Distinct Length Modular Zero-sum Subsequences: A Proof of Graham's Conjecture

Weidong Gao*, Yahya Ould Hamidoune†, Guoqing Wang ‡

Abstract

Let S be a sequence of n nonnegative integers not exceeding $n-1$ such that S takes at least three distinct values. We show that S has two nonempty $\pmod n$ zero-sum subsequences with distinct lengths. This proves a conjecture of R.L. Graham. The validity of this conjecture was verified by Erdős and Szemerédi for all sufficiently large prime n.

1 Introduction and main result

We quote:

" Graham stated the following conjecture:

Let p be a prime and $a_1, \ldots, a_p$ p non-zero residues $\pmod p$. Assume that if $\sum\limits_{i=1}^{p} \epsilon_i a_i$, $\epsilon_i = 0$ or 1 (not all $\epsilon_i = 0$) is a multiple of p then $\sum\limits_{i=1}^{p} \epsilon_i$ is uniquely determined. The conjecture states that there exist only two distinct residues among the a's. We are going to prove this conjecture for all sufficiently large p. In fact we will give a sharper result. To extend our proof for the small values of p would require considerable computation, but no theoretical difficulty. Our proof is surprisingly complicated and we are not convinced that a simpler proof is not possible, but we could not find one. (P. Erdős and E. Szemerédi [4])"

The conviction that a simple proof must exist was restated by Erdős and Graham in [2].

*Center for Combinatorics, LPMC, Nankai University, Tianjin, 300071, Peoples Republic of China
†UPMC Univ Paris 06, E. Combinatoire, Case 189, 4 Place Jussieu, 75005 Paris, France.
‡Center for Combinatorics, LPMC, Nankai University, Tianjin, 300071, Peoples Republic of China

In this work, we prove Graham's conjecture for all moduli, not necessarily prime. Since our proof uses a recent result due to Savchev and Chen (proved by elementary methods, but not very short), it could not be the simple proof whose existence has been suspected by Erdős and Szemerédi. Actually the Erdős-Szemerédi Theorem may be formulated equivalently as a modular zero-sum statement:

Theorem A *(Erdős-Szemerédi [4])Let p be a sufficiently large prime and let S be a sequence of p integers in the interval $[0, p-1]$. If S takes at least three distinct values, then S has two nonempty $\pmod p$ zero-sum subsequences with distinct lengths.*

In this paper, we obtain the following generalization of this result:

Theorem 1.1 *Let n be a positive integer and let S be a sequence of n integers in the interval $[0, n-1]$. If S takes at least three distinct values, then S has two nonempty $\pmod n$ zero-sum subsequences with distinct lengths.*

In the investigation of zero-sum sequences in an abelian group G, it is quite convenient to work with an unordered sequence. This is usually done by identifying a sequence with an element of the free abelian monoid generated by G. This point of view together with the bases of zero-sum theory are presented in the text book of Geroldinger and Halter-Koch [5].

One may also define a sequence as a word. In this case, multiplication is just juxtaposition and thus x^n is the word $\underbrace{x \cdot \ldots \cdot x}_{n}$. We shall present our proofs in such a way to fit with each of these definitions.

We give below examples of sequences with a unique length for modular zero-sum subsequences.

- $S = 1^{n-1}x$, where x is an integer.
- $S = 1^{n-2}(q+1)^2$, where $n = 2q+1$.

2 Preliminaries

Let T be a subsequence of a sequence S. We shall denote by ST^{-1} the sequence obtained by deleting T from S. The sum of elements of S will be denoted by $\sigma(S)$. The maximal repetition of a value of S will be denoted by $h(S)$.

We present below a few tools:

Lemma B *(folklore) A sequence S of n integers in the interval $[0, n-1]$ has a nonempty subsequence with length $\leq h(S)$ and sum $\equiv 0 \pmod{n}$.*

Lemma B is a special case of Conjecture 4 of Erdős and Heilbronn [3]. In a note added in proofs, Erdős and Heilbronn [3] mentioned that Flor proved this conjecture using the Moser-Scherk's Theorem [7].

