Some new results on subset sums

Van H. Vu

Department of Mathematics, Rutgers, Piscataway, NJ 08854, USA

Received 27 July 2006
Available online 15 November 2006
Communicated by Ronald Graham

Abstract

Let \( n \) be a large integer and \( A \) be a subset of \( [n] = \{1, \ldots, n\} \). The set \( S_A \) is the collection of the subset sums of \( A \). In this note, we discuss new results (and proofs) on few well-known problems concerning \( S_A \). In particular, we improve an estimate of Alon and Erdős concerning monochromatic representations.

© 2006 Elsevier Inc. All rights reserved.

1. Introduction

Let \( n \) be a large integer and \( A \) be a subset of \( [n] = \{1, \ldots, n\} \). The set \( S_A \) is the collection of the subset sums of \( A \). In this note, we discuss new results (and proofs) on few well-known problems concerning \( S_A \). These developments are based on the following recent result from [9].

For a set \( A \) of integers and a positive integer \( l \leq |A| \), let \( l^*A \) denote the set of sums of \( l \) different elements of \( A \)

\[
l^*A := \{a_1 + \cdots + a_l \mid a_i \in A, \; a_i \neq a_j\}.
\]

Theorem 1.1. There are positive constants \( c \) and \( C \) such that the following holds. If \( A \) is a subset of \( [n] \) and \( l \) is a positive integer at most \( |A|/2 \) such that \( l|A| \geq Cn \), then \( l^*A \) contains an arithmetic progression of length \( cl|A| \).

1.2. Monochromatic representations. Let \( f(n) \) be the smallest number such that one can color \( \{1, \ldots, n - 1\} \) by \( f(n) \) colors so that \( n \) cannot be represented as sum of numbers with the

---

E-mail address: vanvu@math.rutgers.edu.

1 Van H. Vu is an A. Sloan Fellow and is supported by NSF Career Grant 0635606.

0022-314X/$ – see front matter © 2006 Elsevier Inc. All rights reserved.
doi:10.1016/j.jnt.2006.08.012
same color (monochromatic representation). In other words, \( f(n) \) is the largest number \( k \) such that no matter how one partitions \( \{1, \ldots, n-1\} \) into \( k-1 \) sets \( A_1, \ldots, A_{k-1} \), then one of the \( S_{A_i} \) contains \( n \).

Alon and Erdős [1] proved the following estimate for \( f(n) \):

**Theorem 1.3.** There are positive constants \( c_1 \) and \( c_2 \) such that

\[
c_1 \frac{n^{1/3}}{\log^{4/3} n} \leq f(n) \leq c_2 \frac{n^{1/3}}{\log^{1/3} n}.
\]

It is conjectured in [1] that the exact order of magnitude of \( f(n) \) is closer to the upper bound. Motivated by this conjecture, here we improve the lower bound by a factor \( \Omega(\log^{1/3} n) \).

**Theorem 1.4.** There is a positive constant \( c_1 \) such that

\[
c_1 \frac{n^{1/3}}{\log n} \leq f(n).
\]

To prove this theorem, we will follow the ideas from [1], replacing a lemma in this paper by Theorem 1.1.

1.5. Complete sets. A subset \( A \) of a group \( G \) is complete if \( S_A = G \). We are interested in \( \mathbb{Z}_n \), the cyclic group of rank \( n \). A basic problem here is to give a sufficient condition for completeness. Olson [6] proved

**Theorem 1.6.** If \( n \) is a prime and \( A \) is a subset of at least \( 2n^{1/2} \) elements of \( \mathbb{Z}_n \) then \( A \) is complete.

The constant 2 is sharp (see below). Recently, using a corollary of the modular version of Theorem 1.1, Nguyen, Szemerédi and Vu [5] obtained the following extension of Theorem 1.6 which classifies all incomplete sets of size close to \( 2n^{1/2} \). For an element \( a \in \mathbb{Z}_n \), \( \|a\| \) denotes the distance from \( a \) to 0 (e.g., \( \|n-1\| = 1 \)).

