ACTA ARITHMETICA
119.2 (2005)

Sum-sets of small upper density

by

GUILLAUME BORDES (Bordeaux)

1. Introduction. Let $A \subseteq \mathbb{N}$ be an infinite set of non-negative integers. For $y > x > 0$, we put

$$A(x) := |A \cap [0; x]|, \qquad A(x, y) := |A \cap [x; y]|.$$

We define the *lower asymptotic density* $\underline{d}(A)$ and *upper asymptotic density* $\overline{d}(A)$ by

$$\underline{d}(A) := \liminf_{x \to \infty} \frac{A(x)}{x}, \qquad \overline{d}(A) := \limsup_{x \to \infty} \frac{A(x)}{x}.$$

Unless explicitly stated otherwise, we assume that

$$0 \in A, \quad \gcd(A) = 1. \tag{1}$$

We define the sum $X + Y$ of two sets $X, Y \subset \mathbb{R}$ by

$$X + Y = \{x + y \mid x \in X,\ y \in Y\}.$$

Inverse additive theory describes sets A with "small" sum-set $A + A$. Say, one may ask about sets A of positive lower or upper density with small quotient $\underline{d}(A + A)/\underline{d}(A)$ or $\overline{d}(A + A)/\overline{d}(A)$. For example, let $N \geq 3$ be an integer. Then for the set $A = \{0, 1\} + N\mathbb{N}$ we have

$$\underline{d}(A) = \overline{d}(A) = 2/N, \qquad \underline{d}(A + A) = \overline{d}(A + A) = 3/N,$$

so that the above-mentioned quotients are both equal to $3/2$. (As we shall see in a while, this is the minimal possible value under the assumption (1).)

Kneser [7, 4] gave a complete description of sets A satisfying $\underline{d}(A + A) < 2\underline{d}(A)$. In brief, he showed that A should be "approximately" of the form $K + N\mathbb{N}$, where N is a positive integer and K is a set of residues mod N.

Among other things, Kneser's theorem implies that $\underline{d}(A + A) \geq \frac{3}{2}\,\underline{d}(A)$ when A satisfies (1), and the equality $\underline{d}(A + A) = \frac{3}{2}\,\underline{d}(A)$ is possible only with $|K| = 2$.

2000 *Mathematics Subject Classification*: Primary 11P70.

Extending Kneser's results to upper density seems to be a rather difficult problem. The following example, due to Jin [5], shows that this time one cannot get away with sets of the type $K + N\mathbb{N}$.

EXAMPLE 1.1. Let α be a real number satisfying $0 < \alpha < 1/2$. Let $(T_n)_{n\geq 1}$ be an increasing sequence of positive integers such that $\lim T_{n+1}/T_n = \infty$. Then the set

$$A = \mathbb{N} \cap \bigcup_{n=1}^{\infty} [\lceil (1-\alpha)T_n \rceil, T_n]$$

satisfies $\overline{d}(A) = \alpha$ and $\overline{d}(A + A) = \frac{3}{2}\alpha$.

In what follows, we use the notation

$$\alpha = \overline{d}(A), \qquad \gamma = \overline{d}(A + A).$$

We assume that $0 < \alpha \leq 1/2$ and we put $\sigma = \gamma/\alpha$.

It is not difficult to show that $\sigma \geq 3/2$ (see Lemma 2.1 below), but the structure of sets A with $\sigma = 3/2$ was only recently determined by Jin [6]. He proved that a set A with $\sigma = 3/2$ is "similar" either to $K + N\mathbb{N}$ with $|K| = 2$, or to the set from Example 1.1. For $\sigma > 3/2$ the problem is open.

In the present article we determine the structure of sets A with $3/2 \leq \sigma < 5/3$ subject to the additional assumption $\alpha < \alpha_0$, where α_0 is a small absolute constant. For $\sigma = 3/2$ our result is covered by that of Jin, but for $3/2 < \sigma < 5/3$ our result is new.

Now, we can formulate the main result of this article.

