一个带符号和问题的最优结果

沙柯岑

2026 年 1 月 15 日

摘要

本文研究了 2021 年全国高中数学联赛（加试）中提出的一个符号赋值问题. 我们的主要成果是给出了一个精确的阈值：构造了一个显式序列 c_n，满足 $c_n = \frac{n}{2} + O(1)$，其具体取值由 $n \bmod 8$ 决定，并证明了对任意正整数 n，只要子集 $A \subseteq [n]$ 的大小满足 $|A| > c_n$，就总存在一种对 A 中元素赋予 ± 1 符号的方式，使其带符号和落在集合 $\{-1, 0, 1\}$ 中；同时，该阈值 c_n 是最优的.

此前已有工作在 $n \geq 564$ 时建立了相同结论. 我们首先将这一范围大幅改进至 $n \geq 38$. 对于较小的 n，通过实际可行的计算即可完成验证，从而确认了 c_n 在所有情形下的最优性. 综上，我们完整解决了该问题，并给出了最优结果.

1 背景

本文所研究的问题源自 2021 年全国高中数学联赛（加试），一项享有盛誉的国家级数学竞赛. 原赛题要求找出最小的正常数 $c > 0$，使得对任意整数 $n \geq 4$ 及任意子集 $A \subseteq \{1, 2, \ldots, n\}$，只要满足 $|A| > cn$，就总存在一个符号函数 $f : A \to \{-1, 1\}$，使得

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

官方解答证明了最优常数为 $c = \frac{2}{3}$. 其证明采用了一个巧妙的组合论证，并辅以一个关于有界整数带符号和的引理. 具体而言，该方法根据 $|A|$ 的奇偶性构造一个导出集合 $\bar{A}$：当 $|A|$ 为偶数时，$\bar{A}$ 由相邻元素的差 $a_{i+1} - a_i$ 构成；当 $|A|$ 为奇数时，则在这些差值的基础上再加入最小元素 a_1. 随后通过归纳法证明：若 $\bar{A}$ 满足 $s(\bar{A}) = \sum_{a \in \bar{A}} a \leq 2|\bar{A}|$，则其带符号和可取到 0 或 1. 而条件 $|A| > \frac{2}{3}n$ 恰好保证了 $\bar{A}$ 满足上述不等式，从而完成证明.

此后，汪秉原对该结果作出了重要改进，对所有充分大的 n 精确确定了阈值序列 c_n（见公式 (2)）. 他证明了当 $n \geq 564$ 时，最优阈值恰由 c_n 给出，其渐近密度为 $\frac{1}{2}$，显著优

于竞赛原解中 $\frac{2}{3}$ 的界限. 然而，他的方法仅在 n 足够大时才有效，无法覆盖较小的 n 值，因此在 $n < 564$ 的范围内仍留有较大空白.

在本文中，我们彻底解决了该问题对所有正整数 n 的情形. 我们采用了一种更为精细且能自适应不同情形的策略：不再依赖单一的普适判据，而是针对候选集合的不同组合结构，设计了多套互补的筛选算法。这一多管齐下的方法，将理论约简与有针对性的计算验证相结合，使我们能够穷尽处理所有剩余的有限情形. 最终，我们证明：对于 $n \geq 18$，最优阈值 c_n 与汪秉原所得公式完全一致；而对于所有更小的 n，我们也给出了精确的数值. 由此，我们完成了对该最优阈值序列的完整、非渐近刻画.

2 主要结果

定义序列 $\{c_n\}_{n=1}^{+\infty}$ 如下：当 $n \geq 18$ 时，

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

当 $n \leq 17$ 时，令 $\{c_i\}_{i=1}^{17} = \{0, 1, 2, 2, 3, 4, 4, 5, 5, 5, 6, 6, 7, 7, 7, 8, 8\}$.

本文的核心结果如下：

定理 1. 对任意正整数 $n \geq 1$ 及任意子集 $A \subseteq \{1, 2, 3, \ldots, n\}$，若 $|A| > c_n$，则存在函数 $f : A \to \{1, -1\}$，使得 (1) 成立. 此外，存在某个满足 $|A| = c_n$ 的子集 A，使得不存在这样的函数 f 使 (1) 成立.

此前已有工作在 $n \geq 564$ 时证明了定理 1. 我们首先给出 $n \geq 38$ 时的严格理论证明，随后通过编程对较小的 n 值进行验证，从而完成全部情形的确认.

定理 2 (理论推导部分)**.** 对任意整数 $n \geq 38$ 及任意子集 $A \subseteq \{1, 2, 3, \ldots, n\}$，若 $|A| > c_n$，则存在函数 $f : A \to \{1, -1\}$，使得 (1) 成立.

为证明定理 2，我们首先处理如下情形：集合 A 中包含一个子集 B，使得 $B = \{1\}$ 或 $B = \{x, x+1\}$. 在此类情形下，只需考虑 $A \setminus B$ 即可. 其余有限的例外构型则可通过人工分析逐一处理. 容易验证当 $n \geq 38$ 时，$c_n - 1 \geq \frac{n-5}{2}$.

定理 3. 对任意整数 $n \geq 38$ 及任意子集 $A \subseteq \{1, 2, 3, \ldots, n\}$，若 $|A| \geq \frac{n-5}{2}$，则存在函数 $f : A \to \{1, -1\}$，使得

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

3 记号与定义

我们将分析两类数据结构: 集合与多重集. 一个**集合**仅包含互异的元素, 而一个**多重集**允许元素重复出现，且每个元素具有明确定义的重数. 集合可视为一种特殊的多重集，其中每个元素的重数均为 1. 若无特别说明，本文中所有大写字母 $A, B, C, \ldots$ 均表示多重集.

对任意 $r \geq 1$，记多重集 $I_r = \{1, 1, \ldots, 1\}$ 为包含 r 个 1 的多重集.

