Two combinatorial problems in group theory

by

R. B. Eggleton and P. Erdős (Calgary, Alberta)

Abstract. Sequences of elements from (additive) abelian groups are studied. Conditions under which a nonempty subsequence has sum equal to the group identity 0 are established. For example, an n-sequence with exactly k distinct terms represents 0 if the group has order \( g \leq n + \binom{k}{2} \) and \( n \geq k \binom{k}{2} \).

The least number \( f(k) \) of distinct partial sums is also considered, for the case of k-sequences of distinct elements such that no nonempty partial sum is equal to 0. For example, \( 2k - 1 < f(k) \leq \lfloor \frac{1}{2} k^2 \rfloor + 1 \).

In this paper a sequence is a selection of members of a set, possibly with repetitions, in which order is not important; elements are members of sets, and terms are members of sequences.

Definition. Let \(+\) be a binary operation on a set \( A \), and let \( S = (a_{1})_{i=1}^{n} \) be a sequence of elements from \( A \). \( S \) will be said to represent the element \( x \in A \) if

(i) \( x \) is a term in \( S \), or

(ii) there exist \( y, z \in A \) such that \( x = y + z \), and \( y \) and \( z \) are represented by disjoint subsequences of \( S \).

(Clearly this notion extends to general algebras.)

In particular, if \( \langle G, + \rangle \) is an abelian group and \( S = (a_{i})_{i=1}^{n} \) is a sequence of elements from \( G \), then \( S \) represents \( x \in G \) just if there exists a sequence \( E = (e_{i})_{i=1}^{m} \) of elements from \( \{0, 1\} \), not all 0, such that \( \sum_{i=1}^{m} e_{i} a_{i} = x \).

We resolve here some aspects of the following two related problems.

1. Under what circumstances does an n-sequence of elements from an abelian group represent the zero element?

2. If an n-sequence of distinct elements from an abelian group does not represent the zero element, how many elements does it represent?
Sequences representing zero.

**Theorem 1.** Any \( n \)-sequence \( S = (a_i)_{i=1}^{n} \) of elements from an abelian group \( (G, +) \), exactly \( k \) of which are distinct, represents the group identity 0 if the group has order \( g \leq n + \binom{n}{2} \) and \( n \geq k \binom{n}{2} \).

**Proof.** Suppose on the contrary that \( S \) does not represent 0. Then, none of the elements represented by the first \( m \) terms of \( S \) is 0, and none is equal to any of the \( n - m \) sums of the form \( \sum_{i=1}^{r} a_i \), with \( m + 1 \leq r \leq n \), for otherwise the difference would be a sum equal to 0. Again, none of these latter \( n - m \) sums equals 0, and all are distinct, for otherwise there would be a difference equal to 0, contrary to hypothesis.

We shall show that a suitable choice of \( m \) can be made, such that at least \( m + \binom{m}{2} \) elements are represented by the first \( m \) terms of \( S \), so with the latter \( n - m \) sums a total of at least \( n + \binom{n}{2} \) nonzero elements are represented. This is inconsistent with \( g \leq n + \binom{n}{2} \) so the initial hypothesis is false and the theorem follows.

We may suppose there are \( t \) equal terms in \( S \), say \( a_i = a_1 \) for \( 1 \leq i \leq t \), where \( kt \geq n \). Then \( S \) represents those elements equal to \( sa_1 \), for \( 1 \leq s \leq t \), which are necessarily distinct and different from 0. There are now two cases to consider: either (i) \( S \) has a term not in \( \{a_i\} \), the subgroup of \( G \) generated by \( a_1 \), or (ii) all terms of \( S \) are in \( \{a_i\} \).

Case (i). Suppose \( a_{i_1}, \ldots, a_{i_t} \). Then with \( m = t + 1 \), these first \( m \) terms of \( S \) must represent \( 2t + 1 \) distinct elements. If \( n \geq k \binom{k}{2} \), at least \( m + \binom{m}{2} \) distinct elements are represented, because \( kt \geq n \), which is what we require.

