# 关于纳桑森的三角数现象

Kevin O'Bryant*

2025 年 6 月 28 日

**摘要**

对于有限集合 $A \subseteq \mathbb{Z}$，$h$ 倍和集是 $hA \coloneqq \{x_1 + \cdots + x_h : x_i \in A\}$。我们用格的连续 $L^1$-最小值来解释和集大小序列的开始部分 $(|hA|)_{h=1}^{\infty}$。因此，我们解释了（但没有证明）Nathanson 最近实验中出现三角数的现象。

## 1 介绍

给定一个有限整数集 $A$，根据 Nathanson 的经典定理，和集大小的序列最终是线性的 [3]。在这项工作中，我们确定了该序列是如何开始的。

$A$ 集合包含 $k$ 个整数，其 $h$ 倍和集表示为 $hA$，通过基础计数满足

$$|hA| \le h \cdot \operatorname{diam}(A) + 1, \qquad \text{and} \qquad |hA| \le \binom{h+k-1}{k-1}.$$

。对于较小的 $h$，二项式系数是正确的；而根据 Nathanson 的定理，当 $h$ 足够大时，线性界起决定作用。

例如，考虑 $A = \{0, 2, 18, 25\}$。如 $|hA| \le \binom{h+3}{3}$ 所示，数量 $\binom{h+3}{3} - |hA|$ 计算了“缺失”的和。

| $h$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $\|hA\|$ | 4 | 10 | 20 | 34 | 52 | 74 | 100 | 130 | 162 | 193 | 222 | 249 |
| $\binom{h+3}{3} - \|hA\|$ | 0 | 0 | 0 | 1 | 4 | 10 | 20 | 35 | 58 | 93 | 142 | 206 |
| first differences | | 0 | 0 | 1 | 3 | 6 | 10 | 15 | 23 | 35 | 49 | 64 |

三角数（A000217）出现在最底行令人惊讶，并且本文的主要定理解释了它的出现。在下面的第 5.1 节中，我们重复了 Nathanson[8] 的相关实验，在该实验中三角数出现了。事实上，这项工作直接受到了那项工作的启发。

*On faculty at the City University of New York, both at the College of Staten Island campus and The Graduate Center campus.

### 1.1 符号和约定

我们使用 $\mathbb{N}$ 表示正整数，$\mathbb{N}_0$ 表示集合 $\{0\} \cup \mathbb{N}$。对于每个有限集 $A \subseteq \mathbb{R}$，我们将它的元素表示为 $a_1 < \cdots < a_k$，使得 $|A| = k$ 和 $\operatorname{diam}(A) = a_k - a_1$。我们约定，如果二项式系数的上标小于下标，则该系数为零；并且约定，如果 $X$ 是一个集合且 $k$ 是一个非负整数，那么 $\binom{X}{k}$ 是 $X$ 的所有子集构成的集合，这些子集的基数为 $k$。

集合 $hA$ 是集合 $\{b_1 + \cdots + b_h : b_i \in A\}$。递归地，$1A \coloneqq A$ 和 $hA \coloneqq \{s + b : s \in (h-1)A, b \in A\}$。$hA$ 的大小在 $A$ 的平移和缩放下是不变的，这使我们能够不失去一般性地假设 $\min A = 0$ 并且 $\gcd(A) = 1$。

令

$$\mathcal{X}_{t,k} \coloneqq \{\langle c_1, \ldots, c_k \rangle : c_i \in \mathbb{N}_0, c_1 + \cdots + c_k = t\}.$$

由 $\overline{\boldsymbol{a}} = \langle a_1, a_2, \ldots, a_k \rangle$，我们有

$$hA = \{\overline{\boldsymbol{x}} \cdot \overline{\boldsymbol{a}} : \overline{\boldsymbol{x}} \in \mathcal{X}_{h,k}\}.$$

根据插板法，

$$|\mathcal{X}_{t,k}| = \binom{t+k-1}{k-1}$$

对于每个 $A$，我们令 $\mathcal{L}$ 是在 $\mathbb{Z}^k$ 中与 $\overline{\mathbf{1}}$（全 1 向量）和 $\overline{\boldsymbol{a}}$ 都垂直的向量构成的 $(k-2)$ 维格子。对于 $k \geq 4$，这个格子的维度至少为 2。我们定义 $2h_1, 2h_2$ 为格关于 $L^1$-范数的第一和第二最小值。即，向量空间 $\mathcal{L}$ 中的每个非零向量的 $L^1$-范数至少为 $2h_1$，存在一个 $\overline{\boldsymbol{y}}_1 \in \mathcal{L}$ 使得 $\|y_1\|_1 = 2h_1$ 成立，并且 $\mathcal{L}$ 中与 $\overline{\boldsymbol{y}}$ 线性无关的每个向量的 $L^1$-范数至少为 $2h_2$。

