BOLYAI SOCIETY
MATHEMATICAL STUDIES

Combinatorics,
Paul Erdős is Eighty (Volume 1)
Keszthely (Hungary), 1993, pp. 159–172.

The Minimal Number of Zero Sums

Z. FÜREDI* and D. J. KLEITMAN†

Let $C = c_1, c_2, \ldots, c_n$ be an arbitrary sequence of integers. We asymptotically prove the following conjecture of Bialostocki: C contains at least $\binom{\lfloor n/2 \rfloor}{m} + \binom{\lceil n/2 \rceil}{m}$ subsequences of length m, whose terms add up to 0 modulo m. In the extremal case C contains only two distinct values (mod m), say 0's and 1's.

1. Preliminaries, results

The purpose of this paper is to give an asymptotic proof of a conjecture of Bialostocki concerning the minimal number of zero sum sets. Our starting point is the following old result of Erdős, Ginzburg and Ziv [7]. Let $c_1, c_2, \ldots, c_{2m-1}$ be a sequence of integers. Then there exists an m-element subset $I \subset \{1, 2, \ldots, 2m-1\}$ such that

$$\sum_{i \in I} c_i \equiv 0 \,(\text{mod } m). \tag{1.1}$$

A sequence is called a mod m *zero sum sequence* if the sum of its members is divisible by m. For an arbitrary sequence of integers, C, let $\mathcal{Z}_m(C)$ denote the family of m-element zero sum subsequences. If the value of m is clear,

* This research was done while the author visited the Department of Mathematics at Massachusetts Institute of Technology, Spring 1991.
Research was supported in part by the Hungarian National Science Foundation under grant No. 1812

† This work was supported in part by AF Grant OSR 0271 and NSF Grant DMS–86–06225

we call $\mathcal{Z}_m(C)$ the family of *zero sum sets* and abbreviate it as $\mathcal{Z}$ or $\mathcal{Z}(C)$. Recently, Bialostocki and Dierker [4] raised several interesting problems inspired by (1.1). For a list of the newest problems see Bialostocki [3], and for a survey of the new developments see Caro [6]. A pleasing conjecture of Bialostocki [2], (see in [5]) proposed: If $|C| = n$, then

$$|\mathcal{Z}_m(C)| \geq \binom{\lfloor n/2 \rfloor}{m} + \binom{\lceil n/2 \rceil}{m} \qquad (?) \tag{1.2}$$

Note that $\lfloor n/2 \rfloor$ 0's and $\lceil n/2 \rceil$ 1's show that the right-hand side cannot be increased. The case $m = 2$ is trivial, and the case $m = 3$ was verified by Dierker (see in [5]). The size of $\mathcal{Z}_m(C)$ is denoted by $z_m(C)$, and let $z_m(n) := \min\{z_m(C) : |C| = n\}$. Here we prove the following.

Theorem 1.1. *Conjecture* (1.2) *is true for the following cases:*

(a) *m is prime,*

(b) *for all m if $n > n_0(m)$,*

(c) *$m = 4$, or $n = pq$, where p and q are distinct primes.*

Bialostocki informed us that his conjecture was also proved, independently, by Kisin [9] for the cases $m = p^\alpha$ and $m = p^\alpha q$, where p and q are primes, and α is a positive integer. The first few unsolved cases are $m = 30, 36, 42, 60, \ldots$. Kisin also believes that it is unlikely that (1.2) is true for any m not in this form. However, our result (b) makes plausible that (1.2) is always true.

In the next section we give a short proof for the prime modulus. We do this for completeness, and also because we are going to use later not only the result but the proof, too. In Section 3, we use induction on m, (more exactly, an induction on the number of prime divisors of m), and prove the lower bound

$$z_m(n) \geq 2\binom{\lfloor n/2 \rfloor}{m} - m^2\binom{\lfloor n/2 \rfloor - 1}{m-1} \tag{1.3}$$

for all m and n. The proof is an extension of the original method of Erdős, Ginzburg and Ziv, using a proper matching of the entries in the sequence. In Section 4, with some technical steps, we prove (1.2) for $n > n_0(m)$. Our main technique is to distinguish between the homogeneous and non-homogeneous m-sets. Because of Kisin's result, we cut short the proof of (c) in Section 5, and conclude the paper with some problems and remarks in Section 6.

