# Study on Problem 4 of the 2021 China High School Mathematics League (Add-on Round)

Kecen Sha

September 15, 2021

## 1. Original Problem and Solution

**Original Problem:** Find the smallest positive constant $c$ such that for any integer $n \geq 4$, and any subset $A \subseteq \{1, 2, 3, \ldots, n\}$, if $|A| > cn$, then there exists a function $f : A \to \{1, -1\}$ satisfying

$$\left| \sum_{a \in A} f(a) \cdot a \right| \leq 1.$$

On one hand, we can show that $c \geq \frac{2}{3}$. Conversely, suppose $c < \frac{2}{3}$. Take $n = 6$ and $A = \{1, 4, 5, 6\}$. Then $|A| > cn$. Note that

$$\sum_{a \in A} f(a) \cdot a \equiv \sum_{a \in A} a \equiv 0 \pmod 2,$$

and the condition $\left|\sum_{a \in A} f(a) \cdot a\right| \leq 1$ implies $\sum_{a \in A} f(a) \cdot a = 0$. However, this contradicts the fact that $A$ contains no subset whose elements sum to $\frac{16}{2} = 8$.

We now prove that $c = \frac{2}{3}$ is sufficient. First, we establish the following lemma:

**Lemma 1.** *Let $a_1, a_2, \ldots, a_n$ be nonzero integers with $|a_i| \leq i$ for $i = 1, 2, \ldots, n$. Then there exist signs $\varepsilon_i \in \{-1, 1\}$ such that*

$$\left| \sum_{i=1}^{n} \varepsilon_i a_i \right| \leq 1.$$

*Proof.* We proceed by induction on $n$. The claim is trivial for $n = 1, 2$. Assume it holds for all $n \leq k$, and consider $n = k + 1$. Without loss of generality, assume all $a_i$ are positive (otherwise replace $a_i$ by $-a_i$, which does not affect the conclusion). Since $1 \leq a_{k+1} \leq k + 1$ and $1 \leq a_k \leq k$, we have $|a_{k+1} - a_k| \leq k$.

If $|a_{k+1} - a_k| = 0$, apply the induction hypothesis to $a_1, a_2, \ldots, a_{k-1}$. Otherwise, if $|a_{k+1} - a_k| \leq k$, apply the induction hypothesis to $a_1, a_2, \ldots, a_{k-1}, |a_{k+1} - a_k|$. The lemma follows. $\square$

Now return to the original problem. Let $A = \{a_1, a_2, \ldots, a_t\} \subseteq \{1, 2, \ldots, n\}$ with $a_1 < a_2 < \cdots < a_t$, and $t > \frac{2n}{3}$.

**(1)** Suppose $t$ is even. Consider the $\frac{t}{2}$ positive integers:

$$a_t - a_{t-1},\ a_{t-2} - a_{t-3},\ \dots,\ a_2 - a_1.$$

We claim that for every $m \in \{0, 1, 2, \dots, \frac{t}{2} - 1\}$, at most $m$ of these $\frac{t}{2}$ numbers are $\geq \frac{t}{2} - m + 1$. (*)

Otherwise, suppose there exists $m \in \{0, 1, 2, \dots, \frac{t}{2} - 1\}$ such that at least $m + 1$ of them are $\geq \frac{t}{2} - m + 1$. Then

$$\begin{aligned} a_t &= (a_t - a_{t-1}) + (a_{t-1} - a_{t-2}) + \cdots + (a_3 - a_2) + (a_2 - a_1) + a_1 \\ &\geq (m+1)\left(\frac{t}{2} - m + 1\right) + (t - (m+1)) \cdot 1 \\ &= t + (m+1)\left(\frac{t}{2} - m\right). \end{aligned}$$

Using the monotonicity of the quadratic function in $m$, we obtain

$$t + (m+1)\left(\frac{t}{2} - m\right) \geq \frac{3}{2}t,$$

with equality attained when $m = 0$ or $m = \frac{t}{2} - 1$. Thus $a_t \geq \frac{3}{2}t > n$, a contradiction.

From (*) it follows that for every $m \in \{0, 1, 2, \dots, \frac{t}{2} - 1\}$, at least $\frac{t}{2} - m$ of the differences are $\leq \frac{t}{2} - m$. Let $i = \frac{t}{2} - m$. Then for every $i \in \{1, 2, \dots, \frac{t}{2}\}$, at least $i$ of the $\frac{t}{2}$ numbers are $\leq i$.