对多重集 A 和 B，我们记 $A - B$ 表示 $A \setminus (A \cap B)$. 采用标准记号 $[n] = \{1, 2, \ldots, n\}$. 对任意多重集 X，我们约定如下:

1. $|X|$ 表示其基数，对多重集而言计入重数.
2. $X^{\uparrow} = \{x_i\}_{i=1}^{|X|}$ 表示将 X 中元素按非降序排列所得的序列.
3. $s(X) = \sum_{x \in X} x$ 表示其元素之和，对多重集计入重数.
4. $\mathcal{L}(X, m) = |X \cap [m, +\infty)|$ 表示 X 在水平 m 处的上层集元素个数.

设集合 $A \subset \mathbb{N}^*$. 若 $|A|$ 为偶数，定义差分多重集

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

若 $|A|$ 为奇数，则定义差分多重集

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

接下来，我们引入若干关键定义，用于本文的定理与证明.

定义 1. 称多重集 A 是 r **强表达的**, 如果对任意整数 $0 \leq x \leq s(A \cup I_r)$ 且满足 $x \equiv s(A \cup I_r) \pmod 2$，均存在函数 $f : A \cup I_r \to \{1, -1\}$，使得

$$\sum_{a \in A \cup I_r} f(a) \cdot a = x. \tag{3}$$

记 S_r 为所有 r 强表达多重集构成的集合.

由定义可知，若 $A \in S_0$，则存在函数 $f : A \to \{1, -1\}$ 使得 (1) 成立，并且不难看出，若 $A \in S_r$ 且 $s \leq r$，则 $A \cup I_s \in S_{r-s}$.

定义 2. 称一个 r 强表达的多重集 A 能**吸收**一个正整数 $b \in \mathbb{N}$，如果 $A \cup \{b\}$ 仍是 r 强表达的.

我们给出一个强表达多重集可以吸收一个元素的充要条件.

命题 1. 设 A 是一个 r 强表达的多重集，$b \in \mathbb{N}$，则 A 能吸收 b 当且仅当 $b \leq s(A)+r+1$.

证明. 令 $X = A \cup I_r$，则 X 可表示所有形如 $s(X), s(X)-2, \ldots, \varepsilon$ 的值，其中 $\varepsilon \in \{0,1\}$. 加入 b 后，通过赋符号 $+b$ 可表示区间 $[b+\varepsilon, b+s(X)]$ 中的同奇偶性整数；通过赋符号 $-b$ 则可表示 $[b-\varepsilon, b-s(X)]$. 若 $b \leq s(X)+1$，则这两个区间在整数轴上覆盖了从 ε 到 $b+s(X)$ 的所有同奇偶性整数，从而保证 $X \cup \{b\}$ 仍能表示从 0 或 1 开始的所有合法目标值. 反之，若 $b \geq s(X)+2$，则 $X \cup \{b\}$ 的任何带符号和都无法取到 0 或 1，故不能保持 r 强表达性. □

定义 3. 设 $A \subset \mathbb{N}^*$ 为一个集合，定义其**剩余能量** $E(A)$ 为：

$$E(A) = s(\bar{A}) + s(\tilde{A}) - |A|. \tag{4}$$

剩余能量反映了集合中相邻元素之间间隙的大小. 相邻元素间隔若超过 1，则超出的部分贡献给 $E(A)$. $E(A)$ 越小，相邻元的差值 $a_{i+1} - a_i$ 越小，说明 A 越"稠密"，因而其差分更可能具有强表达性. 特别地，若 $E(A) = 0$, 则 A 是首项为 1，公差为 1 的等差数列. 下述命题给出了剩余能量的上下界估计.

命题 2. 设 $A \subset [n]$，则 $0 \leq E(A) \leq n - |A|$，且对任意 $m, l \geq 0$，有

$$E(A) \geq \mathcal{L}(\bar{A}, |\bar{A}|+1-m) \cdot (|\bar{A}|-m) + \mathcal{L}(\tilde{A}, |\tilde{A}|+1-l) \cdot (|\tilde{A}|-l). \tag{5}$$

证明. 由于 $s(\bar{A}) \geq |\bar{A}|$、$s(\tilde{A}) \geq |\tilde{A}|$ 且 $|\bar{A}| + |\tilde{A}| = |A|$，故 $E(A) \geq 0$. 设 $A^{\uparrow} = \{a_i\}_{i=1}^{t}$，则

$$n \geq a_t = \sum_{i=0}^{t-1}(a_{i+1}-a_i) = \sum_{\substack{i \\ t-i \text{ 为奇}}} (a_{i+1}-a_i) + \sum_{\substack{i \\ t-i \text{ 为偶}}} (a_{i+1}-a_i) = s(\bar{A}) + s(\tilde{A}),$$

因此 $E(A) \leq n - |A|$. 另一方面，由上层集的定义可得

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

$$s(\tilde{A}) \geq \mathcal{L}(\tilde{A}, |\tilde{A}|+1-l)(|\tilde{A}|+1-l) + |\tilde{A}| - \mathcal{L}(\tilde{A}, |\tilde{A}|+1-l).$$

将两式相加并整理即得所需不等式. □

4 若干实用引理

我们首先给出一个关于强表达多重集的实用判据.

引理 1 (r 强表达性的充分条件)**.** 设 $A \subseteq \mathbb{N}^*$ 且 $r \in \mathbb{N}$. 若对所有 $m \in [0, |A|) \cap \mathbb{N}$，均有

$$\mathcal{L}(A, |A| + r + 1 - m) \leq m, \tag{6}$$

则 A 是 r 强表达的.

证明. 记 $A^{\uparrow} = \{a_1, \ldots, a_s\}$. 若存在某个 i 使得 $a_i \geq i + r + 1$，则 $\mathcal{L}(A, i + r + 1) > s - i$，与条件矛盾. 因此对所有 i 均有 $a_i \leq i + r$. 考虑序列 $\{b_i\}_{i=1}^{r+s}$，其中 $b_1 = \cdots = b_r = 1$，$b_{r+i} = a_i$. 易见对任意 k，有 $b_k \leq k \leq 1 + b_1 + \cdots + b_{k-1}$. 因此，从空集出发，可依次吸收每个 b_i. 由定义可知，A 是 r 强表达的. □

下一条引理表明：若 $A \subseteq [n]$ 且 $|A| \sim 0.4n$，则 $\bar{A}$ 几乎是 1 强表达的.

