Available online at www.sciencedirect.com

ScienceDirect

Discrete Mathematics 308 (2008) 1–8

DISCRETE MATHEMATICS

www.elsevier.com/locate/disc

Long n-zero-free sequences in finite cyclic groups

Svetoslav Savchev[1], Fang Chen

Oxford College of Emory University, Oxford, GA 30054, USA

Received 16 April 2006; received in revised form 21 February 2007; accepted 15 March 2007
Available online 24 March 2007

Abstract

A sequence in the additive group $\mathbb{Z}_n$ of integers modulo n is called n-zero-free if it does not contain subsequences with length n and sum zero. The article characterizes the n-zero-free sequences in $\mathbb{Z}_n$ of length greater than $3n/2-1$. The structure of these sequences is completely determined, which generalizes a number of previously known facts. The characterization cannot be extended in the same form to shorter sequence lengths. Consequences of the main result are best possible lower bounds for the maximum multiplicity of a term in an n-zero-free sequence of any given length greater than $3n/2-1$ in $\mathbb{Z}_n$, and also for the combined multiplicity of the two most repeated terms. Yet another application is finding the values in a certain range of a function related to the classic theorem of Erdős, Ginzburg and Ziv.

Keywords: Zero-sum problems; Zero-free sequences; Erdős–Ginzburg–Ziv theorem

1. Introduction

The Erdős–Ginzburg–Ziv theorem [4] states that each sequence of length $2n-1$ in the cyclic group of order n has a subsequence of length n and sum zero. This article characterizes all sequences with length greater than $3n/2-1$ in the same group that do not satisfy the conclusion of the celebrated theorem.

In the sequel, the cyclic group of order n is identified with the additive group $\mathbb{Z}_n=\mathbb{Z}/n\mathbb{Z}$ of integers modulo n. A sequence in $\mathbb{Z}_n$ is called a *zero sequence* or a *zero sum* if its sum is the zero element of $\mathbb{Z}_n$. A sequence is *zero-free* if it does not contain a nonempty zero subsequence. Sequences in $\mathbb{Z}_n$ without zero subsequences of length n will be called *n-zero-free*.

The n-zero-free sequences in $\mathbb{Z}_n$ were given considerable attention. Here we mention only results most closely related to our topic. Yuster and Peterson [11] and, independently, Bialostocki and Dierker [1], determined all n-zero-free sequences of length $2n-2$ in $\mathbb{Z}_n$. These are the sequences containing exactly two distinct elements a and b of $\mathbb{Z}_n$, each of them repeated $n-1$ times, such that $a-b$ generates $\mathbb{Z}_n$. Ordaz and Flores [9] solved the same problem for length $2n-3$. Again, two distinct terms have high combined multiplicity (details can be found in Section 4). In general, the combined multiplicity of the two most represented terms was intensively studied. Gao [6] proved a statement to

[1] No current affiliation.
E-mail address: fchen2@learnlink.emory.edu (F. Chen).

0012-365X/$ - see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.disc.2007.03.049

this effect for n-zero-free sequences of length roughly greater than $7n/4$. A recent work of Gao et al. [8] considered the same question in the case of a prime n, for length roughly greater than $5n/3$.

Based on the main theorem in [6], Bialostocki et al. [2] obtained an explicit characterization of the n-zero-free sequences in $\mathbb{Z}_n$ with length greater than or equal to $2n - 2 - \lfloor n/4 \rfloor$. The core of their proof is essentially present already in the article of Gao and Hamidoune [7].

Our goal is to characterize the n-zero-free sequences in $\mathbb{Z}_n$ of length greater than $3n/2 - 1$. The argument relies on a key structural result from [10] about zero-free sequences of length greater than $n/2$ in $\mathbb{Z}_n$. The description obtained generalizes the one from [2] and cannot be extended in the same shape to shorter sequences. In this sense the range of the characterization is optimal.

Let a be an integer coprime to n and b an element of $\mathbb{Z}_n$. The function $x \mapsto ax + b$ from $\mathbb{Z}_n$ into itself will be called an *affine map*. In particular the *translations* $x \mapsto x + b$ are affine maps, for each $b \in \mathbb{Z}_n$. Suppose that a sequence β in $\mathbb{Z}_n$ can be obtained from a sequence α through an affine map and rearrangement of terms. Then we say that α is *similar* to β and write $\alpha \sim \beta$. Clearly $\sim$ is an equivalence relation. Affine maps preserve zero sums of length n and do not bring in new ones. So it is usual not to distinguish between similar sequences in questions involving n-term zero subsequences. Our characterization will be up to similitude, i.e. up to affine maps and rearrangement of terms.

If $a \in \mathbb{Z}_n$, let $\overline{a}$ denote the unique integer in $[1, n]$ that belongs to the congruence class a modulo n. We call $\overline{a}$ the *least positive representative* of a. For a sequence α in $\mathbb{Z}_n$ we denote by $\overline{\alpha}$ the sequence of its least positive representatives, and by $L(\alpha)$ the sum of $\overline{\alpha}$.