Case (ii). Let \( a_i = r_i a_1 \) for \( 1 \leq i \leq n \), where the sequence \( S' = (r_i)_{i=1}^{n} \) comprises positive integers, exactly \( k \) of which are distinct, and \( r_i = 1 \) for \( 1 \leq i \leq t \). (Since \( S \) does not represent 0, \( S' \) has no zero terms.) Regard \( S' \) as a sequence from the additive group of integers. If no term of \( S' \) exceeds \( t \), then \( S' \) represents all positive integers up to and including \( \sum_{i=1}^{n} r_i \), and this sum is at least as large as the sum of the first \( k \) positive integers together with a further \( n - k \) ones. Thus \( S' \) certainly represents \( g \) if \( g \leq n + \binom{n}{2} \), so \( S \) represents \( y_1 \), which is 0 since \( g \) is a multiple of the order of \( a_1 \). This is a contradiction, so \( S' \) must contain a term which exceeds \( t \), say \( r_{i_1} > t \). Again take \( m = t + 1 \) and repeat the argument of Case (i).

This theorem is best possible in the sense that the bound on \( g \) cannot be improved in general, for if \( a_i = i \) for \( 1 \leq i \leq k \) and \( a_{k+1} = 1 \) for \( k+1 \leq i \leq n \), then \( S \) represents all nonzero elements of the additive group of residue classes modulo \( n + \binom{n}{2} + 1 \), but does not represent 0. On the other hand, it is not clear what the best bound for \( n \) should be. If we take \( G \) to be the additive group of residue classes modulo \( 2s - 1 \times 4s \), where \( s \) is any positive integer, and \( S \) to comprise \( n = 3s \) terms specified by \( a_i = i \) for \( 1 \leq i \leq s \), \( a_{s+1} = 1 \) for \( s + 1 \leq i \leq 2s - 1 \), and \( a_{2s+1} = s - i \) for \( 2s \leq i \leq 3s \), then \( k = 2s + 1 \), and \( S \) does not represent 0. Since \( g = 2s^2 + 4s = n + \binom{n}{2} \) in this case, it follows that the bound on \( n \) in Theorem 1 could not be reduced as far as \( \frac{3}{4}(s - 1) \) in general. We conjecture that the theorem is true for \( n \geq c_k \), where \( c_k \) is some positive constant.

It is desirable to obtain a result corresponding to Theorem 1 for the case in which the exact number of distinct elements appearing in \( S \) is not known, the only relevant information being a lower bound on the number. We can deduce this result by using a theorem first conjectured by Erdős and Hajnal [9] and recently proved by Szemerédi [5], viz.,

**Theorem (Szemerédi).** Any \( h \)-sequence \( S = (a_i)_{i=1}^{n} \) of elements from an abelian group \( (G, +) \), all of which are distinct, represents the group identity 0 if the group has order \( g \geq g_0 \) and \( k \geq c_k \), where \( g_0 \) and \( c_k \) are absolute constants.

Thus, if the \( n \)-sequence \( S \) in Theorem 1 contains \( h \geq k \) distinct elements, and the order of \( G \) satisfies \( g_0 \leq g \leq n + \binom{n}{2} \), then the supposition that \( S \) does not represent 0 implies \( k \geq c_k \sqrt{n + \binom{n}{2}} \). If \( t \) is the number of terms of \( S \) equal to \( a_1 \), we may assume \( kt \geq n \), whence the argument used in Case (i) of the proof of Theorem 1 shows that the first \( m \) terms of \( S \) represent at least \( m + \binom{m}{2} \) distinct elements provided \( n \geq c_k t \), where \( c_k \) is an absolute constant. All other details of the proof carry over, so we have the

**Corollary to Theorem 1.** Any \( n \)-sequence \( S = (a_i)_{i=1}^{n} \) of elements from an abelian group \( (G, +) \), at least \( k \) of which are distinct, represents the group identity 0 if the order of the group satisfies \( g_0 \leq g \leq n + \binom{n}{2} \) and \( n \geq c_k t \), where \( g_0 \) and \( c_k \) are absolute constants.

It is possible to obtain similar results even when the number of distinct elements in \( S \) is not so small in comparison with \( n \).

**Theorem 2.** Any \( n \)-sequence \( S = (a_i)_{i=1}^{n} \) of elements from an abelian group \( (G, +) \), at least \( k \) of which are distinct, represents the group identity 0 if the group has order \( g \leq n + k - 1 \).

