Proc. Indian Acad. Sci. (Math. Sci.) Vol. 115, No. 1, February 2005, pp. 67–77.

On the structure of p-zero-sum free sequences and its application to a variant of Erdös–Ginzburg–Ziv theorem

W D GAO, A PANIGRAHI* and R THANGADURAI*

Department of Computer Science and Technology, University of Petroleum, Changping Shuiku Road, Beijing 102200, China
*School of Mathematics, Harish-Chandra Research Institute, Chhatnag Road, Jhusi, Allahabad 211 019, India
E-mail: wdgao_1963@yahoo.com.cn; anupama@mri.ernet.in; thanga@mri.ernet.in

MS received 8 September 2004

Abstract. Let p be any odd prime number. Let k be any positive integer such that $2 \le k \le \left[\frac{p+1}{3}\right] + 1$. Let $S = (a_1, a_2, \ldots, a_{2p-k})$ be any sequence in $\mathbb{Z}_p$ such that there is no subsequence of length p of S whose sum is zero in $\mathbb{Z}_p$. Then we prove that we can arrange the sequence S as follows:

$$S = (\underbrace{a, a, \ldots, a}_{u \text{ times}}, \underbrace{b, b, \ldots, b}_{v \text{ times}}, a'_1, a'_2, \ldots, a'_{2p-k-u-v})$$

where $u \ge v$, $u + v \ge 2p - 2k + 2$ and $a - b$ generates $\mathbb{Z}_p$. This extends a result in [13] to all primes p and k satisfying $(p+1)/4 + 3 \le k \le (p+1)/3 + 1$. Also, we prove that if g denotes the number of distinct residue classes modulo p appearing in the sequence S in $\mathbb{Z}_p$ of length $2p - k$ ($2 \le k \le [(p+1)/4] + 1$), and $g \ge 2\sqrt{2}\sqrt{k-2}$, then there exists a subsequence of S of length p whose sum is zero in $\mathbb{Z}_p$.

Keywords. Sequences; zero-sum problems; zero-free; Erdös–Ginzburg–Ziv theorem.

1. Introduction

Let n be any positive integer. Let $S = (a_1, a_2, \ldots, a_\ell)$ be a sequence (possibly with repetition) in the cyclic group of order n (denoted by $\mathbb{Z}_n$) of length ℓ. We call a subsequence $T = (b_1, b_2, \ldots, b_r)$ of S to be zero-sum subsequence if $b_1 + b_2 + \cdots + b_r = 0$ in $\mathbb{Z}_n$.

In 1961, Erdös–Ginzburg–Ziv proved the following theorem (which we call the EGZ theorem).

EGZ Theorem [8]. *Given a sequence S in $\mathbb{Z}_n$ of length $2n-1$, one can extract a zero-sum subsequence of length n in $\mathbb{Z}_n$.*

The EGZ theorem is tight in the following sense. If

$$S = (\underbrace{0, 0, \ldots, 0}_{n-1 \text{ times}}, \underbrace{1, 1, \ldots, 1}_{n-1 \text{ times}})$$

is a sequence in $\mathbb{Z}_n$ of length $2n - 2$, then S does not have a zero-sum subsequence of length n.

Many authors studied the characterization of the above extremal example. In particular, Yuster and Peterson [18] and independently Bialostocki and Dierker [1] proved that any sequence S in $\mathbb{Z}_n$ of length $2n-2$ having no zero-sum subsequence of length n will be of the form

$$S = (\underbrace{a, a, \ldots, a}_{n-1 \text{ times}}, \underbrace{b, b, \ldots, b}_{n-1 \text{ times}}),$$

where $a \neq b \in \mathbb{Z}_n$.

Also, Flores and Ordaz [9] proved the following result of this nature. Suppose S is any sequence in $\mathbb{Z}_n$ of length $2n-3$ such that S has no zero-sum subsequence of length n. Then there exists $a, b \in \mathbb{Z}_n$ such that $\mathbb{Z}_n$ is generated by $b-a$ and a appearing $n-1$ times in S and one of the following conditions hold: (i) b appearing exactly $n-2$ times; (ii) b appearing exactly $n-3$ times in S and also, $2b-a$ appearing exactly once in S.

In 1996, Gao [13] proved the generalization of the above two results as follows.

Theorem [13]. *Let n be any positive integer. Let k be any positive integer such that $2 \leq k \leq \left[\frac{n+1}{4}\right] + 2$. Let $S = (a_1, a_2, \ldots, a_{2n-k})$ be any sequence in $\mathbb{Z}_n$ such that there is no subsequence of length n of S whose sum is zero in $\mathbb{Z}_n$. Then we can re-arrange the sequence S as follows*:

$$S = (\underbrace{a, a, \ldots, a}_{u \text{ times}}, \underbrace{b, b, \ldots, b}_{v \text{ times}}, a'_1, a'_2, \ldots, a'_{2n-k-u-v})$$

where $u \geq v$, $u + v \geq 2n - 2k + 2$ and $a - b$ generates $\mathbb{Z}_n$.

