

# A note on Erdős problem 741

23 April 2026

## Abstract

We separate the two density interpretations appearing in Erdős problem #741. For ordinary natural density, the first question has a negative answer. For upper density, the first question has a positive answer. The second question also has a positive answer: there is an additive basis of order 2 for which every bipartition leaves one monochromatic self-sumset with unbounded gaps. For completeness we give short proofs of all three statements, and we compare the explicit basis from [1] with the alternative construction summarized in the forum thread [6].

## 1 Introduction

Erdős asked the following questions (see [3, p. 263] and the current problem page [2]):

Let  $A \subseteq \mathbb{N}$  be such that  $A + A$  has positive density. Can one always decompose

$$A = A_1 \sqcup A_2$$

such that  $A_1 + A_1$  and  $A_2 + A_2$  both have positive density?

Is there a basis  $A$  of order 2 such that if  $A = A_1 \sqcup A_2$  then  $A_1 + A_1$  and  $A_2 + A_2$  cannot both have bounded gaps?

The current problem page notes that, in the first question, Erdős probably intended *upper density* rather than strict existence of the natural density [2]. The forum comments record two further developments: a formal proof by DeepMind of the upper-density version, and a counterexample to the strict natural-density version [6, 5]. The second question is answered affirmatively in the short paper of Alexeev, Putterman, Sawhney, Sellke and Valiant [1].

Our aim is to package these points into a single self-contained note.

**Definition 1.1.** For  $X \subseteq \mathbb{N}$ , we write

$$d(X) := \lim_{N \rightarrow \infty} \frac{|X \cap [1, N]|}{N}$$

when the limit exists, and

$$\bar{d}(X) := \limsup_{N \rightarrow \infty} \frac{|X \cap [1, N]|}{N}.$$

A set  $S \subseteq \mathbb{N}$  is *syndetic* if it has bounded gaps.

Shifting a set by a fixed integer does not change any of the density statements considered below, so in the proof of the strict-density counterexample we shall work in  $\mathbb{N}_0 = \{0, 1, 2, \dots\}$  for notational convenience.

Our main results are the following.

**Theorem 1.2.** *If  $A \subseteq \mathbb{N}$  satisfies  $\bar{d}(A + A) > 0$ , then there is a partition*

$$A = A_1 \sqcup A_2$$

*such that  $\bar{d}(A_1 + A_1) > 0$  and  $\bar{d}(A_2 + A_2) > 0$ .*

**Theorem 1.3.** *There exists  $A \subseteq \mathbb{N}$  such that  $d(A + A) = 1$ , but for every partition  $A = A_1 \sqcup A_2$  it is not the case that both  $d(A_1 + A_1)$  and  $d(A_2 + A_2)$  exist and are positive.*

**Theorem 1.4.** *There exists a set  $A \subseteq \mathbb{N}$  such that  $A$  is an asymptotic basis of order 2, and for every partition  $A = A_1 \sqcup A_2$ , at least one of  $A_1 + A_1$  and  $A_2 + A_2$  is not syndetic.*

Thus the thread [6] is settled as follows:

- the first question is *false* for strict natural density, but *true* for upper density;
- the second question has an affirmative answer.

## 2 The upper-density version is true

We begin with two elementary lemmas.

**Lemma 2.1.** *If  $X \subseteq \mathbb{N}$  and  $\bar{d}(X) > 0$ , then  $\bar{d}(X + X) \geq \bar{d}(X)$ .*

*Proof.* For each  $N$ , put  $X_N := X \cap [1, N]$ . Since  $X_N$  is a finite set of integers, the standard bound

$$|X_N + X_N| \geq 2|X_N| - 1$$

holds. Therefore

$$\frac{|(X + X) \cap [1, 2N]|}{2N} \geq \frac{|X_N + X_N|}{2N} \geq \frac{2|X_N| - 1}{2N} = \frac{|X \cap [1, N]|}{N} - \frac{1}{2N}.$$

Taking the limsup along  $N \rightarrow \infty$  gives  $\bar{d}(X + X) \geq \bar{d}(X)$ . □

