



TRIANGULAR AND TETRAHEDRAL NUMBER DIFFERENCES OF SUMSET SIZES IN ADDITIVE NUMBER THEORY

Melvyn B. Nathanson

Department of Mathematics, Lehman College (CUNY), Bronx, New York

Received: 6/24/25, Accepted: 2/12/26, Published: 9/28/26

Abstract

The study of sums of finite sets of integers has mostly concentrated on sets with very small sumsets (Freiman's theorem and related work) and on sets with very large sumsets (Sidon sets and B_h -sets). This paper considers the full range of sumset sizes of finite sets of integers and an unexpected pattern (related to the triangular and tetrahedral numbers) that appears in the distribution of popular sumset sizes of sets of size 4.

– To Carl Pomerance on his 80th birthday

1. The Sumset Size Problem

The *integer interval* $[u, v]$ is the set of all integers n such that $u \leq n \leq v$.

The h -fold sum of a set A of integers is the set hA of all sums of h not necessarily distinct elements of A :

$$hA = \{a_1 + \cdots + a_h : a_i \in A \text{ for all } i \in [1, h]\}.$$

It is a fundamental unsolved problem in additive number theory, and a problem that strangely seems not to have been previously investigated, to compute and understand the sumset sizes of finite sets of integers.¹ For all positive integers h and k , we define the *range of sumset sizes*

$$\mathcal{R}_{\mathbf{Z}}(h, k) = \{|hA| : A \subseteq \mathbf{Z} \text{ and } |A| = k\}.$$

This set was introduced in Nathanson [7] and studied in [8, 10, 11].

The idea of “sumset” is central in additive number theory and suggests many simple but *a priori* difficult questions. For example, does there exist a set A of size $|A| = 11$ whose 7-fold sum has size $|7A| = 15551$?

DOI: 10.5281/zenodo.23018242

¹This problem was discussed at three recent number theory conferences (Integers 2025 in Athens, Georgia, CANT 2025 in New York, and BIRS-IMAG in Granada, Spain) and no one could cite earlier work. References to previous papers would be appreciated.

The answer is “yes,” and the set

$$A_0 = \{0, 897, 2056, 2441, 2988, 3259, 5294, 6506, 8013, 9391, 9872\}$$

has the desired sumset sum. However, the answer was obtained by cheating: I started with the set A_0 and computed its 7-fold sum. For this set A_0 , Maple quickly calculated the sumsets hA_0 and the sumset sizes $|hA_0|$ for $h \in [1, 7]$. The second line of Table 1 gives the sequence of sumset sizes $|hA_0|$. The maximum size of an h -fold sumset of a set of size k is $\binom{h+k-1}{h}$. The third line of the table gives the maximum size of an h -fold sumset of a set A of size 11 for $h \in [1, 7]$.

h	1	2	3	4	5	6	7
$ hA_0 $	11	66	285	986	2878	7226	15551
$\max hA $	11	66	286	1001	3003	8008	19448

Table 1: Sumset sizes $|hA_0|$

A B_h -set is a set A of integers such that every integer in the h -fold sumset hA has a unique representation as a sum of h increasing elements of A . If A has size k , then A is a B_h -set if and only if $|hA| = \binom{h+k-1}{h}$. A B_2 -set is also called a Sidon set. We see from Table 1 that A_0 is a Sidon set. Every B_3 -set of size 11 satisfies $|3A| = 286$, but the special set A_0 has $|3A_0| = 285$, and so is not a B_3 -set. Moreover, there must be exactly one integer n_0 that has two representations as a sum of three elements of A_0 and no integer that has more than two representations as a sum of three elements of A_0 . We have

$$n_0 = 18782 = 0 + 9391 + 9391 = 897 + 8013 + 9872 \in 3A_0.$$

There are many related questions. The maximum size of a 7-fold sumset of a set of 10 integers is $\binom{16}{7} = 11440$. It follows that if A is a set of integers with $|7A| = 15551$, then $|A| \geq 11$. What is the cardinality of the largest set A such that $|7A| = 15551$?