### 1.2 定理和推论

我们的主定理将格 $\mathcal{L}$ 的最小值与和集大小的序列联系起来。

**定理 1.** 设 $A = \{a_1, \ldots, a_k\}$ 是一组 $k \geq 4$ 整数。设 $\mathcal{L}$ 是垂直于 $\overline{\mathbf{1}}$ 和 $\overline{\boldsymbol{a}} = \langle a_1, \ldots, a_k \rangle$ 的 $\mathbb{Z}^k$ 子格。令 $2h_1, 2h_2$ 为 $\mathcal{L}$ 关于 $L^1$-范数的第一和第二个最小值。如果 $1 \leq h < h_2$，则

$$|hA| = \binom{h+k-1}{k-1} - \binom{h-h_1+k-1}{k-1}.$$

如果 $h < h_1$，则 $\binom{h-h_1+k-1}{k-1} = 0$，因此 $|hA| = \binom{h+k-1}{k-1}$。也就是说，$A$ 是一个 $B_h$-集。进一步，对于 $h > h_1$

$$\begin{aligned}\left(\binom{h+k-1}{k-1} - |hA|\right) &- \left(\binom{(h-1)+k-1}{k-1} - |(h-1)A|\right) \\ &= \binom{h-h_1+k-1}{k-1} - \binom{(h-1)-h_1+k-1}{k-1} = \binom{h-h_1+k-2}{k-2}.\end{aligned}$$

如果 $k=4$，带有 $t=h-h_1+2$，这是 $\binom{t}{2}$，三角数。我们评论说，当使用 $k=4$ 时，$\binom{h+k-1}{k-1}$ 和 $\binom{h-h_1+k-1}{k-1}$ 都是四面体数 A000292。

回到上面的示例，$A=\{0,2,18,25\}$，检查格子 $\mathcal{L}$ 我们发现 $h_1=4$ 和 $h_2=9$。根据我们的定理，$|hA|=\binom{h+3}{3}$ 对于 $1\le h<4$，以及 $|hA|=\binom{h+3}{3}-\binom{h-1}{3}$ 对于 $4\le h<9$。

尽管风格不同，以下结果的证明在概念和风格上与我们对定理 1 的证明有一些重叠。

**定理 2.** 我们有

$$\begin{aligned}\{|hA|:|A|=k,A\subseteq\mathbb{N}\}&=\{|hA|:|A|=k,A\subseteq\mathbb{R}\}\\&=\left\{|A^h|:|A|=k,A\subseteq\mathbb{N}\right\}=\left\{|A^h|:|A|=k,A\subseteq\mathbb{R}\right\}.\end{aligned}$$

## 2 定理 1 的证明

假设我们有不相同的 $\overline{\boldsymbol{c_1}},\overline{\boldsymbol{c_2}}\in\mathcal{X}_{h,k}$ 和 $\overline{\boldsymbol{c_1}}\cdot\overline{\boldsymbol{a}}=\overline{\boldsymbol{c_2}}\cdot\overline{\boldsymbol{a}}$。那么 $\overline{\boldsymbol{c_1}}-\overline{\boldsymbol{c_2}}\in\mathcal{L}$，并且

$$2h_1=\|\overline{\boldsymbol{y}}\|_1\le\|\overline{\boldsymbol{c_1}}-\overline{\boldsymbol{c_2}}\|_1\le\|\overline{\boldsymbol{c_1}}\|_1+\|\overline{\boldsymbol{c_2}}\|_1=2h. \tag{1}$$

如果 $h<h_1$，这是不可能的，因此 $|hA|=|\mathcal{X}_{h,k}|=\binom{h+k-1}{k-1}$，正如所声称的。

我们现在假设 $h_1\le h<h_2$，并取 $y\in\mathcal{L}$ 与 $\|y\|_1=2h_1$。

$h$ 类型的 $A$ 是由 $\overline{\boldsymbol{c_1}}\equiv\overline{\boldsymbol{c_2}}$ 定义的在 $\mathcal{X}_{h,k}$ 上的等价关系，如果 $\overline{\boldsymbol{c_1}}\cdot\overline{\boldsymbol{a}}=\overline{\boldsymbol{c_2}}\cdot\overline{\boldsymbol{a}}$。$h$ 类型的 $A$ 的类数是 $|hA|$。

我们定义了在 $\mathcal{X}_{h,k}$ 上的第二个等价关系，如果 $\overline{\boldsymbol{c_1}}-\overline{\boldsymbol{c_2}}$ 是 $\overline{\boldsymbol{y}}$ 的倍数，则取 $\overline{\boldsymbol{c_1}}\cong\overline{\boldsymbol{c_2}}$。我们将首先证明这两个等价关系 $\cong,\equiv$ 是相同的，然后证明该 $\cong$ 等价关系有 $\binom{h+k-1}{k-1}-\binom{h-h_1+k-1}{k-1}$ 个类。