**Lemma 2.2.** *Let  $X, B \subseteq \mathbb{N}$ , and suppose  $B \cap [1, N] \subseteq [1, K]$ . Then*

$$|(X + B) \cap [1, N]| \leq K |X \cap [1, N]|.$$

*In particular, if  $A = A_1 \sqcup A_2$  and  $A_2 \cap [1, N] \subseteq [1, K]$ , then*

$$|(A + A) \cap [1, N]| \leq |(A_1 + A_1) \cap [1, N]| + K |A \cap [1, N]|.$$

*Proof.* Since  $B \cap [1, N] \subseteq [1, K]$ , we have

$$(X + B) \cap [1, N] \subseteq \bigcup_{b=1}^K ((X + b) \cap [1, N]).$$

Each translate  $X + b$  contributes at most  $|X \cap [1, N]|$  elements to  $[1, N]$ , and the first claim follows. For the second claim, every sum in  $(A + A) \cap [1, N]$  outside  $A_1 + A_1$  uses at least one summand from  $A_2 \cap [1, N]$ , hence lies in  $(A + A_2) \cap [1, N]$ . □

*Proof of Theorem 1.2.* Write

$$a(N) := |A \cap [1, N]|, \quad s(N) := |(A + A) \cap [1, N]|.$$

We shall construct an increasing sequence

$$0 < M_0 < M_1 < M_2 < \dots$$

and then define the alternating-block partition

$$A_1 := A \cap \bigcup_{j \geq 0} (M_{2j}, M_{2j+1}], \quad A_2 := A \setminus A_1.$$

We split into two cases.

**Case 1:**  $\bar{d}(A) > 0$ . Choose  $c > 0$  and infinitely many  $N$  such that  $a(N) \geq cN$ . Recursively choose  $M_{k+1}$  so large that

$$a(M_{k+1}) \geq cM_{k+1} \quad \text{and} \quad a(M_k) \leq \frac{c}{4}M_{k+1}. \quad (1)$$

This is possible because  $M_k$  is fixed while there are arbitrarily large scales with density at least  $c$ .

At the odd scales  $N = M_{2j+1}$  we have

$$|A_1 \cap [1, N]| \geq a(N) - a(M_{2j}) \geq cN - \frac{c}{4}N = \frac{3c}{4}N,$$

so  $\bar{d}(A_1) > 0$ . Likewise, evaluating at the even scales  $N = M_{2j+2}$  gives  $\bar{d}(A_2) > 0$ . By Lemma 2.1, both  $\bar{d}(A_1 + A_1)$  and  $\bar{d}(A_2 + A_2)$  are positive.

**Case 2:**  $\bar{d}(A) = 0$  but  $\bar{d}(A + A) > 0$ . Choose  $c > 0$  and infinitely many  $N$  such that  $s(N) \geq cN$ . Since  $a(N) = o(N)$ , for each fixed  $K$  there are arbitrarily large  $N$  such that

$$Ka(N) \leq \frac{c}{4}N.$$

Recursively choose  $M_{k+1}$  so large that

$$s(M_{k+1}) \geq cM_{k+1} \quad \text{and} \quad M_k a(M_{k+1}) \leq \frac{c}{4}M_{k+1}. \quad (2)$$

Now let  $N = M_{2j+1}$ . Then every element of  $A_2 \cap [1, N]$  is at most  $M_{2j}$ . Applying Lemma 2.2 with  $K = M_{2j}$  gives

$$|(A + A) \cap [1, N]| \leq |(A_1 + A_1) \cap [1, N]| + M_{2j}a(N).$$

Using (2), we obtain

$$|(A_1 + A_1) \cap [1, N]| \geq cN - \frac{c}{4}N = \frac{3c}{4}N.$$

Hence  $\bar{d}(A_1 + A_1) > 0$ . The same argument at the even scales  $N = M_{2j+2}$  yields  $\bar{d}(A_2 + A_2) > 0$ .  $\square$

### 3 The strict natural-density version is false

In this section we prove Theorem 1.3. We work in  $\mathbb{N}_0$  and then shift by 1 at the end.

Define

$$B_1 := \left\{ \sum_{r \geq 0} \varepsilon_r 4^r : \varepsilon_r \in \{0, 1\} \right\}, \quad B_2 := \left\{ \sum_{r \geq 0} 2\delta_r 4^r : \delta_r \in \{0, 1\} \right\}.$$

For each  $k \geq 1$ , put

$$N_k := 4^{3^k}, \quad I_k := [N_k, (k+1)N_k) \cap \mathbb{N}_0.$$

Finally set

$$A := B_1 \cup B_2 \cup \bigcup_{k \geq 1} I_k.$$

The key point is that  $B_1 + B_2 = \mathbb{N}_0$ , while the long intervals  $I_k$  force large local oscillations in any monochromatic self-sumset.

