S$ be a zero-sum-free sequence consisting of three distinct elements in an abelian group $G$. Then

(i) $f(S) \geq 5$.

(ii) If no element in $S$ has order 2 then $f(S) \geq 6$.

Proof of Theorem 5. Let $S = (a_1, \ldots, a_r)$ be a zero-sum-free sequence of $r$ elements in $G$. By Lemma 6 we have $f(S) \geq \sum_{i=1}^r f((a_i)) = r$. If $r \leq m - 1$, let $a$ be an element in $G$ of order $m$. Then $T = (a, \ldots, a)$ is zero-sum-free and $f(T) = r$. Hence $f_G(r) = r$ for every $1 \leq r \leq m - 1$.

We now turn to the case $r = m$, with $(6, m) = 1$ and $G$ noncyclic. Choose $g \in G$ so that $g$ occurs in $S$ a maximal number of times. Write $v(g)$ for the number of occurrences of $g$. We distinguish two cases.

Case 1. $v(g) < \frac{m+2}{3}$. Let $l$ be the maximal integer $t$ such that $S$ contains $t$ disjoint subsets each consisting of three distinct elements. Let $A_1, \ldots, A_l$ be $l$ disjoint 3-subsets such that the residual sequence $T = S \setminus A_1 \setminus \cdots \setminus A_l$ contains as many distinct elements as possible. Clearly, $T$ contains at most two distinct elements. We claim that in fact we have $|T| \leq 2$. Indeed, suppose to the contrary that $|T| \geq 3$. Then:

Subcase 1. $T$ contains exactly two distinct elements. Suppose

$$T = (a, \ldots, a, b, \ldots, b) \quad \text{with } u \geq v \geq 1 \text{ and } u + v = |T|.$$ 

Since $|T| \geq 3$, we have $u \geq 2$. If $a \notin A_i$ for some $1 \leq i \leq l$, take $c \in A_i \setminus \{b\}$ and set $A'_i = (A_i \setminus \{c\}) \cup \{a\}$. Then, $A_1, \ldots, A_{i-1}, A'_i, A_{i+1}, \ldots, A_l$ are $l$ disjoint 3-subsets of $S$ and the residual sequence contains three distinct elements $a, b, c$, contradicting the choice of $A_1, \ldots, A_l$. This shows that $a \in A_i$ holds for every $i = 1, \ldots, l$. Therefore, $a$ occurs at least $l + u \geq \frac{n - 2u}{3} + u = n/3 + u/3 \geq \frac{n+2}{3}$ times in $S$, a contradiction.

Subcase 2. All terms in $T$ are the same. Suppose

$$T = (a, \ldots, a) \quad \text{with } u = |T| \geq 3.$$ 

If $a \notin A_i$ for some $1 \leq i \leq l$, take $b \in A_i$ and set $A'_i = (A_i \setminus \{b\}) \cup \{a\}$. Then $A_1, \ldots, A_{i-1}, A'_i, A_{i+1}, \ldots, A_l$ are $l$ disjoint 3-subsets of $S$, and the residual sequence contains two distinct el-
We may therefore assume that $a$ and $b$, contradicting the choice of $A_1, \ldots, A_l$. This shows that $a \in A_i$ for every $i = 1, \ldots, l$. Therefore $a$ occurs at least $l + u = \frac{n + a}{3} + u \geq \frac{n + 6}{3}$ times in $S$, a contradiction. This proves our claim. So we know that $|T| \leq 2$. Since $T$ is zero-sum-free, we clearly have $f(T) \geq 2|T| - 1$. Since $m$ is odd, $S$ contains no element of order 2. It follows from Lemma 6 that $f(S) = f(T) + \sum_{i=1}^l f(A_i) \geq 2|T| - 1 + 6l = 2m - 1$. This completes the proof in this case.

**Case 2.** $v(g) \geq \frac{m + 2}{3}$. Let $H$ be the cyclic subgroup generated by $g$. Write $S = S_1S_2$ such that all terms of $S_1$ are in $H$ and no term of $S_2$ is in $H$. We clearly have $|S_1| \geq v(g) \geq \frac{n + 2}{3}$ and $|S_2| \geq 1$. Suppose $S_2 = (b_1, \ldots, b_w)$. Let $\phi$ be the projection from $G$ to $G/H$. Then $\ker(\phi) = H$. Set $\phi(S_2) = (\phi(b_1), \ldots, \phi(b_w))$. Put $h = \lfloor f(S_2) \rfloor \{0\}$. We clearly have $f(S) \geq (h + 1)f(T_1) + h$. We distinguish subcases.

