

An elementary proof of Jin's theorem with a bound

Mauro Di Nasso

Dipartimento di Matematica
Università di Pisa, Italy
dinasso@dm.unipi.it

Submitted: Oct 29, 2013; Accepted: Mar 27, 2014; Published: May 22, 2014

Mathematics Subject Classifications: 11B13, 11B05, 11B75

Abstract

We present a short and self-contained proof of Jin's theorem about the piecewise syndeticity of difference sets which is entirely elementary, in the sense that no use is made of nonstandard analysis, ergodic theory, measure theory, ultrafilters, or other advanced tools. An explicit bound to the number of shifts that are needed to cover a thick set is provided. Precisely, we prove the following: If A and B are sets of integers having positive upper Banach densities a and b respectively, then there exists a finite set F of cardinality at most $1/ab$ such that $(A - B) + F$ covers arbitrarily long intervals.

Keywords: Sumsets; upper Banach density; piecewise syndetic set; Jin's theorem

1 Introduction

A beautiful result in additive and combinatorial number theory was proved in 2001 by R. Jin [8] with the tools of nonstandard analysis, namely the property that sumsets $A + B$ are piecewise syndetic whenever $A, B \subseteq \mathbb{N}$ have positive upper Banach density. (See §1 for definitions). Researchers showed interest in that result but were not comfortable with nonstandard analysis. Several papers recently appeared where Jin's theorem was re-proved and extended by using more familiar non-elementary tools; most notably, see the proofs in [4, 1] by means of ergodic theory, and the ultrafilter proof in [2]. Recently, in [5] I gave a different nonstandard proof of the result, where an explicit bound to the number of shifts of $A + B$ that are needed to cover arbitrarily large intervals is given.

Here, the underlying ideas of the main constructions used in [5] are suitably translated into "standard" terms; the goal is to provide a short and elementary proof of Jin's theorem, completely from scratch. (Here, by "elementary" we mean that no use is made

of nonstandard analysis, measure theory and ergodic theory, ultrafilters, or any other advanced tool; and by “from scratch” we mean that the arguments used are self-contained, with no references to other existing results in the literature.)

Notation: By \mathbb{N} we denote the set of *positive* integers. If not specified otherwise, lower-case letters a, b, c, x, y, z, \dots will denote integer numbers, with letters k, n, m reserved for natural numbers, and upper-case letters A, B, C, \dots will denote sets of integers. By $[a, b]$ we denote intervals of integers, *i.e.* $[a, b] = \{x \in \mathbb{Z} \mid a \leq x \leq b\}$; $A - B = \{a - b \mid a \in A, b \in B\}$ is the *difference set* of A and B ; and $A + B = \{a + b \mid a \in A, b \in B\}$ is the *sumset*. With obvious notation, we shall simply write $A - z$ (or $A + z$) to indicate the left-ward shift $A - \{z\}$ (or the right-ward shift $A + \{z\}$, respectively).

2 Jin’s theorem with a bound

Recall that a set A of integers is *thick* if it covers intervals of arbitrary length, *i.e.* if for every k there exists an interval $I = [x + 1, x + k]$ of length k such that $I \subseteq A$; A is *syndetic* if it has bounded gaps, *i.e.* if there exists k such that $A \cap I \neq \emptyset$ for every interval I of length k ; A is *piecewise syndetic* if it covers arbitrarily long intervals of a syndetic set, *i.e.* if $A = B \cap C$ where B is thick and C is syndetic. Useful characterizations are the following: A is syndetic if and only if $A + F = \mathbb{Z}$ for a suitable finite set F ; and A is piecewise syndetic if and only if $A + F$ is thick for a suitable finite set F .