**Lemma 3.1.** *For every  $m \geq 1$ , the map*

$$(B_1 \cap [0, 4^m)) \times (B_2 \cap [0, 4^m)) \rightarrow [0, 4^m), \quad (x, y) \mapsto x + y,$$

*is a bijection. In particular,  $B_1 + B_2 = \mathbb{N}_0$ .*

*Proof.* Every digit  $d \in \{0, 1, 2, 3\}$  has a unique decomposition

$$d = \varepsilon + 2\delta, \quad \varepsilon, \delta \in \{0, 1\}.$$

Applying this independently to each base-4 digit yields a unique representation of every integer in  $[0, 4^m)$  as  $x + y$  with  $x \in B_1 \cap [0, 4^m)$  and  $y \in B_2 \cap [0, 4^m)$ .  $\square$

**Lemma 3.2.** *Let*

$$T_k := \sum_{j < k} |I_j| = \sum_{j < k} j N_j.$$

*Then  $T_k = o(\sqrt{N_k})$ . Consequently,*

$$T_k \sqrt{N_k} + T_k^2 = o(N_k).$$

*Proof.* Since  $N_j = 4^{3^j}$  grows super-exponentially,

$$T_k \leq k^2 N_{k-1} = k^2 4^{3^{k-1}} = k^2 2^{2 \cdot 3^{k-1}} = o(2^{3^k}) = o(\sqrt{N_k}).$$

The second claim follows immediately.  $\square$

**Proposition 3.3.** *Suppose  $A = A_1 \sqcup A_2$  and both natural densities*

$$d(A_1 + A_1) = \alpha_1 > 0, \quad d(A_2 + A_2) = \alpha_2 > 0$$

*exist. Then  $\alpha_1 + \alpha_2 < 1$ .*

*Proof.* Fix  $k$ , and write

$$X_k := A_1 \cap B_1 \cap [0, N_k), \quad Y_k := A_1 \cap B_2 \cap [0, N_k).$$

Set

$$x_k := \frac{|X_k|}{\sqrt{N_k}}, \quad y_k := \frac{|Y_k|}{\sqrt{N_k}}.$$

Then  $A_2$  occupies proportions  $1 - x_k$  and  $1 - y_k$  inside  $B_1 \cap [0, N_k)$  and  $B_2 \cap [0, N_k)$  respectively.

By Lemma 3.1, the cross-sum map between  $B_1 \cap [0, N_k)$  and  $B_2 \cap [0, N_k)$  is bijective, so

$$|X_k + Y_k| = |X_k| |Y_k| = x_k y_k N_k.$$

Next, the only points of  $A$  below  $N_k$  outside  $B_1 \cup B_2$  lie in  $\bigcup_{j < k} I_j$ , whose total size is  $T_k$ . By Lemma 3.2, the number of sums below  $N_k$  using at least one point of  $\bigcup_{j < k} I_j$  is at most

$$2T_k \sqrt{N_k} + T_k^2 = o(N_k).$$

Also,

$$|(B_1 + B_1) \cap [0, N_k)| \leq 3^{3^k} = o(N_k), \quad |(B_2 + B_2) \cap [0, N_k)| \leq 3^{3^k} = o(N_k),$$

because  $B_1 \cap [0, N_k)$  has  $3^{3^k}$  possible digitwise sums and  $B_2 = 2B_1$ .

It follows that

$$\frac{|(A_1 + A_1) \cap [0, N_k)|}{N_k} = x_k y_k + o(1),$$

and similarly

$$\frac{|(A_2 + A_2) \cap [0, N_k)|}{N_k} = (1 - x_k)(1 - y_k) + o(1).$$

Since the two natural densities exist, we obtain

$$x_k y_k \rightarrow \alpha_1, \quad (1 - x_k)(1 - y_k) \rightarrow \alpha_2.$$

Now set

$$p_k := x_k y_k, \quad q_k := (1 - x_k)(1 - y_k).$$

Then

$$1 - p_k - q_k = x_k(1 - y_k) + y_k(1 - x_k).$$

Writing

$$u_k := x_k(1 - y_k), \quad v_k := y_k(1 - x_k),$$

we have  $u_k, v_k \in [0, 1]$  and

$$p_k q_k = u_k v_k \leq u_k + v_k = 1 - p_k - q_k.$$

Hence

$$p_k + q_k \leq 1 - p_k q_k.$$

Taking  $k \rightarrow \infty$  yields

$$\alpha_1 + \alpha_2 \leq 1 - \alpha_1 \alpha_2 < 1. \quad \square$$

**Proposition 3.4.** *Under the hypotheses of Proposition 3.3, one has  $\alpha_1 + \alpha_2 \geq 1$ .*

*Proof.* Let

$$U_k := A_1 \cap I_k, \quad V_k := A_2 \cap I_k.$$

Since  $U_k \sqcup V_k = I_k$ , we have

$$|U_k| + |V_k| = |I_k| = kN_k.$$

For any finite set  $X \subseteq \mathbb{Z}$ ,

$$|X + X| \geq 2|X| - 1.$$

Therefore

$$|U_k + U_k| + |V_k + V_k| \geq (2|U_k| - 1) + (2|V_k| - 1) = 2kN_k - 2.$$

All these sums lie in  $[0, 2(k+1)N_k)$ , so with

$$L_k := 2(k+1)N_k$$

we get

$$\frac{|(A_1 + A_1) \cap [0, L_k]|}{L_k} + \frac{|(A_2 + A_2) \cap [0, L_k]|}{L_k} \geq \frac{2kN_k - 2}{2(k+1)N_k}.$$