引理 2. 设 $A \subseteq [n]$ 且 $|A| > \frac{2(n+3)}{5}$. 则对所有 $m \in \{2, 3, \ldots, |\bar{A}| - 2\}$，有

$$\mathcal{L}(\bar{A}, |\bar{A}| + 2 - m) \leq m. \tag{7}$$

证明. 令 $t = |A|$, $s = |\bar{A}|$. 若 $\mathcal{L}(\bar{A}, |\bar{A}| + 2 - m) \geq m + 1$，则由命题 2 的 (5) 可得

$$n - t \geq (m + 1) \cdot (s + 1 - m) \geq 3s - 3,$$

其中最后一个不等式由二次函数在 m 上的单调性得出. 这与 $s \geq t/2$ 且 $t > \frac{2n+6}{5}$ 矛盾. □

5 定理 3 的证明

我们采用反证法. 假设 A 不满足结论，则由 A 中元素（每个恰好使用一次）通过赋符号所生成的任何多重集都不是 1 强表达的. 例如，A 本身及其导出集 $\bar{A}$ 均不满足该性质.

5.1 情形 1：$\mathcal{L}(\bar{A}, 3) < |\bar{A}|$.

令 $t = |A|$，$s = \lceil t/2 \rceil = |\bar{A}|$. 由于当 $n \geq 38$ 时有 $t \geq \frac{n-5}{2} > \frac{2n+6}{5}$，引理 2 表明 $\mathcal{L}(\bar{A}, s + 1) \geq 1$. 因此存在某个下标 $j \equiv t \pmod 2$，使得

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

不失一般性，设 $d = a_j - a_{j-1} = \max \bar{A}$.

断言 1. $\mathcal{L}(\bar{A} - \{d\}, 2s+1) = \mathcal{L}(\tilde{A}, 2s+1) = 0$.

证明. 否则，有 $n - t \geq E(A) \geq 2s + d - 1 \geq 3s$，这与 $t \geq \frac{n-5}{2}$ 及 $n \geq 26$ 矛盾. □

断言 2. 令 $A_1 = [A \cap (0, a_{j-1})] \cup \{0\}$，$A_2 = A \cap (a_j, n]$. 则存在 $x, y \in A_1$ 或 $x, y \in A_2$，使得

$$d - s \leq |x \pm y| \leq d + s.$$

证明. 假设结论不成立. 记 $A_1^{\uparrow} = \{0 < x_1 < \cdots < x_p\}$，$A_2^{\uparrow} = \{y_1 < \cdots < y_q\}$. 则必有 $x_p \leq d - s - 1$；否则 $x_p \geq d + s + 1$. 但由断言 1 可知，对所有 i 有 $x_{i+1} - x_i \leq 2s + 1$，结合介值原理将导致矛盾. 类似地，也有 $y_q - y_1 \leq d - s - 1$.

设在序列 $\{x_i\}$ 和 $\{y_i\}$ 中，共有 w 个相邻差（包括 $x_1 - 0$ ）不等于 1. 则

$$t - 3 + w = p + q + w - 1 \leq x_p + (y_q - y_1) \leq 2(d - s - 1),$$

从而推出 $d \geq t - 0.5 + 0.5w$. 另一方面，由 $E(A) \geq d - 1 + w$ 可得 $d \leq n + 1 - t - w$.

结合 $n + 1 - t - w \geq d \geq t - 0.5 + 0.5w$，可知 $0 \leq w \leq 4$. 再由

$$n - t \geq E(A) \geq d - 1 + w - 1 + \max\{x_{i+1} - x_i, x_1\} - 1, \quad d \geq t - 0.5 + 0.5w,$$

可得 $D \triangleq \max\{x_{i+1} - x_i, x_1\} \leq 8.5 - 1.5w$.

于是，在 $\{y_i\}$ 中至多有 w 个相邻差不为 1. 在 A_2 中选取 $\lceil \frac{q-w}{2} \rceil$ 个互不相交的相邻对 (y_i, y_{i+1})，使得 $y_{i+1} - y_i = 1$. 注意到 $x_i - x_{i-1}, x_1 \leq D$，因此若 $\lceil \frac{q-w}{2} \rceil \geq D - 2$，则基于这些“1-对”，我们可以依次吸收 $x_1, x_2, \ldots, x_p$. 若所吸收元素之和

$$\sum_{i=1}^{p} x_i + \left\lceil \frac{q-w}{2} \right\rceil \geq \frac{p(p+1)}{2} + \frac{q-w}{2}$$

不小于 $d - 2$，则可进一步吸收 $\bar{A}_2 \cup \{d\}$ 中的所有元素，从而 A 满足结论，矛盾. 因此必有

$$p^2 + p + q - w \leq 2d - 6 \quad \text{或} \quad q \leq 2D + w - 6.$$

若 $x_1 + x_2 \geq d + s + 1$，则 $x_2 > \frac{d+1+s}{2}$，进而

$$y_q > a_2 + d + (t - 3) \geq n \quad (\text{当} n \geq 25),$$

矛盾. 故 $x_1+x_2\le d-s-1$ ，同理可得 $x_p+x_{p-1}\le d-s-1$，从而

$$p\le \frac{n+1-s-t-w}{2},\quad p\le \frac{n+3-D-s-t-w}{2}.$$

若 $q\le 2D+w-6$，则 $p\ge t-2D-w+4$，于是

$$n-1-s-t-w\ge 2t-4D-2w+8,$$

且

$$n+3-D-s-t-w\ge 2t-4D-2w+8,$$

当 $n\ge 32$ 且 $w\ge 3$ 或 $D\le 6$ 时，上述不等式不可能成立；否则有 $p\le\sqrt{2(n-1-1.5t)}$.