THEOREM 1.2. *There exists a positive absolute constant* α_0 *such that the following holds. Let* A *be a set of non-negative integers such that* $0 \in A$ *and* $\gcd(A) = 1$. *Put* $\alpha = \overline{d}(A)$ *and* $\gamma = \overline{d}(A + A)$. *Assume that* $0 < \alpha = \overline{d}(A) \leq \alpha_0$ *and that*

$$\gamma = \sigma\alpha, \quad \textit{where } 3/2 \leq \sigma < 5/3.$$

Then we have one of the following cases:

1. Non-archimedean case: *there exist two positive integers* N *and* t *with* $\gcd(N, t) = 1$ *such that* $A \subseteq \{0, t\} + N\mathbb{N}$, *and*

$$\alpha \geq \frac{6}{(4\sigma - 3)N}.$$

2. Archimedean case: *there exists an increasing sequence* $(y_j)_{j\geq 1}$ *of integers with*

$$\lim_{j\to\infty} \frac{A(y_j)}{y_j} = \alpha,$$

and two sequences $(b_j)_{j\geq 1}$ *and* $(t_j)_{j\geq 1}$ *with* $0 \leq b_j \leq t_j \leq y_j$ *such that*, *if we define*

$$\lambda_j := \frac{b_j}{y_j - t_j}, \qquad r_j := \frac{A(t_j, y_j)}{y_j - t_j + 1},$$

then $A(b_j, t_j) = 0$ *for all* $j \geq 1$ *and*

$$\lim_{j\to\infty} \lambda_j = \lambda, \qquad \lim_{j\to\infty} r_j = r$$

with

$$\lambda \leq \frac{2\sigma - 3}{2\sigma - 2}\left(\frac{1}{2\sigma - 2} - \alpha\right)^{-1}, \qquad r \geq \left(\frac{1}{2\sigma - 2} + \lambda\left(\frac{1}{2\sigma - 2} - \alpha\right)\right).$$

EXAMPLE 1.3. We cannot extend Theorem 1.2 to the case $\overline{d}(A+A) = \frac{5}{3}\overline{d}(A)$. It suffices to consider the set $A := N\mathbb{N} \cup (1 + 2N\mathbb{N})$ which satisfies this condition and for which $\alpha = 3/2N$. Putting $\sigma = 5/3$ in Theorem 1.2 would give $\alpha \geq 18/11N > 3/2N$.

The following example proves that the lower bound obtained in the non-archimedean case of Theorem 1.2 cannot be refined.

EXAMPLE 1.4. Fix $3/2 \leq \sigma < 5/3$. Let $(T_n)_{n\geq 1}$ be an increasing sequence of positive integers such that $\lim_{n\to\infty} T_{n+1}/T_n = \infty$ and set

$$E := \bigcup_{n=1}^{\infty} [\lceil (1-\alpha')T_n \rceil; T_n],$$

where $\alpha' = 3/(4\sigma - 3)$. Let N be a sufficiently large positive integer and

$$A := NE \cup (1 + NE).$$

We can verify that

$$\alpha = \frac{6}{(4\sigma - 3)N} < \alpha_0, \qquad \gamma = 3\,\frac{1+\alpha'}{2N} = \frac{6\sigma}{(4\sigma - 3)N}.$$

2. General results in additive number theory. Before proving the main theorem, let us show why 3/2 is a lower bound for the quotient σ.

LEMMA 2.1. *Let* A *be a set of non-negative integers. Suppose that* $0 \in A$ *and* $\gcd(A) = 1$. *Then* $\gamma \geq \frac{3}{2}\alpha$.

We can easily deduce the lemma from the following

THEOREM 2.2 ([9, p. 23]). *Let* $k \geq 3$ *be an integer. Let* $A = \{a_0, a_1, \ldots \ldots, a_{k-1}\}$ *be a set of non-negative integers such that*

$$0 = a_0 < a_1 < \cdots < a_{k-1}, \qquad \gcd(A) = 1.$$

If $a_{k-1} \geq 2k - 3$, *then* $|A + A| \geq 3k - 3$.

Proof of Lemma 2.1. Since $\overline{d}(A)=\alpha$, there exists an increasing sequence $(y_j)_{j\geq 1}$ of integers such that, for all $\varepsilon>0$, if we define $A_j := A\cap[0;y_j]$ and assume j sufficiently large, we have

$$\alpha-\varepsilon<|A_j|/y_j<\alpha+\varepsilon, \qquad \gcd(A_j)=1.$$

In what follows, we will assume that $y_j\in A_j$.

Under the hypothesis $\alpha<1/2$, we can see that A_j satisfies the hypothesis of Theorem 2.2. Thus,

$$|A_j+A_j|\geq 3|A_j|-3,$$

and therefore,

$$\frac{(A+A)(2y_j)}{2y_j}\geq\frac{|A_j+A_j|}{2y_j}\geq\frac{3}{2}\,\frac{|A_j|}{y_j}-\frac{3}{2y_j}\geq\frac{3}{2}\,\alpha-2\varepsilon.$$

This yields $\overline{d}(A+A)\geq\frac{3}{2}\alpha$ and concludes the proof. ■

In the rest of this section, we give some general results in additive number theory, to be used in the next section.

