FREIMAN
ON BLOCKS OF CONSECUTIVE INTEGERS IN SMALL SUMSETS

AN EXPOSITION BY VSEVOLOD F. LEV

Abstract. This is a concise exposition of a year 2009 paper by Freiman on the existence of a long block of consecutive integers in a small sumset of a finite integer set.

As shown in [L97], if a A is a finite set of integers with $\operatorname{diam}(A) < \frac{3}{2}|A| - 1$, then the sumset $2A$ contains a block of $2|A| - 1$ consecutive integers. The main result of [F09] gives a slightly stronger conclusion under a substantially weaker assumption.

Theorem (Freiman [F09])**.** *Suppose that A is a finite set of integers with $\min A = 0$, $l := \max A > 0$, and $\gcd(A) = 1$, and write $n := |A|$. Let $b := \max\{g \in \mathbb{Z} \colon g < l,\ g \notin 2A\}$ and $c := \min\{g \in \mathbb{Z} \colon g > l,\ g \notin 2A\}$, so that $J := [b+1, c-1] \subseteq 2A$. If $|2A| \le 3n - 4$, then $c > b + l$ and*

$$|J| \ge 2n - 1 + 2|(b, c-l) \setminus A|.$$

Corollary. *Suppose that A is a finite set of integers, not contained in an arithmetic progression with the difference greater than 1. If $|2A| \le 3|A| - 4$, then $2A$ contains a block of at least $2n - 1$ consecutive integers.*

The proof of Freiman's theorem is presented below.

By a *gap* is a finite set of integers B we mean an integer $g \in [\min B, \max B] \setminus B$.

We consider (ordered) pairs of the form $(g, g+l)$ with $g \in [0, l]$. We say that the pair $(g, g+l)$ is *left-empty* if $g \notin 2A$ while $g+l \in 2A$, and that it is *right-empty* if $g \in 2A$ while $g + l \notin 2A$; such pairs will be collectively referred to as *half-empty.*† As an immediate consequence of Freiman's "$(3n-4)$-theorem", the assumption $|2A| \le 3n - 4$ implies that all residue classes modulo l are represented in the sumset $2A$. (The theorem readily shows that $l \le 2n - 4$, whence the canonical image of A in the group $\mathbb{Z}/l\mathbb{Z}$ contains more than half the elements of the group.) As a result, for any $g \in [0, l]$, either $\{g, g+l\} \subseteq 2A$, or the pair $(g, g+l)$ is half-empty. Moreover, the number of half-empty pairs is equal to the number of gaps in $2A$, which is $2l + 1 - |2A|$. Consequently, the number of those $g \in [0, l]$ with $\{g, g+l\} \subseteq 2A$ is $|2A| - l$, and so the number of gaps $g \in [0, l] \setminus A$ with $\{g, g+l\} \subseteq 2A$ is $|2A| - l - n$:

$$\big|\{g \in [0, l] \colon g \notin A,\ \{g, g+l\} \subseteq 2A\}\big| = |2A| - l - n. \qquad (1)$$

†Our terminology reflects the practice to visualize elements of $2A$ as filled circles, and gaps in $2A$ as empty circles. It differs from Freiman's original terminology.

Freiman's critical observation is that every left-empty pair lies to the left of every right-empty pair in the sense that if $(g_l, g_l + l)$ is left-empty and $(g_r, g_r + l)$ is right-empty, then $g_l < g_r$. To prove this, assume for a contradiction that $g_l > g_r$ and that $g_r, g_l + l \in 2A$ while $g_l, g_r + l \notin 2A$. Without loss of generality, we can further assume that the interval $(g_r + 1, g_l - 1)$ does not contain any elements from half-empty pairs; that is, $\{g, g + l\} \subseteq 2A$ for every integer $g \in (g_r, g_l)$. By the pigeonhole principle, from $g_l \notin 2A$ we get

$$|A \cap [0, g_l]| \le \frac{1}{2}\,(g_l + 1) \tag{2}$$

and, similarly, from $g_r + l \notin 2A$,

$$|A \cap [g_r, l]| \le \frac{1}{2}\,(l + 1 - g_r). \tag{3}$$

Adding up these two estimates we obtain

$$n + |A \cap [g_r, g_l]| \le \frac{1}{2}\,(l + (g_l - g_r)) + 1;$$

equivalently,

$$|[g_r, g_l] \setminus A| \ge n - \frac{1}{2}\,l + \frac{1}{2}\,(g_l - g_r).$$

Combining this with the trivial estimate

$$|[g_r, g_l] \setminus A| \le g_l - g_r + 1$$

gives

$$|[g_r, g_l] \setminus A| \ge 2n - l - 1.$$

Comparing this with (1) leads to

$$|2A| - l - n \ge 2n - l - 1,$$

contradicting the assumption $|2A| \le 3n - 4$.

Having shown that every left-empty pair lies to the left of every right-empty pair, we can complete the proof of the theorem. Since $(b, b + l)$ is left-empty and $(c - l, c)$ is right-empty, we have $b < c - l$; that is, $c > b + l$. Notice that this inequality already shows that $2A$ contains a block of at least l consecutive integers. To boost it further, we re-use (2) and (3) with $g_l = b$ and $g_r = c - l$, respectively, to get

$$|A \cap [0, b]| \le \frac{1}{2}\,(b + 1)$$

and

$$|A \cap [c - l, l]| \le \frac{1}{2}\,(2l - c + 1).$$

Also, in a trivial way,

$$|A \cap (b, c - l)| = c - l - b - 1 - |(b, c - l) \setminus A|.$$

Taking the sum yields

$$n \le \frac{1}{2}\,(c - b) - |(b, c - l) \setminus A|,$$

which is equivalent to the assertion of the theorem.

References

[F09] G.A. Freiman, Inverse Additive Number Theory. XI. Long arithmetic progressions in sets with small sumsets, *Acta Arithmetica* **137** (4) (2009), 325–331.

[L97] V. Lev, Optimal Representations by Sumsets and Subset Sums, *Journal of Number Theory* **62** (1997), 127–143.