首先，假设 $\overline{\boldsymbol{c_1}}\equiv\overline{\boldsymbol{c_2}}$ 和 $\overline{\boldsymbol{c_1}}\ne\overline{\boldsymbol{c_2}}$。然后 $(\overline{\boldsymbol{c_1}}-\overline{\boldsymbol{c_2}})\cdot\overline{\boldsymbol{a}}=0$，并且 $(\overline{\boldsymbol{c_1}}-\overline{\boldsymbol{c_2}})\cdot\overline{\boldsymbol{1}}=\overline{\boldsymbol{c_1}}\cdot\overline{\boldsymbol{1}}-\overline{\boldsymbol{c_2}}\cdot\overline{\boldsymbol{1}}=h-h=0$，从而 $\overline{\boldsymbol{c_1}}-\overline{\boldsymbol{c_2}}\in\mathcal{L}$。但由于 $\|\overline{\boldsymbol{c_1}}-\overline{\boldsymbol{c_2}}\|=2h$，以及 $2h<2h_2$，因此必须存在某个 $\alpha\in\mathbb{Z}$ 使得 $\overline{\boldsymbol{c_1}}-\overline{\boldsymbol{c_2}}=\alpha\overline{\boldsymbol{y}}$。因而，$\overline{\boldsymbol{c_1}}\cong\overline{\boldsymbol{c_2}}$。

现在，假设 $\overline{\boldsymbol{c_1}}\cong\overline{\boldsymbol{c_2}}$ 和 $\overline{\boldsymbol{c_1}}\ne\overline{\boldsymbol{c_2}}$。那么 $\overline{\boldsymbol{c_1}}-\overline{\boldsymbol{c_2}}=\alpha\overline{\boldsymbol{y}}$。因此，$(\overline{\boldsymbol{c_1}}-\overline{\boldsymbol{c_2}})\cdot\overline{\boldsymbol{a}}=\alpha\overline{\boldsymbol{y}}\cdot\overline{\boldsymbol{a}}=0$，因为 $\overline{\boldsymbol{y}}\in\mathcal{L}$。从而，$\overline{\boldsymbol{c_1}}\equiv\overline{\boldsymbol{c_2}}$。

只剩下计算 $\cong$-类。每个类都具有形式 $\overline{\boldsymbol{c}},\overline{\boldsymbol{c}}+\overline{\boldsymbol{y}},\overline{\boldsymbol{c}}+2\overline{\boldsymbol{y}},\dots,\overline{\boldsymbol{c}}+\alpha\overline{\boldsymbol{y}}$，其中 $y$ 是某个非负整数。由 [1] 可知类的数量是

$$|\mathcal{X}_{h,k}|-\#\left\{(\overline{\boldsymbol{c_1}},\overline{\boldsymbol{c_2}})\in\mathcal{X}_{h,k}^2:\overline{\boldsymbol{c_2}}-\overline{\boldsymbol{c_1}}=\overline{\boldsymbol{y}}\right\}.$$

令 $\overline{\boldsymbol{y}}=\langle y_1,\dots,y_k\rangle$，并定义

$$\overline{\boldsymbol{u}}=\langle u_1,\dots,u_k\rangle,u_i:=\begin{cases}0,&y_i\le 0;\\y_i,&y_i>0;\end{cases}\qquad\text{and}\qquad\overline{\boldsymbol{v}}=\langle v_1,\dots,v_k\rangle,v_i:=\begin{cases}-y_i,&y_i\le 0;\\0,&y_i>0.\end{cases}$$

[1] 这与对于森林 $G=(V,E)$，其连通分量的数量是 $|V|-|E|$ 的定理相同。

我们有 $\|\overline{\boldsymbol{u}}\|_1 = \|\overline{\boldsymbol{v}}\|_1 = h_1$，因为 $\overline{\boldsymbol{y}} \cdot \overline{\mathbf{1}} = 0$，以及 $\overline{\boldsymbol{y}} = \overline{\boldsymbol{u}} - \overline{\boldsymbol{v}}$。每个 $\overline{\boldsymbol{c}_2} - \overline{\boldsymbol{c}_1} = \overline{\boldsymbol{y}}$ 的解都对应于 $\overline{\boldsymbol{c}_2} - \overline{\boldsymbol{u}} = \overline{\boldsymbol{c}_1} - \overline{\boldsymbol{v}}$，反之亦然。现在，$\overline{\boldsymbol{r}} \coloneqq \overline{\boldsymbol{c}_2} - \overline{\boldsymbol{u}}$ 具有非负整数坐标之和为 $h - h_1$，因此 $\overline{\boldsymbol{r}} \in \mathcal{X}_{h-h_1,k}$。也就是说，$\overline{\boldsymbol{c}_2} - \overline{\boldsymbol{c}_1} = \overline{\boldsymbol{y}}$ 的解与 $\mathcal{X}_{h-h_1,k}$ 的元素之间是一一对应的。我们现在有

$$|hA| = \big(\text{number of} \equiv \text{-classes}\big) = \big(\text{number of} \cong \text{-classes}\big) = |\mathcal{X}_{h,k}| - |\mathcal{X}_{h-h_1,k}|,$$

