(k + 1)-SUMS VERSUS k-SUMS

Simon Griffiths
IMPA, Est. Dona Castorina 110, Jardim Botânico, Rio de Janeiro, Brazil
sgriff@impa.br

Received: 6/7/12, Revised: 8/15/12, Accepted: 9/30/12, Published: 10/8/12

Abstract
A k-sum of a set \( A \subseteq \mathbb{Z} \) is an integer that may be expressed as a sum of \( k \) distinct elements of \( A \). How large can the ratio of the number of \((k + 1)\)-sums to the number of \(k\)-sums be? Writing \( k \wedge A \) for the set of \( k\)-sums of \( A \) we prove that

\[
\frac{|(k + 1) \wedge A|}{|k \wedge A|} \leq \frac{|A| - k}{k + 1}
\]

whenever \( |A| \geq (k^2 + 7k)/2 \). The inequality is tight – the above ratio being attained when \( A \) is a geometric progression. This answers a question of Ruzsa.

1. Introduction

Given a set \( A = \{a_1, \ldots, a_n\} \) of \( n \) integers we denote by \( k \wedge A \) the set of integers which may be represented as a sum of \( k \) distinct elements of \( A \). In this paper we consider the problem of how large the ratio \( |(k + 1) \wedge A|/|k \wedge A| \) can be. The upper bound

\[
\frac{|(k + 1) \wedge A|}{|k \wedge A|} \leq \frac{n}{k + 1}
\]  \hspace{1cm} (1)

is easily obtained using a straightforward double-counting argument.

Ruzsa [1] asked whether this inequality may be strengthened to

\[
\frac{|(k + 1) \wedge A|}{|k \wedge A|} \leq \frac{n - k}{k + 1}
\]

whenever \( n \) is large relative to \( k \). We confirm that this is indeed the case.

**Theorem 1.1.** Let \( A \) be a set of \( n \) integers and suppose that \( n \geq (k^2 + 7k)/2 \). Then

\[
\frac{|(k + 1) \wedge A|}{|k \wedge A|} \leq \frac{n - k}{k + 1}.
\]  \hspace{1cm} (2)
Furthermore, in the case that \( n > (k^2 + 7k)/2 \), equality holds if and only if \( |k \land A| = \binom{n}{k} \) and \( |(k + 1) \land A| = \binom{n}{k+1} \).

Since the ratio \( (n - k)/(k + 1) \) is obtained for all \( k \) in the case that \( A \) is a geometric progression this result is best possible for each pair \( k,n \) covered by the theorem. However, we do not believe that \( n > (k^2 + 7k)/2 \) is a necessary condition for inequality (2). Indeed, we pose the following question.

**Question 1.2.** Does (2) hold whenever \( n > 2k \)?

The inequality \( n > 2k \) is necessary. Indeed, for any pair \( k,n \) with \( n/2 \leq k \leq n - 1 \) the inequality (2) fails for the set \( A = \{1,\ldots,n\} \) (or indeed any arithmetic progression of length \( n \)). To see this note that \( |k \land A| = k(n - k) + 1 \) for each \( k = 1,\ldots,n \), and that the inequality

\[
\frac{(k + 1)(n - k) + 1}{k(n - k) + 1} \leq \frac{n - k}{k + 1}
\]

holds if and only if \( k \leq (n - 1)/2 \). Thus, we have also verified for the case that \( A \) is an arithmetic progression that (2) holds whenever \( n > 2k \). We also note for any set \( A \subseteq \mathbb{Z} \) that (2) holds trivially (and with equality) in the case that \( k = (n - 1)/2 \). Indeed this follows immediately from the symmetry \( |k \land A| = |(n - k) \land A|, k = 1,\ldots,n - 1 \).

### 2. Proof of Theorem 1.1

The proof of Theorem 1.1 is closely related to the double-counting argument one uses to prove (1). We recall that argument now.