类似地，$q-1\le y_q-y_1\le d-s-1\le n+2-t-s-w-D$.

综合以上讨论，要么

$$t-2=p+q\le \frac{3}{2}(n+3-s-t-w-D),\quad 1\le w\le 2, D\ge 7,$$

要么

$$t-2=p+q\le \sqrt{2(n-1-1.5t)}+n+1-t-s-w,\quad w\ge 3,$$

二者均与 $n\ge 22$ 的假设矛盾. □

最后，我们说明此情形不可能发生. 由断言 2，可在 A_1 或 A_2 中选取 x,y （其中一个可以是 0），使得 $z=|d-|x\pm y||\le s$. 在所有可行对中，优先选择含减法（即带负号）的情形，并进一步优先选择使 z 最大的那一对.

考虑集合 $B=A-\{a_{j-1},a_j,x,y\}$，则 $|B|\ge t-4$, $|\bar{B}|\ge s-2$, 且由于 x,y 的选取方式, $\bar{B}$ 中不含形如 a_u-a_v （其中 $u\ge j+1$, $v\le j-2$ ）的元素. 若 $\mathcal{L}(\bar{B},|\bar{B}|+2-m)\ge m+1$，则

$$n-t+2\ge E(A-\{x,y\})\ge d+(|\bar{B}|+1-m)(m+1),$$

在 $n\ge 36$ 的假设下，该式仅当 $m=0,1$ 或 $|\bar{B}|-1$ 时可能成立. 记 $B^{\uparrow}=\{c_i\}_{i=1}^{h}, c_0=0$, $\bar{B}^{\uparrow}=\{b_i\}_{i=1}^{l}$. 则必有以下三种情况之一成立：$b_1\ge 3$，或 $b_{l-1}\ge l+1$，或 $b_l\ge l+2$. 且在 $b_1\le 2$ 时 $X=\{b_1,\dots,b_{l-2}\}$ 是 1 强表达的. 分三种情形讨论如下:

1. $b_1\le 2, b_l\ge l+2, b_{l-1}\le l$. 此时 $\{b_1,\dots,b_{l-1}\}$ 是 1 强表达的. 由 x,y 的选取方式，对任意 b_i，要么 $|b_i-d|\ge s+1$，要么 $|b_i-d|\le z$. 此外，若 $|b_l-z|<s$，则 $X\cup\{|b_l-z|\}$ 是 1 强表达的，矛盾. 故 $|b_l-z|\ge s$，而因 $b_l\ge 1$，必有 $b_l\ge z+l+2$，从而 $z\le\Delta$,

其中

$$\Delta=\sum_{i=1}^{l-1}(b_i-1)+(b_l-l-2)+d-(s+1)+s(\tilde{B}-\{a_{j+1}-a_{j-2}\})-(|\tilde{B}|-1),$$

且 $\Delta\le 7$, 因为

$$n-t+2\ge E(A-\{x,y\})\ge l+1+s+1+\Delta. \tag{8}$$

若存在 $1\le i\le l-1$ 使得 $|b_i-d|\le z$，则 $b_i-1\ge s-z$，从而 $\Delta\ge z+s-z>7$ (当 $n\ge 34$ ，矛盾. 因此对所有 $1\le i\le l-1$，均有 $b_i\le d-s-1$. 若 $|b_l-d|\ge s+1$，则 $b_l-l-2\ge d+s-l-1>\Delta$ （当 $n\ge 30$ ），矛盾. 故 $b_l\le d+z$.

由于 $l\ge 7$ （因 $n\ge 38$ ），存在四个连续元素 $c_i,c_{i+1},c_{i+2},c_{i+3}$，使得 $b_u=c_{i+1}-c_i$，$b_v=c_{i+3}-c_{i+2}$，且 $u,v<l$. 若 $c_{i+2}-c_i\ge d-z$，则 $\Delta\ge d-z-2+z>8$，矛盾. 故 $2\le c_{i+2}-c_i\le d-s-1$，即 $d\ge s+3$. 若 $c_{i+3}-c_i\ge d-z$，则 $\Delta\ge d-z-3+z+d-s-1>8$，矛盾. 故 $3\le c_{i+3}-c_i\le d-s-1$，即 $d\ge s+4$.

若存在某个 $k<l$ 使得 $b_k\ge 4$，则 $b_l\ge 3+\sum_{i=1}^{l-1}b_i\ge l+5$，从而 $\Delta\ge 3+3+3>7$，矛盾. 类似地，至多存在一个 $k<l$ 使得 $b_k=3$.

重新考虑 $c_i,\dots,c_{i+3}$: 从空集开始, 首先吸收除 b_u,b_v 外的所有 b_i（可行, 因其余元素 ≤ 3 且至多一个取等号），此时和至少为 $l-3\ge 4$. 接着吸收 $c_{i+2}-c_{i+1}\le 5$，再吸收 $c_{i+3}-c_i$，其与前者的差 ≤ 6，按此顺序可完成吸收. 此时总和至少为 $l-3+4=l+1$. 由于 A 不满足要求，必有 $b_l\ge l+4$，从而

$$\sum_{i=1}^{l-1}(b_i-1)+s(\tilde{B}-\{a_{j+1}-a_{j-2}\})-(|\tilde{B}|-1)\le\Delta-2-3\le 2.$$

由于 $l\ge 7$，除 $c_i\sim c_{i+3}$ 外，还存在另一组 $c_j\sim c_{j+3}$ （$j<i$ 或 $j>i+3$ ），对应 $b_w=c_{j+1}-c_j$，$b_x=c_{j+3}-c_{j+2}$，且 $w,x<l$. 此时可先吸收除 b_u,b_v,b_w,b_x 外的所有 b_k （可行，因最小值 ≤ 2，最大值 ≤ 3 ），和至少为 $l-5\ge 2$. 再依次吸收 $c_{i+2}-c_i$, $c_{i+3}-c_{i+1}$, $c_{j+2}-c_j$, $c_{j+3}-c_{j+1}$ （它们均 ≤ 4 ），总和至少为 $l+3$，故 $b_l-z\ge l+6$.