Rearranging $a_t - a_{t-1}, a_{t-2} - a_{t-3}, \dots, a_2 - a_1$ in increasing order as $b_1 \leq b_2 \leq \cdots \leq b_{\frac{t}{2}}$, we have $0 < b_i \leq i$. By the lemma, there exist signs $\varepsilon_i \in \{-1, 1\}$ such that

$$\left| \sum_{i=1}^{t/2} \varepsilon_i b_i \right| \leq 1.$$

Expressing each $b_i$ as a difference of elements from $A$, we obtain the desired result.

**(2)** Suppose $t$ is odd. Consider the $\frac{t+1}{2}$ positive integers:

$$a_t - a_{t-1},\ a_{t-2} - a_{t-3},\ \dots,\ a_3 - a_2,\ a_1.$$

A similar argument applies. We claim that for every $m \in \left\{0, 1, 2, \dots, \frac{t-1}{2}\right\}$, at most $m$ of these $\frac{t+1}{2}$ numbers are $\geq \frac{t+1}{2} - m + 1$. ($\triangle$)

Otherwise, suppose there exists $m \in \left\{0, 1, 2, \dots, \frac{t-1}{2}\right\}$ such that at least $m+1$ of them are $\geq \frac{t+1}{2} - m + 1$. Then

$$\begin{aligned} a_t &\geq (m+1)\left(\frac{t+1}{2} - m + 1\right) + (t - (m+1)) \cdot 1 \\ &= t + (m+1)\left(\frac{t+1}{2} - m\right). \end{aligned}$$

Using the monotonicity of the quadratic function, we get

$$t + (m+1)\left(\frac{t+1}{2} - m\right) \geq \frac{3t+1}{2},$$

so $a_t \geq \frac{3t+1}{2} > n$, again a contradiction.

From ($\triangle$), for every $m \in \left\{0, 1, 2, \ldots, \frac{t-1}{2}\right\}$, at least $\frac{t+1}{2} - m$ of the numbers are $\leq \frac{t+1}{2} - m$. Let $i = \frac{t+1}{2} - m$. Then for every $i \in \left\{1, 2, \ldots, \frac{t+1}{2}\right\}$, at least $i$ of the $\frac{t+1}{2}$ numbers are $\leq i$.

Rearranging them as $b_1 \leq b_2 \leq \cdots \leq b_{\frac{t+1}{2}}$, we have $0 < b_i \leq i$. Again, by the lemma, there exist signs $\varepsilon_i \in \{-1, 1\}$ such that

$$\left|\sum_{i=1}^{(t+1)/2} \varepsilon_i b_i\right| \leq 1.$$

Expressing each $b_i$ as a difference of elements from $A$, the proof is complete.

## 2. Optimal Bound for Large $n$

In fact, for sufficiently large $n$, the constant $\frac{2}{3}$ is not optimal. For example, taking $A = \{2, 4, \ldots, n\}$ when $n$ is even shows that $c \geq \frac{1}{2}$. We will show that $c = \frac{1}{2}$ is sufficient—and hence optimal—for large $n$.

Define $c_n$ as follows:

$$c_n = \begin{cases} 4k+1 & \text{if } n = 8k+2 \text{ or } 8k+3, \\ 4k+2 & \text{if } n = 8k+4, \ldots, 8k+7, \\ 4k+3 & \text{if } n = 8k+8 \text{ or } 8k+9. \end{cases}$$

Our strengthened result is stated below:

**Theorem 1.** *There exists an absolute constant $N \leq 166$ such that for any integer $n \geq N$, and any subset $A \subseteq \{1, 2, 3, \ldots, n\}$, if $|A| > c_n$, then there exists a function $f : A \to \{1, -1\}$ satisfying*

$$\left|\sum_{a \in A} f(a) \cdot a\right| \leq 1.$$

It is straightforward to verify that $c_n$ is optimal for each fixed $n$. A known result establishes Theorem 1 with $N = 556$. Our goal is to provide an alternative proof and improve the bound on $N$.

We first handle the case where there exists a subset $B \subset A$ such that either $B = \{1\}$ or $B = \{x, x+1\}$. In such cases, it suffices to consider $A \setminus B$. The remaining exceptional configurations can be dealt with manually.