Fix \( k \in \{0,\ldots,n - 1\} \) and a set \( A = \{a_1,\ldots,a_n\} \) of \( n \) integers. We say that an element \( s \in k \land A \) extends to \( t \in (k + 1) \land A \) if there exist distinct elements \( a_1,\ldots,a_{k+1} \) of \( A \) such that \( s = a_1 + \cdots + a_k \) and \( t = a_1 + \cdots + a_{k+1} \). Define the bipartite graph \( G \) with vertex sets \( U = \{u_s : s \in k \land A\} \) and \( V = \{v_t : t \in (k+1) \land A\} \) and edge set \( E(G) = \{u_s v_t : s \text{ extends to } t\} \). We prove (1) by counting \( e(G) \) in two different ways:

(i) \( e(G) \leq n|U| \), since each vertex \( u_s \in U \) has at most \( n \) neighbours in \( V \).

(ii) \( e(G) \geq (k + 1)|V| \) since each vertex \( v_t \in V \) is adjacent to each vertex \( u_{i-a_i} : i = 1,\ldots,k+1, \) where \( a_1 + \cdots + a_{k+1} \) is a \((k + 1)\)-sum to \( t \).

Since \( |U| = |k \land A| \) and \( |V| = |(k + 1) \land A| \) we obtain that

\[
(k + 1)|k \land A| \leq e(G) \leq n|k \land A|,
\]
completing the proof of (1).
The alert reader will note that the extremal cases of each of (i) and (ii) occur in rather different situations. The inequality \( e(G) \leq n|U| \) may be tight only if each element \( s \in k \wedge A \) extends to \( s + a \) for all \( a \in A \). Equivalently, for each \( s \in k \wedge A \) and \( a \in A \), \( s \) may be represented as a \( k \)-sum that does not use \( a \), i.e., \( s = a_1 + \cdots + a_k \) for distinct \( a_1, \ldots, a_k \in A \setminus \{a\} \). In particular, the inequality in (i) may be tight only if each \( k \)-sum has at least two representations. On the contrary, the second inequality \( e(G) \geq (k + 1)|V| \) may be tight only if each \( t \in (k + 1) \cap A \) may be represented as a \((k + 1)\)-sum in a unique way. This simple observation is the key to our proof.

We put the above observations into action by defining \( Q_k \subseteq k \wedge A \) to be the set of \( s \in k \wedge A \) that have a unique representation as a \( k \)-sum and \( S = (k \wedge A) \setminus Q_k \) to be the set of \( s \) with at least two representations. We immediately obtain a new upper bound on \( e(G) \), namely:

\[
e(G) \leq (n - k)|U_k| + n|S| = (n - k)|k \wedge A| + k|S|.
\]

Correspondingly, one may define \( Q_{k+1} \) to be the set of \( t \in (k + 1) \wedge A \) that are uniquely represented as a \((k + 1)\)-sum and \( T = ((k + 1) \wedge A) \setminus Q_{k+1} \) to be the set of \( t \) with at least two representations. It then follows (using Lemma 2.2 below) that

\[
e(G) \geq (k + 1)|U_{k+1}| + (k + 3)|T| = (k + 1)(k + 1) \wedge A| + 2|T|.
\]

Unfortunately (3) and (4) do not directly imply Theorem 1.1 since it is non-trivial to relate \(|S|\) and \(|T|\). For this reason we define a subgraph \( H \) of \( G \) as follows. Recall that a pair \( u_s v_t \) is an edge of \( G \) if there exists a representation \( s = a_1 + \cdots + a_k \) of \( s \) as a \( k \)-sum of elements of \( A \) and \( a \in A \setminus \{a_1, \ldots, a_k\} \) such that \( t = s + a \). Include an edge \( u_s v_t \) of \( G \) in \( H \) if and only if there exist two representations \( s = a_1 + \cdots + a_k = b_1 + \cdots + b_k \) of \( s \) as a \( k \)-sum of elements of \( A \) and \( a \in A \setminus \{a_1, \ldots, a_k\} \cup \{b_1, \ldots, b_k\} \) such that \( s + a = t \). (Note: if an edge \( u_s v_t \) of \( G \) is included in \( H \) then in particular \( s \in S \) and \( t \in T \).)