The next lemma is just an exercise:

Lemma C *(folklore) A sequence of $n-1$ integers in the interval $[0, n-1]$, assuming at least two distinct values, has a nonempty subsequence with sum $\equiv 0 \pmod{n}$.*

Let $S = a_1 \cdot \ldots \cdot a_t$ be a sequence of integers. We define $m * S = (ma_1) \cdot \ldots \cdot (ma_t)$. When working with a fixed moduli n, the ma_i is taken to be the value $b_i \in [0, n-1]$ such that $b_i \equiv ma_i \pmod{n}$.

The following result is a basic tool in our approach:

Theorem D *([6], [8]) Let $t \geq \frac{n+1}{2}$ be an integer. Let $a_1 \cdot \ldots \cdot a_t$ be integers and put $T = a_1 \cdot \ldots \cdot a_t$. If T has no nonempty subsequence with sum $\equiv 0 \pmod{n}$. Then there exists an integer m co-prime to n and positive integers $b_1, \ldots, b_t \in [1, n-1]$ such that $m * T = b_1 \cdot \ldots \cdot b_t$ and $b_1 + \cdots + b_t < n$.*

3 Proof of the main result

We start with one lemma:

Lemma 3.1 *Let $S = 1^v a_1 \cdot \ldots \cdot a_t$ be a sequence of positive integers with $v + t \geq \frac{n+1}{2}$, $t \geq 1$ and $2 \leq a_1 \leq \cdots \leq a_t \leq v + \sum_{i=1}^{t} a_i \leq n - j$, where j is a positive integer. Then the following hold:*

(i) $v \geq a_t + \cdots + a_{t-j+1} - j + 1$;

(ii) For any integer $k \in [2, v + \sum_{i=1}^{t} a_i]$, there exists a subsequence T of S with $|T| \geq 2$ and $\sigma(T) = k$;

(iii) If $v + \sum_{i=1}^{t} a_i \leq n - 2$, then for every integer $k \in [a_1, v + \sum_{i=2}^{t} a_i]$, there exist two subsequences T_1, T_2 of S with $\sigma(T_1) = \sigma(T_2) = k$ and $|T_1| > |T_2|$.

Proof. We have clearly

$$\begin{aligned} n - j &\geq v + \sum_{i=1}^{t} a_i \\ &\geq v + 2(t-j) + \sum_{i=t-j+1}^{t} a_i \\ &= 2(v+t) - 2j - v + \sum_{i=t-j+1}^{t} a_i \geq n + 1 - 2j - v + \sum_{i=t-j+1}^{t} a_i. \end{aligned}$$

Thus (i) holds.

By (i), we have $a_t \leq v$ and (ii) holds clearly for $k \leq a_t$. So we may assume that $k > a_t$. Let ℓ be the maximal integer of $[1, t]$ such that $\sum_{i=1}^{\ell} a_i \leq k$ and put $k' = k - \sum_{i=1}^{\ell} a_i$. Note that $k' \leq v$. Thus, $(\prod_{i=1}^{\ell} a_i) \cdot 1^{k'}$ is a subsequence of S of length at least two and of sum $= k$, proving (ii).

Let us prove (iii). Assume that $v + \sum_{i=1}^{t} a_i \leq n - 2$. Since $a_t = \sigma(a_t) = \sigma(1^{a_t})$, and since $v \geq a_1$ by (i), we may assume that $t \geq 2$. By (i),

$v \geq a_{t-1} + a_t - 1.$

Let s be the maximal integer of $[1, t]$ such that $\sum_{i=1}^{s} a_i \leq k$.