One of our main theorems in this article is to extend the above result to all primes p and integer k for the range $\frac{p+1}{4} + 3 \leq k \leq \frac{p+1}{3} + 1$. This extension is meaningful for all large primes p. Also, we shall study the problem of how many distinct residue classes modulo p occur in those sequences of length $2p-k$ in $\mathbb{Z}_p$ having a zero-sum subsequence of length p in it. Before we state our main theorems, we shall fix up notations as follows.

For every integer $1 \leq k \leq \ell$, define

$$\sum_k(S) = \left\{a_{i_1} + a_{i_2} + \cdots + a_{i_k} | 1 \leq i_1 < i_2 < \cdots < i_k \leq \ell\right\}$$

and $\sum(S) = \cup_{k=1}^{\ell} \sum_k(S)$. For any subsequence $T = (b_1, b_2, \ldots, b_r)$ of S, we let $\sigma(T) = \sum_{i=1}^{r} b_i$. We denote ST^{-1} by the deleted sequence R which is obtained from S by deleting the elements of T. Also, if $S = (\underbrace{a, a, \ldots, a}_{r \text{ times}}, b_1, b_2, \ldots)$, then we write $S = (a^r, b_1, b_2, \ldots)$. For any $b \in \mathbb{Z}_n$, we denote by $b + S$ the sequence $(b + a_1, b + a_2, \ldots, b + a_\ell)$. For every $x \in \mathbb{Z}_n$, define $\overline{x}$ to be the least positive inverse image under the natural homomorphism from the additive group of integers $\mathbb{Z}$ onto $\mathbb{Z}_n$. For example, $\overline{0} = n$. If $A \subset \mathbb{Z}_n$, then we denote the cardinality of A by $|A|$. If A is a sequence in $\mathbb{Z}_n$, we denote the length of A by $|A|$ (same notation as the cardinality). For any $g \in \mathbb{Z}_n$, we define $v_g(S)$ by the number of times g appears in S. Also, we define $h = h(S) = \max_{g \in \mathbb{Z}_n} v_g(S)$. Gao [13] introduced the following definition.

DEFINITION 1.1

Let $S = (a_1, a_2, \ldots, a_\ell)$ and $T = (b_1, b_2, \ldots, b_\ell)$ be two sequences in $\mathbb{Z}_n$ of length ℓ. We say that S is equivalent to T (written as $S \sim T$) if there exist an integer c coprime to n, an element $x \in \mathbb{Z}_n$, and a permutation π of $\{1, 2, \ldots, \ell\}$ such that $a_i = c(b_{\pi(i)} - x)$ for every $i = 1, 2, \ldots, n$. Clearly, $\sim$ is an equivalence relation; and if $S \sim T$, then $0 \in \sum_n(S)$ if and only if $0 \in \sum_n(T)$.

In this article, we shall prove theorems 3.1 and 3.2.

Theorem 3.1. *Let p be any odd prime number. Let k be any positive integer such that $2 \le k \le \left[\frac{p+1}{3}\right] + 1$. Let $S = (a_1, a_2, \ldots, a_{2p-k})$ be any sequence in $\mathbb{Z}_p$ such that $0 \notin \sum_p(S)$. Then*

$$S \sim (0^u, 1^v, a'_1, a'_2, \ldots, a'_{2p-k-u-v}),$$

where $u \ge v$ and $u + v \ge 2p - 2k + 2$.

Using the information in Theorem 3.1, we consider the following problem of variant of EGZ theorem as follows. Before we state our theorem, we recall the following definition which was introduced in [3] and state the known results.

DEFINITION 1.2

Let n, k be positive integers, $1 \le k \le n$. Denote by $f(n, k)$ the least positive integer g for which the following holds: If $S = (a_1, a_2, \ldots, a_g)$ is a sequence of elements of $\mathbb{Z}_n$, the cyclic group of order n, of length g such that the number of distinct a_i's is equal to k, then there are n indices $i_1, i_2, \ldots, i_n$ belonging to $\{1, 2, \ldots, g\}$ such that $a_{i_1} + a_{i_2} + \cdots + a_{i_n} = 0$.