**Proof.** Suppose on the contrary that \( S \) does not represent 0. We may take the first \( k \) terms of \( S \) to be distinct. As will be shown in the
second part of this paper these \( k \) terms represent at least \( f(k) \) distinct elements, and for any \( k, f(k) \geq 2k - 1 \). None of the elements they represent can be equal to any of the \( n-k \) sums \( \sum_{i=1}^{r} a_{i} \), where \( k+1 \leq r \leq n \), for otherwise the corresponding difference equals 0 and \( S \) would represent 0. Similarly, no two of the latter \( n-k \) sums can be equal, so \( S \) represents at least \( n-k+1 \) distinct elements. Since \( S \) does not represent 0, this constitutes a contradiction if \( g \leq n-k+1 \), so the theorem follows. Indeed, it holds if \( g \leq n+f(k)-k \). \( \blacksquare \)

In a sense, the upper bound on \( g \) in Theorem 2 is low because of the structure of cyclic groups. This is clarified by the next result.

**Theorem 3.** Any \( n \)-sequence \( S = (a_{i})_{i=1}^{n} \) of elements from a noncyclic abelian group \( (G, +) \) represents the group identity 0 if the group has order \( g \leq 2n-1 \).

This may readily be deduced from the following result of Olson [1]:

**Theorem (Olson).** If \( H, K \) are abelian groups of order \( h, k \) respectively, and \( h\mid k \), then any \( n \)-sequence \( S = (a_{i})_{i=1}^{n} \) of elements from their direct sum \( G = H \oplus K \) represents the identity 0 of \( G \) if \( n \geq h+k-1 \).

**Proof of Theorem 3.** If \( G \) is a noncyclic abelian group of finite order, there is a direct sum decomposition

\[ G \cong \bigoplus_{e_{j} \in \mathbb{Z}} C_{e_{j}} \]

where \( C_{e_{j}} \) is the cyclic group of order \( e_{j}, m \geq 2 \) and \( e_{j+1} \mid e_{j} \), for \( 1 \leq j \leq m-1 \). With

\[ H \cong \bigoplus_{e_{j} \in \mathbb{Z}, m-1} C_{e_{j}} \quad \text{and} \quad K \cong C_{e_{m}} \]

we have

\[ h = \prod_{j=1}^{m-1} e_{j} \quad \text{and} \quad k = e_{m}, \quad \text{so} \quad k \mid h. \]

By Olson's theorem, \( S \) represents 0 \( \in G \) if \( n \geq h+k-1 \). Thus, it suffices to see that \( 2n-1 \geq g = hk \) ensures \( n \geq h+k-1 \). This is easy; for if at least one of \( h, k \) is even, we require \( \frac{1}{2} hk \geq h+k-2 \), so \( (h-2)(k-2) \geq 0 \), and if both \( h \) and \( k \) are odd, we require \( (h-2)(k-2) \geq 1 \), which conditions are satisfied because \( k \geq 2 \). \( \blacksquare \)

**Sequences not representing zero.** Let \( S = (a_{i})_{i=1}^{k} \) be a sequence of \( k \) distinct elements from an abelian group \( (G, +) \), such that \( S \) does not represent 0, and let \( f(k) \) denote the minimum number of elements which can be represented by \( S \), i.e.,

\[ f(k) = \min_{S, \sigma} |\{ x \in S : x \text{ is represented by } S \}|. \]

**Theorem 4.** \( f(k) \geq 2k-1 \) for \( k \geq 1 \).

**Proof.** Clearly \( f(1) = 1 \). For some \( k \geq 1 \), suppose \( f(k) \geq 2k-1 \), and let \( S = (a_{i})_{i=1}^{k+1} \) be a \((k+1)\)-sequence of distinct elements from \( G \) which does not represent 0.

**Case (i).** There is a term in \( S \) which is not represented by the remaining \( k \) terms. Then without loss of generality we assume \( a_{k+1} \) is such a term. The \( 2k-1 \) elements (or more) which are represented by the first \( k \) terms of \( S \) do not include \( a_{k+1} \), nor do they include \( \sum_{i=1}^{k} a_{i} \), for otherwise the difference between this sum and some other representation of the same element is 0, contradicting the fact that \( S \) does not represent 0. Hence \( S \) represents at least \( 2k+1 \) elements.

**Case (ii).** Every term in \( S \) is represented by the remaining \( k \) terms.

To resolve this case we use the theorem of Moser and Scherk [3], viz.

**Theorem (Moser and Scherk).** If \( A, B \) are finite sets of elements from an abelian group \( (G, +) \), such that \( 0 \not\in A, 0 \not\in B \), and \( a+b = 0 \), \( a \in A, b \in B \) implies \( a = -b \), then \( |A+B| \geq |A|+|B| - 1 \), where \( A+B = \{ a+b : a \in A, b \in B \} \).