If $s = t$, note that $k - \sum_{i=2}^{t} a_i \leq (v + \sum_{i=2}^{t} a_i) - \sum_{i=2}^{t} a_i = v$. Thus, $(\prod_{i=1}^{t} a_i) \cdot 1^{k - \sum_{i=1}^{t} a_i}$ and $(\prod_{i=2}^{t} a_i) \cdot 1^{k - \sum_{i=2}^{t} a_i}$ are two subsequences of S with sum k and of distinct lengths.

Now assume that $s < t$, we have

$$k - \sum_{i=1}^{s-1} a_i \leq (\sum_{i=1}^{s+1} a_i - 1) - \sum_{i=1}^{s-1} a_i = a_s + a_{s+1} - 1 \leq a_{t-1} + a_t - 1 \leq v.$$

Thus, $(\prod_{i=1}^{s-1} a_i) \cdot 1^{k - \sum_{i=1}^{s-1} a_i}$ and $(\prod_{i=1}^{s} a_i) \cdot 1^{k - \sum_{i=1}^{s} a_i}$ are two subsequences of S with sum k and of distinct lengths. ∎

Proof of Theorem 1.1:

Suppose to the contrary of the theorem that all the nonempty $\pmod{n}$ zero-sum subsequences have the same length r (say). Then 0 is not among the values of S, otherwise

$S \cdot 0^{-1}$ would be a modular zero-sum free subsequence of S with length $n-1$, and hence $S \cdot 0^{-1}$ assumes only one value by Lemma C, a contradiction.

We distinguish two cases.

Case 1. $r \geq \frac{n}{2}$.

By Lemma B, we have $r \leq h(S)$.

Let $v = h(S)$ and write $S = a^v a_1 \cdot \ldots \cdot a_t$, where $a_i \neq a$ for $i = 1, \ldots, t$.

By our assumption, we have $t \geq 2$.

Assume first that $\gcd(a,n) > 1$. Thus $h(S) \geq r \geq \frac{n}{2} \geq \frac{n}{\gcd(a,n)}$. It forces that $r = \frac{n}{\gcd(a,n)} = \frac{n}{2}$ and $\gcd(a,n) = 2$. It follows that $2 \nmid a_i$ for $i = 1, \ldots, t$. Otherwise, $a_i \equiv \ell a \pmod n$ for some positive integer $2 \leq \ell \leq \frac{n}{\gcd(a,n)} = \frac{n}{2}$ and $a^{\frac{n}{2}-\ell} \cdot a_i$ would be a modular zero-sum subsequence with length $< r$, a contradiction.

Since $\gcd(a,n) = 2$ and all a_i are odd, we have that $a_i + a_j \equiv s_{ij} a \pmod n$ for any $i \neq j \in \{1, 2, \ldots, t\}$ and some integer $s_{ij} \in [0, \frac{n}{2}-1]$. Now $\frac{n}{2} = r = |(a_i \cdot a_j) a^{\frac{n}{2}-s_{ij}}| = 2 + \frac{n}{2} - s_{ij}$. It follows that $s_{ij} = 2$. Therefore $a_i + a_j \equiv 2a \pmod n$ for any $i \neq j \in \{1, 2, \ldots, t\}$.

If $t \geq 3$, since $a_i + a_j \equiv 2a \pmod n$ and $a_i \in [1, n-1]$ we infer that $a_1 = a_2 = \ldots = a_t$, a contradiction. So, we have $t = 2$. But now we have $a^{n-2}(a_1 a_2)$ is zero-sum modulo n, also a contradiction.

Therefore, we assume that $\gcd(a,n) = 1$. Thus for some m coprime to n, we have $m * S = R = 1^v b_1 \cdot \ldots \cdot b_t$, and $2 \leq b_1 \leq \ldots \leq b_t \leq n-1$. Clearly, every modular zero-sum subsequence of R has length $= r$.