2. Proof of the prime case

To prove the theorem in the case m is a fixed prime we use induction on n. Obviously, $z_m(n) = 0$ for $n \leq 2m-2$. We may suppose that for every entry of $c \in C$, $0 \leq c < m$. The multiplicity of $c \in C$ is denoted by μ_c (or $\mu_c(C)$). Consider, first, the case when some multiplicity $\mu := \mu_c \geq n/2$. Then, such an element $c = c_i \in C$ belongs to at least $\binom{\mu-1}{m-1}$ zero sum sets. There are at least $z_m(n-1)$ zero sum sets avoiding c_i. We obtain

$$\begin{aligned} z_m(C) &\geq \binom{\mu-1}{m-1} + z_m(n-1) \\ &\geq \binom{\lceil n/2 \rceil - 1}{m-1} + \binom{\lfloor (n-1)/2 \rfloor}{m} + \binom{\lceil (n-1)/2 \rceil}{m} \\ &= \binom{\lceil n/2 \rceil}{m} + \binom{\lfloor n/2 \rfloor}{m}. \end{aligned}$$

Consider now the case when all multiplicities are at most $n/2$. Then, it is very easy to see, there exists a partition, $\mathcal{M}$, of C into $\lfloor n/2 \rfloor$ pairs and a singleton (in case of n is odd), such that every pair contains two distinct values. Such a partition is called a *proper* matching. A set R is *orthogonal* to the (disjoint) sets $M_1, M_2, \ldots$ if $|R \cap M_i| \leq 1$ for all i and $R \subset \cup M_i$. The following claim is the crucial step in the original proof of (1.1).

Claim 2.1. *Suppose that $\{x\} \cup M_1 \cup \ldots \cup M_{p-1} = S$ is a proper partition of $2p-1$ integers, where p is a prime. Then there is a zero sum set from $\mathcal{Z}_p(S)$ containing x and orthogonal to all parts.*

The proof of this claim follows from the following (easy) special case of the Cauchy-Davenport theorem (see, e.g., in [1]). For any subset $S \subset Z_p$, of the p-element cyclic group and distinct elements $s, t \in Z_p$ one has

$$|S + \{s,t\}| \geq \min\{p, |S|+1\}. \tag{2.1}$$

Proof of Claim 2.1. Suppose that every pair M_i consists of distinct integers mod p. Then for all integer y there is a p-element set, S, orthogonal to these pairs and containing x such that $\sum\{s : s \in S\} \equiv y \pmod{p}$. I.e., these 2^{p-1} sets produce all the p possible subset sums, including the 0. Let the set S_i be defined as all the possible 2^i sums mod p in the form $x_0 + x_1 + \ldots + x_i$, where $x_0 = x$ and $x_j \in M_i$. The inequality (2.1) easily implies by induction on i that

$$|S_i| \geq i+1, \tag{2.2}$$

hence $|S_{p-1}| = p$. ■

Claim 2.1 induces the required number of zero sum sets, all orthogonal to $\mathcal{M}$. Indeed, every m pairs from the matching contains at least 2 of them, this provides $2\binom{\lfloor n/2\rfloor}{m}$ distinct zero sum sets. If n is odd, the singleton from $\mathcal{M}$ carries additional $\binom{\lfloor n/2\rfloor}{m-1}$ zero sum sets. ■

3. Proof of the asymptotic result

The aim of this section is to prove (1.3) for all, not necessarily prime, m. We first give some definitions. Let C be a collection of integers. A partition of C is called an *almost proper* matching if it contains $\lfloor |C|/2\rfloor$ pairs and either it is a proper partition or all the $\lceil |C|/2\rceil$ parts have a common element i. A set $S \subset C$ is called *homogeneous* mod p (or briefly p-homogeneous) if $s_1 \equiv s_2 \pmod p$ for all $s_1, s_2 \in S$.

Proof of (1.3). We use induction on m. Suppose that p is a prime and $p|m$ (i.e., p divides m). Let C be a sequence of length n with minimal number of zero sum sets, $\mathcal{Z}_m(C) = z_m(n)$. Let $\mathcal{M}$ be an almost proper matching mod m with pairs $M_1, M_2, \dots, M_{\lfloor n/2\rfloor}$ such that $\mathcal{M}$ is almost proper mod p, too. This can be done as follows. Rearrange C such that that identical entries form intervals (i.e., C consists of at most m subintervals). Suppose further, that the p-homogeneous entries form (larger) intervals, and within such an interval the elements are in increasing order. (Finally, in the case n is odd, make sure that the longest p-homogeneous interval contains the middle element $c_{\lceil n/2\rceil}$.) Form the pairs (c_i, c_j) if $j = i + \lceil n/2\rceil$.

Let $\mathcal{P}$ be a family of $\lfloor n/2p\rfloor$ disjoint p-sets of pairs. One can choose exactly

$$\binom{\lfloor n/2\rfloor}{p}\binom{\lfloor n/2\rfloor - p}{p}\cdots\binom{\lfloor n/2\rfloor - p(\lfloor \frac{n}{2p}\rfloor - 1)}{p}\frac{1}{(\lfloor n/2p\rfloor)!} \tag{3.1}$$

such families $\mathcal{P}$.