Theorem. *We have*

1. $f(n, k) \le 2n - 1$ *for all n and for all $1 \le k \le n$ (By EGZ theorem).*
2. $f(n, n) = \begin{cases} n, & \text{if } n \text{ is odd} \\ n + 1, & \text{if } n \text{ is even} \end{cases}$ [10].
3. $f(n, k) = n + 2$, *for all $n \ge 5$ and $1 + n/2 < k \le n - 1$* [5, 10].
4. $f\left(n, \frac{n}{2} + 1\right) = n + 3$ *for all $n \in 2\mathbb{N}$* [12].
5. $f(n, k) = 2n - ((k-1)/2)^2 - 1$ *for all $n \ge (k-1)^2 - 4$ for an odd $k \ge 5$* [19].
6. $f(n, k) = 2n - k(k-2)/4 - 1$ *for all $n \ge k(k-2) - 4$ for an even $k \ge 6$* [19].
7. $f(n, 2) = 2n - 1$, $f(n, 3) = 2n - 2$ *and* $f(n, 4) = 2n - 3$ *for all n* [3].
8. $f(n, k) \le 2n - k + 1$ *for all $2 \le k \le n$* [16].
9. $f(p, k) \le 2p - 3k + 11$ *for all $5 \le k \le (p + 15)/3$* [17].

Other than these results many authors (for instance [11], [3] and [2]) consider some lower bounds for $f(n, k)$ for various k.

In this article, we shall prove the following result.

Theorem 3.2. *Let p be any odd prime number. Let k be any positive integer such that $2 \le k \le \left[\frac{p+1}{3}\right] + 1$. Then $f(p, \ell) \le 2p - k$ for all $\ell \ge 2\sqrt{2}\sqrt{k-2}$.*

2. Preliminaries

We shall start this section with a well-known fundamental inequality of subsets as follows.

Cauchy–Davenport inequality [6, 7]. Let p be any prime number. Let $A_1, A_2, \ldots, A_t$ be non-empty subsets of $\mathbb{Z}_p$. Then

$$|A_1 + A_2 + \cdots + A_t| \geq \min\left\{p, \sum_{i=1}^{t} |A_i| - t + 1\right\}.$$

Theorem 2.1 [4]. *Let n and k be any positive integers such that $n - 2k \geq 1$. If $S = (a_1, a_2, \ldots, a_{n-k})$ is a sequence in $\mathbb{Z}_n$ such that $0 \notin \sum(S)$, then there exists $a \neq 0 \in \mathbb{Z}_n$ which appear at least $n - 2k + 1$ times in S.*

The following Theorem is crucial for the proof of Theorem 3.1.

Theorem 2.2. *Let p be any prime number and $1 \leq k \leq p-2$. Let S be a sequence in $\mathbb{Z}_p$ of length $p + k$. If $0 \notin \sum_p(S)$, then $h(S) \geq k + 1$.*

Proof. When $k = 1$, the result follows from the Pigeon hole principle. So, we can assume that $k \geq 2$. If possible, we assume that $h(S) \leq k$. Then, we can distribute the elements of S into a union $A_1 \sqcup A_2 \sqcup \cdots \sqcup A_k$, so that in each A_i, an element occurs only once. By the Cauchy–Davenport theorem, we see that

$$\left|\sum_{i=1}^{k} A_i\right| \geq \min\left\{p, \sum_{i=1}^{k} |A_i| - k + 1\right\}$$
$$= \min\{p, p + k - k + 1 = p + 1\} = p.$$

Therefore, $A_1 + A_2 + \cdots + A_k = \mathbb{Z}_p$. In particular, $\sigma(S) \in \sum_k(S)$. Without loss of generality we shall assume that $\sigma(S) = a_1 + a_2 + \cdots + a_k$. Then we have $a_{k+1} + a_{k+2} + \cdots + a_{p+k} = 0$ which implies $0 \in \sum_p(S)$ as $|S| = p + k$. This contradicts the assumption that $0 \notin \sum_p(S)$. Therefore, $h(S) \geq k + 1$. □

Theorem 2.3 [14]. *Let n be any positive integer. Let $1 \leq k \leq \left[\frac{n+1}{3}\right]$, and let S be a sequence in $\mathbb{Z}_n$ of length $n - k$ such that $0 \notin \sum(S)$. Then*

$$S \sim (1^{n-2k+1}, x_1, x_2, \ldots, x_{k-1})$$

with $\sum_{i=1}^{k-1} \overline{x_i} \leq 2k - 2$.

Lemma 2.4. *Let p be any odd prime and $1 \leq k \leq \left[\frac{p+1}{3}\right]$. Let $S = (1^{p-2k+1}, x_1, x_2, \ldots, x_{k-1})$ be a sequence in $\mathbb{Z}_p \setminus \{0\}$ of length $p - k$ such that $\sum_{i=1}^{k-1} \overline{x_i} \leq 2k - 2$. Then, for any $x \in \mathbb{Z}_p$ satisfying $p - 2k + 1 \leq \overline{x} \leq p - 2k + 1 + \sum_{i=1}^{k-1} \overline{x_i}$, there exists a subsequence T of S such that $|T| \geq p - 2k + 1$ with $\sigma(T) = x$.*

Proof. Let $x \in \mathbb{Z}_p$ such that $p - 2k + 1 \leq \overline{x} \leq p - 2k + 1 + \sum_{i=1}^{k-1} \overline{x_i}$. If $x = p - 2k + 1$, then $x = \sum_{i=1}^{p-2k+1} 1$ and we are done; otherwise, we have

$$p - 2k + 2 \leq \overline{x} \leq p - 2k + 1 + \sum_{i=1}^{k-1} \overline{x_i} \leq p - 1.$$

Therefore, we have $1 \leq \overline{x} - (p - 2k + 1) \leq \sum_{i=1}^{k-1} \overline{x_i}$.