**Theorem 2.** *There exists an absolute constant $N \leq 166$ such that for any integer $n \geq N$, and any subset $A \subseteq \{1, 2, 3, \ldots, n\}$, if $|A| > \frac{n-7}{2}$, then there exists a function $f : A \to \{1, -1\}$ satisfying*

$$\left| \sum_{a \in A} f(a) \cdot a \right| \leq 2.$$

### 2.1. Notation and Definitions

We begin by introducing some standard notation. Uppercase letters such as $A, B, C, \ldots$, along with their primed variants $A', B', C', \ldots$ and double-primed versions $A'', B'', C'', \ldots$, denote sets. Barred or tilded letters $\bar{A}, \bar{B}, \bar{C}, \tilde{A}, \tilde{B}, \tilde{C}, \ldots$ denote multisets.

For sets $A$ and $B$, we write $A - B$ to mean $A \setminus (A \cap B)$. We use the standard notation $[n] = \{1, 2, \ldots, n\}$. For any set $X$ or multiset $\bar{X}, \tilde{X}$, we adopt the following conventions:

1. $|X|$, $|\bar{X}|$, or $|\tilde{X}|$ denotes its cardinality;
2. $A^{\uparrow} = \{a_i\}_{i=1}^{|A|}$ or $\bar{A}^{\uparrow} = \{a_i\}_{i=1}^{|\bar{A}|}$ denotes the sequence of elements of $A$ or $\bar{A}$ sorted in non-decreasing order;
3. $s(A) = \sum_{x \in A} x$, $s(\bar{A}) = \sum_{x \in \bar{A}} x$ denotes the sum of elements, with multiplicities counted for multisets.

Throughout this work, all sets are assumed to be subsets of $\mathbb{N}$. For a set $A$ (or multiset $\bar{A}, \tilde{A}$) and an integer $m$, define

$$\mathcal{L}(A, m) = |A \cap [m, +\infty)|, \quad \mathcal{L}(\bar{A}, m) = |\bar{A} \cap [m, +\infty)|, \quad \mathcal{L}(\tilde{A}, m) = |\tilde{A} \cap [m, +\infty)|.$$

Denote $d(A) = \max_{x,y \in A} |x - y|$. If $|A|$ is even, define

$$\bar{A} = \{a_2 - a_1, a_4 - a_3, \ldots, a_{|A|} - a_{|A|-1}\}, \quad \tilde{A} = \{a_3 - a_2, \ldots, a_{|A|-1} - a_{|A|-2}\};$$

otherwise, define

$$\bar{A} = \{a_1, a_3 - a_2, \ldots, a_{|A|} - a_{|A|-1}\}, \quad \tilde{A} = \{a_2 - a_1, \ldots, a_{|A|-1} - a_{|A|-2}\}.$$

For any sequence $\{a_i\}$, we set $a_l = 0$ if $l \leq 0$.

Next, we introduce key definitions relevant to our theorems and proofs.

**Definition 1.** *We say that $A$ (or $\bar{A}$) is $s$-**excellent** if there exists a function $f : A \to \{1, -1\}$ (or $f : \bar{A} \to \{1, -1\}$) such that*

$$\left| \sum_{a \in A} f(a) \cdot a \right| \leq s \quad \left(\text{or } \left| \sum_{a \in \bar{A}} f(a) \cdot a \right| \leq s\right).$$

*If equality holds, we say that $A$ (or $\bar{A}$) is **exactly** $s$-**excellent**.*

**Definition 2.** *We say that $A$ (or $\bar{A}$) is* ***strongly expressive*** *if for every integer $x \le s(A)$ with $x \equiv s(A) \pmod 2$, there exists a function $f : A \to \{1, -1\}$ (or $f : \bar{A} \to \{1, -1\}$) such that*

$$\sum_{a \in A} f(a) \cdot a = x \quad \text{(or } \sum_{a \in \bar{A}} f(a) \cdot a = x\text{).}$$

By definition, if $A$ or $\bar{A}$ is strongly expressive, then it is 1-excellent. The following proposition is straightforward to verify:

**Proposition 1.** *The following statements hold:*

1. *If $A, B \subset \mathbb{N}$ are disjoint, $A$ is $s$-excellent and $B$ is exactly $t$-excellent, then $A \cup B$ is $|s - t|$-excellent.*
2. *If $\bar{A}$ is $s$-excellent, then $A$ is $s$-excellent.*
3. *If $\bar{A}$ is strongly expressive and $b \le s(\bar{A})$, then $\bar{A} \cup \{b\}$ is strongly expressive.*

*Proof.* The second statement follows directly from the definition. Let $f$ and $g$ be the corresponding sign functions for $A$ and $B$ in Definition 1. Defining $h : A \cup B \to \{1, -1\}$ by $h|_A = f$ and $h|_B = \pm g$ proves the first statement.