Sets A and B are *affinely equivalent* if there are numbers $\lambda \neq 0$ and μ such that

$$B = \lambda * A + \mu = \{\lambda a + \mu : a \in A\}.$$

Affinely equivalent sets have the same sumset sums. Sets A and B are *affinely inequivalent* if they are not affinely equivalent. How many affinely inequivalent sets A satisfy $|A| = 11$ and $|7A| = 15551$? How many affinely inequivalent finite sets A satisfy $|7A| = 15551$? The computation complexity inequality (2) implies that all these questions can be answered in finite time.

We have

$$\max \mathcal{R}_{\mathbf{Z}}(h, k) = \binom{h+k-1}{k-1}$$

and sets with this sumset size (B_h -sets and Sidon sets) have been extensively studied ([5, 6, 9, 12]). Similarly,

$$\min \mathcal{R}_{\mathbf{Z}}(h, k) = h(k - 1) + 1$$

and sets with this sumset size (called arithmetic progressions) and sets with very small sumset size have also been extensively studied (Freiman's theorem and related work [4, 17]). Thus,

$$\mathcal{R}_{\mathbf{Z}}(h, k) \subseteq \left[h(k - 1) + 1, \binom{h + k - 1}{k - 1} \right]. \quad (1)$$

Because the set $\mathcal{R}_{\mathbf{Z}}(h, k)$ is finite, we can define the integer $N(h, k)$ as the smallest integer N such that

$$\mathcal{R}_{\mathbf{Z}}(h, k) = \{|hA| : A \subseteq [0, N - 1] \text{ and } |A| = k\}.$$

The integer $N(h, k)$ determines the computational complexity of the sumset size problem. It is proved in [11] that, for all $h \geq 3$ and $k \geq 3$,

$$N(h, k) < 4(8h)^{k-1}. \quad (2)$$

Thus, for fixed k , the set $\mathcal{R}_{\mathbf{Z}}(h, k)$ can be computed in polynomial time.

Here are some simple results. For all h and k ,

$$\mathcal{R}_{\mathbf{Z}}(h, 1) = \{1\}, \quad \mathcal{R}_{\mathbf{Z}}(1, k) = \{k\}, \quad \text{and} \quad \mathcal{R}_{\mathbf{Z}}(h, 2) = \{h + 1\}.$$

For all k , the set $\mathcal{R}_{\mathbf{Z}}(2, k)$ is the interval of integers:

$$\mathcal{R}_{\mathbf{Z}}(2, k) = \left[2k - 1, \binom{k + 1}{2} \right].$$

However, the set $\mathcal{R}_{\mathbf{Z}}(3, 3)$ is not an interval of integers. We have

$$\mathcal{R}_{\mathbf{Z}}(3, 3) \subseteq [7, 10]$$

but

$$\mathcal{R}_{\mathbf{Z}}(3, 3) = \{7, 9, 10\}.$$

The number 8 missing. For $k = 3$ and all $h \geq 1$, Nathanson [7] proved that

$$\mathcal{R}_{\mathbf{Z}}(h, 3) = \left\{ \binom{h + 2}{2} - \binom{t}{2} : t \in [1, h] \right\}.$$

and so $|\mathcal{R}_{\mathbf{Z}}(h, 3)| = h$.

Problem 1. For $k = 4$, we have

$$\mathcal{R}_{\mathbf{Z}}(h, 4) \subseteq \left[3h + 1, \binom{h+3}{3} \right].$$

Compute $\mathcal{R}_{\mathbf{Z}}(h, 4)$ for all positive integers h .

Problem 2. For all positive integers h and k , compute the set of sumset sizes $\mathcal{R}_{\mathbf{Z}}(h, k)$. Explain the integers “missing” from the interval on the right side of (1). This is a core problem in additive number theory.