By Claim 2.1, for every $P \in \mathcal{P}$, $\cup P$ contains at least two zero sum sets mod p (not mod m, yet), orthogonal to $\mathcal{M}$. Denote two of these by $S(P,1)$ and $S(P,2)$, and let their sums be $s(P,1)$ and $s(P,2)$, respectively, i.e., $s(P,\varepsilon) := \sum\{c : c \in S(P,\varepsilon)\}$. There are

$$2^{\lfloor n/2p\rfloor} \tag{3.2}$$

choices of $\varepsilon \in \{1,2\}^{\mathcal{P}}$ to obtain a sequence of p-element zero sum sets $S(P,\varepsilon(P)) : P \in \mathcal{P}$. Consider the following sequence of integers of the same length: $s(P,\varepsilon(P))/p : P \in \mathcal{P}$. It contains at least $z_{m/p}(\lfloor n/2p \rfloor)$ zero sum sets mod m/p. These zero sum sets induce m-element zero sum sets mod m in C, which are unions of the sets $S(P,\varepsilon)$. In this way, altogether, each family determined by a given $\mathcal{P}$ and ε induces at least

$$z_{m/p}(\lfloor n/2p \rfloor) \tag{3.3}$$

zero sum sets from $\mathcal{Z}_m(C)$. An m-element set which is orthogonal to $\mathcal{M}$ can be obtained in at most

$$\begin{aligned}&2^{\lfloor n/2p \rfloor-(m/p)} \times \binom{m}{p}\binom{m-p}{p}\cdots\binom{p}{p}\frac{1}{(m/p)!}\times \\ &\binom{\lfloor n/2 \rfloor - m}{p}\binom{\lfloor n/2 \rfloor - m - p}{p}\cdots\binom{\lfloor n/2 \rfloor - p(\lfloor \frac{n}{2p} \rfloor - 1)}{p}\times \\ &\frac{1}{(\lfloor n/2p \rfloor - (m/p))!}\end{aligned} \tag{3.4}$$

different ways by the above method. Dividing the products of (3.1)–(3.3) by (3.4) we obtain

$$\begin{aligned} z_m(n) = |\mathcal{Z}_m(C)| &\geq 2^{m/p}\binom{\lfloor n/2 \rfloor}{m}\frac{z_{m/p}(\lfloor n/2p \rfloor)}{\binom{\lfloor n/2p \rfloor}{m/p}} \\ &= \binom{\lfloor n/2 \rfloor}{m} \times \frac{z_{m/p}(\lfloor n/2p \rfloor)}{\binom{\lfloor n/4p \rfloor}{m/p}} \times \frac{2^{m/p}\binom{\lfloor n/4p \rfloor}{m/p}}{\binom{\lfloor n/2p \rfloor}{m/p}} \end{aligned} \tag{3.5}$$

Using the induction hypothesis, i.e., (1.3), (with m/p instead of m) we obtain that for a fixed m the limit of the middle factor is 2 whenever $n \to \infty$, while the limit of the third factor is 1. The rest of the proof of (1.3) is a simple calculation. ■

4. The exact value for large n

Suppose that p is a prime and $p|m$. An m-element set, A, has $m!/(p!^{m/p}(m/p)!)$ partitions into p-sets. If the set A itself is not p-homogeneous, then not all of these partitions consist of zero sum sets modulo p. Denote by $\delta(S,p)$ the fraction of partitions having a non zero sum set part, and let $\delta(m,p) = \min\{\delta(S,p)\}$, where the minimum is taken over all integer sequences of length m. Obviously, $\delta > 1/m!$, but one can get a better bound ($\delta > 1/m^2$) by observing that each partition of zero sum sets can be spoiled by exchanging 2 elements.

Split $\mathcal{Z}_m(C)$ into three parts. Let $\mathcal{R}^{p-\hom}$ denote those members $A \in \mathcal{Z}_m(C)$ which are orthogonal to the matching $\mathcal{M}$ and homogeneous mod p; let $\mathcal{R}'$ denote the non p-homogeneous zero sum sets mod m orthogonal to $\mathcal{M}$, and let $\mathcal{R} = \mathcal{R}(\mathcal{M}) = \mathcal{R}^{p-\hom} \cup \mathcal{R}'$; finally, let $\mathcal{Q} = \mathcal{Z} \setminus \mathcal{R}$. Suppose that

$$|\mathcal{Z}_m(C)| \le \binom{\lfloor n/2 \rfloor}{m} + \binom{\lceil n/2 \rceil}{m}, \tag{4.1}$$

otherwise there is nothing to prove. We are going to show that equality holds in (4.1). In the calculations in the previous Section we considered only $\mathcal{R}$, and proving the lower bound we have counted the members of $\mathcal{R}'$ by a weight $1-\delta$ only. We had

$$|\mathcal{R}^{p-\hom}| + |\mathcal{R}'|(1-\delta) \ge 2\binom{\lfloor n/2 \rfloor}{m} - m^2\binom{\lfloor n/2 \rfloor - 1}{m-1}.$$