Thus, if we let \( A = B = \{ 0, a_{1}, a_{2}, \ldots, a_{k+1} \} \), then \( |A+B| \geq 2k+3 \).

Under the assumptions of case (ii), every expression of the form \( 2a_{i} \) is expressible in the form \( a_{j} + \sum_{i=1}^{k} c_{i}a_{i} \), where \( c_{i} \in \{ 0, 1 \} \) for \( 1 \leq i \leq k+1 \), and not all the \( c_{i} \) are zero, but \( c_{k+1} = 0 \). This shows that every element of \( A+B \) other than 0 is represented by \( S \), yielding a total of at least \( 2k+2 \) elements represented by \( S \). The theorem now follows by induction on \( k \). \( \blacksquare \)

**Attainment of the bound for \( f(k) \) in Theorem 4, with \( k = 1, 2, 3 \), is shown by \( 1 \) (mod 2); \( 1, 2 \) (mod 4); \( 1, 3, 4 \) (mod 6).**

**Theorem 5.** \( f(k) \geq 2k+1 \) for \( k \geq 4 \).

**Proof.** The proof of Theorem 5 shows that in Case (i) if the first \( k \) terms of \( S \) represent at least \( 2k \) elements of \( G \), then \( S \) represents at least \( 2k+2 \) elements, while in Case (ii) this conclusion is invariably valid. Thus, the present theorem follows by induction on \( k \), provided it can be shown in Case (i) that if \( k = 3 \) and the first 3 terms of \( S \) represent only 5 elements of \( G \), nevertheless \( S \) represents at least 8 elements. Under these circumstances we may assume that \( a_{1}, a_{2}, a_{3}, a_{1}+a_{2}, a_{1}+a_{3}, a_{2}+a_{3} \) are all different, and \( a_{i} = a_{j}+a_{k} \), \( a_{i}+a_{j}+a_{k} \) so \( 2a_{i} = 0 \). (It is not possible to have further independent restrictions consistent with the conditions that \( a_{1}, a_{2}, a_{3} \) are distinct and do not represent 0.) Also, we may assume \( a_{1} \) is not represented by \( a_{2}, a_{3}, a_{1} \). Then, as before, \( a_{1}, a_{2}, a_{1}+a_{2}+a_{3}+a_{4} \) are distinct and are not represented by \( a_{1}, a_{2}, a_{3} \); the same is true for \( a_{1}+a_{2}+a_{4} \); for in particular \( a_{1}+a_{2}+a_{3} \) would imply
a_1 + a_2 + a_3 + a_4 = 0$, contradicting hypotheses concerning $S$. Thus, $S$ represents at least 8 elements of $\mathcal{B}$. The theorem follows. \[ \square \]

Attainment of the bound for $f(k)$ in Theorem 5, with $k = 4$, is shown by $1, 3, 4, 7 \pmod{9}$. In general, precise evaluation of $f(k)$ is increasingly laborious, even though entirely elementary. We have shown $f(5) = 13$. The proof is available as an appendix in [1]. Furthermore, $f(6) \leq 19$, and equality seems likely. (Computations in this direction are in progress.)

Szemerédi [5] can show $f(k) \geq 6k$, where $c$ is some positive constant.

On the other hand, $f(k) \leq \lfloor \frac{k}{3} \rfloor + 1$, as shown by the following two examples (where $s$ is any positive integer):

1. $a_i = i$ for $1 \leq i \leq s$, $a_i = s^2 + i$ for $s + 1 \leq i \leq 2s + 1 \pmod{2s^2 + 2s + 2}$, where $k = 2s + 1$, and the number of elements represented is $\frac{1}{2}k^2 + \frac{3}{4};$

2. $a_i = i$ for $1 \leq i \leq s$, $a_i = s^2 - s + 1$ for $s + 1 \leq i \leq 2s \pmod{2s^2 + 2}$, where $k = 2s$, and the number of elements represented is $\frac{1}{2}k^2 + 1.$

It is interesting to note that in all resolved cases, $f(k)$ can be achieved within the class of cyclic groups. We conjecture this to be the case for all $k$.

Finally, we remark that our theorems perhaps carry over to non-abelian groups, but we have no results in this direction.

References


The University of Calgary
Calgary, Alberta, Canada

Received on 6. 4. 1971