Claim. For any positive integer k, if $S' = (x_1, x_2, \ldots, x_k)$ be a sequence in $\mathbb{Z}_p \setminus \{0\}$ such that $|S'| = k$ and $\sum_{i=1}^{k} \overline{x_i} \leq 2k$, then, for every $x \in \mathbb{Z}_p$ satisfying $1 \leq \overline{x} \leq \sum_{i=1}^{k} \overline{x_i}$, either $x \in \sum(S')$ or $x + 1 \in \sum(S')$.

If the claim is proven, then, we get, either $x - (p - 2k + 1)$ or $x - (p - 2k + 1) + 1$ in $\sum((x_1, x_2, \ldots, x_{k-1}))$. That is, either $x = \underbrace{1 + 1 + \cdots + 1}_{p-2k+1} + y$ or $x = \underbrace{1 + 1 + \cdots + 1}_{p-2k} + y$ where $y \in \sum((x_1, x_2, \ldots, x_{k-1}))$. So, to end the proof of this lemma, it is enough to prove this claim.

When $k = 1, 2$, the claim is trivially true. So, we let $k \geq 3$. Assume the result is true for $k - 1$ and we shall prove for k. If necessary by renaming the indices, without loss of generality, we can assume that $S' = (x_1, x_2, \ldots, x_k)$ with $\overline{x_1} \leq \overline{x_2} \leq \cdots \leq \overline{x_k}$. Suppose $\overline{x_{k-1}} = 1$. Then, we have $x_1 = x_2 = \cdots = x_{k-1} = 1$. As $\sum_{i=1}^{k} \overline{x_i} \leq 2k$, we see that $\overline{x_k} \leq 2k - (k - 1) = k + 1$. Therefore, we see that

$$\sum(S') = \begin{cases} \{1, 2, \ldots, x_k + k - 1\}, & \text{if } \overline{x_k} \leq k, \\ \{1, 2, \ldots, x_k + k - 1\} \setminus \{k\}, & \text{if } \overline{x_k} = k + 1 \end{cases}$$

which clearly implies the claim. Thus, now, we can assume that $2 \leq \overline{x_{k-1}} \leq \overline{x_k}$. If $\overline{x} \leq \overline{x_k} + \sum_{i=1}^{k-2} \overline{x_i}$, then by induction, either x or $x + 1$ in $\sum((x_1, x_2, \ldots, x_{k-2}, x_k))$, and we are through; otherwise, we have, $\overline{x_k} + \sum_{i=1}^{k-2} \overline{x_i} \leq \overline{x} < \sum_{i=1}^{k} \overline{x_i}$. Therefore, we have

$$k - 2 \leq \overline{x_k} - \overline{x_{k-1}} + \sum_{i=1}^{k-2} \overline{x_i} \leq \overline{x - x_{k-1}} \leq \overline{x_k} + \sum_{i=1}^{k-2} \overline{x_i}.$$

Therefore, by the induction hypothesis, we see that either $x - x_{k-1}$ or $x - x_{k-1} + 1$ in $\sum((x_1, x_2, \ldots, x_{k-2}, x_k))$ and hence, we have either x or $x + 1$ in $\sum(S')$. □

3. Proof of Theorems 3.1 and 3.2

Proof of Theorem 3.1. Let S be a sequence in $\mathbb{Z}_p$ of length $2p-k$ where $2 \leq k \leq [\frac{p+1}{3}]+1$. Given that $0 \notin \sum_p(S)$. Without loss of generality we can assume that 0 (if necessary, by translating by an element) appears maximum number of, say u, times in S. By Theorem 2.2, it is clear that $u \geq p-k+1$. Therefore, $S = (0^u, a_1, a_2, \ldots, a_{2p-k-u})$ where $a_i \in \mathbb{Z}_p \setminus \{0\}$. Let $S_1 = (a_1, a_2, \ldots, a_{2p-k-u})$ be a subsequence of S. Since $u \geq p - k + 1$, we have $2p - k - u \leq 2p - k - p + k - 1 = p - 1$. That is, $|S_1| \leq p - 1$. Let $|S_1| = p - m$ for some positive integer m. Note that $p - m + u = 2p - k$ which implies $u + k - p = m$. As $0 \notin \sum_p(S)$, we have $u \leq p - 1$. Therefore, $m = u + k - p \leq p - 1 + k - p = k - 1$.