This inequality and (4.1) give, that

$$|\mathcal{R}'| + |\mathcal{Q}| < (1/\delta)(m^2+1)\binom{\lfloor n/2 \rfloor - 1}{m-1} = O(n^{m-1}). \tag{4.2}$$

Lemma 4.1. *Let r_1, r_2, r_3 and k be four distinct nonnegative integers such that $k \ge 3$ and $0 \le r_i < k$ for all i. Then, there exist positive integers a_1, a_2, a_3 such that $a_1 + a_2 + a_3 = k$ and $a_1 r_1 + a_2 r_2 + a_3 r_3 \equiv 0 \pmod{k}$.*

Proof of the Lemma. (Easy.) The statement is independent from mod k shifting, i.e., instead of r_1, r_2 and r_3 we can consider $r_1 + r$, $r_2 + r$, $r_3 + r$ mod k. So we may suppose that one of the r_i's is 0, say $r_3 = 0$. For $r_1 < r_2$ define the a_i's as follows, $a_1 = k - r_2$, $a_2 = r_1$, $a_3 = r_2 - r_1$. ■

Let $W(j) = \{c \in C : c \equiv j \pmod{m}\}$. The next Proposition will enable us to prove that most of these residue classes are relatively small.

Proposition 4.2. *Suppose that $W(r_1)$, $W(r_2)$ and $W(r_3)$ are residue classes of C such that the set $\{r_1, r_2, r_3\}$ is not p-homogeneous. Then $\min_{1 \le j \le 3} |W(r_j)| < mn^{(m-1)/m}$.*

Proof. Suppose, on the contrary, that for $j = 1, 2, 3$ we have $|W(r_j)| > m^2 n^{(m-1)/m}$. Apply Lemma 4.1 for the (distinct) numbers r_1, r_2, r_3 and

$k =: m$. We get the positive integers a_1, a_2, a_3 such that $a_1 + a_2 + a_3 = m$ and $\sum_j a_j r_j \equiv 0 \pmod m$. Consider the m-sets containing exactly a_j elements from the corresponding residue classes. This way we have got at least

$$\binom{|W(r_1)|}{a_1}\binom{|W(r_2)|}{a_2}\binom{|W(r_3)|}{a_3} > m^4\binom{\lfloor n/2\rfloor}{m-1}$$

zero sum sets mod m. These sets are not p-homogeneous, all belong to $\mathcal{R}' \cup \mathcal{Q}$. Thus the above lower bound contradicts the upper bound in (4.2). ∎

Lemma 4.3. *Let V be a sequence of integers of length v. Suppose that $1 < d \leq m$ are integers, $d|m$, and V is d-homogeneous. Then*

$$|\mathcal{Z}_m(V)| > 2^{1+m-(m/d)}(1 - \frac{m^3}{v})\binom{\lfloor v/2\rfloor}{m} \tag{4.3}$$

Proof. Almost identical to the proof of (1.3). Make an arbitrary matching, but now every d pairs carries 2^d orthogonal zero sum sets mod d (instead of only two; the sets $S(P, \varepsilon)$ in that proof). For simplicity, suppose that $v/4d$ is an integer. The product of (3.1)–(3.3) divided by (3.4) becomes

$$\frac{\dfrac{(v/2)!}{(d!)^{v/2d}(v/2d)!} \times (2^d)^{v/2d} \times z_{m/d}(v/2d)}{(2^d)^{(v/2d)-(m/d)} \times \dfrac{m!}{(d!)^{m/d}(m/d)!} \times \dfrac{(v/2-m)!}{(d!)^{(v/2d)-(m/d)}((v/2d)-(m/d))!}}$$

Using the lower bound (1.3) for $z_{m/d}$ we get (4.3). The details are left to the reader. ∎

Let $V(i) = \{c \in C : c \equiv i \pmod p\}$ and let $|V(a)|$ be the largest among these. Our next aim is to prove that $V(a)$ does not cover almost all of C, we have

$$|V \setminus V(a)| > n/4. \tag{4.4}$$

We prove this by showing that $|V(a)| < 2^{1/p}n/2 + O(m^3)$. Indeed, apply Lemma 4.3 to the sequence $V = V(a)$, $v = |V(a)|$, $d := p$; we obtain at least $(1 - o(1))2^{1-(m/p)}v^m/m!$ zero sum sets mod m, then (4.1) implies the desired upper bound on the size of $V(a)$.

Let $W(x) \subset V(a)$ be the largest residue class mod m in it, let $W(y)$ be the largest residue class mod m such that $W(y) \cap V(a) = \emptyset$ and let $W = W(x) \cup W(y)$. We claim, that almost all elements of C belong to $W(x)$ or $W(y)$,

$$|C \setminus W| \leq m^2 n^{(m-1)/m} = O(n^{(m-1)/m}). \tag{4.5}$$

Indeed, $|W(x)| \geq |V(a)|/m \geq n/m$ by definition, and $W(y)$ has at least $n/4m$ elements by (4.4). Then Proposition 4.2 implies that all the other residue classes are small.