The right-hand side tends to 1 as  $k \rightarrow \infty$ . Since the two natural densities exist, the left-hand side tends to  $\alpha_1 + \alpha_2$ , proving  $\alpha_1 + \alpha_2 \geq 1$ .  $\square$

*Proof of Theorem 1.3.* By Lemma 3.1, we have  $A + A = \mathbb{N}_0$ , so  $d(A + A) = 1$ . If there were a partition  $A = A_1 \sqcup A_2$  for which both  $d(A_1 + A_1)$  and  $d(A_2 + A_2)$  existed and were positive, then Propositions 3.3 and 3.4 would give the contradiction

$$1 \leq \alpha_1 + \alpha_2 < 1.$$

Therefore no such partition exists. Finally, shifting the whole construction by 1 gives a set in  $\mathbb{N}$  with the same property.  $\square$

## 4 A basis of order 2 with no syndetic split

We now prove Theorem 1.4. The construction and proof are essentially those of Alexeev, Putterman, Sawhney, Sellke and Valiant [1, Section 3]; we reproduce them here for completeness.

For  $k \geq 1$ , define

$$c_k := 4 \cdot 5^{k-1}, \quad B_k := [5 \cdot 5^{k-1}, 6 \cdot 5^{k-1} - 1], \quad F_k := [10 \cdot 5^{k-1} - 1, 15 \cdot 5^{k-1}].$$

Set

$$A := [2, 3] \cup \bigcup_{k \geq 1} (\{c_k\} \cup B_k \cup F_k),$$

and for  $k \geq 0$  put

$$A_k := [2, 3] \cup \bigcup_{i=1}^k (\{c_i\} \cup B_i \cup F_i).$$

**Lemma 4.1.** *For every  $k \geq 0$ ,*

$$[4, 6 \cdot 5^k] \subseteq A_k + A_k.$$

*Proof.* We argue by induction on  $k$ . The case  $k = 0$  is immediate, since

$$[2, 3] + [2, 3] = [4, 6].$$

Assume  $k \geq 1$  and the claim holds for  $k - 1$ . Put  $Q := 5^{k-1}$  and

$$I := [2Q, 3Q].$$

Then  $I \subseteq A_k$ : for  $k = 1$  this is just  $[2, 3]$ , while for  $k \geq 2$  we have

$$F_{k-1} = [2Q - 1, 3Q],$$

so indeed  $I \subseteq F_{k-1} \subseteq A_k$ .

Now the following intervals all lie in  $A_k + A_k$ :

$$\begin{aligned} I + I &= [4Q, 6Q], & I + \{c_k\} &= [6Q, 7Q], & I + B_k &= [7Q, 9Q - 1], \\ \{c_k\} + B_k &= [9Q, 10Q - 1], & B_k + B_k &= [10Q, 12Q - 2], & I + F_k &= [12Q - 1, 18Q], \\ B_k + F_k &= [15Q - 1, 21Q - 1], & F_k + F_k &= [20Q - 2, 30Q]. \end{aligned}$$

These intervals overlap consecutively, hence together they cover  $[4Q, 30Q]$ . By the inductive hypothesis,

$$[4, 6Q] \subseteq A_{k-1} + A_{k-1} \subseteq A_k + A_k.$$

Since  $4Q \leq 6Q$ , we conclude that

$$[4, 30Q] = [4, 6 \cdot 5^k] \subseteq A_k + A_k. \quad \square$$

**Lemma 4.2.** *For each  $k \geq 1$ , let*

$$J_k := [9 \cdot 5^{k-1}, 10 \cdot 5^{k-1} - 1].$$

*If  $n \in J_k$  and  $n = a + b$  with  $a, b \in A$ , then one of  $a, b$  is  $c_k$  and the other lies in  $B_k$ .*

*Proof.* Write  $Q := 5^{k-1}$ , so  $J_k = [9Q, 10Q - 1]$ . Let  $C_k$  be the union of  $[2, 3]$  and all stages strictly before  $k$ . The largest element of  $C_k$  is  $15 \cdot 5^{k-2} = 3Q$  (for  $k = 1$ , interpret  $C_1 = [2, 3]$ ). Every element from a stage later than  $k$  is at least

$$c_{k+1} = 20Q.$$

Thus no representation of an element of  $J_k$  can use a later stage.