For the third statement, note that $\bar{A}$ can represent all values $s(\bar{A}), s(\bar{A}) - 2, \ldots, \varepsilon$, where $\varepsilon \in \{0, 1\}$. Adding $b$ allows us to represent $b + \varepsilon, \ldots, b + s(\bar{A})$, while adding $-b$ yields $b - \varepsilon, \ldots$. Since $b \le s(\bar{A})$, this sequence covers all residues modulo 2 down to 0 or 1, depending on $b + s(\bar{A}) \pmod 2$. □

### 2.2. Useful Lemmas

Certain multisets are $(r + 1)$-excellent, as shown in the following lemmas.

**Lemma 2.** *Let $r \in \mathbb{N}$, and let $a_1, a_2, \ldots, a_n$ be nonzero integers with $|a_i| \le i + r$ for $i = 1, 2, \ldots, n$. Then there exist signs $\varepsilon_i \in \{-1, 1\}$ such that*

$$\left| \sum_{i=1}^{n} \varepsilon_i a_i \right| \le r + 1.$$

*Proof.* We proceed by induction on $n$. The claim is trivial for $n = 1, 2$. Assume it holds for all $n \le k$, and consider $n = k + 1$. Without loss of generality, assume all $a_i > 0$. Since $1 \le a_{k+1} \le k + 1 + r$ and $1 \le a_k \le k + r$, we have $|a_{k+1} - a_k| \le k + r$.

If $|a_{k+1} - a_k| = 0$, apply the induction hypothesis to $a_1, \ldots, a_{k-1}$. Otherwise, apply it to $a_1, \ldots, a_{k-1}, |a_{k+1} - a_k|$. The lemma follows. □

We now present a useful criterion for excellence.

**Lemma 3** (Sufficient condition for an excellent set)**.** *Let $A \subseteq \mathbb{N}^*$ and $r \in \mathbb{N}$. If for all $m \in [0, |A|) \cap \mathbb{N}$, we have*

$$\mathcal{L}(\bar{A}, |\bar{A}| + r + 1 - m) \le m,$$

*then $A$ is $(r + 1)$-excellent.*

*Proof.* Let $\bar{A}^{\uparrow} = \{a_1, \ldots, a_{|\bar{A}|}\}$. If $a_i \geq i+r+1$, then $\mathcal{L}(\bar{A}, i+r+1) > |\bar{A}|-i$, a contradiction. Hence $a_i \leq i+r$ for all $i$, and Lemma 2 implies $\bar{A}$ is $(r+1)$-excellent. Therefore, $A$ is also $(r+1)$-excellent. □

The next two lemmas show that if $|A| > 0.4n + 1.2$, then the condition in Lemma 3 nearly holds for $r = 1, 2$, suggesting that $A$ is likely 2- or 3-excellent.

**Lemma 4.** *Let $A \subseteq [n]$ and $|A| > \frac{2(n+3)}{5}$. Then for all $m \in \{2, 3, \ldots, |\bar{A}| - 2\}$, we have*

$$\mathcal{L}(\bar{A}, |\bar{A}| + 2 - m) \leq m. \tag{1}$$

*Proof.* Let $t = |A|$ and $A^{\uparrow} = \{a_i\}_{i=1}^{t}$.

If $t = 2p$ is even, consider the $p$ positive integers:

$$\bar{A} = \{a_t - a_{t-1}, a_{t-2} - a_{t-3}, \ldots, a_2 - a_1\}.$$

Suppose there exists $m \in \{2, \ldots, p-2\}$ such that $\mathcal{L}(\bar{A}, p+2-m) \geq m+1$. Then

$$\begin{aligned} a_t &\geq (m+1)(p-m+2) + (t-(m+1)) \cdot 1 \\ &= t + (m+1)(p-m+1). \end{aligned} \tag{2}$$

Using the monotonicity of the quadratic in $m$, we find

$$t + (m+1)(p-m+1) \geq \frac{5}{2}t - 3 > n,$$

with minimum at $m = 2$ or $m = p-2$. Thus $a_t > n$, a contradiction.