**Subcase 1.** $h \geq 5$. Then $f(S) \geq 6f(T_1) + 5 \geq 6\frac{m + 2}{3} + 5 > 2m - 1$.

**Subcase 2.** $h \leq 4$. It follows from Lemma 7 that $\phi(S_2)$ contains no zero-sum-free subsequence of length at least 5. If $T$ is a 4-subsequence of $S_2$ such that $\lfloor \sum(\phi(T)) \rfloor \{0\} \geq 4$ then, since $\phi(S_2)$ contains no zero-sum-free subsequence of length at least 5, one can find disjoint subsequences $T_1, T_2, \ldots, T_l$ of $S_2 \setminus T$ such that $|T_i| \leq 5$ and $\sigma(\phi(T_i)) = 0$ for $i = 1, \ldots, l$ and $\phi(S \setminus T \setminus T_1 \setminus \cdots \setminus T_l)$ is zero-sum-free. Therefore $|S \setminus T \setminus T_1 \setminus \cdots \setminus T_l| \leq 4$. Hence $l \geq \frac{|S_1| - |T| - 4}{5} = \frac{|S_2| - 8}{5}$. Now, note that $\sigma(T_i) \in H$. Thus the sequence $U = S_1(\sigma(T_1), \ldots, \sigma(T_l))$ is a zero-sum-free sequence of elements in $H$. So $f(S) \geq (\lfloor \sum(\phi(T)) \rfloor \{0\} + 1)f(U) + \lfloor \sum(\phi(T)) \rfloor \{0\} \geq 5f(U) + 4 \geq 5|U| + 4 \geq 5(|S_1| + \frac{|S_2| - 8}{5} + 4 = 5|S_1| + \frac{m - |S_1| - 8}{5} + 4 = 4|S_1| + m - 4 \geq \frac{4m + 2}{3} + m - 4 \geq 2m - 1$.

We may therefore assume that

$$\left\lfloor \sum(\phi(T)) \right\rfloor \{0\} \leq 3$$

holds for every 4-subsequence $T$ of $S_2$. (3)

It follows from Lemma 6 and (3) that no 4-subsequence of $\phi(S_2)$ is zero-sum-free.

If $W$ is a 3-subsequence of $S_2$ such that $\lfloor \sum(\phi(W)) \rfloor \{0\} \geq 3$, then by (3) one can find disjoint subsequences $W_1, \ldots, W_t$ of $S_2 \setminus W$ such that $|W_i| \leq 4$ and $\phi(W_i)$ is zero-sum-free for $i = 1, \ldots, t$. Similarly to the above one can prove that

$$f(S) \geq \left(\lfloor \sum(\phi(W)) \rfloor \{0\} + 1\right)(|S_1| + t) + \lfloor \sum(\phi(W)) \rfloor \{0\} \geq 4\left(|S_1| + \frac{m - |S_1| - 3 - 3}{4}\right) + 3 = 3|S_1| + m - 3 \geq \frac{m + 2}{3} + m - 3 = 2m - 1.$$

Therefore we may assume that

$$\left\lfloor \sum(\phi(W)) \right\rfloor \{0\} \leq 2$$

holds for every 3-subsequence $W$ of $S_2$. (4)

If $|S_2| \geq 3$, let $(a, b, c)$ be an 3-subsequence of $S_2$. By (4), we may assume that $\phi(a) = \phi(b)$. If $\phi(c) = \phi(a) = \phi(b)$, by (4) we obtain that $2\phi(a) = 0$ or $3\phi(a) = 0$, and this together with (6, $m$) = 1 implies that $\phi(a) = 0$. Thus $a \in H$, contradicting the definition of $S_2$. If $\phi(c) \neq \phi(a) = \phi(b)$ then by (4) we obtain that $\phi(c) + \phi(a) = 0$ and $2\phi(a) = 0$ or $2\phi(a) = \phi(c) = -\phi(a)$. Therefore $2\phi(a) = 0$ or $3\phi(a) = 0$. Similarly to the above one can derive a contradiction. This proves that $|S_2| \leq 2$.

If $|S_2| = 2$, suppose that $S_2 = (a, b)$. If $h \geq 2$, then $f(S) \geq (h + 1)f(S_1) + h \geq 3|S_1| + 2 = 3(m - 2) + 2 \geq 2m - 1$ when $m \geq 3$. For $m = 1, 2$ one can check the theorem directly. So we may
assume that $h = 1$, and hence $\phi(a) = \phi(b)$ and $2\phi(a) = 0$, again a contradiction. Therefore we may assume that $|S_2| = 1$. Thus $h = 1$ and $f(S) \geq (h + 1)f(S_1) + h \geq 2f(S_1) + 1 \geq 2|S_1| + 1 = 2m - 1$.

We do not know what happens if $(6, m) \neq 1$—it would be very interesting to work out what happens then.

References

41–43.