Problem 3. Let $k \geq 4$. Prove that “most” numbers in the interval on the right side of (1) are *not* sumset sizes, that is, prove that

$$|\mathcal{R}_{\mathbf{Z}}(h, k)| = o_k(h^{k-1})$$

or, better, that

$$|\mathcal{R}_{\mathbf{Z}}(h, k)| = O_k(h^{k-2}).$$

Problem 4. Is there some kind of relation or reciprocity between the sets $\mathcal{R}_{\mathbf{Z}}(h, k)$ and $\mathcal{R}_{\mathbf{Z}}(k, h)$? For example, is there a function $f : \mathbf{N} \rightarrow \mathbf{N}$ such that, if $t \in \mathcal{R}_{\mathbf{Z}}(h, k)$, then $f(t) \in \mathcal{R}_{\mathbf{Z}}(k, h)$? Is there a function $g : \mathbf{N} \rightarrow \mathbf{N}$ such that $t \in \mathcal{R}_{\mathbf{Z}}(h, k)$ if and only if $g(t) \in \mathcal{R}_{\mathbf{Z}}(k, h)$? Are there functions α, β, γ such that, if $t \in \mathcal{R}_{\mathbf{Z}}(h, k)$, then $\gamma(t) \in \mathcal{R}_{\mathbf{Z}}(\alpha(h), \beta(k))$?

For related work on “sumset size races,” see [2, 3, 14].

2. 5-fold Sums of Sets of Size 4

In this section we describe some computational experiments on 5-fold sumsets of sets of size 4. The calculations were programmed in Maple.

Experiment 1. Let $q = 100$. There are

$$\binom{100}{4} = 3,921,225$$

subsets of size 4 contained in the integer interval $[0, q - 1]$. Their 5-fold sumsets are contained in $[16, 56]$. For all $t \in [16, 56]$, let $S(t)$ denote the number of sets $A \subseteq [0, q - 1] = [0, 99]$ such that $|A| = 4$ and $|5A| = t$. Table 2 records the distribution of sumset sizes. The relation $S(16) = 1617$ is equivalent to the statement that the interval $[0, 99]$ contains exactly 1617 arithmetic progressions of length 4. The relation $S(56) = 1867410$ means that the interval $[0, 99]$ contains exactly 1867410 B_5 -sets.

t	$S(t)$	t	$S(t)$	t	$S(t)$
16	1617	30	4655	44	10574
17	0	31	3444	45	18862
18	0	32	5273	46	450320
19	0	33	1906	47	6706
20	2400	34	5846	48	6120
21	1200	35	5400	49	0
22	0	36	280850	50	15002
23	0	37	1152	51	27332
24	4750	38	3176	52	460734
25	0	39	4612	53	7628
26	950	40	0	54	30688
27	4704	41	10012	55	667836
28	0	42	5888	56	1867410
29	1568	43	2610		

Table 2: Data for $\mathcal{R}(5, 4)$ with $q = 100$

Experiment 2. Let $q = 1,000$. Maple randomly chose 150,000 sets of size 4 in the interval $[0, q - 1]$ and computed their 5-fold sumsets. The set of experimentally observed sumset sizes is

$$\{20, 27, 30, 32, 34, 35, 36, 39, 41, 42, 44, 45, 46, 47, 50, 51, 52, 53, 54, 55, 56\}.$$

Table 3 lists these elements t of the sumset size set $\mathcal{R}(5, 4)$ and the number $S(t)$ of sets A in the sample space with $|5A| = t$.

t	$S(t)$	$ 5A $	$S(t)$
20	1	45	6
27	2	46	1738
30	2	47	1
32	2	50	6
34	2	51	7
35	1	52	1957
36	982	53	2
39	2	54	9
41	2	55	3251
42	1	56	142020
44	6		

Table 3: Data for $\mathcal{R}(5, 4)$ with $q = 1000$

Experiment 3. Let $q = 10,000$. Maple randomly chose 1,000,000 sets of size 4 in the interval $[0, q - 1]$ and computed their 5-fold sumsets. The set of experimentally observed sumset sizes is

$$\{21, 36, 38, 39, 42, 44, 45, 46, 50, 51, 52, 54, 55, 56\}.$$

Table 4 lists these elements t of the set $R(5, 4)$ and the number $S(t)$ of sets A in the sample space with $|5A| = t$.