If $0 \notin \sum(S_1)$, then by Theorem 2.1, we know that there exists an element $a \in \mathbb{Z}_p \setminus \{0\}$ such that $v_a(S_1) \geq p - 2m + 1$. Therefore, $S = (0^u, a^v, b_1, b_2, \ldots, b_{2p-k-u-v})$ and $2p - k - u - v \leq m - 1 \leq k - 2$ which implies $2p - 2k + 2 \leq u + v$ and we are done.

Thus, we can assume that $0 \in \sum(S_1)$. Let W be the maximal zero-sum subsequence of S_1 of length w. Moreover, since $0 \notin \sum_p(S)$ and S_1 is a sequence in $\mathbb{Z}_p \setminus \{0\}$ and

$u \geq p-k+1$, we have

$$2 \leq w \leq p-u-1 \implies 2 \leq w \leq k-2. \tag{1}$$

Also note that $k+u+w \geq k+p-k+1+w \geq p+1$. Put $\ell = k+u+w-p$. Therefore, $2p-k-u-w = p-\ell$. By the definition of W, we have $0 \notin \sum(S_1W^{-1})$ and $|S_1W^{-1}| = p-\ell$. Let $T = S_1W^{-1}$. Also, by the inequality (1), we see that $\ell = k+u+w-p \leq k+u+p-u-1-p = k-1 \leq \left[\frac{p+1}{3}\right]$. Therefore, by Theorem 2.3, we see that

$$T \sim (1^{p-2\ell+1}, x_1, x_2, \ldots, x_{\ell-1}) \quad \text{and} \quad \sum_{i=1}^{\ell-1} \overline{x_i} \leq 2\ell-2.$$

Thus, the given sequence $S = 0^u S_1 = 0^u TW$ is equivalent to the following sequence:

$$S \sim (0^u, 1^{p-2\ell+1}, x_1, x_2, \ldots, x_{\ell-1}, z_1, z_2, \ldots, z_w)$$

where all the $x_i \neq 0$ satisfying $\sum_{i=1}^{\ell-1} \overline{x_i} \leq 2\ell-2$ and $W \sim (z_1, z_2, \ldots, z_w)$ is the maximal zero-sum subsequence of S_1.

Without loss of generality, we shall replace '$\sim$' by '$=$' above. Also, we denote the number of 1's appearing in the sequences $(x_1, x_2, \ldots, x_{\ell-1})$ and $(z_1, z_2, \ldots, z_w)$ by r and t respectively. Put $v = p-2\ell+1+r+t$.

To end the proof of this theorem, it is enough to prove that $u+v \geq 2p-2k+2$.

If $2 \leq \overline{z_i} \leq p-2\ell+1$ for some i satisfying $1 \leq i \leq w$, then as there are $p-2\ell+1$ number of 1's in T, we can write $z_i = \sigma(L_1)$ where $L_1 = (1^{\overline{z_i}})$ with $|L_1| \geq 2$. If $p-2\ell+2 \leq \overline{z_i} \leq p-2\ell+1+\sum_{j=1}^{\ell-1} \overline{x_i}$ holds for some $1 \leq i \leq w$, then by Lemma 2.4, there exists a subsequence L_1 of T such that $z_i = \sigma(L_1)$ and $|L_1| \geq 2$. By letting $W_1 = L_1 W z_i^{-1}$, we see that $\sigma(W_1) = 0$ and $|W_1| \geq w+1$ which contradicts the maximality of W. Hence

$$p-1 \geq \overline{z_i} \geq p-2\ell+2+\sum_{i=1}^{\ell-1} \overline{x_i} \quad \text{for each } z_i \neq 1. \tag{2}$$

Since $\sum_{i=1}^{\ell-1} \overline{x_i} \leq 2\ell-2$, we have

$$2\ell-2 \leq \sum_{i=1}^{\ell-1} \overline{x_i} + r. \tag{3}$$

Therefore, by the inequalities (2) and (3), we get

$$p-1 \geq \overline{z_i} \geq p-r \quad \text{for each } z_i \neq 1. \tag{4}$$

By rearranging the indices and renaming them, if necessary, we can assume that for $0 \leq q \leq w$, we have

$$\overline{z_i} \neq 1 \quad \text{for } 1 \leq i \leq q \quad \text{and} \quad \overline{z_i} = 1 \quad \text{for } q+1 \leq i \leq w. \tag{5}$$

Case i. $(w = 2)$

In this case, by the definition of W, we have $z_1 + z_2 = 0$. Therefore, there are two cases, namely, $z_1 = 1$ and $z_2 = -1$ or $z_1 \neq 1$ and hence $z_2 \neq -1$. When $z_1 \neq 1$, by the inequality (4), we have $p - 2 \geq \overline{z_1} \geq p - r$ and in particular, we have $r \geq 2$. Since $2 \leq r \leq p - 2\ell + 1$, we have a zero-sum subsequence $Z = (\overline{z_1}, 1^{p-\overline{z_1}})$ which has length > 2 which is a contradiction to the maximality of W. Thus, $z_1 = 1$ and $z_2 = -1$. In this case, $v \geq 2p - k - u - \ell + r$. Therefore, $u + v = u + 2p - k - u - \ell + r \geq 2p - k - (k-1) + r = 2p - 2k + 1 + r \geq 2p - 2k + 3$. We are done in this case.