如所声称的。

# 3 表格类型

## 3.1 一些“加性”组合数论的符号

设 $G$ 是一个群（例如，$\mathbb{R}$），设 $\varphi : G^h \to G^i$ 是任意函数（例如，$\sigma_h(\overline{\boldsymbol{x}}) = x_1 + \cdots + x_h$）。令 $X \subseteq G$（例如，$X = \mathbb{N}$，正整数）。如果 $i = 1$，我们定义

$$\begin{aligned} \varphi X &\coloneqq \{\varphi(\overline{\boldsymbol{g}}) : \overline{\boldsymbol{g}} \in X^h\} \\ \mathcal{R}_X(\varphi, k) &\coloneqq \{|\varphi A| : |A| = k, A \subseteq X\} \\ m_X(\varphi, k) &\coloneqq \min\{|A| : A \subseteq X, |A| = k\} \\ M_X(\varphi, k) &\coloneqq \max\{|A| : A \subseteq X, |A| = k\} \end{aligned}$$

如果 $A \subseteq X$ 具有 $|A| = k$ 和 $|\varphi A| = M_X(\varphi, k)$，那么我们称 $A$ 为 $\varphi$-西顿集。如果 $A \subseteq X$ 具有 $|A| = k$ 和 $|\varphi A| = m_X(\varphi, k)$，那么我们称 $A$ 为一个 $\varphi$ 进展。

当 $X$ 从上下文中很明显时，它将被省略。当 $\varphi(\overline{\boldsymbol{a}}) = a_1 + \cdots + a_h$ 时，我们在符号中将其简写为 $h$：集合 $hA$ 是 $h$ 倍的和集 $A$，并且 $\mathcal{R}_{\mathbb{N}}(h, k)$ 是集合 $k$ 正整数的 $h$ 倍和集大小的集合。

| $\varphi(\overline{\boldsymbol{a}})$ | $X$ | $m_X(\varphi, k)$ | $M_X(\varphi, k)$ | $\mathcal{R}_X(\varphi, k)$ | $\varphi$-Sidon | $\varphi$-progression |
|---|---|---|---|---|---|---|
| $a_1 + a_2$ | $\mathbb{Z}$ | $2k - 1$ | $\binom{k+1}{2}$ | $[m, M]$ | Sidon set | arithmetic prog. |
| $a_1 - a_2$ | $\mathbb{Z}$ | $2k - 1$ | $k^2 - k + 1$ | unknown | Sidon set | arithmetic prog. |
| $a_1 a_2$ | $\mathbb{N}$ | $2k - 1$ | $\binom{k+1}{2}$ | $[m, M]$ | multiplicative Sidon set | geometric prog. |
| $a_1 + a_2 + a_3$ | $\mathbb{Z}$ | $3k - 2$ | $\binom{k+2}{3}$ | unknown | $B_3$-set | arithmetic prog. |
| $a_1 + \cdots + a_h$ | $\mathbb{Z}$ | $hk - h + 1$ | $\binom{h+k-1}{k-1}$ | unknown | $B_h$-set | arithmetic prog. |
| $a_1 + 2a_2$ | $\mathbb{Z}$ | $3k - 2$ | $k^2$ | unknown | unnamed | ??? |

有一些有趣的关于 $i \geq 2$ 的问题。例如，与 $\varphi(x_1, x_2) = (x_1 + x_2, x_1 - x_2)$ 相关的问题涉及集合和差集的相对大小，比如 [2]。[9] 的工作全面讨论了 $\varphi(x_1, x_2) = (u_1x_1 + u_2x_2, v_1x_1 + v_2x_2)$。函数 $\varphi(x_1, x_2) = (x_1 + x_2, x_1x_2)$ 导致了著名的加乘猜想 [1, 10]。

### 3.2 表格类型

如果 $G$ 是一个可排序群并且 $\varphi : G^h \to G$，我们可以将 $A$ 的元素列举为 $\{a_1 < \cdots < a_k\}$，并定义 $\varphi$ 表作为函数 $T$ 从 $[k]^h \to G$

$$T(\overline{\boldsymbol{x}}) = \varphi(a_{x_1}, a_{x_2}, \ldots, a_{x_h}).$$

例如，对于 $\varphi(\overline{\boldsymbol{x}}) = x_1 + x_2$，该表为 $T(i,j) = a_i + a_j$。

我们定义类型的（弱）$\varphi$ 表为由 $\varphi$ 引起的在 $[k]^h$ 上的等价关系，即当且仅当 $T(\overline{\boldsymbol{x}}) = T(\overline{\boldsymbol{y}})$ 时有 $\overline{\boldsymbol{x}} \overset{\varphi,A}{\equiv} \overline{\boldsymbol{y}}$。

我们将 $\mathcal{T}_X(\varphi, k)$ 设定为 $A \subseteq X$ 的所有类型 $\varphi$ 表的集合，带有 $|A| = k$。

我们注意到 $|\varphi A|$ 是 $\varphi$ 表类型等价类的数量，因此如果 $A$ 和 $B$ 有相同的 $\varphi$ 表类型，则 $|\varphi A| = |\varphi B|$。

我们评论说，这里定义的"类型"比在 [10] 中密切相关的"类型"要弱一些，并且列举于 A378609 中。在那里，"类型" 包括了等价类的排序。本工作中的版本理论上更容易处理，而排序中额外的信息在计算上是有用的。