**Theorem 1.7.** Let \( n \) be a large prime and \( A \) be an incomplete subset of \( \mathbb{Z}_n \) of size at least \( 1.99n^{1/2} \). Then there is a non-zero element \( b \in \mathbb{Z}_n \) such that

\[
\sum_{a \in bA} \|a\| \leq n + O(n^{1/2}).
\]

The error term \( O(n^{1/2}) \) is best possible, as showed by Deshouillers [2]. A weaker error term \( O(n^{3/4}) \) was obtained earlier by Deshouillers and Freiman [3].

It is natural to generalize Theorem 1.6 for arbitrary \( n \). As a relatively simple corollary of Theorem 1.1, we obtained

**Theorem 1.8.** There is a positive constant \( C \) such that the following holds. If \( A \) is a subset of at least \( Cn^{1/2} \) elements of \( \mathbb{Z}_n \) which are coprime with \( n \) then \( A \) is complete.
A very interesting question is to find the sharp value of $C$. It is easy to see that $C$ should be at least 2. Indeed, for any $C < 2$ and sufficiently large $n$, there are $Cn^{1/2}$ numbers between 1 and $2n^{1/2}$ coprime to $n$. The sum of these numbers is less than $n$ and so they form an incomplete set.

Conjecture 1.9. For any $C > 2$ the following holds for all sufficiently large $n$. If $A$ is a subset of at least $Cn^{1/2}$ elements of $\mathbb{Z}_n$ which are coprime with $n$ then $A$ is complete.

Notation. The asymptotic notation is used under the assumption that $n$ goes to infinity. For two sets $A$ and $B$, $A + B := \{a + b, \ a \in A, \ b \in B\}$. We omit unnecessary floors and ceilings.

2. A simple lemma

We will need the following simple lemma.

Lemma 2.1. Let $n$ be a positive integer and $A$ be a multi-set of $n$ integers co-prime to $n$. Then $S_A$ contains every residue modulo $n$.

In a multi-set we allow possible repetitions among the elements; in a set the elements are different.

Proof of Lemma 2.1. Assume that $a_1, a_2, \ldots, a_n$ are the elements of $A$. We are going to prove, by induction, that $|S_{A_i}| \geq i$, where $A_i = \{a_1, \ldots, a_i\}$. The case $i = 1$ is trivial. Assume that the statement holds for $i - 1$. Let $b_1, \ldots, b_{i - 1}$ be $i - 1$ different elements (modulo $n$) of $S_{A_{i-1}}$. Since the statement is invariant under dilation, we can assume that $a_i = 1$. Consider the elements

$$b_1, \ldots, b_{i-1}, 1, 1 + b_1, \ldots, 1 + b_{i-1}. $$

At least $i$ of the above must be different (modulo $n$) and this concludes the proof. \(\square\)

3. Proof of Theorem 1.4

Let $f(n) = \frac{c}{20} n^{1/3} \log^{-1} n$, for some sufficiently small positive constant $c$. Set $L = n^{2/3}$ and consider the set $P$ of primes in the interval $I = [L, 2L]$. Since $P \geq 0.9 n^{2/3} \log^{-1} n$, there will be a monochromatic set $A$ consisting of at least

$$\frac{0.9 n^{2/3} \log^{-1} n}{(c/20)n^{1/3} \log^{-1} n} = \frac{18}{c} n^{1/3}$$

primes.

Thus $A$ contains three disjoint subsets $A_1, A_2, A_3$ of the same cardinality $\frac{6}{c} n^{1/3}$. We are going to show that $S_{A_1} + S_{A_2} + S_{A_3}$ contains $n$, regardless the choice of the elements in these subsets. Set $l = \frac{1}{2} n^{1/3}$. Set $C = 1/c$, we have

$$l | A_1 | = \frac{3}{c} n^{2/3} = 3 C n^{2/3} \geq C |I|.$$
Thus, by Theorem 1.1, \( l^* A_1 \) contains an arithmetic progression \( P_1 \) of length
\[ c|A_1| = 3n^{2/3}. \]

Assume that \( P_1 = \{a + d, \ldots, a + md\} \), where \( m = 3n^{2/3} \). As \( P_1 \) is a subset of the set \( lI \) which has cardinality at most \( lL \), the difference \( d \) is at most
\[ \frac{lL}{3n^{2/3}} = \frac{1}{6} n^{1/3}. \]

Furthermore, the largest element of \( P_1 \) is at most
\[ lL = n/2. \]