The sequences considered are written multiplicatively, and multiplicities of sequence terms are indicated by using exponents. The length of the sequence α is denoted by $|\alpha|$. The *union* of two sequences α and β, denoted $\alpha \cup \beta$, is the sequence formed by appending the terms of β to α. Also, $1 - \beta$ is the sequence obtained by replacing each term b of β by $1 - b$.

Now the main result in the article, Theorem 5, can be stated as follows:

> *A sequence of length greater than $3n/2 - 1$ in $\mathbb{Z}_n$ is n-zero-free if and only if it is similar to the union of two sequences α and β in $\mathbb{Z}_n$ such that $L(\alpha) < n$ and $L(1 - \beta) < n$.*

Once such a characterization is available, certain basic questions about sufficiently long n-zero-free sequences in cyclic groups receive satisfactory answers.

The preliminaries needed for the key proof are included in Section 2. The main result is proven in Section 3. It is preceded by some properties of sequences of the form $\alpha \cup \beta$, where α and β are sequences in $\mathbb{Z}_n$ satisfying $L(\alpha) < n$ and $L(1 - \beta) < n$. In Section 4 we study the maximum multiplicity of a term in an n-zero-free sequence of length $n - 1 + k$, where $n/2 < k < n$, and also the combined multiplicity of the two most repeated terms. Best possible lower bounds are established in both cases. The main theorem enables us to determine, in Section 5, the values in a certain range of a function related to a variant of the Erdős–Ginzburg–Ziv theorem.

2. Preliminaries

For sequences α and β in $\mathbb{Z}_n$, we say that α is *equivalent* to β if β can be obtained from α through multiplication by an integer coprime to n and rearrangement of terms. Such multiplication is an affine map preserving all zero sums in $\mathbb{Z}_n$, not just the ones of length n. In particular equivalent sequences are similar. Our characterization rests on the following result from [10]:

Theorem 1. *Each zero-free sequence of length greater than $n/2$ in the cyclic group $\mathbb{Z}_n$ is equivalent to a sequence whose sum of the least positive representatives is less than n.*

A restatement of a fact from [5] is also used in the main proof.

Proposition 2. *A sequence in an abelian group of order n is such that the multiplicity of each term is at most the multiplicity of 0. Then each subsequence sum of length greater than n equals a subsequence sum of length exactly n.*

One more statement is necessary for the main argument.

Proposition 3. *Let α be a sequence with positive integer terms of length ℓ and sum S, where $2\ell > S$. Then*:

(a) *α contains at least $2\ell - S$ terms equal to* 1;
(b) *each integer in the interval $[2\ell - S, S]$ is representable as the sum of a subsequence of α with length at least $2\ell - S$.*

Proof. Part (a) is straightforward. If α contains x terms equal to 1 then each of the remaining $\ell - x$ terms is at least 2, hence $S \geqslant x + 2(\ell - x) = 2\ell - x$. This implies $x \geqslant 2\ell - S$. For part (b), fix $2\ell - S$ ones in α. The remaining $\ell - (2\ell - S) = S - \ell$ terms add up to $S - (2\ell - S) = 2(S - \ell)$, so their average is 2. Label these terms $a_1, \ldots, a_{S-\ell}$, assuming that $1 \leqslant a_1 \leqslant \cdots \leqslant a_{S-\ell}$. Due to this nondecreasing order, the sequence a_1, $(a_1 + a_2)/2$, $(a_1 + a_2 + a_3)/3, \ldots$ of arithmetic means is nondecreasing, hence these means are all at most 2. In other words, $a_1 + \cdots + a_i \leqslant 2i$ for all $i = 1, \ldots, S - \ell$.

Suppose that $b_1, \ldots, b_k$ are positive integers such that $b_1 + \cdots + b_i \leqslant 2i$ for all $i = 1, \ldots, k$. Denoting $S_k = \sum_{i=1}^{k} b_i$, we prove by induction on k that the sumset of the sequence $1b_1 \ldots b_k$ is $\{1, 2, \ldots, S_k + 1\}$. By *sumset* we mean the set of integers representable as a nonempty subsequence sum. The base $k = 1$ is clear. For the inductive step, let Σ_{k-1} and Σ_k be the sumsets of $1b_1 \ldots b_{k-1}$ and $1b_1 \ldots b_{k-1}b_k$, respectively. Now $\Sigma_{k-1} = \{1, 2, \ldots, S_{k-1} + 1\}$ by the induction hypothesis, hence $\Sigma_k = \{1, 2, \ldots, S_{k-1} + 1\} \cup \{b_k, b_k + 1, \ldots, b_k + S_{k-1} + 1\}$. Since $b_k + S_{k-1} = S_k$, it suffices to check that $b_k \leqslant S_{k-1} + 2$ which is equivalent to $S_k \leqslant 2S_{k-1} + 2$. The latter is true as $2S_{k-1} + 2 \geqslant 2(k-1) + 2 = 2k \geqslant S_k$. The induction is complete.