A familiar notion in number theory is that of *upper asymptotic density* $\bar{d}(A)$ of sets of natural numbers $A \subseteq \mathbb{N}$:

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

The (upper) *Banach density* $\text{BD}(A)$ refines the upper density to sets of integers by considering arbitrary intervals instead of just initial intervals:

$$\text{BD}(A) = \lim_{n \rightarrow \infty} \left(\max_{x \in \mathbb{Z}} \frac{|A \cap [x + 1, x + n]|}{n} \right).$$

Remark that the above limit always exists, and in fact the following holds:¹

$$\text{BD}(A) = \inf_{n \in \mathbb{N}} \left(\max_{x \in \mathbb{Z}} \frac{|A \cap [x + 1, x + n]|}{n} \right).$$

Trivially $\text{BD}(A) \geq \bar{d}(A)$ for every $A \subseteq \mathbb{N}$; on the other hand, there exist sets $A \subseteq \mathbb{N}$ with $\bar{d}(A) = 0$ and $\text{BD}(A) = 1$. Basic properties of Banach density are the following: $\text{BD}(A) = 1$ if and only if A is thick; the family of sets with null Banach density is closed under finite unions; the Banach density is invariant under shifts, *i.e.* $\text{BD}(A - z) = \text{BD}(A)$. Although Banach density is not additive, the additivity property holds for families of shifts, *i.e.* $\text{BD}(\bigcup_{i=1}^k A - z_i) = k \cdot \text{BD}(A_i)$ whenever the shifts $A - z_i$ are pairwise disjoint.

¹ This is proved by a straight application of *Fekete’s lemma*: “Every sequence $\{a_n\}$ which is subadditive (*i.e.* $a_{n+m} \leq a_n + a_m$) converges to $\inf_n a_n/n$.” Indeed, it is readily seen that $a_n = \max_{x \in \mathbb{Z}} |A \cap [x + 1, x + n]|$ is subadditive.

The following general property of difference sets was first proved (but not published) by P. Erdős and Sárközy for sets of positive upper density.²

Proposition 1. *If $BD(A) > 0$ then $A - A$ is syndetic.*

The proof consists of a *pigeonhole principle* argument. Since $A - A$ is symmetric around 0, if it was not syndetic then one could pick a thick set $T \subseteq \mathbb{N}$ that does not meet $A - A$. Then one can construct a sequence $X = \{x_1 < x_2 < \dots\}$ such that $x_j - x_i \in T$ for all $j > i$. The shifts $A - x_i$ are pairwise disjoint because $(X - X) \cap (A - A) = \emptyset$, and so we would have $BD(\bigcup_{i=1}^k A - x_i) = \sum_{i=1}^k BD(A - x_i) = k \cdot BD(A) > 1$ for sufficiently large k , a contradiction.

Remark that the above property does *not* extend to the general case $A - B$; indeed, it is not hard to find thick sets A, B, C such that also their complements A^c, B^c, C^c are thick, and $A - B \subset C$. However, Jin's theorem showed that $A - B$ is in fact syndetic on a suitable thick set.

Theorem 2 (Jin, with a bound). *Let $A, B \subseteq \mathbb{Z}$ have positive upper Banach densities $BD(A) = \alpha$ and $BD(B) = \beta$, respectively. Then there exists a finite set F of cardinality $|F| \leq 1/\alpha\beta$ such that $(A - B) + F$ is thick (and hence $A - B$ is piecewise syndetic).*

Notice that the above result includes the original Jin's Theorem about sumsets $A + B$ of sets $A, B \subseteq \mathbb{N}$, because trivially $BD(B) = BD(-B)$. Before giving our elementary proof, let us show that $A - B$ is thick when the two sets have sufficiently large densities.

Proposition 3. *Let $A, B \subseteq \mathbb{Z}$, and assume that there exists α such that:*

- $\sup\{a_n - \alpha \cdot n \mid n \in \mathbb{N}\} = +\infty$, where $a_n = \max_{x \in \mathbb{Z}} |A \cap [x + 1, x + n]|$.
- $BD(B) \geq 1 - \alpha$.

Then $A - B$ is thick.