Let A be a finite set of integers. It is easy to see that $|A+A|\geq 2|A|-1$, and $|A+A|=2|A|-1$ if and only if A is an arithmetical progression.

Freiman [9, p. 21] generalized this fact.

THEOREM 2.3 (Freiman). *Let A be a finite set of non-negative integers such that $|A|\geq 3$ and $\min(A)=0$. Denote by a_k the greatest element of A. If*

$$a_k\leq 2|A|-3,$$

then

$$|2A|\geq|A|+a_k.$$

This result has been generalized to distinct sets by V. F. Lev and P. Y. Smeliansky in [8] and was improved by Y. V. Stanchescu in [10]. We will use the following version:

THEOREM 2.4 (Lev, Smeliansky). *Let A and B be two finite sets of non-negative integers such that $0\in A\cap B$. Denote by $l(A):=\max(A)-\min(A)$ the length of A and by $h(A):=l(A)-|A|+1$ the number of holes in A. If*

$$\max(l(A),l(B))\leq|A|+|B|-3,$$

then

$$|A+B|\geq(|A|+|B|-1)+\max(h(A),h(B)).$$

Now, let us introduce some notions taken from [2].

DEFINITION 2.5. Let A and B be two abelian groups and $K\subset A$, $L\subset B$. A map $\varphi:K\to L$ is said to be a *Freiman homomorphism* or an F_2-*homomorphism* if, for all $(x,y,x',y')\in K^4$, we have

$$x+y=x'+y' \;\Rightarrow\; \varphi(x)+\varphi(y)=\varphi(x')+\varphi(y').$$

Such a φ is said to be an F_2-*isomorphism* if it is invertible and if φ^{-1} is also an F_2-homomorphism.

REMARK 2.6. In what follows, we will use some affine maps in $\mathbb{Z}^2$ which are clearly F_2-isomorphisms.

The following proposition is clear:

PROPOSITION 2.7. *An F_2-isomorphism $\varphi : K \to L$ induces a bijective map $K + K \to L + L$.*

REMARK 2.8. Similarly, for any positive integer i we can define the notion of F_i-homomorphism. We say that $\varphi : K \to L$ is an F_i-*homomorphism* if for all $(x_1, \ldots, x_i, x'_1, \ldots, x'_i) \in K^{2i}$,

$$x_1 + \cdots + x_i = x'_1 + \cdots + x'_i \;\Rightarrow\; \varphi(x_1) + \cdots + \varphi(x_i) = \varphi(x'_1) + \cdots + \varphi(x'_i).$$

Clearly, an F_i-homomorphism is an F_2-homomorphism for any $i \geq 2$.

DEFINITION 2.9. A subset P of an abelian group is called a *generalized arithmetical progression* of dimension m if it can be written as

$$\begin{aligned} P &= P(x_0; x_1, \ldots, x_m; b_1, \ldots, b_m) \\ &= \{x_0 + \beta_1 x_1 + \cdots + \beta_m x_m : \beta_i = 0, \ldots, b_i - 1\} \end{aligned} \tag{2}$$

where $x_0, \ldots, x_m$ are elements of the group and $b_1, \ldots, b_m$ are positive integers.

We say that P is an F_2-*progression* if the map

$$\begin{aligned} \theta : \{0, \ldots, b_1 - 1\} \times \cdots \times \{0, \ldots, b_m - 1\} \subset \mathbb{Z}^m \to P, \\ (\beta_1, \ldots, \beta_m) \mapsto x_0 + \beta_1 x_1 + \cdots + \beta_m x_m, \end{aligned}$$

is an F_2-isomorphism.

We will heavily use the following fundamental theorem due to G. Freiman whose proof can be found in [2] and whose version below is taken from [1]:

THEOREM 2.10 (Freiman). *Let σ be a positive real number, and A a finite set of non-negative integers such that $0 \in A$ and $|A| > k(\sigma)$ where $k = k(\sigma)$ is a fixed constant depending only on σ. If*

$$|A + A| \leq \sigma |A|,$$

then A is a subset of an F_2-progression

$$P = P(0; x_1, \ldots, x_m; b_1, \ldots, b_m)$$

of dimension $m \leq \lfloor \sigma - 1 \rfloor$ and whose length is bounded from above: $|P| \leq C_1(\sigma)|A|$.