Going back to the proof of (b), we infer from the previous paragraph that the sequence $1a_1 \ldots a_{S-\ell}$ has sumset $\{1, 2, \ldots, 2(S - \ell) + 1\}$. Take an arbitrary $x \in [2\ell - S, S]$ and set $y = x - (2\ell - S - 1)$. Since $1 \leqslant y \leqslant 2(S - \ell) + 1$, one can express y as a nonempty subsequence sum of $1a_1 \ldots a_{S-\ell}$. In view of (a), adding $2\ell - S - 1$ to both sides of this representation shows that x equals the sum of at least $2\ell - S$ terms of the original sequence α. □

3. The main result

We are about to characterize all sufficiently long n-zero-free sequences in $\mathbb{Z}_n$. Up to similitude, a sequence of length greater than $3n/2 - 1$ is n-zero-free if and only if it can be divided into two sequences α and β satisfying $L(\alpha) < n$ and $L(1 - \beta) < n$. (Recall that $L(\omega)$ denotes the sum of the least positive representatives of the sequence ω.) There exist sequences of any length less than $2n - 1$ that are "separable" in the sense just described. We discuss them before the main theorem in order to indicate that most of their basic properties do not depend on whether or not the sequence is "long."

A couple of technical remarks will be necessary. Let α and β be sequences in $\mathbb{Z}_n$ such that $L(\alpha) < n$ and $L(1 - \beta) < n$. Because $\overline{0} = n$, note that $a \neq 0$ for $a \in \alpha$ and $b \neq 1$ for $b \in \beta$. We will need the observations that

$$\overline{-b} = n - \overline{b} \quad \text{and} \quad \overline{1 - b} = 1 + \overline{-b} \quad \text{for each } b \in \mathbb{Z}_n,\ b \neq 0. \tag{1}$$

By (1), for each sequence β in $\mathbb{Z}_n$ one can write

$$L(1 - \beta) = \sum_{b \in \beta, b \neq 0} \overline{-b} + |\beta|. \tag{2}$$

In what follows, the empty sequence is assumed to have sum 0, both in $\mathbb{Z}$ and in $\mathbb{Z}_n$.

Proposition 4. *Let n and k be integers such that $0 < k < n$. Suppose that the sequences α and β in $\mathbb{Z}_n$ satisfy the conditions $|\alpha| + |\beta| = n - 1 + k$, $L(\alpha) < n$ and $L(1 - \beta) < n$. Then*:

(a) *The union $\alpha \cup \beta$ is n-zero-free.*
(b) *$k \leqslant |\alpha| < n$, $k \leqslant |\beta| < n$ and $\overline{b} - \overline{a} \geqslant k$ for all $a \in \alpha$, $b \in \beta$. In particular $a \neq b$ for all $a \in \alpha$, $b \in \beta$.*
(c) *The multiplicities u and v of* 1 *and* 0 *in $\alpha \cup \beta$ satisfy*

$$u + v \geqslant 2k, \quad \max(u, v) \geqslant k, \quad \min(u, v) \geqslant 2k - n + 1.$$

The equality $u + v = 2k$ is attained if and only if $\alpha = 1^{2p-n+1}2^{n-1-p}$ and $\beta = 0^{2q-n+1}(-1)^{n-1-q}$, for integers p and q such that $(n-1)/2 \leqslant p < n$, $(n-1)/2 \leqslant q < n$ and $p + q = n - 1 + k$. The equality $\max(u, v) = k$ is attained if and only if n and k have different parity and $\alpha = 1^k 2^{(n-1-k)/2}$, $\beta = 0^k(-1)^{(n-1-k)/2}$.
(d) *For $k \geqslant (n-1)/2$, the highest multiplicity of a term in $\alpha \cup \beta$ is $\max(u, v)$.*

Proof. (a) Consider a zero subsequence γ of $\alpha\cup\beta$. Let γ contain r terms $a_1,\ldots,a_r$ from α, s nonzero terms $b_1,\ldots,b_s$ from β, and several zeros, from β again. Because the sum of γ is zero in $\mathbb{Z}_n$, the integers $\sum_{i=1}^{r}\overline{a_i}$ and $\sum_{j=1}^{s}\overline{-b_j}$ are congruent modulo n. Also $0\leqslant\sum_{i=1}^{r}\overline{a_i}\leqslant L(\alpha)<n$ and, by (2),

$$0\leqslant\sum_{j=1}^{s}\overline{-b_j}\leqslant\sum_{b\in\beta,b\neq0}\overline{-b}=L(1-\beta)-|\beta|<n-|\beta|\leqslant n.$$

Hence $\sum_{i=1}^{r}\overline{a_i}=\sum_{j=1}^{s}\overline{-b_j}$. Therefore, $r\leqslant\sum_{i=1}^{r}\overline{a_i}=\sum_{j=1}^{s}\overline{-b_j}<n-|\beta|$, implying $r+|\beta|<n$. Since $|\gamma|\leqslant r+|\beta|$, we infer that $\alpha\cup\beta$ is n-zero-free.