We show next that

$$b_t \leq n - v - 1. \tag{1}$$

Suppose to the contrary that

$$b_t \geq n - v.$$

We must have $b_1 \leq n-v-1$, since otherwise, $b_1 \cdot 1^{n-b_1}$ and $b_t \cdot 1^{n-b_t}$ would be two modular zero-sum subsequences of R of distinct lengths. Since $b_t \cdot 1^{n-b_t}$ is a modular zero-sum subsequence of R, we have that $n - b_t + 1 = |b_t \cdot 1^{n-b_t}| \geq \frac{n}{2} \geq n - v$, and so $b_t \leq v + 1$. Notice that $b_1 + b_t \leq (n - v - 1) + (v + 1) = n$. Thus, $b_1 \cdot b_t \cdot 1^{n-b_1-b_t}$ and $b_t \cdot 1^{n-b_t}$ are two modular zero-sum subsequences of R of distinct lengths, a contradiction. This proves (1).

Choose a subsequence T of R with $\sigma(T) \equiv 0 \pmod n$ and with the maximal number of distinct values. Put $T = 1^\tau \cdot x_1 \cdot \ldots \cdot x_u$ and $RT^{-1} = 1^\gamma \cdot y_1 \cdot \ldots \cdot y_w$.

We shall assume that $2 \leq x_1 \leq \cdots \leq x_u$ and that $2 \leq y_1 \leq \cdots \leq y_w$.

We must have

$$x_1 \geq \gamma + 1,$$

otherwise $\sigma(1^{x_1+\tau} \cdot x_2 \cdot \ldots \cdot x_u) \equiv 0 \pmod{n}$, a contradiction.

Similarly,

$y_1 \geq \tau + 1.$

Clearly, $u \geq 1$. By (1) and since $|T| \leq v$, we have

$$\begin{aligned} w &= |RT^{-1}| - \gamma \\ &= n - |T| - \gamma \geq n - v - \gamma \\ &\geq n - x_1 - v + 1 \\ &\geq n - b_t - v + 1 \geq 2. \end{aligned}$$

By (1) and since $v \geq \frac{n}{2}$, we have that $b_t < \frac{n}{2}$. It follows that

$$n > x_u + y_w \geq x_1 + y_1 \geq \gamma + 1 + \tau + 1 = v + 2 > n - v.$$

Since $x_1 y_1 1^{n-x_1-y_1}$ and $x_u y_w 1^{n-x_u-y_w}$ are modular zero-sum subsequences, we conclude that $x_1 = \cdots = x_u$ and $y_1 = \cdots = y_w$. Note that S has at least 3 distinct values, we derive that $x_1 \neq y_1$. Thus, $T' = 1^{n-x_1-y_1} \cdot x_1 \cdot y_1$ is a modular zero-sum subsequence of R, with more distinct values than T, contradicting the choice of T.

Case 2. $r < \frac{n}{2}$.

Choose a modular zero-sum subsequence T of S. Then ST^{-1} is a modular zero-sum free subsequence with $|ST^{-1}| > \frac{n}{2}$. By Theorem D, for some positive integer m coprime to n, we have $m * (ST^{-1}) = 1^{\gamma} \cdot y_1 \cdot \ldots \cdot y_w$, where $2 \leq y_1 \leq \cdots \leq y_w < \gamma + \sum_{i=1}^{w} y_i \leq n-1$. Put $R = m * S$. Clearly every modular zero-sum subsequence of R has length $= r$. So without loss of generality, we may take $m = 1$. Also, put $T = 1^{\tau} \cdot x_1 \cdot \ldots \cdot x_u$, where $2 \leq x_1 \leq \cdots \leq x_u \leq n-1$.

We first note that

$$x_1 \geq \gamma + 1,$$

otherwise $1^{x_1} \cdot (x_1^{-1} T)$ is a zero-sum sequence of length larger than $|T|$, a contradiction.