**问题 3.** *有多少种加法表？也就是说，什么是 $|\mathcal{T}_{\mathbb{Z}}(h,k)|$？*

# 4 类型的应用

Nathanson [4] 提出了一个问题：$\mathcal{R}_{\mathbb{N}}(h,k) = \mathcal{R}_{\mathbb{R}}(h,k)$ 是否成立。我们给出了更强的等式类型。

**定理 4.** *$\mathcal{T}_{\mathbb{N}}(h,k) = \mathcal{T}_{\mathbb{R}}(h,k)$。特别是，$\mathcal{R}_{\mathbb{N}}(h,k) = \mathcal{R}_{\mathbb{R}}(h,k)$*

**定义 5.** 对于 $\varphi : \mathbb{R}^h \to \mathbb{R}$，集合 $A$ 的分离是

$$\operatorname{sep}_\varphi(X) \coloneqq \min\left\{\left|\phi(\overline{\boldsymbol{a}}) - \phi(\overline{\boldsymbol{b}})\right| : \overline{\boldsymbol{a}}, \overline{\boldsymbol{b}} \in X^h, \varphi(\overline{\boldsymbol{a}}) \neq \varphi(\overline{\boldsymbol{b}})\right\}.$$

如果 $\varphi(\overline{\boldsymbol{x}}) = \sigma_h(\overline{\boldsymbol{x}}) = x_1 + \cdots + x_h$，我们在符号中用 "$h$" 代替 "$\varphi$"，并注意 $\operatorname{sep}_1(X) \geq \operatorname{sep}_2(X) \geq \cdots$。

*定理 4 的证明。.* 对于 $h = 1$ 这是显然的，所以我们取 $h \geq 2$。设 $X \subseteq \mathbb{R}$ 与 $|X| = k$。我们构建一个集合 $A \subseteq \mathbb{N}_0$，其中 $|A| = k$ 和 $A, X$ 具有相同的类型 $h$-倍加法表。通过仿射不变性，我们可以假设 $X = \{0 = x_1 < \cdots < x_k = 1\}$。设 $q_0 \in \mathbb{N}$ 足够大以至于 $\operatorname{sep}_h(X) \cdot q_o \geq 2$。

考虑分数部分序列 $d_q = \langle \{qq_0x_2\}, \ldots, \{qq_0x_{k-1}\} \rangle$，它是 $x_2, \ldots, x_{k-1}$ 的扩张。我们的鸽子是 $d_q$ 对于 $0 \leq q \leq (k-2)^{2h}$。我们的抽屉是 $(k-2)$ 重的笛卡尔积，由 $\left[\frac{i}{2h}, \frac{i+1}{2h}\right)$ 和 $0 \leq i < 2h$ 构成。我们有 $(k-2)^{2h} + 1$ 只鸽子，$(k-2)^{2h}$ 个鸽巢，所以有 $q', q''$ 和

$$0 \leq q' < q'' \leq (k-2)^{2h},$$

以及 $d_{q'}, d_{q''}$ 在同一个鸽巢中。也就是说，对于 $Q \coloneqq q''q_0 - q'q_0 \in [q_0, q_0(k-2)^{2h}]$ 和 $a_i$，$Qx_i$ 最接近的整数，我们有

$$-\frac{1}{2h} < Qx_i - a_i < \frac{1}{2h}, \qquad q_0 \le Q \le q_0(k-2)^{2h}$$

同时适用于 $1 \le i \le k$。设置 $\epsilon_i \in (-1/(2h), 1/(2h))$ 使得 $Qx_i = a_i + \epsilon_i$。注意 $\epsilon_1 = \epsilon_k = 0$。

我们声称 $A = \{a_1, \ldots, a_k\}$ 具有与 $|A| = |X|$ 和 $A, X$ 相同类型的加法表。由于 $x_1 < \cdots < x_k$ 和 $Q \ge q_0 \ge 1$，我们显然有 $a_1 \le \cdots \le a_k$。更准确地说，我们有

$$a_{i+1} - a_i = Q(x_{i+1} - x_i) + (\epsilon_i - \epsilon_{i+1}) \ge \operatorname{sep}_1(X)Q - \frac{1}{h} \ge \operatorname{sep}_h(X)q_0 - \frac{1}{2} > 1$$

使得 $a_{i+1} > a_i$。特别是，$|A| = |X|$。

我们将论证 $h$ 类型的 $X$ 与 $h$ 类型的 $A$ 相同，从而得到 $|hX| = |hA|$。首先，假设 $(b_1, \ldots, b_h) \overset{h,X}{\equiv} (c_1, \ldots, c_h)$，从而有

$$x_{b_1} + \cdots + x_{b_h} = x_{c_1} + \cdots + x_{c_h}.$$

因此有

$$a_{b_1} + \cdots a_{b_h} = a_{c_1} + \cdots + a_{c_h} + (\epsilon_{c_1} + \cdots + \epsilon_{c_h}) - (\epsilon_{b_1} + \cdots + \epsilon_{b_h}).$$