(b) The first two inequalities are immediate, because $|\alpha|+|\beta|=n-1+k$, $|\alpha|\leqslant L(\alpha)<n$ and $|\beta|=|1-\beta|\leqslant L(1-\beta)<n$. To show that $\overline{b}-\overline{a}\geqslant k$ for $a\in\alpha$ and $b\in\beta$, denote $M=\max_{a\in\alpha}\overline{a}+\max_{b\in\beta}\overline{1-b}$. Then

$$2(n-1)\geqslant L(\alpha)+L(1-\beta)\geqslant M+(|\alpha|-1)+(|\beta|-1)=M+n-3+k.$$

This yields $M\leqslant n+1-k$; thus $\overline{a}+\overline{1-b}\leqslant n+1-k$ for all $a\in\alpha$, $b\in\beta$. If $b\neq0$ then $\overline{1-b}=1+\overline{-b}=1+n-\overline{b}$ by (1), so $\overline{a}+\overline{1-b}\leqslant n+1-k$ becomes $\overline{b}-\overline{a}\geqslant k$. The same conclusion holds if $b=0$, as then $\overline{b}=n$, $\overline{1-b}=1$.

(c) We have $n-1\geqslant L(\alpha)\geqslant u+2(|\alpha|-u)=2|\alpha|-u$, since $\overline{a}\geqslant2$ for $a\neq1$. Similarly, $\overline{1-b}\geqslant2$ for $b\neq0$, so that $n-1\geqslant L(1-\beta)\geqslant v+2(|\beta|-v)=2|\beta|-v$. Adding up yields $2(n-1)\geqslant2(|\alpha|+|\beta|)-(u+v)=2(n-1+k)-(u+v)$. It follows that $u+v\geqslant2k$, so $\max(u,v)\geqslant k$. Clearly $\max(u,v)\leqslant n-1$ by (b), which implies $\min(u,v)\geqslant2k-n+1$.

The equality $u+v=2k$ occurs if and only if $n-1=L(\alpha)=2|\alpha|-u$ and $n-1=L(1-\beta)=2|\beta|-v$. These conditions imply $u=2|\alpha|-n+1$, $v=2|\beta|-n+1$; also $\overline{a}=2$ for $a\in\alpha$, $a\neq1$ and $\overline{1-b}=2$ for $b\in\beta$, $b\neq0$. In particular $|\alpha|\geqslant(n-1)/2$, $|\beta|\geqslant(n-1)/2$. So setting $p=|\alpha|$, $q=|\beta|$ and taking (b) into account, we obtain $(n-1)/2\leqslant p<n$, $(n-1)/2\leqslant q<n$, $p+q=n-1+k$ and $\alpha=1^{2p-n+1}2^{n-1-p}$, $\beta=0^{2q-n+1}(-1)^{n-1-q}$. The converse is easy to check; we note only that the last two sequences are well defined whenever $(n-1)/2\leqslant p<n$, $(n-1)/2\leqslant q<n$.

If $\max(u,v)=k$ then $u=v=k$ in view of $u+v\geqslant2k$, so $u+v=2k$. The conclusions of the last paragraph imply $\alpha=1^k2^{(n-1-k)/2}$, $\beta=0^k(-1)^{(n-1-k)/2}$. These sequences are well-defined only if $k\not\equiv n\pmod 2$. The converse is clear.

(d) We have $u+v\geqslant2k$ by (c), so the number of terms different from 1 and 0 in $\alpha\cup\beta$ is at most $(n-1+k)-2k=n-1-k$. Now it suffices to observe that $n-1-k\leqslant k$ for $k\geqslant(n-1)/2$, and that $\max(u,v)\geqslant k$ by (c). □

One can see that sequences "separable" in the sense of Proposition 4 are not just n-zero-free but have an interesting general structure. This is unexpected at first glance as α and β do not seem to be related in any way. While the properties listed in Proposition 4 are fairly simple to derive, it is less trivial to establish that each sufficiently long n-zero-free sequence in $\mathbb{Z}_n$ is "separable." The next theorem proves that length greater than $3n/2-1$ is enough to guarantee this. Moreover, shorter n-zero-free sequences are not necessarily "separable." These conclusions form the essential part of the article.

Theorem 5. *A sequence of length greater than $3n/2-1$ in the cyclic group $\mathbb{Z}_n$ does not contain an n-term zero subsequence if and only if it is similar to the union of two sequences α and β in $\mathbb{Z}_n$ such that*

$$L(\alpha)<n\quad\text{and}\quad L(1-\beta)<n.$$

Proof. The sufficiency follows from Proposition 4(a). For the necessity, let γ be an n-zero-free sequence of length greater than $3n/2-1$ in $\mathbb{Z}_n$. Translations in $\mathbb{Z}_n$ do not affect sums of length n, so one may assume that 0 is a term of γ with maximum multiplicity v. Then Proposition 2 shows that each zero subsequence of γ has length less than n. In particular $v<n$.

Select a zero subsequence σ of γ with nonzero terms and of maximum length; σ may be the empty sequence. This choice implies that the remaining nonzero terms of γ form a zero-free sequence τ. By the remark above, the lengths of σ and τ satisfy $|\sigma|<n-v$ and $|\tau|>(3n/2-1)-(n-1)=n/2$. Therefore Theorem 1 applies to the zero-free sequence τ.