Furthermore, if $b_1 \leq \cdots \leq b_m$, then

$$i > \lfloor \log_2 \sigma \rfloor \;\Rightarrow\; b_i \leq C_2(\sigma).$$

Here $C_1(\sigma)$ and $C_2(\sigma)$ are constants depending only on σ.

Our strategy of proof is simple. First, we are going to transform the *infinite* problem into a *finite* one. Then we will use Theorem 2.10 to obtain the structure of finite sets. Finally, we will come back to the set A using asymptotic arguments.

In the following, Theorem 2.10 will be used with $\sigma < 4$ so that it will give rise to F_2-progressions of dimension at most 2. Then, in view of Definition 2.9, it will be natural to use results concerning addition of sets in $\mathbb{Z}^2$, particularly the following one whose proof can be found in [3, p. 28]:

THEOREM 2.11 (Freiman). *Let $A \subset \mathbb{Z}^2$ be a set of at least twelve elements not on the same line. Assume that*

$$|A+A| < \frac{10}{3}|A| - 5.$$

Then A is contained in a set F_2-isomorphic to

$$A^0 = \{(0,0),(0,1),\ldots,(0,l_1-1)\} \cup \{(1,0),(1,1),\ldots,(1,l_2-1)\}$$

with $l_1, l_2 \geq 1$ and $l_1 + l_2 = |A+A| - 2|A| + 3$.

3. Proof of the main theorem. With a view to use the theorems of the previous section, let us transform our problem into a problem on finite sets.

Let $\varepsilon > 0$. We can choose $y_1 \in \mathbb{N}$ sufficiently large and a strictly increasing sequence $(y_j)_{j\geq 1}$ of positive integers such that for all j,

$$(A+A)(2y_j) \leq (\gamma+\varepsilon)\cdot 2y_j, \qquad (\alpha-\varepsilon)y_j \leq A(y_j) \leq (\alpha+\varepsilon)y_j.$$

We will use the notation

$$A_j := \{a \in A : a \leq y_j\}.$$

In what follows, all the notations will depend on the sequence $(y_j)_{j\geq 1}$. Every change of the sequence will naturally change the sets A_j and all related objects. We will denote by $O(\varepsilon)$ any positive function of ε bounded above by $C\varepsilon$ where C is a constant only depending on the set A.

Now, we are able to determine the structure of the sets A_j. We have

$$\begin{aligned}\frac{|A_j+A_j|}{|A_j|} &= \frac{|A_j+A_j|}{2y_j}\cdot 2\cdot\frac{y_j}{|A_j|}\\ &\leq \frac{(A+A)(2y_j)}{2y_j}\cdot 2\cdot\frac{y_j}{|A_j|}\\ &\leq 2\cdot\frac{\gamma+\varepsilon}{\alpha-\varepsilon} \leq 2\sigma+\varepsilon' < 4,\end{aligned} \tag{3}$$

where $\varepsilon' = O(\varepsilon)$. Thus, for ε sufficiently small, we can apply the fundamental Theorem 2.10 of Freiman to the sets A_j. By a simple calculation, we obtain $m \leq 2$ and $b_2 \leq C_2$. First, we are going to exclude the case where A_j is a subset of an arithmetical progression of dimension $m = 1$ for infinitely many values of j.

Suppose this is the case. Then, for j sufficiently large, $A_j \subseteq P_j$ where P_j is an arithmetical progression of difference 1 (for $\gcd(A) = 1$) and first term 0. We can assume it has minimal length. Then, by Theorem 2.10 and since $\{0, y_j\} \subseteq P_j$, we have

$$|P_j| \geq y_j, \tag{4}$$

$$|P_j| \leq C_1|A_j|. \tag{5}$$

Now we combine (4) and (5) to find a lower bound for α:

$$\begin{aligned} \alpha &\geq \frac{1}{\sigma}\gamma \geq \frac{1}{\sigma}\left(\frac{|A_j + A_j|}{2y_j} - \varepsilon\right) \geq \frac{1}{\sigma}\left(\frac{2|A_j| - 1}{2y_j} - \varepsilon\right) \\ &\geq \frac{1}{\sigma}\left(\frac{1}{2y_j}\left(\frac{2y_j}{C_1} - 1\right) - \varepsilon\right) \geq \alpha_0, \end{aligned} \tag{6}$$

for an absolute constant α_0 (remember that ε can be chosen sufficiently small). Thus, we can exclude this case under hypothesis $\alpha < \alpha_0$ of Theorem 1.2.