t	$S(t)$	t	$S(t)$
21	584	46	1429
36	805	50	1
38	1	51	1
39	1	52	1654
42	1	54	1
44	1	55	2559
45	1	56	992961

Table 4: Data for $\mathcal{R}(5, 4)$ with $q = 10,000$

Experiment 4. Let $q = 1,000,000$. Maple randomly chose 250,000 sets of size 4 in the interval $[0, q - 1]$ and computed their 5-fold sumsets. The set of experimentally observed sumset sizes is

$$\{36, 46, 52, 55, 56\}.$$

Table 5 lists these elements t of the set $R(5, 4)$ and the number $S(t)$ of sets A in the sample space with $|5A| = t$.

t	$S(t)$
36	1
46	4
52	2
55	4
56	249989

Table 5: Data for $\mathcal{R}(5, 4)$ with $q = 1,000,000$

Problem 5. In all four experiments, the sets with sumset size 56 predominate and these are the B_5 -sets. The most popular sumset sizes are 56, 55, 52, 46, and 36. Their differences are the consecutive triangular numbers 1, 3, 6, and 10. Why? Determine the pattern and frequencies of sumset sizes $S(t)$ for values of t between the most popular sumset sizes.

Problem 6. Let $N = N(5, 4) = 4 \cdot 40^3 = 256,000$. Inequality (2) implies that $\mathcal{R}_{\mathbf{Z}}(5, 4)$ is the set of all 5-fold sums of 4-element sets of integers contained in the interval $[0, N - 1]$. Compute $\mathcal{R}_{\mathbf{Z}}(5, 4)$.

3. A Mystery about Popular Sumset Sizes

Maple was used to generate random sets of 4 integers contained in the intervals $[0, 999]$ and $[0, 9999]$, to calculate their h -fold sumsets for $h \in [2, 9]$, and to find the “popular” sumset sizes. Table 6 shows the results.

h	most popular sumset sizes	successive differences
2	9, 10	1
3	16, 19, 20	3, 1
4	25, 31, 34, 35	6, 3, 1
5	36, 46, 52, 55, 56	10, 6, 3, 1
6	49, 64, 74, 80, 83, 84	15, 10, 6, 3, 1
7	64, 85, 100, 110, 116, 119, 120	21, 15, 10, 6, 3, 1
8	81, 109, 130, 145, 155, 161, 164, 165	28, 21, 15, 10, 6, 3, 1
9	100, 136, 164, 185, 200, 210, 216, 219, 220	36, 28, 21, 15, 10, 6, 3, 1

Table 6: Popular sumset sizes for $\mathcal{R}(5, 4)$

Problem 7. For all $h \in [2, 9]$, the sumset size set $\mathcal{R}(h, 4)$ has h popular sizes and the differences between consecutive popular sumset sizes are the first $h - 1$ triangular numbers. Why?

Recall that the j th *tetrahedral number* $\text{Tet}_j = \binom{j+2}{3}$ is the sum of the first j triangular numbers. The popular sumset sizes appear to be the differences between $\binom{h+3}{3}$, which is the size of a 4-element B_h -set, and the h consecutive tetrahedral numbers $\text{Tet}_0, \text{Tet}_1, \dots, \text{Tet}_{h-1}$.

Problem 8. Let $h \geq 3$. Prove that

$$\left\{ \binom{h+3}{3} - \text{Tet}_j : j \in [1, h-1] \right\} \subseteq \mathcal{R}_{\mathbf{Z}}(h, 4).$$

Problem 9. Let h and q be positive integers. Let

$$\mathcal{A}(4, q) = \{A \subseteq [0, q-1] : |A| = 4\}$$

$$\mathcal{A}^*(4, q, h) = \{A \in \mathcal{A}(4, q) : A \text{ is not a } B_h\text{-set}\}$$

and

$$\mathcal{B}^*(4, q, h) = \left\{ A \in \mathcal{A}^*(4, q, h) : |hA| \in \left\{ \binom{h+3}{3} - \text{Tet}_j : j \in [1, h-1] \right\} \right\}$$