Hence multiplying τ by a certain integer coprime to n yields an equivalent sequence with sum of the least positive representatives less than n. We multiply by the same integer all remaining terms of γ, which preserves the zero sums of any length. So there is no loss of generality in assuming that $\gamma=0^v\sigma\tau$, where σ is a zero subsequence of γ with nonzero terms and of maximum length, and τ is a zero-free sequence of length greater than $n/2$ satisfying $L(\tau)<n$.

Let $\sigma = 1^w b_1 \dots b_q$ where $b_1, \dots, b_q$ are all terms of σ different from 1. The following inequality stronger than $L(\tau) < n$ implies the conclusion directly:

$$L(\tau) + \sum_{j=1}^{q} \overline{-b_j} < n. \tag{3}$$

Indeed, assume (3) is true, and let $\alpha = 1^w \tau$, $\beta = 0^v b_1 \dots b_q$. Then $\gamma = \alpha \cup \beta$; in addition, $L(\alpha) < n$ and $L(1-\beta) < n$. Firstly, $w \equiv \sum_{j=1}^{q} \overline{-b_j} \pmod{n}$, since σ has sum zero. Also $0 \leqslant \sum_{j=1}^{q} \overline{-b_j} < n$ by (3), and clearly $0 \leqslant w < n$. Therefore $w = \sum_{j=1}^{q} \overline{-b_j}$. So (3) can be written as $L(\tau) + w < n$, which is the inequality $L(\alpha) < n$. Furthermore, we obtain $|\sigma| = w + q = \sum_{j=1}^{q} \overline{-b_j} + q$, and because $|\sigma| < n - v$, it follows that $\sum_{j=1}^{q} \overline{-b_j} + q < n - v$. This is the same as $\sum_{b \in \beta, b \neq 0} \overline{-b} + |\beta| < n$. By (2), the latter means that $L(1-\beta) < n$.

So it suffices to prove (3) which is clear if σ is the empty sequence. Hence let $\sigma \neq \emptyset$, implying $q \geqslant 1$ (σ cannot have only terms equal to 1). For the sake of clarity, denote $|\tau| = \ell > n/2$, $L(\tau) = S < n$ and $\overline{-b_j} = v_j$, $j = 1, \dots, q$. Note that $2\ell - S \geqslant 2\ell - (n-1) \geqslant 2$ as $\ell > n/2$. Thus, Proposition 3 applies to the sequence $\overline{\tau}$ of the least positive representatives of τ. Also, $1 \leqslant v_j < n - 1$ by the choice of $b_1, \dots, b_q$. The proof of (3) is based on the following observation:

> *Suppose that m terms $v_{j_1}, \dots, v_{j_m}$ of the sequence $v_1 \dots v_q$ are such that the integer $T = n - (v_{j_1} + \dots + v_{j_m})$ satisfies $1 < T \leqslant S$. Then $m \geqslant 2\ell - S$ if $2\ell - S \leqslant T \leqslant S$ and $m \geqslant T$* if $1 < T < 2\ell - S$.

Indeed, if T represents the congruence class t modulo n then $t = \sum_{i=1}^{m} b_{j_i}$. Let $2\ell - S \leqslant T \leqslant S$. By Proposition 3, there is a subsequence ω of τ with length at least $2\ell - S$ such that $T = \sum_{c \in \omega} \overline{c}$. Hence $\sum_{c \in \omega} c = t = \sum_{i=1}^{m} b_{j_i}$. This implies $m \geqslant |\omega| \geqslant 2\ell - S$ as $m < |\omega|$ would yield a zero subsequence of γ with nonzero terms which is longer than σ, obtained upon replacing $b_{j_1} \dots b_{j_m}$ by ω. Similarly, if $1 < T < 2\ell - S$ then T can be expressed as the sum of T terms equal to 1 of $\overline{\tau}$. (There are at least $2\ell - S$ ones in $\overline{\tau}$ by Proposition 3.) Now the same argument as above gives $m \geqslant T$, by the maximum choice of σ.

It follows from the observation that $n - v_j > S$, $j = 1, \dots, q$. Indeed, if $1 < n - v_j \leqslant S$ for some j then $1 \geqslant 2\ell - S$ or $1 \geqslant n - v_j$, both of which are not true. Therefore $1 \leqslant v_j < n - S$, $j = 1, \dots, q$.

Passing on to the proof of (3), suppose that it is false. Then there are subsequences of $v_1 \dots v_q$ whose sum is at least $n - S$, for instance $v_1 \dots v_q$ itself. Without loss of generality, let $v_1 \dots v_m$ be such a (nonempty) subsequence of minimum length m. So $T = n - \sum_{j=1}^{m} v_j \leqslant S$ but $T + v_j > S$ for all $j = 1, \dots, m$. Note that $T > 1$ in view of the previous paragraph, because $v_m < n - S$ yields $T > S - v_m > S - (n - S) = 2S - n \geqslant 2\ell - n \geqslant 1$.