If $t = 2p+1$ is odd, consider

$$\bar{A} = \{a_t - a_{t-1}, \ldots, a_2 - a_1, a_1\}.$$

Suppose there exists $m \in \{2, \ldots, p-1\}$ such that $\mathcal{L}(\bar{A}, p+3-m) \geq m+1$. Then

$$\begin{aligned} a_t &\geq (m+1)(p+3-m) + (t-(m+1)) \cdot 1 \\ &= t + (m+1)(p+2-m). \end{aligned} \tag{3}$$

Again, the quadratic in $m$ gives

$$t + (m+1)(p+2-m) \geq \frac{5}{2}t - \frac{3}{2} > n,$$

so $a_t > n$, a contradiction. □

**Lemma 5.** *Let $A \subseteq [n]$ and $|A| > \frac{2(n+3)}{5}$. Then for all $m \in \{3, \ldots, |\bar{A}| - 1\}$, we have*

$$\mathcal{L}(\bar{A}, |\bar{A}| + 3 - m) \leq m - 1. \tag{4}$$

*Proof.* Let $t = |A|$ and $A^{\uparrow} = \{a_i\}_{i=1}^{t}$.

If $t = 2p$, suppose $\mathcal{L}(\bar{A}, p+3-m) \geq m$ for some $m \in \{3, \ldots, p-1\}$. Then

$$a_t \geq m(p-m+3) + (t-m) \cdot 1 = t + m(p-m+2) \geq \frac{5}{2}t - 3 > n, \tag{5}$$

a contradiction.

If $t = 2p+1$, a similar computation yields $a_t > n$, again a contradiction. □

### 2.3. Proof of Theorem 2

We argue by contradiction. Suppose $A$ is not 2-excellent.

#### 2.3.1 Case 1: $\mathcal{L}(\bar{A},3) < |\bar{A}|$.

Let $t = |A|$ and $s = \lceil t/2 \rceil = |\bar{A}|$. Since $t > \frac{2n+6}{5}$ for $n \geq 61$, Lemmas 2 and 4 imply that $\mathcal{L}(\bar{A}, s+1) \geq 1$. Consequently, there exists an index $j \equiv t \pmod 2$ such that

$$a_j - a_{j-1} \geq s + 1.$$

We claim that there exist two consecutive integers $v, v+1 \in A$ with $v, v+1 \notin \{a_j, a_{j-1} - 1\}$. Otherwise, we would have

$$a_t = (a_t - a_j) + (a_j - a_{j-1}) + a_{j-1} \geq [2(t-j) - 1] + (s+1) + [2(j-1) - 1] = \frac{5t}{2} - 3 > n,$$

a contradiction.

Hence, the set $B = A \setminus \{v, v+1\}$ cannot be 3-excellent. Note that $|\bar{B}| = s - 1$. By our choice of $v$, the difference $d = a_j - a_{j-1}$ must appear in $\bar{B}$ as an element generated by the pair $(a_{j-1}, a_j)$. Applying Lemma 5, we obtain

$$\mathcal{L}(\bar{B}, s + 2 - m) \leq m - 1, \quad \forall\, m \geq 3. \tag{6}$$

Moreover, we have $\mathcal{L}(\bar{B}, s) + \mathcal{L}(\tilde{B}, s) \leq 2$ and $\mathcal{L}(\tilde{B}, 2s) = 0$; otherwise,

$$a_t \geq 3s + 1 + (t - 5) \geq \frac{5t}{2} - 4 > n.$$

Since $d = a_j - a_{j-1} \geq s + 1$, inequality (6) implies that for any subset $\bar{X} \subseteq \bar{B} \setminus \{d\}$ and any $y \in \mathbb{N}^*$,

$$\mathcal{L}(\bar{X} \cup \{y\}, s + 1 - m) \leq m, \quad \forall\, m \geq 2. \tag{7}$$

We now establish several auxiliary lemmas by contradiction.