根据 $\epsilon_i$ 的定义，我们有

$$-1 < (\epsilon_{c_1} + \cdots + \epsilon_{c_h}) - (\epsilon_{b_1} + \cdots + \epsilon_{b_h}) < 1$$

并且由于 $a_i$ 的整性，这必须为 0。那即是，

$$a_{b_1} + \cdots a_{b_h} = a_{c_1} + \cdots + a_{c_h}$$

进而 $(b_1, \ldots, b_h) \overset{h,A}{\equiv} (c_1, \ldots, c_h)$。

现在，假设 $(b_1, \ldots, b_h) \overset{h,A}{\equiv} (c_1, \ldots, c_h)$，由此我们得到

$$a_{b_1} + \cdots a_{b_h} = a_{c_1} + \cdots + a_{c_h}.$$

正如 $a_i = Qx_i - \epsilon_i$，我们有

$$x_{b_1} + \cdots + x_{b_h} = x_{c_1} + \cdots + x_{c_h} - \left(\frac{\epsilon_{b_1} + \cdots + \epsilon_{b_h} - \epsilon_{c_1} - \cdots - \epsilon_{c_h}}{Q}\right).$$

如

$$\left|\frac{\epsilon_{b_1} + \cdots + \epsilon_{b_h} - \epsilon_{c_1} - \cdots - \epsilon_{c_h}}{Q}\right| < \frac{1}{Q} \le \frac{\operatorname{sep}_h(X)}{2},$$

必须是

$$x_{b_1} + \cdots + x_{b_h} = x_{c_1} + \cdots + x_{c_h}.$$

因此，$(b_1, \ldots, b_h) \overset{h,X}{\equiv} (c_1, \ldots, c_h)$。 □

**定理 6.** 令 $\sigma(\overline{\boldsymbol{x}}) = \sum_{i=1}^{h} x_i$ 和 $\pi(\overline{\boldsymbol{x}}) = \prod_{i=1}^{h} x_i$。我们有

$$\mathcal{T}_{\mathbb{N}}(\sigma, k) = \mathcal{T}_{\mathbb{N}}(\pi, k).$$

特别是，$\mathcal{R}_{\mathbb{N}}(\sigma, k) = \mathcal{R}_{\mathbb{N}}(\pi, k)$.

*证明.* 假设 $S$ 是一个包含 $k$ 个正整数的集合，并设 $P \coloneqq \{2^s : s \in S\}$。那么，$a_1 + \cdots + a_h = b_1 + \cdots + b_h$ 当且仅当 $2^{a_1} \cdots 2^{a_h} = 2^{b_1} \cdots 2^{b_h}$。因此，$\mathcal{T}_{\mathbb{N}}(\sigma, k)$ 中的每一种类型也在 $\mathcal{T}_{\mathbb{N}}(\pi, k)$ 中。

现在，假设 $P$ 是一个由正整数组成的集合，并具有 $|P| = k$。我们必须展示一个由正整数组成的集合 $A$，其 $\sigma$ 类型与 $P$ 的 $\pi$ 类型相同。

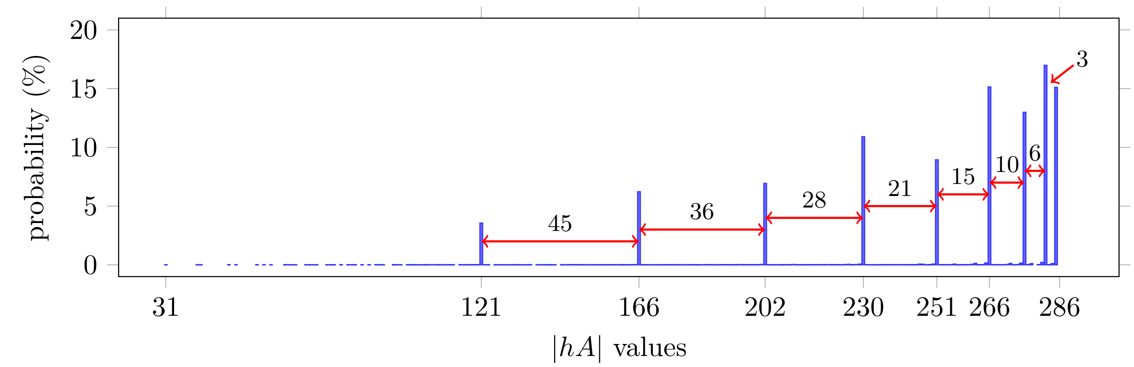
显然，$\pi$ 类型的 $P$ 与 $\sigma$ 类型的 $L_1 \coloneqq \{\log_2 p : p \in P\}$ 相同，但 $L$ 通常不是一个整数集。对于 $c > 0$，我们设定 $L_c \coloneqq \{c \log_2 p : p \in P\}$，它具有与 $L_1$ 相同的 $\sigma$ 类型和与 $P$ 相同的 $\pi$ 类型。我们将证明可以选择 $c$，使得 $A_c \coloneqq \{c + [c \log_2 p] : p \in P\} \subseteq \mathbb{N}$（其中 $[x]$ 是最接近 $x$ 的整数）具有与 $L_c$ 相同的 $\sigma$ 类型。