We must have $w \geq 1$. Otherwise, $\gamma = |ST^{-1}| \geq \frac{n+1}{2}$ and hence $x_1 \geq \gamma + 1 > n - \gamma$. Therefore, $1^{n-x_i} \cdot x_i$ is a modular zero-sum subsequence of S for every $i = 1, \ldots, u$. This forces that $x_1 = \cdots = x_u$, a contradiction.

We must have $u \geq 2$, since otherwise (observing that $u \neq 0$) $u = 1$ and

$x_1 = n - \tau = n - |T| + 1 = |ST^{-1}| + 1 \leq \gamma + \sum_{i=1}^{w} y_i.$

By Lemma 3.1 (ii) with $j = 1$, there is a subsequence U of ST^{-1} with $|U| \geq 2$ such that $x_1 = \sigma(U)$. Now $1^\tau x_1$ and $1^\tau U$ are modular zero-sum subsequences with distinct lengths, a contradiction.

Thus,

$$w \geq 1 \quad \text{and } u \geq 2. \tag{2}$$

Let X_ℓ be the unique integer of $[0, n-1]$ such that

$$X_\ell \equiv \sum_{i=1}^{\ell} x_i \pmod{n}$$

for $\ell = 1, \ldots, u$.

Applying Lemma 3.1 (ii), we have that

$$x_1 \geq \gamma + \sum_{i=1}^{w} y_i + 1, \tag{3}$$

and so

$$\gamma + \sum_{i=1}^{w} y_i \leq x_1 - 1 \leq n - 2.$$

By Lemma 3.1 (iii), we have that

$$\sum(T) \cap [y_1, \gamma + \sum_{i=2}^{w} y_i] = \emptyset, \tag{4}$$

where $\sum(T)$ denotes the set of the sums of the nonempty subsequences of T.

By (3), we have that

$$x_i \geq x_1 \geq \gamma + \sum_{i=1}^{w} y_i + 1 \geq |ST^{-1}| + 1 > \frac{n}{2} + 1 \tag{5}$$

for $i = 1, \ldots, u$, which implies

$$x_{i_1} + x_{i_2} \not\equiv 1, 2 \pmod{n} \tag{6}$$

for any $1 \leq i_1 < i_2 \leq u$. By Lemma 3.1 (i), we see that

$$\gamma \geq y_1. \tag{7}$$

Now by combining (4), (6) and (7) we conclude that $X_2 \notin [1, \gamma + \sum\limits_{i=2}^{w} y_i]$, i.e., $x_1 + x_2 - n = X_2 \geq \gamma + \sum\limits_{i=2}^{w} y_i + 1 \geq \gamma + w = |ST^{-1}| \geq \frac{n+1}{2}$. It follows that $x_2 \geq \frac{x_1+x_2}{2} \geq \frac{3n+1}{4} = n - \frac{n-1}{4}$, i.e.,

$$x_2 \geq \lceil n - \frac{n-1}{4} \rceil = n - \lfloor \frac{n-1}{4} \rfloor. \tag{8}$$

Now we shall show that

$$x_u \leq n - \tau - 3. \tag{9}$$

Since $u \geq 2$, we have $x_u \leq n - \tau - 1$. Suppose $x_u \in \{n - \tau - 1, n - \tau - 2\}$. Then $X_{u-1} \in \{1, 2\}$. By (5) and (6), we have $u - 1 \geq 3$. By (7), $\gamma \geq y_1 \geq 2$, thus, $T \cdot (\prod\limits_{i=1}^{u-1} x_i)^{-1} \cdot 1^{X_{u-1}}$ is a zero-sum subsequence of S with length $|T| - (u-1) + X_{u-1} \leq |T| - 3 + 2 = |T| - 1$, a contradiction. Therefore, $x_u \leq n - \tau - 3$.