Let $2\ell - S \leqslant T \leqslant S$. Then $m \geqslant 2\ell - S$ by the observation above. Hence

$$S + 1 \leqslant n - \sum_{j=1}^{m-1} v_j \leqslant n - (m-1) \leqslant n - (2\ell - S - 1) = (n - 2\ell) + S + 1,$$

implying $n \geqslant 2\ell$ which is a contradiction.

Next, let $1 < T < 2\ell - S$. Now the observation gives $m \geqslant T$. Recalling that $T + v_j > S$, we have $v_j \geqslant S + 1 - T > 0$ for $j = 1, \dots, m$, implying

$$n = T + \sum_{j=1}^{m} v_j \geqslant T + m(S + 1 - T) \geqslant T + T(S + 1 - T) = T(S + 2 - T).$$

Consider the quadratic function $g(t) = t^2 - (S+2)t + n$. We obtained $g(T) \geqslant 0$ for some $T \in \{2, \dots, 2\ell - S - 1\}$. But the maximum of g on $\{2, \dots, 2\ell - S - 1\}$ is $g(2) = n - 2S$, and $n - 2S \leqslant n - 2\ell < 0$. This is a contradiction again; claim (3) follows, concluding the main argument. □

Theorem 5 establishes the desired characterization, in a form hopefully providing general insight into the structure of n-zero-free sequences. On the other hand, the practically important consequence of the theorem is that each n-zero-free sequence of length $n - 1 + k$ in $\mathbb{Z}_n$, where $n/2 < k < n$, is similar to a sequence satisfying the conclusions of Proposition 4. Both Theorem 5 and Proposition 4 are needed for a really clear picture of the "long" n-zero-free sequences in $\mathbb{Z}_n$. The next observation adds one more detail to this picture.

The affine map $x \mapsto 1 - x$ interchanges 0 and 1 and transforms arbitrary sequences α and β into $\alpha_1 = 1 - \alpha$ and $\beta_1 = 1 - \beta$, respectively. If the inequalities $L(\alpha) < n$ and $L(1 - \beta) < n$ hold true, they can be written as $L(1 - \alpha_1) < n$ and $L(\beta_1) < n$. So if a sequence γ is similar to $\alpha \cup \beta$, it is also similar to $\alpha_1 \cup \beta_1$, a sequence with all properties from Proposition 4, in which the multiplicities of 0 and 1 are interchanged. Therefore one can assume additionally that $u \leqslant v$. For $k \geqslant (n-1)/2$, Proposition 4(d) then implies that 0 is a term of highest multiplicity in $\alpha \cup \beta$.

The conditions $L(\alpha) < n$ and $L(1 - \beta) < n$ can be expanded to obtain an explicit form of the characterization established in Theorem 5. Up to certain details, this explicit description has the same shape as the one in [2] of the n-zero-free sequences with length $n-1+k$, for k roughly greater than $3n/4$. It is worth noting that the range $n/2 < k < n$ for k is the natural scope of such a characterization. There are n-zero-free sequences of length $n - 1 + \lfloor n/2 \rfloor$ that do not obey the conclusion of Theorem 5.

Here are examples. For an odd $n \geqslant 9$ and an even $n \geqslant 6$, consider the sequence $0^{n-1}2^{(n-5)/2}3^2$ and $0^{n-1}2^{n/2-1}3$, respectively. Both of them n-zero-free and have length $n - 1 + \lfloor n/2 \rfloor$. Suppose that either of them is similar to a union $\alpha \cup \beta$ where $L(\alpha) < n$, $L(1 - \beta) < n$. Because $k \geqslant (n-1)/2$, α and β can be chosen so that 0 is a term of highest multiplicity v in $\alpha \cup \beta$. Then $v = n - 1$, so $\beta = 0^{n-1}$. It follows that $2^{(n-5)/2}3^2$ or $2^{n/2-1}3$ is equivalent to α, a sequence satisfying $L(\alpha) < n$. However, one can check that the latter is not true.

4. Terms of high multiplicity

Let $n/2 < k < n$, and let γ be an n-zero-free sequence of length $n - 1 + k$ in $\mathbb{Z}_n$. It follows from Theorem 5 and Proposition 4 that γ has a term of multiplicity at least k, and two distinct terms of combined multiplicity at least $2k$. Now we obtain precise forms of these statements.

Let $\alpha \cup \beta$ be a sequence similar to γ and satisfying the conclusions of Proposition 4. We may also assume that 0 is a term of maximum multiplicity v in $\alpha \cup \beta$, as explained in the previous section.

By Proposition 4(c), the equality $v = k$ holds if and only if $k \not\equiv n \pmod 2$ and $\alpha \cup \beta = 0^k 1^k 2^{(n-1-k)/2}(-1)^{(n-1-k)/2}$. This is the unique sequence satisfying $v = k$, up to affine maps and rearrangement of terms.

Suppose that $k \equiv n \pmod 2$. Then $v \geqslant k + 1$ by Proposition 4(c) again. The equality $v = k + 1$ can be attained, for instance for the sequence $0^{k+1}1^{k-1}2^{(n-k)/2}(-1)^{(n-k-2)/2}$ which is well defined and n-zero-free by Proposition 4(a) (setting $\alpha = 1^{k-1}2^{(n-k)/2}$, $\beta = 0^{k+1}(-1)^{(n-k-2)/2}$). Thus the following corollary is proved.