在此情形下，为使等式 (8) 成立，必须满足 $d=s+4,\Delta=7,z=0$，并且除了 d 与 b_l 之外，其他项均不对集合 $A\setminus\{x,y\}$ 的剩余能量产生贡献.

记 $b_l=c_g-c_{g-1}$，并设可能存在某个整数 $r\ge 1$，使得 $c_{r-1}<a_{j-1}$ 且 $c_r>a_j$. 那么，只要 $k\ne g,r$，就有 $c_k-c_{k-1}=1$. 此外，对于任意 $u,v\in A_1$ 或 A_2，若 $|u-v|<d$,

则必有 $|u-v|\le 3$. 因此，对所有 $1\le i\le h$，区间 $[i,i+3]$ 必须包含 g 或 r 中的至少一个，即 $[i,i+3]\cap\{g,r\}\ne\varnothing$. 然而，当 $n\ge 38$ 时，有 $h\ge 13$. 此时考察区间 $[1,4]$、$[5,8]$ 和 $[9,12]$，它们互不重叠，每个长度为 4 的区间都必须包含 g 或 r 中的某个，但 $\{g,r\}$ 仅有至多两个元素，无法同时覆盖这三个互不相交的区间，从而导致矛盾.

2. $b_1\le 2, b_{l-1}\ge l+1$. 此时若 $b_l-b_{l-1}>l$，则

$$n-t+2\ge E(A-\{x,y\})\ge 3l+1+d,$$

当 $n\ge 31$ 时矛盾. 故 $b_l-b_{l-1}\le l$，从而 $X\cup\{b_l-b_{l-1}\}$ 是 1 强表达的. 若 $z\le b_1+\cdots+b_{l-2}+b_l-b_{l-1}+2$, 则 A 满足要求, 矛盾. 因此 $s\ge z\ge l+1+(b_l-b_{l-1})\ge s-1$.
若 $b_l=b_{l-1}$，则除非 $z=b_l=b_{l-1}=l+1$ 且 $X=I_{l-2}$，或 $b_{l-1}\ge l+2$，否则 $b_l+b_{l-1}-z\le l$ 将导致 $X\cup\{b_l+b_{l-1}-z\}$ 是 1 强表达的，矛盾. 但后者将导致

$$n-t+2\ge E(A-\{x,y\})\ge 2l+2+d,$$

当 $n\ge 38$ 时矛盾.

若 $b_l-b_{l-1}=1$，则必有 $z=s=l+2$，$X=I_{l-2}$，同理可得 $b_l=l+2$，$b_{l-1}=l+1$.
因此必有 $X=I_{l-2}$，且 $\max\{z,b_l,b_{l-1}\}\le l+2$. 在 X 的 b_i 中，必存在两个元素形如 $b_u=c_i-c_{i-1}$，$b_v=c_{i+2}-c_{i+1}$，且 $c_{i+1}-c_i=1$；否则

$$n-t+2\ge E(A-\{x,y\})\ge 2l+d+l-4,$$

当 $n\ge 38$ 时矛盾. 令 $b'_u=c_{i+1}-c_{i-1}$，$b'_v=c_{i+2}-c_i$，并用它们替换 b_u,b_v. 则更新后的 X' 是 1 强表达的，且 $s(X')=l$，故 X' 可吸收 z,b_l,b_{l-1}，矛盾.

3. $b_1\ge 3$. 令

$$\Delta=\sum_{i=1}^{l}(b_i-3)+d-(s+1)+s(\tilde{B}-\{a_{j+1}-a_{j-2}\})-(|\tilde{B}|-1),$$

则

$$n-t+2\ge E(A-\{x,y\})\ge 2l+s+1+\Delta,$$

当 $n\ge 34$ 时，$\Delta\le 2$. 于是 $\tilde{B}-\{a_{j+1}-a_{j-2}\}$ 中至少有 $l-2-\Delta>0$ （当 $n\ge 30$ ）个单位元. 任取其一作为基础集 X_0，则 X_0 可吸收 $\bar{B}$ 中所有值 ≤ 3 且未被“断开”的元素，至少吸收 $l-2-\Delta>1$ （当 $n\ge 34$ ）个元素，此时 $s(X_0)\ge 7$. 其余 $\bar{B}$ 中

元素均 ≤ 5，而被断开的间隙 ≤ 9，故 X_0 可吸收 $\bar{B}$ 全体及该间隙. 于是

$$s(X_0) \ge 1 + 3(l-2) + 7 \ge 3s - 4 \ge \max\{d, z\},$$

从而 A 满足结论，矛盾.

5.2 情形 2：$\mathcal{L}(\bar{A}, 3) \ge |\bar{A}|$.

在此情形下，对所有满足 $i \equiv t \pmod 2$ 的指标 i，均有 $a_{i+1} - a_i \ge 3$. 令

$$\Delta = s(\bar{A}) - 3|\bar{A}| + s(\tilde{A}) - |\tilde{A}| = E(A) - 2|\bar{A}|.$$

则有

$$n - t \ge E(A) \ge \Delta + 2s,$$

从而 $\Delta \le 5$. 这表明 $\tilde{A}$ 中至多有 2 个元素大于 2. 设 $\tilde{A}$ 中恰有 h 个元素大于 2，则存在某个 $i \le 2h+2$，使得 $a_i - a_{i-1} \le 2$. 取满足该条件的最小 i. 于是对所有 $1 \le j \le i-1$，均有 $a_j - a_{j-1} \ge 3$，这些差值在 Δ 中至少贡献了 $2\lfloor\frac{i-1}{2}\rfloor$.