Almost all 4-element subsets of $[0, q-1]$ are B_h -sets (Nathanson [5, 6]):

$$\lim_{q \rightarrow \infty} \frac{|\mathcal{A}^*(4, q, h)|}{|\mathcal{A}(4, q)|} = 0.$$

Is it true that the sumset sizes of almost all sets $A \in \mathcal{A}(4, q)$ that are not B_h -sets (that is, sets in $\mathcal{A}^*(4, q)$) are the popular sumset sizes? Equivalently, is it true that

$$\lim_{q \rightarrow \infty} \frac{|\mathcal{B}^*(4, q, h)|}{|\mathcal{A}^*(4, q, h)|} = 1?$$

Problem 10. Discover the pattern of popular sumset sizes for sets of size $k \geq 5$.

4. A Toy Problem for Physics

For positive integers h and k , let $\mathcal{X}_{h,k}$ be the set of k -tuples $\mathbf{x} = (x_1, \dots, x_k)$ of nonnegative integers such that $\sum_{i=1}^k x_i = h$. Let $A = \{a_1, \dots, a_k\}$ be a set of k integers and let $\vec{A} = (a_1, \dots, a_k)$ be the associated lattice point in $\mathbf{Z}^k$. For all $\mathbf{x} = (x_1, \dots, x_k) \in \mathcal{X}_{h,k}$, we define

$$\mathbf{x} \cdot \vec{A} = \sum_{i=1}^k x_i a_i.$$

Then

$$hA = \left\{ \mathbf{x} \cdot \vec{A} : \mathbf{x} \in \mathcal{X}_{h,k} \right\}. \quad (3)$$

Because $\mathcal{X}_{h,k}$ is invariant with respect to the permutation group S_k (that is, $\mathbf{x} = (x_1, \dots, x_k) \in \mathcal{X}_{h,k}$ if and only if $\sigma\mathbf{x} = (x_{\sigma(1)}, \dots, x_{\sigma(k)}) \in \mathcal{X}_{h,k}$ for all $\sigma \in S_k$), relation (3) is independent of the ordering of the coordinates in the vector $\vec{A}$.

The *representation function* $r_{A,h}(n)$ counts the number of representations of n as a sum of h elements of A , that is, the number of k -tuples $\mathbf{x} = (x_1, \dots, x_k) \in \mathcal{X}_{h,k}$ such that $\mathbf{x} \cdot \vec{A} = \sum_{i=1}^k x_i a_i = n$. Analogous to collisions of molecules in a gas, k -tuples $\mathbf{x} \in \mathcal{X}_{h,k}$ and $\mathbf{y} \in \mathcal{X}_{h,k}$ “collide” with respect to the set A if the h -fold sums $\mathbf{x} \cdot \vec{A}$ and $\mathbf{y} \cdot \vec{A}$ are equal. Moreover, like multiple collisions of molecules, multiple collisions of vectors in $\mathcal{X}_{h,k}$, that is, integers $n \in hA$ such that $r_{A,h}(n) \geq 3$, should usually be much rarer than simple collisions with $r_{A,h}(n) = 2$, but not always. For example, with $k \geq 5$, the integer interval $A = [0, k-1]$ satisfies $2A = [0, 2k-2]$ and

$$r_{A,2}(n) = \begin{cases} \lfloor \frac{n}{2} \rfloor + 1 & \text{if } 0 \leq n \leq k-1 \\ \lfloor \frac{2k-2-n}{2} \rfloor + 1 & \text{if } k-1 \leq n \leq 2k-2. \end{cases}$$

Thus,

$$\begin{aligned} r_{A,2}(n) &= 1 && \text{if } n \in \{0, 1, 2k-3, 2k-2\} \\ r_{A,2}(n) &= 2 && \text{if } n \in \{2, 3, 2k-5, 2k-4\} \\ r_{A,2}(n) &\geq 3 && \text{if } 4 \leq n \leq 2k-6 \end{aligned}$$

and so

$$\text{card}(n \in \mathbf{Z} : r_{A,h}(n) \geq 3) > \text{card}(n \in \mathbf{Z} : r_{A,h}(n) = 2) \quad (4)$$

for all $k \geq 7$. Inequality (4) should be atypical. A quantitative version of this remark is the following.