如果 $c$ 足够大，那么显然 $A_c$ 将会是 $|P|$ 个不同的正整数的集合。如同上述论证，我们可以取 $c$ 为任意大的值，并且使得 $\{c + [c \log_2 p] : p \in P\}$ 具有任意大的间隔，同时每个 $c + [c \log p]$ 都在某个整数的 $1/(2h)$ 范围内。证明过程如上。 $\square$

# 5 实验思想

## 5.1 纳桑森的实验

考虑 Nathanson [8] 进行的以下实验。我们首先生成 $10^7$ 个大小为 4 的随机子集 $\{1, 2, \ldots, 1000\}$。我们有 $|10A|$ 取 213 个不同的值，范围在 31 和 286 之间，包括这两个端点。大约 78% 个是 $B_{10}$ 集合，即 $|hA| = \binom{10+4-1}{4-1} = 286$。在我们样本中的其他 2 202 261 集合中，我们绘制了具有每种可能大小的集合的比例 $|hA|$。



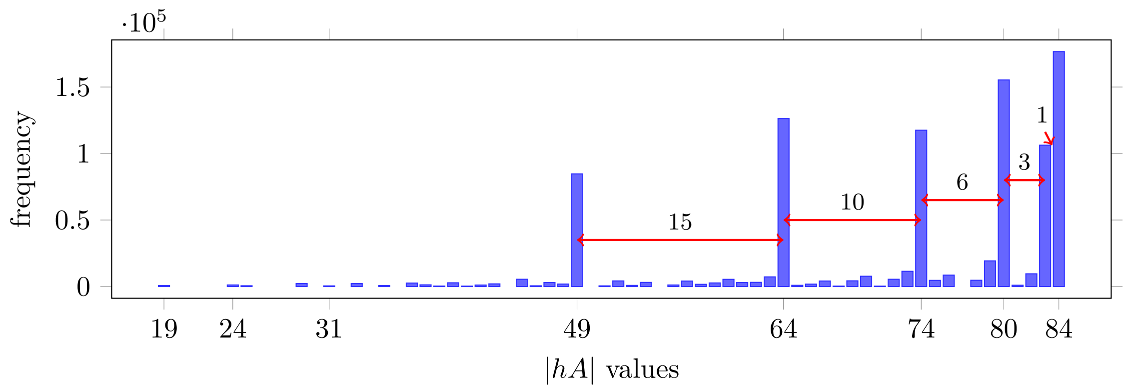



图 1: 6-倍和集大小的柱状图，对于所有 4 元素子集 $A \subsetneqq [70] := \{1, 2, \ldots, 70\}$。正如 Nathanson 观察到的，分布强烈集中在形式为 $\binom{h+3}{3} - \binom{t+2}{3}$ 的大小上，对于 $0 \le t < h$。

我们发现 97% 的 2 202 261 非-$B_{10}$-集 [2] 在 9 个大小中具有 $|hA|$

$$285, 282, 276, 266, 251, 230, 202, 166, 121.$$

列表中小于 286 的 $|10A|$ 公共值的第一差是

$$45, 36, 28, 21, 15, 10, 6, 3,$$

即，三角数。这是个谜团——三角数的出现——我们在本工作中对其给出了部分解释。

我们注意到三角数的部分和是四面体数，由 $\binom{t+2}{3}$ 给出。纳桑森的实验表明，一个随机的 4 元子集 $A$ 在 $[n]$ 中很可能具有 $|hA| = \binom{h+3}{3}$，但如果它没有，则很可能在 $\binom{h+3}{3} - \binom{t+2}{3}$ 中具有 $|hA|$，其中包含 $1 \le t < h$。

### 5.2 极小值的分布

该作者认为，Nathanson 实验中三角数和四面体数的出现是定理 1 的结果以及尚未被证明的事实：大约 80% 个 $\binom{[1000]}{4}$ 具有 $h_1 > 6$，而在剩下的 20% 中，几乎所有都具有 $h_2 > 6$。更一般地说，以下猜想结合定理 1 将会是一个令人满意的解释。

**猜想 7.** 固定 $h \ge 3$，并令 $\epsilon > 0$。存在一个 $n_0$，使得对于所有的 $n \ge n_0$，$\binom{[n]}{4}$ 的子集中至少有 $(1-\epsilon)\binom{n}{4}$ 个包含 $h_1 > h$，并且在那些含有 $h_1 \le h$ 的集合中，包含 $h_2 \le h$ 的比例至少为 $1 - \epsilon$。

[2]预期我们的样本大部分是 $B_{10}$-集：在 $h, k$ 固定的情况下，如果 $n$ 足够大，则 $\binom{[n]}{k}$ 中的几乎所有集合都是 $B_h$-集。