此时，单元素集 $\{a_i - a_{i-1}\}$ 能够吸收 $\bar{A}$ 中那些值为 3 且由区间 $(a_i, a_t]$ 内元素生成的元素. 这类元素至少有 $\lfloor\frac{t-i}{2}\rfloor - x_1$ 个，其中 x_1 表示例外项的个数. 吸收完成后，所得和至少为 $a_i - a_{i-1} + 3\left(\lfloor\frac{t-i}{2}\rfloor - x_1\right)$. 对所有可能的 i 及满足 $x_1 \le 5 - 2\lfloor\frac{i-1}{2}\rfloor$ 的 x_1，当 $n \ge 37$ 时，可验证

$$3\left(\left\lfloor\frac{t-i}{2}\right\rfloor - x_1\right) \ge 4 + \left(5 - 2\left\lfloor\frac{i-1}{2}\right\rfloor - x_1\right).$$

因此，该和还能进一步吸收形如 $a_{i+1} - a_{i-2}, a_{i-3} - a_{i-4}, \ldots$ 或 $a_{i-2} - a_{i-3}, \ldots$ 的剩余差值，因为所有这些差值均不超过 $a_i - a_{i-1} + 11 - 2\left\lfloor\frac{i-1}{2}\right\rfloor - x_1$. 我们的构造表明 A 满足所要求的性质，与假设矛盾.

6 定理 2 的证明

若 $1 \in A$，或存在 $x, y \in A$ 使得 $|x-y| \le 1$，则移除 $\{x, y\}$ （或 $\{1\}$ ）后所得集合的大小满足

$$|A \setminus \{x, y\}|,\ |A \setminus \{1\}| \ge c_n - 1 \ge \frac{n-5}{2}.$$

由定理 3 即可完成证明.

余下的情形是：A 中任意相邻元素之间的距离至少为 2，即对所有 i 有 $a_i - a_{i-1} \ge 2$. 此时 $1 \notin A$，而当 $n = 8k+1, 8k+2, \ldots, 8k+5$ 时，这种构型不可能存在. 此外，我们

已验证当 $n = 25$ 时，所有此类集合（见表 3）均存在带符号和的绝对值小于 2. 因此，对 $n \geq 26$，只需分析 $n = 8k+6, 8k+7, 8k+8, 8k+9$ 的情形，并在归纳假设下设 $n \in A$ 且 $|A| = \lfloor \frac{n}{2} \rfloor$ 或 $\lfloor \frac{n+1}{2} \rfloor$.

- **情形** $n = 8k+6$**:** 此时 $|A| = 4k+3$，且由于 $A \subseteq \{2, 3, \ldots, 8k+6\}$ 且元素间距至少为 2，唯一可能的选择是全部偶数：$A = \{2, 4, 6, \ldots, 8k+6\}$, 其带符号和可取到 0.

- **情形** $n = 8k+7$**:** 此时 $|A| = 4k+3, k \geq 3$，所有可能的集合具有如下形式：

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

  其中 m 为某非负整数.

  - 若 $m \geq 7$，可进行如下化简操作：$(2m-2, 2m-4) \to 2, \quad (2, 2m, 2m+3) \to 1$. 由此得到一个新多重集 $\{1, 2, \cdots, 2m-6, 2m+5, \cdots\}$. 从该多重集最小元开始，可按从小到大的顺序吸收所有不超过多重集中的元素.
  - 若 $2 \leq m \leq 6$，先让 $(2, 2m, 2m+3) \to 1$, 然后合并较大的元素：

    $$(8k+7, 8k+5) \to 2, \quad (8k+1, 8k-1) \to 2, \quad (8k+3, 8k-3) \to 6,$$

    最终化简为形如 $\{1, 2, 2, 6 \ldots, 4, \ldots, 2m-2, 2m+5, \ldots, 8k-5\}$ 的多重集，同样可以从小到大进行吸收.
  - 若 $m = 0$ 或 1, 将大于 7 的相邻奇数两两配对. 最后手动验证即可.

- **情形** $n = 8k+8$ **:** 此时 $|A| = 4k+4$，在最小间距约束下唯一可行的集合为

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

  其带符号和可取到 0.

- **情形** $n = 8k+9$ **:** 此时 $|A| = 4k+4, k \geq 3$，唯一可能的结构为

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

  其构造方式与 $n = 8k+7$ 的情形类似，通过类似的化简步骤，最终可得到一个 0 强表达的集合.

综上所述，所有情形均已覆盖，定理得证.

7 借助计算机完成定理 1 的证明

假设 $|A| > c_n$，当 $18 \le n \le 37$ 时，根据 c_n 的具体形式，我们只需考虑如下情形[1]：

n	19	23	25	27	31	33	35
$\lvert A\rvert$	10	11	12	14	15	16	18
	11	12	13	15	16	17	19

表 1: 待检验的配置（$n, |A|$）

基于命题 2 及前述证明中的讨论，我们采用算法 1 来筛选潜在的反例. 若算法输出为 **false**，则输入集合 A 可能是一个反例. 该算法融合了四种互补的筛选策略：(1) 通过 0-强表达性直接验证（见算法 2，取 $r = 0$）；(2) 在移除关键元素（如 1 或相邻对）后进行松弛验证（见算法 2，取 $r = 1$）；(3) 对具有大间隙的集合采用结构化约简启发式方法（见算法 3）；(4) 贪心吸收与差值构造过程（见算法 4）. 该策略兼顾了较高的计算效率与筛选率.

7.1 算法描述

Algorithm 1 IsLikelyGood

Require: 集合 $A \subset \mathbb{N}$

```
若 IsStronglyExpressive(Ā, 0) 返回 true，则返回 true
A′ ← A，removed ← false
if 1 ∈ A′ then
    从 A′ 中移除 1，removed ← true
else if 存在相邻对 (a_i, a_{i+1}) 满足 a_{i+1} − a_i = 1 then
    从 A′ 中移除这两个元素，removed ← true
end if
if removed then
    若 IsStronglyExpressive(Ā′, 1) 返回 true，则返回 true
end if
若 SecondaryScreen(A) 返回 true，则返回 true
若 GreedyScreen(A) 返回 true，则返回 true
return false
```

[1]对其余 n, 有 $c_n = c_{n+1}$, 故不妨设 $A \subset [n+1]$ 或更宽松，又注意无需额外检验 $|A| \ge c_n + 3$, 因为此时可以去掉 A 中的相邻对化归到 $|A|$ 更小的情形.