Problem 11. For every sufficiently large integer q , let

$$\mathcal{A}(k, q) = \{A \subseteq [0, q-1] : |A| = k\}.$$

Prove that

$$\text{card}(n \in \mathbf{Z} : r_{A,h}(n) \geq 3) < \text{card}(n \in \mathbf{Z} : r_{A,h}(n) = 2)$$

for almost all $A \in \mathcal{A}(k, q)$.

Computing the patterns of sumset sizes of finite sets of integers can be considered as a toy problem for certain much deeper questions in mathematical physics, such as the Hilbert program of deducing macroscopic fluid mechanical phenomena from mesoscopic and microscopic motions. A significant part of recent work of Yu Deng, Zaher Hani, and Xiao Ma [1, 16] on Hilbert’s sixth problem (the derivation of the Navier-Stokes equation from the Boltzman equation from elastic collisions of a gas) is combinatorial and has a number theoretical flavor. It would be of interest to examine the combinatorial results of Deng, Hani, and Ma, and the similarities between phenomena of combinatorial additive number theory and fluid motion.

Acknowledgements. I thank Christian Elsholtz, Harald Helfgott, Noah Kravitz, Kevin O’Bryant, and Steven Senger for helpful remarks. Noah Kravitz (personal communication) and Kevin O’Bryant [13] have obtained results related to some of the problems in this paper.

References

- [1] Yu Deng, Zaher Hani, and Xiao Ma, Hilbert’s sixth problem: derivation of fluid equations via Boltzmann’s kinetic theory, preprint, [arXiv:2503.01800](#).
- [2] J. Fox, N. Kravitz, and S. Zhang, Finer control on relative sizes of iterated sumsets, preprint, [arXiv:2506.05691](#).

- [3] N. Kravitz, Relative sizes of iterated sumsets, *J. Number Theory* **272** (2025), 113–128,
- [4] M. B. Nathanson, *Additive Number Theory: Inverse Problems and the Geometry of Sumsets*, Springer-Verlag, New York, 1996.
- [5] M. B. Nathanson, On the ubiquity of Sidon sets, in *Number Theory (New York, 2003)*, Springer, New York, 2004, 263–272.
- [6] M. B. Nathanson, B_h -sets of real and complex numbers, *Canadian Math. Bull.* **69** (2026), no. 1, 137–141.
- [7] M. B. Nathanson, Problems in additive number theory, VI: Sizes of sumsets of finite sets, *Acta Math. Hungar.* **176** (2025), 498–521.
- [8] M. B. Nathanson, Inverse problems for sumset sizes of finite sets of integers, *Fibonacci Quarterly* (2026), 13 pp.
doi.org/10.1080/00150517.2025.2480773
- [9] M. B. Nathanson, $\mathbf{Q}$ -independence and the construction of B_h -sets of integers and lattice points, *Discrete Math.* **349** (2026), no. 2, Paper No. 114726.
- [10] M. B. Nathanson, Explicit sumset sizes in additive number theory, *Canadian Math. Bull.* (2025), 12 pp.
doi.org/10.4153/S0008439525101549
- [11] M. B. Nathanson, Compression and complexity for sumset sizes in additive number theory, *J. Number Theory* **281** (2026), 321–343
- [12] K. O’Bryant, A complete annotated bibliography of work related to Sidon sequences, *Electron. J. Combin.* **DS11** (2004), 39 pp.
- [13] K. O’Bryant, On Nathanson’s triangular number phenomenon, *Integers*, to appear.
preprint: **arXiv: 2506.20836**
- [14] P. Péringuey and A. de Roton, A note on iterated sumsets races, preprint, **arXiv:2505.11233**.
- [15] V. Schinina, On the sumset of sets of size k , preprint **arXiv:2505.07679**.
- [16] L. Sloman, Epic effort to ground physics in math opens up the secrets of time, *Quanta Magazine*, June 11, 2025.
- [17] T. Tao and V. Vu, *Additive Combinatorics*, Cambridge Univ. Press, 2006.