All the m-subsets of $W(x)$ and $W(y)$ belong to $\mathcal{Z}$, so (4.1) implies that $\binom{|W(x)|}{m} + \binom{|W(y)|}{m} \leq \binom{\lfloor n/2 \rfloor}{m} + \binom{\lceil n/2 \rceil}{m}$. From (4.5) we have that $|W(x)| + |W(y)|$ is almost n. Using the inequality

$$\binom{t+s}{m} + \binom{t-s}{m} \geq \left(2 + \frac{s^2 m(m-1)}{t^2}\right)\binom{t}{m}$$

which holds for all $t > 2s \geq 0$, after a short calculation we obtain that both residue classes have sizes about $n/2$, more exactly

$$|W(x) - \frac{n}{2}| < \frac{1}{2} n^{1-(1/2m)} \tag{4.6}$$

and the same holds for $|W(y)|$, too.

The family $\mathcal{Z}_m(C)$ is unchanged if we add mod m the same amount to each member of C, or if we multiply each entry by the same number b, where g.c.d.$(b, m) = 1$ (greatest common divisor is 1, i.e., relatively prime numbers). So we may suppose, that, say $x = 0$. As we have seen in the proof of (4.4), a d-homogeneous set (where $d > 1$ is a divisor of m) cannot be larger than $3n/4$. As $|W(0) \cup W(y)| = (1 - o(1))n$, by (4.5), we obtain that y and m are relatively prime.

It is well-known, that there exists a b, relatively prime to m, too, such that $by \equiv 1 \pmod{m}$. Multiply each entry in C by this b; so from now on we may suppose that $x = 0$ and $y = 1$.

Considering the zero sum sets in W we obtain

$$\begin{aligned} |\mathcal{R}^{p-\text{hom}}| &\geq \binom{|W(0)|}{m} + \binom{|W(1)|}{m} \\ &\geq \binom{\lfloor n/2 \rfloor}{m} + \binom{\lceil n/2 \rceil}{m} - (n - w_0 - w_1)\binom{\lfloor n/2 \rfloor}{m-1}, \end{aligned} \tag{4.7}$$

where $w_j = |W(j)|$. Suppose that $C \setminus W \neq \emptyset$, and let $c \in C \setminus W$, we have $2 \leq c \leq m-1$. If we select additional $m - c$ elements from $W(1)$ and $c - 1$ elements from $W(0)$, we get a zero sum set mod m. Using (4.6) we obtain

$$\begin{aligned} \binom{w_1}{m-c}\binom{w_0}{c-1} &> (1 - o(1))\binom{\lfloor n/2 \rfloor}{m-c}\binom{\lfloor n/2 \rfloor}{c-1} \\ &\sim (1 - o(1))\binom{\lfloor n/2 \rfloor}{m-1}\binom{m-1}{c-1} \\ &> \frac{m}{2}\binom{\lfloor n/2 \rfloor}{m-1} \end{aligned}$$

zero sum sets containing the element c and having $m-1$ elements in W. Considering all $c \in C \setminus W$ we get, that the number of not p-homogeneous zero sum sets is at least $(n-w_0-w_1)(m/2)\binom{\lfloor n/2 \rfloor}{m-1}$. This and (4.7) imply that

$$|\mathcal{Z}| \geq \binom{\lfloor n/2 \rfloor}{m} + \binom{\lceil n/2 \rceil}{m} + (\frac{m}{2}-1)(n-w_0-w_1)\binom{\lfloor n/2 \rfloor}{m-1}.$$

This inequality and (4.1) give that $n - w_0 - w_1 = 0.$, i.e., C has only 0's and 1's. ■

5. Exact results for some composite moduli

We consider only the case n is even, the odd n requires only a few more technical steps. Suppose that the elements of the sequence C can be arranged into a matching $\mathcal{M}$ of size $|C|/2$, such that every m pairs contain 2 orthogonal zero sum sets mod m. Then, obviously, $z_m(C) \geq 2\binom{n/2}{m}$. Call such a matching *m-perfect.*

Theorem 5.1. *There is an m-perfect matching of any sequence C of even length, n, for $m = 4$, for $m = p$, and for $m = pq$, where p and q are distinct primes.*

Proof. We consider the case $m = pq$ only. Define the matching $\mathcal{M}$ of size $n/2$ as follows. Suppose that $\mathcal{M}$ contains the minimal number of identical pairs, and among these matchings suppose that $\mathcal{M}$ contains the maximum number of pairs with relatively prime entries. We claim that such an $\mathcal{M}$ is pq-perfect. Consider any m pairs, $\mathcal{P} = \{(a_i, b_i) : 1 \leq i \leq m\}$. The construction of $\mathcal{M}$ is hereditary, $\mathcal{P}$ has the same properties, i.e., among the $(2m)!2^{-m}(m!)^{-1}$ matchings formed from the sequence $a_1, \ldots, b_m$ $\mathcal{P}$ has the minimal number of identical pairs, and beside that the maximum number of pairs with relatively prime entries.