Proof. Given k , let m be such that $a_m - \alpha \cdot m > k$, and pick an interval I of length m with $a_m = |A \cap I|$. For every $i = 1, \dots, k$ we have that:

$$|(A - i) \cap I| \geq |A \cap I| - i > \alpha \cdot m + k - i \geq \alpha \cdot m.$$

Recall that $BD(B) = \inf_n b_n/n$ where $b_n = \max\{|B \cap J| \mid J \text{ interval of length } n\}$. Then, by the hypothesis on $BD(B)$, there exists an interval J of length m such that $|B \cap J| = b_m \geq (1 - \alpha) \cdot m$. Finally, pick x such that $x + J = I$. We claim that the interval $[x + 1, x + k] \subseteq A - B$; this will give the thesis because k is arbitrary. For every $i = 1, \dots, k$ one has

$$|(A - i) \cap I| + |(B + x) \cap I| = |(A - i) \cap I| + |B \cap J| > \alpha \cdot m + (1 - \alpha) \cdot m = m = |I|.$$

So, $(A - i) \cap (B + x) \cap I \neq \emptyset$, and there exist $a \in A$ and $b \in B$ such that $a - i = b + x$, and hence $x + i \in A - B$. \square

² This fact is mentioned in I.Z. Ruzsa's paper [11] of 1974. That result have been then generalized in several directions; see, e.g., [7] by H. Hegyvári where iterated difference sets are studied in the context of groups, and [10] by S. Révész in the context of locally compact abelian groups.

Notice that $\text{BD}(A) > \alpha$ implies $\sup\{a_n - \alpha \cdot n \mid n \in \mathbb{N}\} = +\infty$, which in turn implies $\text{BD}(A) \geq \alpha$; however, we remark that neither implication can be reversed. The fact that $A - B$ is thick whenever $\text{BD}(A) + \text{BD}(B) > 1$ was first proved by M. Beiglböck, V. Bergelson and A. Fish in [1]; in fact, their proof actually shows the (slightly) stronger property stated in the above proposition.

The results mentioned in this section are just a hint of the rich combinatorial structure of sumsets and sets of differences, whose investigation seems still far from being completed (see *e.g.* the monographies [14, 13]).

3 The elementary proof of Jin's theorem

We prove here our main Theorem 2, that was stated in the previous section. By the definition of upper Banach density, we can pick two sequences of integers $\langle x_n \mid n \in \mathbb{N} \rangle$ and $\langle y_n \mid n \in \mathbb{N} \rangle$ such that, if we let:

- $A_n = A \cap [x_n + 1, x_n + n^2]$
- $B_n = B \cap [y_n + 1, y_n + n]$

then $\lim_{n \rightarrow \infty} |A_n|/n^2 = \alpha$ and $\lim_{n \rightarrow \infty} |B_n|/n = \beta$. As the first step in our proof, for every n we shall find a suitable shift of A_n that meets B_n on a set whose relative density approaches $\alpha\beta$ as n goes to infinity. To this end, we shall use the following inequality, obtained as a simple application of the *pigeonhole principle*.³

Lemma 4. *Let $N, n \in \mathbb{N}$, $C \subseteq [1, N]$, $D \subseteq [1, n]$. Then there exists z such that*

$$\frac{|(C - z) \cap D|}{n} \geq \frac{|C|}{N} \cdot \frac{|D|}{n} - \frac{|D|}{N}.$$

Proof. Let $\vartheta : \mathbb{N} \rightarrow \{0, 1\}$ be the characteristic function of C . Then:

$$\begin{aligned} \sum_{x=1}^N \left(\sum_{d \in D} \vartheta(x+d) \right) &= \sum_{d \in D} \left(\sum_{x=1}^N \vartheta(x+d) \right) = \sum_{d \in D} |C \cap [d+1, N]| \\ &\geq \sum_{d \in D} (|C| - d) \geq |D| \cdot (|C| - n). \end{aligned}$$

By the *pigeonhole principle*, there must be at least one z such that

$$\frac{|(C - z) \cap D|}{n} = \frac{1}{n} \sum_{d \in D} \vartheta(z+d) \geq \frac{1}{n} \cdot \frac{|D| \cdot (|C| - n)}{N} = \frac{|C|}{N} \cdot \frac{|D|}{n} - \frac{|D|}{N}.$$

□

³ It is Lemma 5.1 of [5], but we re-prove it here to keep the paper self-contained.