Case ii. $(w \geq 3)$

Since W is a zero-sum sequence, $q \neq 0$. So, we have $1 \leq q \leq w$. When $q = 1$, from the inequality (4), we get

$$p - 1 \geq \overline{z_1} \geq p - r. \tag{6}$$

When $q = 2$, we have

$$\begin{aligned} 2p - 2 \geq \overline{z_1} + \overline{z_2} &\geq 2p - 2r \geq 2p - 2(\ell - 1) \geq 2p - 2(k - 2) \\ &= 2p - 2k + 4. \end{aligned}$$

Since $k \leq \left[\frac{p+1}{3}\right] + 1$, it is clear that $p \geq 3k - 4$ and hence

$$2p - 2 \geq \overline{z_1} + \overline{z_2} > p \implies p - 2 \geq \overline{z_1} + \overline{z_2} - p > p - 2\ell + 2.$$

Therefore, it follows that

$$p - 2 \geq \overline{z_1 + z_2} > p - 2\ell + 2.$$

If $\overline{z_1 + z_2} \leq p - 2\ell + 1 + \sum_{i=1}^{\ell-1} \overline{x_i}$, then $z_1 + z_2 = \sigma(L_2)$ for some subsequence L_2 of T with $|L_2| \geq p - 2\ell + 1$ (by Lemma 2.4). If we let $W_2 = L_2 W z_1^{-1} z_2^{-1}$, then $\sigma(W_2) = 0$ and $|W_2| = |L_2| + w - 2 \geq p - 2\ell + 1 + w - 2 = w + p - (2\ell + 1) > w$ (as $\ell < k - 1 \leq (p+1)/3$) which contradicts the maximality of W. Therefore, we have $\overline{z_1 + z_2} \geq \sum_{i=1}^{\ell-1} \overline{x_i} + p - 2\ell + 2$. Thus, by the inequality (3), we have

$$p - 2 \geq \overline{z_1 + z_2} \geq p - r. \tag{7}$$

Now, we shall assume that $q \geq 3$. Let $a = \min\{q, w - 2\}$. Then we claim the following.

Claim 1. For $q \geq 3$ and for every $s = 1, 2, \ldots, a$, we have

$$p - s \geq \left(\sum_{i=1}^{s} \overline{z_i}\right) - (s-1)p = \overline{\sum_{i=1}^{s} z_i} \geq p - r.$$

By inequalities (4) and (7), the Claim 1 is true for $s = 1$ and 2. Now, we shall assume that claim 1 is true for $s - 1$ and we prove for s. By the inequality (4) and induction hypothesis, we have

$$\begin{aligned} p - s \geq \left(\sum_{i=1}^{s} \overline{z_i}\right) - (s-1)p &= \left(\sum_{i=1}^{s-1} \overline{z_i}\right) - (s-2)p + \overline{z_s} - p \\ &\geq p - r - r = p - 2r \geq p - 2\ell + 2 \geq p - 2k + 4 > 0. \end{aligned}$$

Hence,

$$p - 2\ell + 2 \le \overline{\sum_{i=1}^{s} z_i}.$$

If $\overline{\sum_{i=1}^{s} z_i} \le p-2\ell+1+\sum_{i=1}^{\ell-1} \overline{x_i}$, then by Lemma 2.4, there exists a subsequence L_3 of T with $|L_3| \ge p-2\ell+1$ such that $\sum_{i=1}^{s} z_i = \sigma(L_3)$. If we let $W_3 = L_3 W z_1^{-1} z_2^{-1} \dots z_s^{-1}$, then we get $\sigma(W_3) = 0$. Since $w \le k-2$, $\ell \le k-1$ and $p \ge 3k-1$, we have

$$\begin{aligned} |W_3| &= w + |L_3| - s \ge w + p - 2\ell + 1 - (w-2) \\ &\ge w + p - 2k + 4 + 1 - k + 4 = w + p - (3k-9) > w. \end{aligned}$$

This contradicts the fact that W is the maximal zero-sum subsequence of S_1. Therefore, we have

$$p - s \ge \overline{\sum_{i=1}^{s} z_i} \ge p - 2\ell + 2 + \sum_{i=1}^{\ell-1} \overline{x_i}$$

and by the inequality (3), we get Claim 1.

Claim 2. $q \le w - 2$.