If there are m identical entries, say $a_1 = \ldots = a_m$, then these a's form the first zero sum set. Another one, S, is contained in the $2m-1$ element set $(\cup\mathcal{P}) \setminus \{a_1\}$, by (1.1). If S is not orthogonal to $\mathcal{P}$, then we can replace it by an orthogonal one, S', such that $S \cap \{b_1, \ldots, b_m\} = S' \cap \{b_1, \ldots, b_m\}$. From now on, we may suppose that $a_i \neq b_i$ for all i (in $\mathcal{P}$).

Proposition 5.2. *Let A be a sequence of integers of length 4. Then either A is p-homogeneous, or q-homogeneous, or contains two members whose difference is relatively prime to pq.*

Proof. (Easy.) Define the graph $\mathcal{G}^r$ ($r \in \{p, q\}$) with vertex set A and a pair x, y joined by an edge if $x \equiv y \pmod r$. If this graph is connected, then A is r-homogeneous. If neither $\mathcal{G}^p$, nor $\mathcal{G}^q$ are connected, then the union of their edge sets does not cover all the 6 pairs from A. ■

Split $\mathcal{P}$ into three classes, $\mathcal{P} = \mathcal{P}^p \cup \mathcal{P}^q \cup \mathcal{P}^1$, where $\mathcal{P}^r = \{(a_i, b_i) :$ g.c.d.$(a_i - b_i, m) = r$. Consider two pairs, $(a_i, b_i), (a_j, b_j) \in \mathcal{P}^p$. If there is a repetition, say $a_i = a_j$, then all the four elements are p-homogeneous. Otherwise, by Proposition 5.2, one find a pair, say (a_i, a_j), such that g.c.d.$(a_i - a_j, m) = 1$. Replacing the original pairs by (a_i, a_j) and (b_i, b_j) we get a matching $\mathcal{Q}$ with $|\mathcal{Q}^1| > |\mathcal{P}^1|$, a contradiction. So we may suppose, that there are numbers $h^{(p)}$ and $h^{(q)}$ such that $x \equiv h^{(r)} \pmod r$ all $x \in \cup\mathcal{P}^r$ for $r \in \{p, q\}$.

Suppose that $\mathcal{P}^q \neq \emptyset$, $(a_1, b_1) \in \mathcal{P}^q$. We claim, that

$$\text{either } x \equiv h^{(q)} \pmod q \text{ or } y \equiv h^{(q)} \pmod q \tag{5.1}$$

holds for all pairs $(x, y) \in \mathcal{P}$.

Indeed, if not, then, for $(x, y) \in \mathcal{P}^p$ in the case of four distinct elements (namely, a_1, b_1, x and y) by exchanging two entries one can create a new pair for $\mathcal{P}^1$ by Proposition 5.2. In the case $(x, y) \in \mathcal{P}^1$ only two disjoint pairs between (a_1, b_1) and (x, y) can belong to $\mathcal{G}^p$, so the other matching consists of two pairs from $\mathcal{P}^1$, our final contradiction proving (5.2). So we may suppose that the elements a_i form a q-homogeneous set.

First, suppose that $|\mathcal{P}^q| \geq p$, $(a_i, b_i) \in \mathcal{P}^q$ for $1 \leq i \leq p$. Then, one can find two zero sum sets mod m orthogonal to these p pairs and containing the rest of the a's, i.e., the set $\{a_i : p < i \leq m\}$. To see this we need to consider the $p - 1$ pairs formed by $a_i' = (a_i - h^{(q)})/q$ and $b_i' = (b_i - h^{(q)})/q$ and the element $x = \sum_{p \leq i \leq m}(a_i - h^{(q)})/q$. Then Claim 2.1 implies the existence of a zero sum set mod p which corresponds to a zero sum set mod pq containing the element a_p. One can construct another zero sum set containing b_p, so this case is done.

Finally, suppose that $|\mathcal{P}^q| \leq p - 1$, and similarly, $|\mathcal{P}^p| \leq q - 1$. Then $|\mathcal{P}^1| \geq p + q - 2$ (because $(p-2)(q-2) \geq 0$). Reorder the pairs in such a way, that the pair (a_i, b_i) belongs to $\mathcal{P}^1$ for all i with $p|i$ and $q|i$. Then, by an induction on i, as in the proof of Claim 2.1, one can see that (2.2) holds, i.e., the 2^{m-1} m-sets orthogonal to $\mathcal{P}$ (and containing a fixed element, say a_m) produce all the m subset sums. This includes a zero sum set, too. This completes the proof of case $m = pq$. ■

Unfortunately, this pairing method cannot lead to a full solution of (1.2), (without a further idea), as it is shown by the following counterexamples.