For every n , apply the above lemma where $C = A_n - x_n \subseteq [1, n^2]$ and $D = B_n - y_n \subseteq [1, n]$. Clearly, $|C| = |A_n|$ and $|D| = |B_n|$. Then we can pick a suitable sequence $\langle z_n \mid n \in \mathbb{N} \rangle$ such that

$$\frac{|(A_n - x_n - z_n) \cap (B_n - y_n)|}{n} \geq \frac{|A_n|}{n^2} \cdot \frac{|B_n|}{n} - \frac{|B_n|}{n^2}.$$

Now let:

- $E_n = (A_n - x_n - z_n) \cap (B_n - y_n) \subseteq [1, n]$.

By passing the previous inequality to the limit, we obtain $\lim_{n \rightarrow \infty} |E_n|/n \geq \alpha\beta$.

Next, we shall use a property of difference sets of a sequence of subsets of $[1, n]$, which can be seen as a refinement of Proposition 1.

Lemma 5. *For $n \in \mathbb{N}$, let $E_n \subseteq [1, n]$. If $\lim_{n \rightarrow \infty} |E_n|/n = \gamma > 0$, then there exists a finite set F of cardinality $|F| \leq 1/\gamma$ such that for every m :*

$$(\star) \quad [1, m] \subseteq (E_n - E_n) + F \text{ for infinitely many } n.$$

Proof. We inductively define a finite increasing sequence $\sigma = \langle m_i \mid i = 1, \dots, k \rangle$. Set $m_1 = 0$. If property (\star) is satisfied by $F = \{0\}$, then put $\sigma = \langle m_1 \rangle$, and stop. Otherwise, let $m_2 \in \mathbb{N}$ be the least counterexample. So, $\Gamma_1 = \{n \mid [1, m_2 - 1] \subseteq (E_n - E_n) + m_1\}$ is infinite, but $\Lambda_1 = \{n \in \Gamma_1 \mid m_2 \in (E_n - E_n) + m_1\}$ is finite. Notice that for every $n \in \Gamma_1 \setminus \Lambda_1$ one has $(E_n + m_1) \cap (E_n + m_2) = \emptyset$. If for every $m \in \mathbb{N}$ the set of all $n \in \Gamma_1$ such that $[1, m] \subseteq \bigcup_{i=1}^2 (E_n - E_n + m_i)$ is infinite, then put $\sigma = \langle m_1, m_2 \rangle$ and stop. Otherwise, let $m_3 \in \mathbb{N}$ be the least counterexample. So, the set $\Gamma_2 = \{n \in \Gamma_1 \mid [1, m_2 - 1] \subseteq \bigcup_{i=1}^2 (E_n - E_n + m_i)\}$ is infinite, but $\Lambda_2 = \{n \in \Gamma_2 \mid m_3 \in \bigcup_{i=1}^2 (E_n - E_n) + m_i\}$ is finite. Notice that for every $n \in \Gamma_2 \setminus \Lambda_2$ one has $(E_n + m_i) \cap (E_n + m_3) = \emptyset$ for $i = 1, 2$. Iterate this process. We claim that we must stop at a step $k \leq 1/\gamma$. To see this, we show that whenever $m_1 < \dots < m_k$ are defined, one necessarily has $k \leq 1/\gamma$. This is trivial for $k = 1$, so let us assume $k \geq 2$.