**Lemma 6.** $\mathcal{L}(\bar{B}, s) = 1$.

*Proof.* Suppose $\mathcal{L}(\bar{B}, s) = 2$. Choose $u \in B \setminus \{d\}$ with $u \geq s$. If $|u - d| > s$, then $u + d \geq 3s + 1$, and hence

$$a_t \geq 3s + 1 + (t - 2) \geq \frac{5t}{2} - 1 > n,$$

a contradiction. Thus $|u - d| \leq s$.

Now let $\bar{X} = \bar{B} \setminus \{d, u\}$ and $y = |u - d|$. One verifies that inequality (7) holds for all $m \geq 0$, which yields

$$\mathcal{L}(\bar{X} \cup \{y\}, |\bar{X} \cup \{y\}| + 3 - m) \leq m, \quad \forall\, m \geq 0. \tag{8}$$

By Lemma 2, the set $\bar{X} \cup \{y\}$ is 3-excellent, implying that $B$ is 3-excellent—contradicting our earlier conclusion. $\square$

**Lemma 7.** *Let $B_1 = B \cap (0, a_{j-1})$ and $B_2 = B \cap (a_j, n]$. Then there exist $x, y \in B_1$ or $x, y \in B_2$ such that*
$$d - s \leq |x \pm y| \leq d + s.$$

*Proof.* Assume the contrary. Write $B_1^{\uparrow} = \{x_1 < \cdots < x_p\}$ and $B_2^{\uparrow} = \{y_1 < \cdots < y_q\}$. Then $x_p - x_1 \leq d - s - 1$; otherwise $x_p - x_1 \geq d + s + 1$. But since $\mathcal{L}(\bar{B}, 2s) = 0$ and $\mathcal{L}(\bar{B}, s) = 1$, we have $x_{i+1} - x_i \leq 2s$ for all $i$, and the intermediate value principle yields a contradiction. Similarly, $y_q - y_1 \leq d - s - 1$.

Therefore,

$$t - 6 = p + q - 2 \leq (x_p - x_1) + (y_q - y_1) \leq 2(d - s - 1),$$

which implies $d \geq t - 2$. On the other hand, $d \leq n + 1 - t$.

If $x_1 + x_2 \geq d + s + 1$, then $x_2 > \frac{d+1+s}{2}$, and consequently

$$y_q > a_2 + d + (t - 6) \geq n \quad (\text{for } n \geq 28),$$

a contradiction. Hence $x_1 + x_2 \leq d - s - 1$, and by the same reasoning, $x_p + x_{p-1} \leq d - s - 1$, which gives $p \leq \frac{n+1-s-t}{2}$.

Similarly, $q - 1 \leq y_q - y_1 \leq d - s - 1 \leq n - t - s$. Combining these,

$$t - 4 = p + q \leq \frac{3}{2}(n + 1 - s - t) \leq \frac{3n+3}{2} - \frac{9}{4}t,$$

which contradicts the assumption $n \geq 122$. □

**Lemma 8.** *Write $B^{\uparrow} = \{b_i\}_{i=1}^{t-2}$. Then there exists an index $l$ such that $b_l = a_j$ and $b_{l-1} = a_{j-1}$, and there exist indices $k, r$ with $1 \leq r \leq 4$ such that either $k \leq l - 1$ or $k \geq l + r$, and*
$$b_k - b_{k-r} \geq s - 1.$$

*Proof.* By Lemma 7, choose $x, y \in B_1$ or $B_2$ such that $z = |d - |x \pm y|| \leq s$. Define a set $Y$ as follows:

$$Y = \begin{cases} B \setminus \{x, y, a_j, a_{j-1}\}, & \text{if } z = 0, \\ B \setminus \{x, y, a_j, a_{j-1}, z\}, & \text{if } z > 0 \text{ and } z \in B \setminus \{x, y, a_j, a_{j-1}\}, \\ (B \setminus \{x, y, a_j, a_{j-1}\}) \cup \{z\}, & \text{if } z > 0 \text{ and } z \notin B \setminus \{x, y, a_j, a_{j-1}\}. \end{cases}$$

Then $|Y| \geq t - 7 > \frac{2n+6}{5}$ for $n \geq 113$, so by Lemma 2,

$$\mathcal{L}(\bar{Y}, |\bar{Y}| + 3 - m) \leq m - 1, \quad \forall\, m \geq 3. \tag{9}$$

Note that $|\bar{Y}| \geq s - 4$, and if $z \notin B \setminus \{x, y, a_j, a_{j-1}\}$, then $|\bar{Y}| \geq s - 3$.