Example 5.1. *If m has 3 distinct primę divisors p, q and r, or if m is divisible by p^2q^2, then, for $n \geq 16m$, there exists an integer sequence C of length n such that C has no m-perfect matching.*

Let C consist of four distinct entries 0, pq, pr and qr, with multiplicities $\lfloor n/4 \rfloor$ and $\lceil n/4 \rceil$. We claim that this C has no any m-perfect matching. Suppose, on the contrary, that the matching $\mathcal{M}$ is m-perfect. We distinguish two cases.

If the pair $(0,0)$ has multiplicity at least $m-1$ in $\mathcal{M}$, then there exists a pair (x,y) with $x, y \neq 0$. Then the m element matching containing (x,y) and $m-1$ pairs of $(0,0)$ has no orthogonal zero sum set mod m.

Otherwise, there is another value from C, say pq, such that the multiplicity of $(0,pq)$ is at least $m-1$. There exists a pair in $\mathcal{M}$, (x,y), such that $(x,y) \cap (0,pq) = \emptyset$. Then the m element matching containing (x,y) and $m-1$ pairs of $(0,pq)$ has no orthogonal zero sum set mod m. This is so, because the equation $ipq + pr \equiv 0 \pmod{pqr}$ has no integer solution for i (since q does not divide pr), and neither has the equation $ipq + qr \equiv 0 \pmod{pqr}$.

In the case $p^2q^2|m$, we can consider a similar sequence, $n/4$ 0's, p^2's, q^2's and pq's. The proof is similar and omitted. ■

6. Generalizations and Questions

The first problem is to decrease the value of $n_0(m)$. We did not make a serious effort, and got something like $n_0(m) < m^{6m}$. We strongly believe, that conjecture (1.2) is true for all $n > 4m$.

Let C be a sequence of integers of length n, and suppose that it is not m-homogeneous. What is the *maximum* number of m-element zero sum sets?

Suppose that $p|m$ and A is a not a p-homogeneous sequence of length m. What is the minimum of $\delta(A,p)$, i.e., in other words, what is the maximum number of partitions consisting of p-element zero sum sets mod p? We think that δ should be close to 1, $\delta > 1 - O(1/m)$.

Consider the following polynomial of n variables

$$p(x_1, \dots, x_m) = \sum \binom{x_{i_1}}{j_1} \times \dots \times \binom{x_{i_t}}{j_t},$$

where the sum is taken for all $1 \le i_k \le m$, $1 \le j_k \le m$, such that $\sum j_k = m$ and $\sum i_k j_k \equiv 0 \pmod{m}$. Clearly, the minimum of $p(x_1, \dots, x_m)$ over $\sum x_i = n$, where every x_i is a non-negative integer, is exactly our desired $z_m(n)$. This was the way Dierker (see in [5]) established the case $m = 3$. This approach might lead to a full solution, for example one might use the fact that $p(x_1, x_2, \dots, x_m) = p(x_2, x_3, \dots, x_1)$.

In Sections 3 and 4 we, in fact, characterized the extremal sequences. Namely, for $n > n_0(m)$, if equality holds in (1.2), then the sequence C consists of two values only, $\lfloor n/2 \rfloor$ a's and $\rceil n/2 \rceil$ b's such that $b - a$ and m are relatively prime.

The above questions can also be formulated for other finite groups, not just cyclic, for abelian and non-abelian. For example, Olson ([12], see e.g. in [1]) proved the following generalization of the Erdős-Ginzburg-Ziv Theorem. Let $g_1, g_2, \dots, g_{2m-1}$ be a sequence of $2m - 1$ elements of a finite (but not necessarily abelian) group of order m. Then there is sequence of m terms such that $g_{i_1} + g_{i_2} + \dots + g_{i_m} = 0$, however, here not necessarily $i_1 < i_2 < \dots < i_m$. He conjectures, that one can find a subsequence, too.

The Erdős-Ginzburg-Ziv theorem was rediscovered, even published, several times. For a recent example, see [16].

Concerning zero sum sets of arbitrary sizes, Olson [11] proved the following. Let G be an abelian group such that every sequence $g_1, g_2, \dots, g_{\ell+1}$ of length $\ell + 1$ contains a (non-empty) zero sum sets. Then, every sequence of length $\ell + t$ contains at least $2^t - 1$ non-empty zero sum sets. He also determined $\ell = \ell(G)$ for a large class of groups.

These questions seem to be somewhat related to the following conjecture of Manickam and Miklós [10]. Let $C = c_1, \dots, c_n$ be a sequence of integers with $\sum c_i \ge 0$. Let $A(C, m)$ denote the number of m-element subsets with nonnegative sums, and let $A(n, m) = \min A(C, m)$. If m is not a divisor of n, then $A(n, m) = \min\{\binom{n-1}{m-1}, \binom{n - \lfloor n/m \rfloor}{m}\}$, and $A(n, m) = \binom{n-1}{m-1}$ for $m | n$. (In particular, they conjecture $A = \binom{n-1}{m-1}$ for all $n \ge 4m$.) They proved this for $n > m^{m+1}$.