The set $X = \Gamma_k \setminus (\Lambda_1 \cup \dots \cup \Lambda_k)$ is infinite, and since $X \subseteq \Gamma_i \setminus \Lambda_i$ for all i , the sets in the family $\{E_N + m_i \mid i = 1, \dots, k\}$ are pairwise disjoint for every $N \in X$. Now, every $E_N + m_i \subseteq [1, N + m_k]$, and so we obtain:

$$N + m_k \geq \left| \bigcup_{i=1}^k (E_N + m_i) \right| = \sum_{i=1}^k |E_N + m_i| = k \cdot |E_N| \implies \frac{|E_N|}{N} \leq \frac{1}{k} + \frac{m_k}{kN}.$$

By taking limits as $N \in X$ approaches infinity, one gets the desired inequality $\gamma \leq 1/k$. Finally, observe that, by the definition of $\sigma = \langle m_i \mid i = 1, \dots, k \rangle$, for every $n \in \Gamma_k$ and for every $m \in \mathbb{N}$ we have the inclusion $[1, m] \subseteq \bigcup_{i=1}^k (E_n - E_n + m_i)$. This shows that property (\star) is fulfilled by setting $F = \{m_1, \dots, m_k\}$. \square

By the above lemma where $\gamma = \alpha\beta > 0$, we can pick a finite set F of cardinality $|F| \leq 1/\alpha\beta$ and such that property (\star) is satisfied by the sets

$$E_n = (A_n - x_n - z_n) \cap (B_n - y_n).$$

So, for every m there exists n (in fact, infinitely many n) such that:

$$[1, m] \subseteq (E_n - E_n) + F \subseteq (A_n - x_n - z_n) - (B_n - y_n) + F \subseteq (A - B) + F - t_n,$$

and hence $[t_n + 1, t_n + m] \subseteq (A - B) + F$, where we denoted $t_n = x_n - y_n + z_n$. This completes the proof that $(A - B) + F$ is thick.

4 Final remarks and open problems

R. Jin's *nonstandard* proof in [8] was grounded on the following observation. Consider the set ${}^*\mathbb{N}$ of nonstandard natural numbers; then a set $A \subseteq \mathbb{N}$ is piecewise syndetic if and only its nonstandard extension ${}^*A \subseteq {}^*\mathbb{N}$ is *somewhere dense* (i.e. not nowhere dense) in a suitable order-topology on ${}^*\mathbb{N}$, as determined by the *additive cut* of (finite) natural numbers. A few years later, R. Jin's himself [9] elaborated a "standard" proof of his theorem, where the nonstandard reasonings were directly translated into the language of elementary combinatorics. Unfortunately, in this way "certain degree of intuition and motivation are lost" (*ibid.*), also because of the partly topological nature of the original arguments.

In a note with comments on a previous version of this paper, I.Z. Ruzsa pointed out that by similar arguments as the ones used here, an elementary proof of Jin's theorem can also be obtained by utilizing results from his 1978 paper [12]. Precisely, by referring to terminology and numbered theorems as presented there, one can proceed as follows.⁴ Given $A, B \subseteq \mathbb{Z}$ with positive upper Banach densities $d^*(A) = \alpha$ and $d^*(B) = \beta$, take the corresponding *homogeneous systems* $h_1(A)$ and $h_1(B)$. By Theorem 3, their densities as homogeneous systems – that we denote here by δ to avoid ambiguities – are the same as the original sets: $\delta(h_1(A)) = \alpha$ and $\delta(h_1(B)) = \beta$. Now consider the homogeneous system $H = h_1(A) \cap h_1(B)$ given by the intersection. By Theorem 5, $\delta(H) \geq \alpha \cdot \beta$ and so, by Theorem 4, there exists a set E with asymptotic density $d(E) \geq \delta(H)$, and such that $h_1(E) \subseteq H$. By a *pigeonhole principle* argument, it is shown that the positive density of E implies the existence of a finite set F with cardinality $|F| \leq 1/d(E) \leq 1/\alpha\beta$ and such that $(E - E) + F = \mathbb{Z}$. Now notice that, since $h_1(E) \subseteq h_1(A)$ and $h_1(E) \subseteq h_1(B)$, for every finite $G \subset E$ there exist x, y such that $x + G \subset A$ and $y + G \subset B$, and hence $(x + y) + (G - G) + F \subset (A - B) + F$. Since $(E - E) + F = \mathbb{Z}$, it follows that every finite set has a shift included in $(A - B) + F$, which is therefore thick.