We claim $\mathcal{L}(\bar{Y}, |\bar{Y}| + 2) \leq 1$; otherwise,

$$\max \bar{Y} \geq 2(s - 2) + d + (t - 10) > n \quad (\text{for } n \geq 88),$$

a contradiction. Similarly, $\mathcal{L}(\bar{Y}, |\bar{Y}| + 1) \leq 2$.

Since $Y$ cannot be 3-excellent, Lemma 2 forces $\mathcal{L}(\bar{Y}, |\bar{Y}|+3) = 1$. If $z \in (a_{j-1}, a_j)$, then $z \notin B \setminus \{x, y, a_j, a_{j-1}\}$, and $\mathcal{L}(\bar{Y}, s) = 1$. But $z \leq s$, so the only way $z$ contributes to $\mathcal{L}(\bar{Y}, s)$ is if $z = s$ and $a_{j-1} = 0$. Even then, $|\bar{Y}| \geq s-2$ and $\mathcal{L}(\bar{Y}, s+1) = 1$, yet $z$ cannot contribute to $\mathcal{L}(\bar{Y}, s+1)$; thus all contributions come from elements of $B$. If $z \notin (a_{j-1}, a_j)$, then contributions to $\mathcal{L}(\bar{Y}, s-1)$ also lie in $B$. Restoring the removed elements to $Y$ completes the proof. □

Now define the interval
$$I = [b_l - 1, b_l] \cup [b_{k-r}, b_k].$$

**Lemma 9.** *There are at least $d-18$ elements in $\bar{B}$ equal to 1 that are not generated by elements within $[b_{k-r}, b_k]$.*

*Proof.* Let $g$ denote the number of such elements. The gap $b_k - b_{k-r} \geq s-1$ involves at most two elements from $\bar{B}$ and two from $\tilde{B}$. Hence,

$$n \geq b_{t-2} \geq g + 2(s-4-g) + d + s - 1 + (t-2-s+1-2),$$

which simplifies to $g \geq d + t + 2s - 12 - n \geq d - 18$. □

**Lemma 10.** *For every $0 \leq i \leq t-4$ such that $(b_i, b_{i+2}) \cap I = \varnothing$, we have*

$$b_{i+2} - b_i \leq 15 \quad \textit{and} \quad d \leq s + 14.$$

*More precisely,*
$$(b_{i+2} - b_i - 2) + (d - (s+1)) \leq 13.$$

*Proof.* If some $i$ violates the first inequality, then

$$n \geq b_{t-2} \geq d + s - 1 + (t-9) + 16 > n,$$

a contradiction. If $d \geq s + 15$, then

$$n \geq b_{t-2} \geq (s+15) + s - 1 + (t-7) > n,$$

again a contradiction. The refined bound follows by a similar estimation. □

We now conclude this case. By Lemmas 9 and 10, and since $d - 18 \geq 14$ for $n \geq 130$, there are at least 14 copies of the value 1 in $\bar{B}$ that are "free" from the interval $I$.

Write $\bar{B} = \{x_1, x_2, \ldots, x_{s-1}\}$, where $x_i = b_{t-2i} - b_{t-2i-1}$. Let $\Delta = d - s - 1$. Select $14 - \Delta$ of the free unit elements (including possibly $v + 1 - v = 1$), and remove any elements associated with the gap $b_l - b_{l-1}$ or the interval $[b_{k-r}, b_k]$. Denote the remaining multiset by $\bar{C}$. Then
$$|\bar{C}| \geq s - 18 + \Delta.$$

The indices of the $b_i$ involved in $\bar{C}$ lie in at most three disjoint subintervals of $(0, n] \setminus I$. Within each such block, we can pair adjacent differences $(x_w, x_{w+1}) = (b_{i+1} - b_i, b_{i+2} - b_{i+1})$ and replace them with
$$(y_w, y_{w+1}) = (b_{i+2} - b_i, b_{i+3} - b_{i+1}),$$

effectively "shifting" the pairing. Let $\bar{D}$ be the resulting multiset consisting of all such $y_w, y_{w+1}$ and any unpaired $x_i$.