Assume, on the contrary that $q \ge w-1$. Then $q = w-1$ or $q = w$. If $q = w-1$, then we have $p-(w-2) \ge \overline{z_1+z_2+\cdots+z_{w-2}} \ge p-r$, $p-1 \ge \overline{z_{w-1}} \ge p-r$ and $z_w = 1$. Therefore,

$$\begin{aligned} 2p > 2p - w + 2 &\ge \overline{z_1+z_2+\cdots+z_{w-2}} + \overline{z_{w-1}} + \overline{z_w} \\ &\ge 2p - 2r + 1 > p \end{aligned}$$

which is a contradiction to $\sigma(W) = 0$. Hence $q \ne w-1$.

If $q = w$, then $p-(w-2) \ge \overline{z_1+z_2+\cdots+z_{w-2}} \ge p-r$, $p-1 \ge \overline{z_{w-1}}$, $\overline{z_w} \ge p-r$. Therefore,

$$\begin{aligned} 3p > 3p - w &\ge \overline{z_1+z_2+\cdots+z_{w-2}} + \overline{z_{w-1}} + \overline{z_w} \\ &\ge 3p - 3r \ge 3p - 3(k-2) \ge 3p - 3k + 6 = 2p + p - 3k + 6 > 2p, \end{aligned}$$

(as $r \le \ell - 1 \le k-2$ and $p \ge 3k-4$) which is also a contradiction to $\sigma(W) = 0$. Hence $q \ne w$. Thus Claim 2 is true.

From Claims 1 and 2, we see that s varies from 1 to q. Since we have $p - s \ge p - r$ which implies $r \ge s$. In particular, when $s = q$, we get

$$q \le r. \tag{8}$$

But by the definition of q, we have $q = w - t$ which implies that $w = q + t$. Therefore, by the inequality (8), we have $r + t \ge q + t = w$. Thus

$$\begin{aligned} u + v &= u + p - 2\ell + 1 + r + t \ge u + p - 2\ell + 1 + w \\ &= 2p - k - (\ell - 1) \ge 2p - 2k + 2, \end{aligned}$$

as desired. □

Proof of Theorem 3.2. Let S be a given sequence in $\mathbb{Z}_p$ of length $2p-k$. Suppose the number of distinct residue classes appearing in S is $g \geq 2\sqrt{2}\sqrt{k-2}$. If possible, we assume that $0 \notin \sum_p(S)$. Then by Theorem 3.1, $S = 0^u TW$ (notations as in the proof of Theorem 3.1). Now, we shall count the number of distinct residue classes modulo p appearing in T and in W separately.

We recall that $T = (1^{p-2\ell+1}, x_1, x_2, \ldots, x_{\ell-1})$ with $\sum_{i=1}^{\ell-1} \overline{x_i} \leq 2\ell - 2$ and $r = v_1((x_1, x_2, \ldots, x_{\ell-1}))$. Also, $W = (z_1, z_2, \ldots, z_q, \underbrace{1, 1, \ldots, 1}_{w-q \text{ times}})$ where $z_i \neq 1$. Note that by Claim 2 of Theorem 3.1, we have $1 \leq q \leq w-2$ and by (6) and (8) we have $q \leq r$ and $r \geq 2$.

Let g_1 (respectively, g_2) denote the number of distinct residue classes modulo p appearing in T (respectively, in W). Thus, including 0, the total number of distinct residue classes modulo p appearing in S is $g = g_1 + g_2 + 1 - 1 = g_1 + g_2$ because the residue 1 is calculated twice in g_1 and g_2. So, to end the proof of this theorem, it is enough to estimate $g = g_1 + g_2$.

Since $\sum_{i=1}^{\ell-1} \overline{x_i} \leq 2\ell - 2$ and r number of 1's appearing in (x_i)s, we have

$$1 + 2 + \cdots + g_1 \leq 2\ell - 2 - (r-1)$$

$$g_1^2 + g_1 \leq 4\ell - 4 - 2(r-1) \leq 4\ell - 4 - 2 = 2(\ell - 3).$$

Therefore, since $\ell \leq k-1$, we have

$$g_1^2 + g_1 \leq 2(k-4) \implies g_1 \leq \sqrt{2}\sqrt{k-4} < \sqrt{2}\sqrt{k-2}. \tag{9}$$

Now, note that $\overline{-z_i} = p - \overline{z_i}$. Therefore by Claim 1 of Theorem 3.1, we get $\sum_{i=1}^{q} \overline{-z_i} \leq r$. Thus,

$$1 + 2 + \cdots + g_2 \leq r \implies g_2 \leq \sqrt{2r}.$$

Since $r \leq \ell - 1 \leq k-2$, we have

$$g_2 \leq \sqrt{2k-4} = \sqrt{2}\sqrt{k-2}. \tag{10}$$

Thus, from the inequalities (9) and (10) and counting 0, we have

$$g_1 + g_2 < \sqrt{2}\sqrt{k-2} + \sqrt{2}\sqrt{k-2} = 2\sqrt{2}\sqrt{k-2},$$

a contradiction. Hence the theorem. □

We shall end this section with the following open problems.