If both summands lie in  $C_k$ , then their sum is at most  $6Q < 9Q$ . If one summand lies in  $C_k$  and the other lies in stage  $k$ , then the sum is either at most  $7Q$  (if the stage- $k$  point is  $c_k$ ), at most  $9Q - 1$  (if it lies in  $B_k$ ), or at least  $10Q + 1$  (if it lies in  $F_k$ ). Hence no representation of an element of  $J_k$  uses  $C_k$  at all.

Among sums formed entirely inside stage  $k$ , the only ones meeting  $J_k$  are

$$c_k + B_k = [9Q, 10Q - 1] = J_k.$$

Indeed,

$$2c_k = 8Q, \quad c_k + F_k \geq 14Q - 1, \quad B_k + B_k \subseteq [10Q, 12Q - 2],$$

and all other stage- $k$  sums are even larger. This proves the claim.  $\square$

*Proof of Theorem 1.4.* By Lemma 4.1,

$$[4, \infty) \subseteq A + A,$$

so  $A$  is an asymptotic basis of order 2.

Now let  $A = A_1 \sqcup A_2$  be any partition. One of the two colors contains infinitely many of the points  $c_k$ ; without loss of generality, suppose  $c_k \in A_1$  for infinitely many  $k$ . For each such  $k$ , Lemma 4.2 implies that every representation of any  $n \in J_k$  uses  $c_k$ , and hence

$$J_k \cap (A_2 + A_2) = \emptyset.$$

But

$$|J_k| = 5^{k-1} \rightarrow \infty,$$

so  $A_2 + A_2$  has arbitrarily long gaps and therefore is not syndetic. This proves the theorem.  $\square$

## 5 Remarks on the forum constructions

The proofs above are self-contained, but it is worth recording how they compare with the constructions described in the current forum thread [6].

**Remark 5.1.** The strict-density counterexample summarized in the thread also starts from the carry-free base-4 pair  $(B_1, B_2)$  and then adds sparse “fat” intervals [6]. Our proof uses the same underlying mechanism, but chooses the intervals  $I_k$  with length  $kN_k$  rather than a fixed multiple of  $N_k$ . This makes the second half of the contradiction especially transparent: along the scales  $2(k+1)N_k$ , the two monochromatic self-sumsets must have combined local density tending to 1.

**Remark 5.2.** For the bounded-gap question, the thread records a different DeepMind construction in terms of “forbidden zones”  $Z_k$  around scales  $P_k = 100^k$ , together with a single “oasis” point  $x_k = 10P_k$  in each zone [6]. The role of the oasis is analogous to the role of  $c_k$  in the APSSV construction: there is a long interval  $J_k$  of target sums such that every representation of a point of  $J_k$  must use that distinguished point. Whichever color misses it consequently leaves a monochromatic gap of length  $|J_k|$ . The APSSV version packages this rigidity into the stage triple  $(\{c_k\}, B_k, F_k)$  and has the advantage that the basis property follows from the short inductive covering argument in Lemma 4.1.

**Remark 5.3.** Nothing in this note addresses the lower-density variant. The point is only that the ambiguity in Erdős’s phrase “positive density” matters here: the natural-density and upper-density readings have different answers.

## References

- [1] B. Alexeev, M. Putterman, M. Sawhney, M. Sellke, and G. Valiant, *Short proofs in combinatorics and number theory*, arXiv:2603.29961, 2026.
- [2] T. F. Bloom, *Erdős Problem #741*, <https://www.erdosproblems.com/741>, accessed 2026-04-23.
- [3] P. Erdős, *Some problems in number theory, combinatorics and combinatorial geometry*, *Mathematica Pannonica* **5** (1994), 261–269.

- [4] P. Erdős and M. B. Nathanson, *Oscillations of bases for the natural numbers*, Proc. Amer. Math. Soc. **53** (1975), 253–258.
- [5] Google DeepMind, *Formal Conjectures*, file *ErdosProblems/741.lean*, <https://github.com/google-deepmind/formal-conjectures/blob/main/FormalConjectures/ErdosProblems/741.lean>, accessed 2026-04-23.
- [6] *Erdős Problem #741 — Discussion thread*, <https://www.erdosproblems.com/forum/thread/741>, accessed 2026-04-23.