Let $t \in [1, u]$ be the largest integer such that $X_i > \lceil \frac{n-1}{4} \rceil$ for every $i \in \{1, \ldots, t\}$. By (8) and (9), we see that $n - \lfloor \frac{n-1}{4} \rfloor \leq x_i \leq n - 3$ for $i = 2, \ldots, u$. It follows that

$$\lceil \frac{n-1}{4} \rceil < X_\ell = X_{\ell-1} + x_\ell - n \leq X_{\ell-1} - 3 \tag{10}$$

for $\ell = 2, 3, \ldots, t$.

Put

$$q = \min(\lceil \frac{u+1}{3} \rceil, t).$$

We shall show that

$$X_q \leq \gamma + \sum_{i=2}^{w} y_i.$$

If $q = \lceil \frac{u+1}{3} \rceil \leq t$, then by (10) $X_q \leq X_1 - 3(q-1) = x_1 - 3(q-1) \leq n - \tau - 3 - 3\lceil \frac{u+1}{3} \rceil + 3 \leq n - \tau - u - 1 = |ST^{-1}| - 1 \leq \gamma + \sum\limits_{i=2}^{w} y_i$. If $q = t < \lceil \frac{u+1}{3} \rceil \leq u$, then $X_{t+1} \leq \lceil \frac{n-1}{4} \rceil$, which implies that $X_t = X_{t+1} + (n - x_{t+1}) \leq X_{t+1} + \lfloor \frac{n-1}{4} \rfloor \leq \lceil \frac{n-1}{2} \rceil \leq |ST^{-1}| - 1 \leq \gamma + \sum\limits_{i=2}^{w} y_i$.

Thus, by (4) and (7), we have that $X_q < y_1 \leq \gamma$. Therefore, $T \cdot (\prod\limits_{i=1}^{q} x_i)^{-1} \cdot 1^{X_q}$ is a zero-sum subsequence of S with length $|T| - q + X_q > |T| - \lceil \frac{u+1}{3} \rceil + \lceil \frac{n-1}{4} \rceil \geq |T| - \lceil \frac{|T|+1}{3} \rceil + \lceil \frac{n-1}{4} \rceil \geq |T| - \lceil \frac{n+1}{6} \rceil + \lceil \frac{n-1}{4} \rceil \geq |T|$, a contradiction. ∎

Acknowledgement. This research has been supported in part by the 973 Project, the PCSIRT Project of the Ministry of Education, the Ministry of Science and Technology, the National Science Foundation of China. This paper was completed partly during a visit by the first author to University of P. et M. Curie in France. He would like to thank the host institution for its kind hospitality.

References

[1] N. Alon, *Subset sums*, J. Number Theory, 27(1987) 196-205.

[2] P. Erdős and R.L. Graham, *Old and new problems and results in combinatorial number theory*. Monographies de L'Enseignement Mathmatique, 28. Universit de Genve, L'Enseignement Mathmatique, Geneva, 1980. 128 pp. (Reviewer: L. C. Eggan)

[3] P. Erdős and H. Heilbronn, *On the Addition of residue classes mod p*, Acta Arith. 9(1964), 149-159.

[4] P. Erdős and E. Szemerédi, *On a problem of Graham*, Publ. Math. Debrecen, 23(1976), no. 1-2, 123-127.

[5] A. Geroldinger, F. Halter-Koch, *Non-unique factorizations. Algebraic, combinatorial and analytic theory. Pure and Applied Mathematics (Boca Raton)*, 278. Chapman & Hall/CRC, Boca Raton, FL, 2006. xxii+700 pp.

[6] S. Savchev and F. Chen, *Long zero-free sequences in finite cyclic groups*, Discrete Mathematics, 307(2007) 2671-2679.

[7] P. Scherk, Distinct elements in a set of sums, Amer. Math. Monthly, 62 (1955), pp. 46–47.

[8] P.Z. Yuan, *On the index of minimal zero-sum sequences over finite cyclic groups*, J. Combin. Theory Ser. A, 114(2007) 1545-1551.