Open Problem. Let n and k be two positive integers such that $k \leq n-2$. Determine the constant defined by

$$h(n,k) = \min\{h(S) \mid |S| = n+k\},$$

where S runs over all sequences in $\mathbb{Z}_n$ of length $n+k$ such that $0 \notin \sum_n(S)$.

It is proved in [1] and [18] that $h(n, n-2) = n-1$ and proved in [9] that $h(n, n-3) = n-1$. Theorem 2.2 shows that $h(p, k) \geq k+1$ for all $1 \leq k \leq p-2$. The main result in [13] implies that $h(n, k) \geq k+1$ whenever $n - [(n+1)/4] - 1 \leq k \leq n-2$. It is natural to ask if $h(n, k) \geq k+1$ for every positive integer n and every k such that $1 \leq k \leq n-2$. However, the answer is 'no' in general. Recently, in [15] we provided a counter example for k satisfying $p \leq k \leq n/p - 2$. We conjectured the following.

Conjecture [15]. Let $n > 1$ be any positive integer and let p be the smallest prime divisor of n. Let k be an integer such that $k \geq (n/p) - 1$. Then $h(n, k) \geq k+1$.

In [15], it is proved that Conjecture 1 is true for $n = p^\ell$ for any prime p. Also, it is not known whether Conjecture 1 is true for $k < p/3$.

Acknowledgements

The first author is supported by NSFC with grant Nos 19971058 and 10271080.

References

[1] Bialostocki A and Dierker P, On Erdös–Ginzburg–Ziv theorem and the Ramsey numbers for stars and matchings, *Disc. Math.* **110** (1992) 1–8
[2] Bialostocki A, Dierker P, Grynkiewicz D and Lotspeich M, On some developments of the Erdös–Ginzburg–Ziv theorem II, *Acta Arith.* **110(2)** (2003) 173–184
[3] Bialostocki A and Lotspeich M, Some developments on the Erdös–Ginzburg–Ziv theorem I, Sets, graphs and numbers (Budapest, 1991) Colloq. Math. Soc. János Bolyai, (Amsterdam: North-Holland) (1992) Vol. 60, pp. 97–117
[4] Bovey J D, Erdös P and Niven I, Conditions for zero-sum modulo n, *Canad. Math. Bull.* **18** (1975) 27–29
[5] Brakemeier W, Eine Anzahlformel von Zahlen modulo n, *Monatsh. Math.* **85** (1978) 277–282
[6] Cauchy A L, Recherches sur les nombers, *J. Ecole Polytechniques* **9** (1813) 99–123
[7] Davenport H, On the addition of residue classes, *J. London Math. Soc.* **22** (1947) 100–101
[8] Erdös P, Ginzburg A and Ziv A, Theorem in the additive number theory, *Bull. Res. Council Israel* **10F** (1961) 41–43
[9] Flores C and Ordaz O, On sequences with zero sum in Abelian groups. Volume in homage to Dr. Rodolfo A Ricabarra (Spanish) Vol. Homenaje, 1 (Baha Blanca: Univ. Nac. del Sur) (1995) 99–106
[10] Gallardo L, Grekos G and Pihko J, On a variant of Erdös–Ginzburg–Ziv theorem, *Acta Arith.* **89(4)** (1999) 331–336
[11] Gallardo L and Grekos G, On Brakemeier's variant of the Erdös–Ginzburg–Ziv problem, *Tatra Mt. Math. Publ.* **20** (2000) 91–98
[12] Gallardo L, Grekos G, Habsieger L, Hennecart F, Landreau B and Plagne A, Restricted addition in $\mathbb{Z}/n\mathbb{Z}$ and an application to the Erdös–Ginzburg–Ziv problem, *J. London Math. Soc. (2)* **65(3)** (2002) 513–523
[13] Gao W D, An addition theorem for finite cyclic groups, *Disc. Math.* **163** (1997) 257–265
[14] Gao W D, Zero sums in finite cyclic groups, *Integers* **0** (2000) #A 12, pp. 1–7
[15] Gao W D, Thangadurai R and Zhuang J, Addition theorems on the cyclic groups $\mathbb{Z}_{p^\ell}$ (2004) preprint
[16] Hamidoune Y O, Ordaz O and Ortuño A, On a combinatorial theorem of Erdös, Ginzburg and Ziv, *Combin. Prob. Comput.* **7(4)** (1998) 403–412

[17] Panigrahi A and Thangadurai R, A variant of Erdös–Ginzburg–Ziv Theorem (2004) preprint
[18] Peterson B and Yuster T, A generalization of an addition theorem for solvable groups, *Can. J. Math.* **XXXVI(3)** (1984) 529–536
[19] Wang C, Note on a variant of the Erdös–Ginzburg–Ziv problem, *Acta Arith.* **108(1)** (2003) 53–59