Algorithm 2 IsStronglyExpressive (X_1, r)

Require: 多重集 X_1，参数 $r \in \mathbb{N}$

$s \leftarrow 0$
for 每个 $x \in X_1^{\uparrow}$ **do**
  if $x > s + 1 + r$ **then**
    return false
  end if
  $s \leftarrow s + x$
end for
return true

Algorithm 3 SecondaryScreen (X_2)

Require: 已排序集合 $X_2^{\uparrow} = [a_1, a_2, \ldots, a_t]$，其中 $t \geq 4$，并设 $a_0 = 0$

$d_{\max} \leftarrow \max(\bar{X}_2)$
定位索引 j ，使得 X_2 中对应元素在 $\bar{X}_2$ 中生成 $d_{\max}$
if 不存在这样的 j **then**
  return false
end if
将 $X_2 \backslash \{a_j, a_{j-1}\}$ 划分为 $A_1 = X_2 \cap (0, a_{j-1})$ 和 $A_2 = X_2 \cap (a_j, n]$
初始化候选列表 $\mathcal{C} \leftarrow \emptyset$
for 所有满足 $x > y$ 的对 $(x, y) \in (A_1 \times A_1) \cup (A_2 \times A_2)$ **do**
  将候选 $(z = x + y,\ x,\ y,\ \delta = |d_{\max} - z|)$ 加入 $\mathcal{C}$
  将候选 $(z = x - y,\ x,\ y,\ \delta = |d_{\max} - z|)$ 加入 $\mathcal{C}$
end for
if $\mathcal{C} = \emptyset$ **then**
  return false
end if
按 δ 升序对 $\mathcal{C}$ 排序
for $\mathcal{C}$ 中前 $K = \min(5, |\mathcal{C}|)$ 个候选 **do**
  设当前候选为 (z, x, y, δ)
  构造 $Y \leftarrow X_2 \setminus \{a_{j-1}, a_j, x, y\}$
  若 $\delta > 0$，将 δ 添加至 Y
  若 IsStronglyExpressive$(Y, 0)$ 返回 **true**，则返回 **true**
end for
return false

Algorithm 4 GREEDYSCREEN (X_3)

Require: 集合 $X_3 \subset \mathbb{N}$

初始化集合 $\mathcal{R} \leftarrow X_3$ 和 $s \leftarrow 0$
while $\mathcal{R} \neq \emptyset$ **do**
  寻找 $\mathcal{R}$ 中最大的 a 使得 $a \le s+1$
  if 存在这样的 a **then**
    $s \leftarrow s+a$，$\mathcal{R} \leftarrow \mathcal{R} \setminus \{a\}$
    continue
  end if
  寻找对 $(x,y) \in \mathcal{R} \times \mathcal{R}$ （$x > y$ ），使得 $d = x - y \le s+1$ 且 d 最大
  if 不存在这样的对 **then**
    return false
  end if
  $s \leftarrow s+d$，$\mathcal{R} \leftarrow \mathcal{R} \setminus \{x,y\}$
end while
return true

7.2 筛选后剩余的潜在反例

对表 1 中所有可能配置执行算法筛选后，最终剩余的潜在反例极少．所有结果如表 2 至表 5 所示．我们对每种情形均进行了人工验证．

如上表所示，当 $n \ge 19$ 时，这些例子均不含长度为 1 的间隙．我们已通过手工验证，确认所有这些例子均非真正的反例[2]．当 $n \le 18$ 时，通过算法筛选亦可确认不存在反例．

最后，我们给出当 $|A| = c_n$ 时的显式反例．对每个 n，以下集合 A_n 满足 $|A_n| = c_n$ 且均为反例：

$$A_1 = \varnothing,\ A_2 = \{2\},\ A_3 = A_4 = \{1,3\},\ A_5 = \{1,2,5\};$$

$$A_6 = A_7 = \{1,4,5,6\},\ A_8 = A_9 = A_{10} = \{1,2,6,7,8\},\ A_{11} = A_{12} = \{1,7,8,9,10,11\};$$

$$A_{13} = A_{14} = A_{15} = \{1,2,9,10,11,12,13\};$$

$$A_{16} = A_{17} = \{1,2,3,12,13,14,15,16\};$$

$$A_n = \{2,4,\cdots,2m\},\ m \text{ 是满足} 2m \le n \text{ 且} 4 \nmid m(m+1) \text{ 的最大整数}, \quad n \ge 18.$$

至此，我们最终确立了主要结论，并为这一问题画上了圆满的句号．

[2]根据定理 2 对无 1 间隙情况的证明，只需对表 2, 3 验证

2	4	6	8	10	12	14	16	18	20	22
2	4	6	8	10	12	14	16	18	21	23
2	4	6	8	10	12	14	17	19	21	23
2	4	6	8	10	13	15	17	19	21	23
2	4	6	9	11	13	15	17	19	21	23
2	5	7	9	11	13	15	17	19	21	23

表 2: $n = 19, 23$ 时的潜在反例.

2	4	6	8	10	12	14	16	18	20	22	24
2	4	6	8	10	12	14	16	18	20	23	25
2	4	6	8	10	12	14	16	19	21	23	25
2	4	6	8	10	12	15	17	19	21	23	25
2	4	6	8	11	13	15	17	19	21	23	25
2	4	7	9	11	13	15	17	19	21	23	25
2	5	7	9	11	13	15	17	19	21	23	25

表 3: $n = 25, 27$ 时的潜在反例.