Corollary 6. *Let n and k be integers satisfying $n/2 < k < n$. Each n-zero-free sequence of length $n - 1 + k$ in $\mathbb{Z}_n$ contains a term of multiplicity at least k, if n and k are of different parity, and at least $k + 1$, if n and k are of the same parity. Both estimates are best possible.*

The sum of the two highest multiplicities was probably the most widely explored question concerning n-zero-free sequences in $\mathbb{Z}_n$. We are now in the position to resolve this question completely for each length $n - 1 + k$ where $n/2 < k < n$. Indeed, the lower bound $u + v \geqslant 2k$ for this combined multiplicity follows from above. Now let us take another look at the examples for the maximum multiplicity of a single term. In both possible cases, $k \not\equiv n \pmod 2$ and $k \equiv n \pmod 2$, it is easy to see that 0 and 1 are the two terms with highest combined multiplicity, and the value of this multiplicity is $2k$. So we proceed with one more structural conclusion.

Corollary 7. *Let n and k be integers satisfying $n/2 < k < n$. Each n-zero-free sequence of length $n - 1 + k$ in $\mathbb{Z}_n$ contains two terms of combined multiplicity at least $2k$, and this estimate is best possible.*

Naturally, some well-known results about the structure of the n-zero-free sequences with a certain length are now immediate. For example, let us consider the lengths $2n - 2$ (as in [11,1]) and $2n - 3$ (as in [9]). By the discussion above, any n-zero-free sequence of length $2n - 2$ (i.e. $n - 1 + k$ with $k = n - 1$) is similar to $0^v 1^u$, where $v = n - 1$ (as $v \geqslant k$) and $u = n - 1$ (as $u + v \geqslant 2k$). Here we assume $n > 2$ to ensure that $k > n/2$; however, the same conclusion holds true for $n = 2$ as well. Similarly, for $n > 4$, any n-zero-free sequence of length $2n - 3$ (i.e. $n - 1 + k$ with $k = n - 2$) is similar to $0^v 1^u \gamma$, where v is the maximum multiplicity of a term and $u + v \geqslant 2k = 2n - 4$. Since $k = n - 2 \equiv n \pmod 2$, Corollary 6 implies $v \geqslant k + 1 = n - 1$. Hence $v = n - 1$ and $u = n - 2$ or $u = n - 3$. Now it is easy to infer that each n-zero-free sequence of length $2n - 3$, $n > 4$, is similar to $0^{n-1}1^{n-2}$ or $0^{n-1}1^{n-3}2$. For $n = 3, 4$, this conclusion can be checked directly. For $n = 2$, the only n-zero-free sequence of length $2n - 3 = 1$ is similar to the one-term sequence 0.

5. The $g(n, k)$ function

For positive integers n and k, $k \leqslant n$, let $g(n, k) \geqslant k$ be the least integer such that each sequence in $\mathbb{Z}_n$ with at least k distinct terms and length $g(n, k)$ contains an n-term zero sum. The function $g(n, k)$ was introduced by Bialostocki and Lotspeich in [3]. The structural results about n-zero-free sequences of lengths $2n-2$ and $2n-3$ (such as mentioned after Corollary 7) imply $g(n, 2) = 2n-1$ for $n \geqslant 2$ and $g(n, 3) = 2n-2$ for $n \geqslant 4$. It is easy to see that $g(3, 3) = 3$.

The values of $g(n, k)$ for $4 \leqslant k \leqslant \sqrt{n+4}+1$ were found in [2]: If $k \geqslant 4$ is even and $n \geqslant k^2 - 2k - 4$, or if $k \geqslant 5$ is odd and $n \geqslant k^2 - 2k - 3$, then

$$g(n, k) = 2n - 1 - \left\lfloor \left(\frac{k-1}{2}\right)^2 \right\rfloor.$$

For the lower bound, the following examples were used. In the case of an even $k \geqslant 2$, consider the sequence

$$-\left(\frac{k-2}{2}\right) \ldots (-1)(0)^{n-(k^2+2k)/8}(1)^{n-(k^2+2k)/8}(2) \ldots \left(\frac{k}{2}\right);$$

if $k \geqslant 3$ is odd, consider the sequence

$$-\left(\frac{k-3}{2}\right) \ldots (-1)(0)^{n-(k^2-1)/8}(1)^{n-(k^2+4k+3)/8}(2) \ldots \left(\frac{k+1}{2}\right).$$

These examples are valid under the weaker restrictions $n \geqslant (k^2+2k)/8+1$ when k is even, and $n \geqslant (k^2+4k+3)/8+1$ when k is odd. The multiplicities of 0 and 1 in both sequences are positive integers. By Proposition 4(a), both sequences are n-zero-free, and each one contains k distinct terms.

Here we prove that $g(n, k)$ obeys the same formula as above under the weaker constraints $4 \leqslant k \leqslant \sqrt{2n-1}+1$. In this range the examples above still provide the lower bound $g(n, k) \geqslant 2n - 1 - \lfloor ((k-1)/2)^2 \rfloor$.