A system of m-element subsets (called *blocks*) of an n-element set C is called a Turán (n,k,m)-*system* if every k-element subset of C contains at least one of the blocks. The Turán number $T(n,k,m)$ is the minimal size of such a system. The Erdős-Ginzburg-Ziv theorem implies

$$z_m(n) \geq T(n, 2m-1, m). \tag{6.1}$$

Indeed, $T(n,3,2) = \binom{\lfloor n/2 \rfloor}{2} + \binom{\lceil n/2 \rceil}{2}$, by Turán's theorem about the triangles. Turán conjectured (in about 1961) that $T(n,5,3) = \binom{\lfloor n/2 \rfloor}{3} + \binom{\lceil n/2 \rceil}{3}$. However, this was disproved for $n = 9$ by Surányi [15] and for all odd $n > 9$ by Sidorenko and Kostochka [13]. For even n's, the conjecture might be true. For $m > 3$, the real order of magnitude of $\lim_{n\to\infty} T(n, 2m-1, m)\binom{n}{m}^{-1}$ is not even conjectured. A recent survey of this topic is [14], and about the more general Turán type problems [8].

Acknowledgement. The authors are indebted to A. Bialostocki (Moscow, Idaho) and J. Kahn (Rutgers, NJ) for several fruitful conversations.

References

[1] N. Alon, Tools from higher algebra, to appear in: *Handbook in Combinatorics* (eds.: R. L. Graham, A. Grötschel and L. Lovász), North-Holland.

[2] A. Bialostocki, Some combinatorial number theory aspects of Ramsey theory, *research proposal,* 1989.

[3] A. Bialostocki, Zero sum trees: a survey of results and open problems, submitted to the *Proc. Conf. of Finite and Infinite Combinatorics in Sets and Logic,* Banff, Canada, 1991.

[4] A. Bialostocki and P. Dierker, Zero sum Ramsey theorems, *Lecture on 20th Southeastern Conf. on Combinatorics, Graph Theory and Computing,* Boca Raton, Fla., 1989, *Congressus Numerantium* **70**(1990), 119-130.

[5] A. Bialostocki and M. Lotspeich, Some developments of the Erdős-Ginzburg-Ziv theorem I., in: *Sets, Graphs and Numbers,* (eds.: G. Halász et al.), Colloq. Math. Soc. J. Bolyai, **60**, North-Holland, Amsterdam, 1992, 97–117.

[6] Y. Caro, Zero-sum problems – a survey, manuscript, April, 1991.

[7] P. Erdős, A. Ginzburg, and A. Ziv, A theorem in additive number theory, *Israel (State) Research and Development National Council Bulletin, Section F (Mathematics and Physics),* **10**(1961), 41–43.

[8] Z. Füredi, Turán type problems, in: *Surveys in Combinatorics. 1991,* (Proc. of the 13th British Combinatorial Conference), (ed.: A. D. Keedwell), *London Math. Soc. Lecture Note Series,* **166** Cambridge Univ. Press, (1991), 253–300.

[9] M. Kisin, The number of zero sums modulo m in a sequence of length n.

[10] N. Manickam and D. Miklós, On the number of non-negative partial sums of a non-negative sum, in: *Combinatorics,* (eds.: A. Hajnal et al.), *Colloq. Math. Soc. J. Bolyai,* **52**, North-Holland, Amsterdam 1988, 385–392.

[11] J. E. Olson, A combinatorial problem on finite Abelian groups, II, *J. Number Theory* **1**(1969), 195–199.

[12] J. E. Olson, On a combinatorial problem of Erdős-Ginzburg-Ziv, *J. Number Theory* **8**(1976), 52–57.

[13] A. F. Sidorenko, The Turán problem for 3-graphs, (Russian), in: *Combinatorial Analysis,* **6**(1983), 51–57.

[14] A. F. Sidorenko, What we know and what we do not know about Turán numbers, *Graphs and Combinatorics,* to appear.

[15] J. Surányi, Some combinatorial problems of geometry, (Hungarian), *Mat. Lapok* **22**(1971), 215–230.

[16] see in *Amer. Math. Monthly* **96**(1989), 240–242.

Zoltán Füredi

Dept. Math. , Univ. Illinois,
Urbana, IL 61801–2917,
USA

e-mail: zoltan@math.uiuc.edu

and

Math. Inst. of the
Hungarian Academy of Sciences,
P.O.B. 127, Budapest 1364
Hungary

e-mail: h1154fur@ella.hu

Daniel J. Kleitman

Dept. of Math.,
Massachusetts Inst. of Tech.
Cambridge, MA 02139–4307,
USA

e-mail: djk@math.mit.edu