这是一个关于“随机格子”的问题，该格子垂直于 $\overline{\mathbf{1}}$ 和 $\overline{\boldsymbol{a}}$。实验表明，如果从 $\binom{[n]}{4}$ 中均匀选择 $A$，则 $h_1(A)$ 的期望值接近 $\frac{1}{2}n^{1/2}$，标准偏差约为 $\frac{1}{5}n^{1/2}$。如果 $A$ 是从 $\binom{[n]}{5}$ 中均匀选择的，期望值似乎接近 $\frac{5}{8}n^{1/3}$，标准差在 $\frac{1}{5}n^{1/3}$ 附近。

示例：设 $A = \{1, 5, 96, 100\}$。集合 $A$ 不是 $B_2$-集，因为 $1 + 100 = 5 + 96$。用格论术语来说，$h_1 = 2$ 和 $\overline{\boldsymbol{y}}_1 = \langle 1, -1, -1, 1\rangle$。然而，$A$ 确实有

$$|hA| = \binom{h+3}{3} - \binom{h+1}{3}$$

对于 $2 \le h \le 94$。但对于 $h = 95$，我们遇到了新的等式：

$$95 \cdot 96 = 91 \cdot 100 + 4 \cdot 5.$$

用格的语言来说，$h_2 = 95$ 和 $y_2 = \langle 0, -4, 95, -91\rangle$。

这项实验工作将从对该问题的肯定回答中受益：

**问题 8.** *有没有一种方法可以更高效地生成 $k$ 元子集的 $[n]$，这些子集不是 $B_h$ 集，而不仅仅是生成 $k$ 元子集的 $[n]$ 并测试 $B_h$ 属性？*

## 5.3 纳桑森的计划

在近期的研究中，Nathanson [8, 7, 6, 5, 4] 提出了一个计划，以深化对 $|hA|$ 的理解，而不仅仅是其极值。当前的工作符合他的框架。

**文献中的一个空白。** 在 [4] 中，Nathanson 写道“尚不清楚是否为 $\mathcal{R}_{\mathbb{N}_0}(h,k) = \mathcal{R}_{\mathbb{R}}(h,k)$。”我们的定理 2 填补了文献中的这一空白。

**纳桑森的一个问题。** 在 [8] 中，问题 8 是要证明

$$\left\{\binom{h+3}{3} - \binom{j+2}{3} : 1 \le j \le h\right\} \subseteq \mathcal{R}_{\mathbb{Z}}(h,4).$$

这是定理 1 的结果，在集合 $A = \{0, 1, (a-1)(b-1)+1, (a-1)(b-1)+a\}$ 中，其中 $b \ge a \ge 2$。得到的 $\mathcal{L}$ 具有最小值 $2a, 2b$。

**纳桑汉的另一个问题。** 在 [8] 中，问题 9 是“发现大小为 $k \ge 5$ 的集合的流行和集大小模式”。当前的工作表明，流行的值将是

$$\left\{\binom{h+k-1}{k-1} - \binom{j+k-1}{k-1} : 0 \le j < h\right\},$$

前提是 $n$ 在 $h, k$ 方面足够大。适度的实验与这一建议相符。

# 参考文献

[1] Andrew Granville and József Solymosi. Sum-product formulae. In *Recent trends in combinatorics*, volume 159 of *IMA Vol. Math. Appl.*, pages 419–451. Springer, [Cham], 2016. https://doi.org/10.1007/978-3-319-24298-9_18.

[2] Greg Martin and Kevin O'Bryant. Many sets have more sums than differences. In *Additive combinatorics*, volume 43 of *CRM Proc. Lecture Notes*, pages 287–305. Amer. Math. Soc., Providence, RI, 2007. https://doi.org/10.1090/crmp/043/16.

[3] Melvin B. Nathanson. Sums of finite sets of integers. *Amer. Math. Monthly*, 79:1010–1012, 1972. https://doi.org/10.2307/2318072.

[4] Melvyn B. Nathanson. Compression and complexity for sumset sizes in additive number theory, 2025. https://arxiv.org/abs/2505.20998.

[5] Melvyn B. Nathanson. Explicit sumset sizes in additive number theory, 2025. https://arxiv.org/abs/2505.05329.

[6] Melvyn B. Nathanson. Inverse problems for sumset sizes of finite sets of integers, 2025. https://arxiv.org/abs/2412.16154.

[7] Melvyn B. Nathanson. Problems in additive number theory, VI: Sizes of sumsets, 2025. https://arxiv.org/abs/2411.02365.

[8] Melvyn B. Nathanson. Triangular and tetrahedral number differences of sumset sizes in additive number theory, 2025. https://arxiv.org/abs/2506.15015.

[9] Melvyn B. Nathanson, Kevin O'Bryant, Brooke Orosz, Imre Ruzsa, and Manuel Silva. Binary linear forms over finite sets of integers. *Acta Arith.*, 129(4):341–361, 2007. https://doi.org/10.4064/aa129-4-5.

[10] Kevin O'Bryant. Visualizing the sum-product conjecture, 2025. https://arxiv.org/abs/2411.08139.