Theorem 8. *Let $n \geqslant k$ be integers such that $4 \leqslant k \leqslant \sqrt{2n-1}+1$. Then*

$$g(n, k) = 2n - 1 - \left\lfloor \left(\frac{k-1}{2}\right)^2 \right\rfloor.$$

Proof. As already mentioned, we need to prove only the upper bound. The condition $k \leqslant \sqrt{2n-1}+1$ is equivalent to $n - \lfloor ((k-1)/2)^2 \rfloor > n/2$. Also $k \geqslant 4$, so the integer $\ell = n - \lfloor ((k-1)/2)^2 \rfloor$ satisfies $n/2 < \ell < n$. Consider any n-zero-free sequence γ of length $n-1+\ell$ in $\mathbb{Z}_n$. It suffices to prove that the number of distinct terms in γ is less than k; then the definition of $g(n, k)$ implies $g(n, k) \leqslant n - 1 + \ell = 2n - 1 - \lfloor ((k-1)/2)^2 \rfloor$.

Let $\alpha \cup \beta$ be a sequence similar to γ, where α and β satisfy the conditions in Proposition 4, with k replaced by ℓ. Let there be x distinct terms in α and y distinct terms in β. Then Proposition 4(b) shows that the number of distinct terms in γ is $z = x + y$. The sum $L(\alpha)$ does not increase upon replacing x distinct summands in it by the least possible values $1, 2, \ldots, x$, and all remaining summands by 1. Therefore

$$1 + 2 + \cdots + x + (|\alpha| - x) \leqslant L(\alpha) \leqslant n - 1,$$

which gives $(x^2 - x)/2 \leqslant n - 1 - |\alpha|$. Likewise, noticing that there are y distinct terms in the sequence $1 - \beta$, we obtain $(y^2 - y)/2 \leqslant n - 1 - |\beta|$. Hence

$$\frac{1}{2}(x^2 - x) + \frac{1}{2}(y^2 - y) \leqslant 2(n-1) - (|\alpha| + |\beta|).$$

Since $|\alpha| + |\beta| = n - 1 + \ell$, the right-hand side expression equals $n - 1 - \ell = \lfloor ((k-1)/2)^2 \rfloor - 1$. On the other hand,

$$\frac{1}{2}(x^2 - x) + \frac{1}{2}(y^2 - y) \geqslant \left(\frac{x+y}{2}\right)^2 - \frac{x+y}{2} = \left(\frac{z}{2}\right)^2 - \frac{z}{2} = \left(\frac{z-1}{2}\right)^2 - \frac{1}{4}.$$

Thus $((z-1)/2)^2 - 1/4 \leqslant \lfloor ((k-1)/2)^2 \rfloor - 1$ which implies the desired $z < k$ and completes the proof. □

Acknowledgments

The first author extends his gratitude to Patricia Fauring and Flora Gutiérrez from Olimpíada Matemática Argentina for making possible his work on this article.

References

[1] A. Bialostocki, P. Dierker, On the Erdős–Ginzburg–Ziv theorem and the Ramsey numbers for stars and matchings, Discrete Math. 110 (1–3) (1992) 1–8.
[2] A. Bialostocki, P. Dierker, D. Grynkiewicz, M. Lotspeich, On some developments of the Erdős–Ginzburg–Ziv Theorem II, Acta Arith. 110 (2) (2003) 173–184.
[3] A. Bialostocki, M. Lotspeich, Some developments of the Erdős–Ginzburg–Ziv theorem I, in: Sets, Graphs and Numbers, Budapest, 1991, pp. 97–117, Colloq. Math. Soc. János Bolyai 60 (1992).
[4] P. Erdős, A. Ginzburg, A. Ziv, Theorem in the additive number theory, Bull. Res. Council Israel 10F (1961) 41–43.
[5] W.D. Gao, Addition theorems for finite abelian groups, J. Number Theory 53 (2) (1995) 241–246.
[6] W.D. Gao, An addition theorem for finite cyclic groups, Discrete Math. 163 (1–3) (1997) 257–265.
[7] W.D. Gao, Y.O. Hamidoune, Zero sums in abelian groups, Combin. Probab. Comput. 7 (3) (1998) 261–263.
[8] W.D. Gao, A. Panigrahi, R. Thangadurai, On the structure of p-zero-sum free sequences and its application to a variant of Erdős–Ginzburg–Ziv theorem, Proc. Indian Acad. Sci. Math. Sci. 115 (1) (2005) 67–77.
[9] O. Ordaz, C. Flores, On sequences with zero sum in abelian groups.Volume in Homage to Dr. Rodolfo A. Ricabarra, vol. Homenaje, 1, Univ. Nac. del Sur, Bahía Blanca, 1995, pp. 99–106 (in Spanish).
[10] S. Savchev, F. Chen, Long zero-free sequences in finite cyclic groups, Discrete Math., 2007, in press, doi:10.1016/j.disc.2007.01.012.
[11] T. Yuster, B. Peterson, A generalization of an addition theorem for solvable groups, Canad. J. Math. 36 (3) (1984) 529–536.