Every element of $\bar{D}$ is at most $15-\Delta$, and together with $15-\Delta$ copies of 1, the multiset $\bar{D}$ is strongly expressive. Moreover,

$$s(\bar{D}) \geq s(\bar{C}) + 2 \cdot \frac{|\bar{C}|-3}{2} \geq 2s - 39 + 2\Delta \geq s + 1 + \Delta \quad (\text{for } n \geq 166).$$

Thus, $\bar{D}$ can absorb all remaining elements of $\bar{B}$, implying that $\bar{A}$ is 1-excellent—a contradiction.

#### 2.3.2 Case 2: $\mathcal{L}(\bar{A}, 3) \geq |\bar{A}|$.

Here, $a_{i+1} - a_i \geq 3$ for all $i \equiv t \pmod 2$. Let $h$ be the number of 1's in $\tilde{A}$. Then

$$n \geq a_t \geq 3s + h + 2(t - s - h) = 2t + s - h,$$

so $h \geq \frac{t}{2} - 6 \geq 11$. Also, $a_{i+1} - a_i \leq 10$, otherwise $a_t > n$.

Consider $B = A \setminus \{a_1\}$. Then $\bar{B}$ contains at least 10 ones and no element exceeds 10. By induction, $\bar{B}$ can represent all $M \leq 10$ with $M \equiv s(\bar{B}) \pmod 2$. Since $a_1 \leq 10$, it follows that $A$ is 2-excellent—a contradiction.

### 2.4. Proof of Theorem 1

If $1 \in A$ or there exist $x, y \in A$ such that $|x - y| \leq 1$, then removing $\{x, y\}$ (or $\{1\}$) leaves a set whose size satisfies

$$|A \setminus \{x, y\}|,\ |A \setminus \{1\}| > c_n - 2 \geq \frac{n-7}{2}.$$

By the induction hypothesis, this remaining set is 2-excellent, and hence $A$ itself is 1-excellent.

It remains to consider the case where all consecutive elements of $A$ are at least distance 2 apart; that is, $a_i - a_{i-1} \geq 2$ for all $i$. In this situation, $1 \notin A$. As observed in Bingyuan Wang's solution, such a configuration is impossible when $n = 8k+1, 8k+2, \ldots, 8k+5$. Thus, we only need to analyze the cases $n = 8k+6, 8k+7, 8k+8, 8k+9$, under the assumption (by induction) that $n \in A$ and $|A| = \lfloor \frac{n}{2} \rfloor$ or $\lfloor \frac{n+1}{2} \rfloor$.

- **Case $n = 8k + 6$:** Then $|A| = 4k + 3$, and since $A \subseteq \{2, 3, \ldots, 8k + 6\}$ with gaps of at least 2, the only possible choice is the even numbers:

  $$A = \{2, 4, 6, \ldots, 8k + 6\},$$

  which is 0-excellent.

- **Case $n = 8k + 7$:** The only admissible sets have the form

  $$A = \{2, 4, \ldots, 2m, 2m + 3, 2m + 5, \ldots, 8k + 7\}$$

  for some $m$.

– If $m \geq 8$, we can perform the following reductions:

$$(2m, 2m-2) \to 2, \quad (2m-2, 2m-4) \to 2, \quad (2, 2m, 2m+3) \to 1.$$

This yields a new set containing 1 along with smaller even numbers and the tail starting from $2m+5$. By the induction hypothesis, this set can generate all integers up to $n$.

– If $1 \leq m \leq 7$, we instead combine large elements:

$$(8k+7, 8k+5) \to 2, \quad (8k+3, 8k+1) \to 2, \quad (8k-1, 8k-3) \to 2,$$

eventually reducing to a set of the form

$$\{3, 2, \ldots, 2, 2m-2, 2m+5, \ldots, 8k-5\},$$

which again, by induction, generates all numbers $\leq n$.

- **Case** $n = 8k+8$**:** Here $|A| = 4k+4$, and the only feasible set with minimal spacing is

$$A = \{2, 4, 6, \ldots, 8k+8\},$$

which is 0-excellent.

- **Case** $n = 8k+9$**:** Then $|A| = 4k+4$, and the only possible structure is

$$A = \{2, 4, \ldots, 2m, 2m+3, \ldots, 8k+9\}.$$

The construction proceeds analogously to the $n = 8k+7$ case, using similar reduction steps to eventually produce a 1-excellent set.

Combining all these cases completes the proof.