Next, we use \( A_2 \) to create a complete residue modulo \( d \). Notice that
\[ |A_2| \geq \frac{6}{c} n^{1/3} > n^{1/3} > d. \]

Furthermore, the elements of \( A_2 \) are primes, so they are co-primes to \( d \) (with the exception of at most one element). By Lemma 2.1, \( S_{A_2} \) contains a set \( \{b_1, \ldots, b_d\} \) which contains every residue modulo \( d \). Moreover, the \( b_i \) are relatively small. Indeed,
\[ \max_i b_i \leq dL \leq dn^{2/3}. \]

Consider a number \( x \), where \( \max_i b_i \leq x - a \leq 3n^{2/3} d \). We claim that \( x \) belongs to \( S_{A_1} + S_{A_2} \). Indeed, there is a number \( b_i \) such that \( x - a - b_i \) is non-negative and divisible by \( d \). Moreover, \( (x - a - a_i)/d \leq n^{2/3} < |P_1| \). Thus
\[ x \in P_1 + \{a_1, \ldots, a_d\} \subset S_{A_1} + S_{A_2}, \]
as claimed.

It thus follows that the sum \( S_{A_1} + S_{A_2} \) contains an interval \( J \) of length at least
\[ 3n^{2/3} d - \max_i a_i \geq 3n^{2/3} d - dn^{2/3} \geq 2n^{2/3} = 2L. \]

Furthermore, the smallest element of this interval is at most
\[ a + \max_i a_i < n. \]

Finally we make use of \( A_3 \). Let \( A_3 = \{b_1, \ldots, b_k\} \). The set \( N = \{a_1, a_1 + a_2, \ldots, a_1 + a_2 + \cdots + a_k\} \) is a subset of \( S_{A_3} \). On the other hand, this set is an \( L \)-net (the distance between two consecutive elements is at most \( L \)) whose smallest element is \( a_1 \leq 2L \) and largest element is
\[ a_1 + \cdots + a_k \geq Lk \geq L \cdot \frac{6}{c} n^{1/3} > n. \]

It is easy to see that \( N + J \) contains \( n \). Since \( N + J \) is a subset of \( S_{A_1} + S_{A_2} + S_{A_3} \subset S_A \), our proof is complete.
Remark. By restricting to the primes, we paid a log $n$ factor, which is the remaining gap between the upper and lower bounds.

4. Proof of Theorem 1.8

We are going to use the following (direct) corollary of Theorem 1.1. This corollary improves upon earlier results of Freiman [4] and Sárközy [8].

Corollary 4.1. There is a positive constant $C$ such that the following holds. For every sufficiently large integer $n$ and a subset $A$ of $[n]$ of cardinality at least $Cn^{1/2}$, $S_A$ contains an arithmetic progression of length $n$.

Assume that $A$ has at least $2\lceil Cn^{1/2} \rceil$ elements, where $C$ is the constant in Corollary 4.1. For convenience, we think of the elements of $A$ as positive integers between one and $n - 1$. We are going to prove that $S_A$ contains every residue modulo $n$.

Let $A'$ be a subset of $A$ of $\lceil Cn^{1/2} \rceil$ elements. Apply Corollary 4.1 to $A'$ to get an arithmetic progression $P'$ of length $n$. If the difference $d'$ of $P'$ is co-prime to $n$, then $P'$ contains every residue modulo $n$ and we are done. If $d'$ is not co-prime to $n$, set $d = \gcd(d', n)$.

Since the largest element in $S_{A'}$ is less than $Cn^{3/2}$, $d$ is less than $Cn^{1/2}$. The set $B = A \setminus A'$ has at least $\lceil Cn^{1/2} \rceil > d$ elements, each of which is co-prime to $n$ (and thus co-prime to $d$). By Lemma 2.1, we conclude that $S_B$ contains every residue modulo $d$.

The set $S_{A'} + S_B$ thus contains every residue modulo $n$. But this set is clearly a subset of $S_A$, completing the proof.

Remark. Notice that the proof requires that the elements of $B$ are co-prime to $n$; but for $A'$, it is enough to assume that its elements are non-zero modulo $n$.

Acknowledgments

We thank T. Tao for useful discussions. We also thank M. Nathanson for pointing out Ref. [7].

References