Remark 3.1. The value of C_1 (one can find an estimate in [2]) implies a very small value for the bound α_0. What happens for $\alpha > \alpha_0$ is an open question.

Thus, for infinitely many integers j, the set A_j is a subset of an arithmetical progression of dimension $m = 2$. By extracting a subsequence, we can assume that this is the case for all A_j. Then, for all $j \geq 1$, there is an F_2-isomorphism θ_j between a subset of $\mathbb{Z}^2$ and A_j (see Definition 2.9). By Proposition 2.7, the sets $\theta_j^{-1}(A_j)$ satisfy the inequality

$$|\theta_j^{-1}(A_j) + \theta_j^{-1}(A_j)| \leq (2\sigma + \varepsilon')|\theta_j^{-1}(A_j)|.$$

At this point, using the assumption $\sigma < 5/3$, we can apply Theorem 2.11 to $\theta_j^{-1}(A_j)$. Composing isomorphisms, we see that, for all $j \geq 1$, there exists an F_2-isomorphism $\varphi_j : \mathbb{Z}^2 \to \mathbb{N}$ such that $A_j \subseteq \varphi_j(A_j^0)$ where $A_j^0 = \{(0,0), (0,1), \ldots, (0, l_{1,j} - 1)\} \cup \{(1,0), (1,1), \ldots, (1, l_{2,j} - 1)\}$. Combining, if necessary, those isomorphisms with suitable affine maps, we can assume that $\varphi_j((0,0)) \in A_j$ and $\varphi_j((1,0)) \in A_j$. Furthermore, we have $l_{1,j} + l_{2,j} = |A_j + A_j| - 2|A_j| + 3$.

Notice that the number of elements of $\varphi^{-1}(A_j)$ in each line cannot be bounded, since otherwise, for all $\varepsilon > 0$, we could obtain $\overline{d}(A + A) > (2 - \varepsilon)\overline{d}(A)$ by considering the sequence $(A + A)(y_j)/y_j$.

We set $d_{1,j} := \varphi_j((1,0)) - \varphi_j((0,0))$ and $d_{2,j} := \varphi_j((0,1)) - \varphi_j((0,0))$.

Then we can give the explicit F_2-isomorphism

$$\varphi_j : \mathbb{Z} \times \{0,1\} \to \mathbb{N}, \qquad (x,y) \mapsto a_j + x d_{1,j} + y d_{2,j}, \tag{7}$$

where $a_j = \varphi_j((0,0))$.

Since $A \subseteq \mathbb{N}$, the number $d_{1,j}$ has to be positive for infinitely many values of j which we again extract. We can also assume, by switching the lines if necessary, that the differences $d_{2,j}$ are positive.

Lemma 3.2. *The sequence $(d_{1,j})_{j\geq 1}$ is bounded.*

Proof. Assume the contrary. Then there exists an index j such that $A(d_{1,j}) > 3$ and, consequently, there exist distinct $a, b \in A \cap [0; d_{1,j}]$ such that $\varphi_j^{-1}(a)$ and $\varphi_j^{-1}(b)$ lie on the same line. We deduce from (7) that $|b-a| = kd_{1,j}$ where k is a positive integer. This is impossible since $|b-a| < d_{1,j}$. ∎

Since the sequence $(d_{1,j})_{j\geq 1}$ is bounded, there exists a positive integer N such that $d_{1,j} = N$ for infinitely many j. We choose the largest N with this property, and, again extracting a subsequence, we assume that $d_{1,j} = N$ for all j.

3.1. *The non-archimedean case.* In this case, we assume $N > 1$. We show that the sequence $(d_{2,j})_{j\geq 1}$ can then be supposed to be constant.

Lemma 3.3. *There exist a positive integer t and a sequence $(y_j)_{j\geq 1}$ such that $d_{2,j} = t$ for all $j \geq 1$.*

Proof. Each A_j is included in two residue classes mod N. Since those sets satisfy $A_j \subseteq A_k$ for $j < k$, the whole set A is included in two residue classes. If we denote by t the smallest term of the part of A not congruent to 0 mod N, we can choose, for each $j \geq 1$, the isomorphism φ_j such that $\varphi_j((0,0)) = 0$ and $\varphi_j((1,0)) = t$. ∎