We begin with two lemmas.

**Lemma 2.1.** Let the set \( S \subseteq k \wedge A \) and the graph \( H \subseteq G \) be as defined above. Then \( e(H) \geq (n - 2k)|S| \).

**Proof.** For each \( s \in S \) the vertex \( u_s \) has degree at least \( n - 2k \) in \( H \). Indeed, writing \( s = a_1 + \cdots + a_k = b_1 + \cdots + b_k \) we have that \( u_s v_t \in E(H) \) for each \( t \in \{s + a : a \in A \setminus \{a_1, \ldots, a_k\} \cup \{b_1, \ldots, b_k\}\} \). \( \square \)

**Lemma 2.2.** Let the set \( T \subseteq (k + 1) \wedge A \) be as defined above. Then \( d_G(v_t) \geq k + 3 \) for all \( t \in T \).

**Proof.** An element \( t \in T \) has at least two representations \( t = a_1 + \cdots + a_{k+1} = b_1 + \cdots + b_{k+1} \) as a \((k + 1)\)-sum of elements of \( A \). Furthermore the sets \( \{a_1, \ldots, a_{k+1}\} \) and
\{$b_1, \ldots, b_{k+1}\}$ cannot have precisely $k$ common elements (as in that case they would have different sums). It follows that the set $B = \{a_1, \ldots, a_{k+1}\} \cup \{b_1, \ldots, b_{k+1}\}$ has cardinality at least $k + 3$. The proof is now complete since $u_s v_t$ is an edge of $G$ for each $s \in \{t - b : b \in B\}$.

Combining Lemma 2.2 with the trivial bound $d_G(v_t) \geq d_H(v_t)$ for each $t \in T$, we deduce that
\[
d_G(v_t) \geq \frac{k + 1}{k + 3}(k + 3) + \frac{2}{k + 3}d_H(v_t) = k + 1 + \frac{2d_H(v_t)}{k + 3}.
\]

Consequently,
\[
e(G) \geq (k + 1)|Q_{k+1}| + \sum_{t \in T} \left( (k + 1) + \frac{2d_H(v_t)}{k + 3} \right) = (k + 1)((k + 1) \wedge A) + \frac{2e(H)}{k + 3}.
\]

The proof of the theorem is now nearly complete. Indeed, applying Lemma 2.1 we obtain the bound
\[
e(G) \geq (k + 1)((k + 1) \wedge A) + \frac{2(n - 2k)|S|}{k + 3},
\]
which combined with (3) yields
\[
(k + 1)((k + 1) \wedge A) + \frac{2(n - 2k)|S|}{k + 3} \leq (n - k)|k \wedge A| + k|S|.
\]

Now, since $n \geq (k^2 + 7k)/2$ the second term on the left hand side is at least the second term on the right hand side. Thus, $(k + 1)((k + 1) \wedge A) \leq (n - k)|k \wedge A|$, completing the proof of the inequality stated in Theorem 1.1.

In the case that $n > (k^2 + 7k)/2$, it follows from (6) that the equality $(k + 1)((k + 1) \wedge A) = (n - k)|k \wedge A|$ may only occur if $|S| = 0$. In this case the $k$-sums of $A$ are distinct, and so $|k \wedge A| = \binom{n}{k}$ and $|(k + 1) \wedge A| = (n - k)|k \wedge A|/(k + 1) = \binom{n}{k+1}$. This completes the proof of Theorem 1.1.

Acknowledgements This work began during the course of Imre Ruzsa on “Sumsets and Structure” [2] at the CRM, Barcelona. Thanks are due to Imre for his excellent course and stimulating problems, and to the organizers and the CRM for their hospitality.

References

[1] I. Ruzsa, Open problem included as Problem 30 (page 202) of the following reference.