We close this paper by posing a couple of open questions. Recall that a family \mathcal{F} of sets of integers is *partition regular* if it is closed under supersets and it satisfies the Ramsey property that $A \cup B \in \mathcal{F}$ implies $A \in \mathcal{F}$ or $B \in \mathcal{F}$ (three relevant examples are

⁴ The contents of I.Z. Ruzsa's note are reproduced here under his permission.

given by the infinite sets, by the sets with positive upper density, and by the piecewise syndetic sets). We remark that, by exactly the same proof, the property of Lemma 5 also holds if one replaces (\star) with

$$(\star)_{\mathcal{F}} \quad \{n \in \mathbb{N} \mid [1, m] \subseteq (E_n - E_n) + F\} \in \mathcal{F}$$

where \mathcal{F} is any partition regular family of infinite sets. Can one derive a stronger version of Jin's theorem by using that stronger version of Lemma 5?

In §1, we recalled the property that $A - B$ is thick whenever $\text{BD}(A) + \text{BD}(B) > 1$. Can one combine this fact with similar arguments as the ones presented in this paper, and prove interesting structural properties involving the three difference sets $A - B$, $A - C$, $B - C$, under the assumption that $\text{BD}(A) + \text{BD}(B) + \text{BD}(C) > 1$?

References

- [1] M. Beiglböck, V. Bergelson, and A. Fish. Sumset phenomenon in countable amenable groups, *Adv. Math.*, 223:416–432, 2010.
- [2] M. Beiglböck. An ultrafilter approach to Jin's theorem. *Israel J. Math.*, 185:369–374, 2011.
- [3] M. Kiyomi, T. Saitoh, and R. Uehara. Reconstruction of interval graphs. In *Computing and Combinatorics*, volume 5609 of *Lecture Notes in Comput. Sci.*, pages 106–115. Springer, 2009.
- [4] V. Bergelson, H. Fürstenberg, and B. Weiss. Piece-wise sets of integers and combinatorial number theory. In *Topics in Discrete Mathematics*, volume 26 of *Algorithms Combin.*, pages 13–37. Springer, 2006.
- [5] M. Di Nasso. Embeddability properties of difference sets. *Integers*. To appear.
- [6] H. Fürstenberg. *Recurrence in Ergodic Theory and Combinatorial Number Theory*. Princeton University Press, 1981.
- [7] N. Hegyvári. On iterated difference sets in groups. *Period. Math. Hungar.*, 43:105–110, 2001.
- [8] R. Jin. The sunset phenomenon. *Proc. Amer. Math. Soc.*, 130:855–861, 2002.
- [9] R. Jin. Standardizing nonstandard methods for upper Banach density problems. In *Unusual Applications of Number Theory* (M. Nathanson, ed.), volume 64 of *DIMACS Series*, pages 109–124, 2004.
- [10] S. Révész. On asymptotic uniform upper density in locally compact abelian groups. In *Real Analysis Exchange*, Summer Symposium 2011, pages 24–31.
- [11] I.Z. Ruzsa. On difference-sequences. *Acta Arith.*, XXV:151–157, 1974.
- [12] I.Z. Ruzsa. On difference sets. *Studia Sci. Math. Hungar.*, 13:319–326, 1978.
- [13] I.Z. Ruzsa. Sumsets and structure. Part II of *Combinatorial Number Theory and Additive Group Theory* (A. Geroldinger and I.Z. Ruzsa, eds.), pages 87–210. Birkhäuser, 2009.
- [14] T. Tao and V.H. Vu. *Additive Combinatorics*. Cambridge University Press, 2006.