2	4	6	8	10	12	14	16	18	20	22	24	26	28	30
2	4	6	8	10	12	14	16	18	20	22	24	26	29	31
2	4	6	8	10	12	14	16	18	20	22	25	27	29	31
2	4	6	8	10	12	14	16	18	21	23	25	27	29	31
2	4	6	8	10	12	14	17	19	21	23	25	27	29	31
2	4	6	8	10	13	15	17	19	21	23	25	27	29	31
2	4	6	9	11	13	15	17	19	21	23	25	27	29	31
2	5	7	9	11	13	15	17	19	21	23	25	27	29	31

表 4: $n = 31$ 时的潜在反例.

2	4	6	8	10	12	14	16	18	20	22	24	26	28	30	32
2	4	6	8	10	12	14	16	18	20	22	24	26	28	31	33
2	4	6	8	10	12	14	16	18	20	22	24	27	29	31	33
2	4	6	8	10	12	14	16	18	20	23	25	27	29	31	33
2	4	6	8	10	12	14	16	19	21	23	25	27	29	31	33
2	4	6	8	10	12	15	17	19	21	23	25	27	29	31	33
2	4	6	8	11	13	15	17	19	21	23	25	27	29	31	33
2	4	7	9	11	13	15	17	19	21	23	25	27	29	31	33
2	5	7	9	11	13	15	17	19	21	23	25	27	29	31	33

表 5: $n = 33, 35$ 时的潜在反例.

8 定理 3 证明逻辑的深度解析：能量约束下的结构收敛

定理 3 的证明是全文的技术核心。其本质在于建立了一套基于 **"总量预算—局部结构—动态吸收"** 的博弈模型，论证了在集合 A 的基数 $|A|$ 超过特定阈值时，任何破坏符号和平衡的尝试都将导致逻辑自洽性的崩溃。

8.1 全局约束：剩余能量 $E(A)$ 的资源属性

证明的基点在于对"剩余能量" $E(A)$ 的资源化理解。由定义 $E(A) = s(\overline{A}) + s(\tilde{A}) - |A|$ 及命题 2 可知：

$$E(A) \le n - |A|$$

在定理 3 的条件下，$n - |A|$ 构成了一个极其狭窄的 **"能量预算"**。由于剩余能量本质上累加了集合元素间隙偏离"连续整数"的程度，这一不等式强制要求集合 A 在整体分布上必须保持高度的稠密性。

8.2 局部表征：差分多重集与上层集的结构探测

为了判定一个集合是否能通过符号赋值达到平衡，我们通过**差分多重集** $\overline{A}$ 刻画了集合内部的微观步进。而**上层集** $\mathcal{L}(\overline{A}, m)$ 则充当了探测器，用以衡量结构中"大跨度间隙"的密度。

- **引理 1（强表达性判据）**建立了一个结构标准：若一个多重集的"大元素"分布满足特定比例，则它具有足够的灵活性（强表达性）来构造任何所需的数值。
- **反证逻辑**：若假设集合 A 不满足命题要求，则根据引理 1，其差分多重集的上层集必须显著超过阈值，即结构中必须包含大量互不兼容的高跨度间隙。

8.3 核心机制：能量预算与结构代价的对消

证明的精彩之处在于利用**命题 2** 构建了从"结构"到"能量"的单向传导机制。

- **代价函数**：若对手试图通过构造"大间隙"来破坏强表达性，每一处大间隙都会在 $\mathcal{L}(\overline{A}, m)$ 中留下记录，并根据命题 2 的下界估算公式，成倍地消耗 $E(A)$。
- **逻辑坍缩**：当 $|A| \ge \frac{n-5}{2}$ 时，总能量预算被锁定在低位。计算表明，即使是最小限度的"结构破坏"，其所需的能量代价也会迅速穿透预算上限。这种矛盾迫使所有可能的反例必须向几种极度受限的稀疏情形收敛。

8.4 动态协同：吸收算法与离散介值原理

在具体的构造环节，证明采用了 **“吸收策略”**：

1. **基础池构建**：利用差分集合中的小元素（1 或 2）构建一个能够覆盖连续区间的“基础表达池”。

2. **动态扩充**：只要新引入的较大元素 b 满足 $b \leq$ 表达池元素和 $+1$，该元素即可被“吸收”，从而以指数级速度扩大表达范围。

3. **断裂弥合**：对于吸收算法无法直接覆盖的巨大断层，作者退回到原集合 A 中，利用离散介值原理证明，总能找到特定的元素组合来“桥接”这些断层，除非集合 A 本身违背了能量守恒。

8.5 结构微扰与矛盾导出：极端刚性情形的兜底逻辑

对于极少数差分多重集结构极其“刚性”、导致常规吸收算法失效的极端情形，证明并未止步于分类讨论，而是引入了**结构微扰（Perturbation）**的策略。当集合 A 的差分序列呈现出某种高度规整但非强表达的特征时，作者通过对差分项进行细微的局部调整（例如考察差分序列中相邻项的交换，或利用原集合中特定元素的微小偏移）来诱导逻辑冲突：

- **矛盾诱导**：证明指出，若要维持“非强表达”的状态，差分结构必须满足极度苛刻的排他性条件；

- **逻辑坍缩**：然而，这种排他性要求与前述的低能量预算（$E(A) \leq n - |A|$）是直接冲突的。通过对差分序列局部性质的细致推演，我们证明了即使是微小的结构扰动也会迫使集合跨越阈值，从而反向证明了在给定基数下，这种“顽固”的非平衡结构在数学上是不存在的。

这一技巧为定理 3 提供了最后的逻辑闭环，确保了证明能够覆盖从“随机分布”到“极度规整”的所有结构谱系。

8.6 结论

定理 3 并非简单的分类讨论，而是一场 **“结构与能量的极限博弈”**。它证明了：在稠密性受限的系统中，局部的无序（不平衡）由于需要消耗过高的全局能量（间隙空间），在逻辑上是不可能持续存在的。