Hence, we can assume that $d_{2,j} = t$ for all $j \geq 1$ and we can exhibit an F_2-isomorphism φ between $\mathbb{Z}^2$ and $\mathbb{N}$ such that $\varphi_{|A_j} = \varphi_j$:

$$\varphi : \mathbb{Z} \times \{0,1\} \to \mathbb{N}, \qquad (x,y) \mapsto xN + yt.$$

By hypothesis (1), we must have $\gcd(t, N) = 1$ and A is included in two residue classes mod N which we denote by B and C:

$$B = \{a \in A : a \equiv 0 \bmod N\}, \qquad C = \{a \in A : a \equiv t \bmod N\}.$$

We define $B_j := B(y_j)$ and $C_j := C(y_j)$ and we assume, choosing y_1 sufficiently large, that those sets are non-empty. We define $t_0 := \min(C)$, $b_j := \max(B_j)$ and $c_j := \max(C_j)$. We may assume that $b_j = y_j$, replacing if necessary A by $A - t_0$ and extracting a subsequence of $(y_j)_{j\geq 1}$.

LEMMA 3.4. *There exists a sequence* $(y_j)_{j\ge 1}$ *such that*, *for all* $\varepsilon > 0$ *and for* j *sufficiently large*,

$$|A_j| \ge \frac{1}{(2\sigma - 2 + \varepsilon)N}\,(b_j + c_j).$$

Proof. Remember that t_0 is the smallest element of A not divisible by N. We define $S_j := b_j + c_j - t_0 + 2$.

Let $\varepsilon > 0$. We have, using Theorem 2.11,

$$\frac{S_j}{N} \le |A_j + A_j| - 2|A_j| + 3 \le (2\sigma - 2 + \varepsilon')|A_j| + 3 \le (2\sigma - 2 + \varepsilon'')|A_j|,$$

where $\varepsilon' = O(\varepsilon)$ and $\varepsilon'' = O(\varepsilon)$. It suffices to choose j sufficiently large to obtain the result. ■

Below, $(y_j)_{j\ge 1}$ is a sequence of integers as in the last lemma.

Now, we are going to refine the last results. We define

$$X_j := \frac{c_j}{b_j}, \qquad \lambda_j := \frac{N|A_j|}{b_j + c_j}.$$

LEMMA 3.5. *There exists a sequence* $(y_j)_{j\ge 1}$ *such that* $\lim_{j\to\infty} X_j = 1$.

Proof. We will only use the definition of the upper asymptotic density of A. Given $\varepsilon > 0$, for all sufficiently large j we have

$$\frac{A(c_j)}{c_j} \le \frac{A(b_j)}{b_j} + \varepsilon.$$

Furthermore,

$$A(c_j) \ge A(b_j) - \frac{b_j - c_j}{N}.$$

Putting together the last two relations, we obtain

$$N\varepsilon + \frac{\lambda_j(b_j + c_j)}{b_j} \ge \frac{\lambda_j(b_j + c_j)}{c_j} - \frac{b_j}{c_j} + 1.$$

This yields the following polynomial inequality:

$$\lambda_j X_j^2 - (1 - N\varepsilon)X_j - (\lambda_j - 1) \ge 0. \tag{8}$$

It remains to determine the discriminant and the roots. We obtain

$$\Delta = (2\lambda_j - 1)^2 + \varepsilon(N^2\varepsilon - 2N).$$

Thus, using Lemma 3.4 to bound λ_j from below, we see that the roots $X_j' < X_j''$ satisfy

$$X_j' = \frac{1}{2\lambda_j}\,(1 - N\varepsilon - \sqrt{\Delta}) = \frac{1}{\lambda_j} - 1 + O(\varepsilon),$$

$$X_j'' = \frac{1}{2\lambda_j}\,(1 - N\varepsilon + \sqrt{\Delta}) = 1 - O(\varepsilon).$$

Clearly, $X_j < 1/\lambda_j - 1 + O(\varepsilon)$ is impossible, since the lower bound on λ_j obtained in Lemma 3.4 would imply

$$X_j \le 1/\lambda_j - 1 + O(\varepsilon) \le 2\sigma - 3 + O(\varepsilon) < 1/3$$

for ε sufficiently small, and hence

$$\frac{(A+A)(b_j+t_0)}{b_j+t_0} \ge \frac{|B_j| + |B_j| + |C_j + C_j|}{b_j+t_0} \ge 2\alpha - O(\varepsilon),$$

which contradicts the main hypothesis of Theorem 1.2. Thus, $X_j \ge 1 - O(\varepsilon)$, which is the conclusion of the lemma. ∎

Now we combine the results of the last two lemmas and apply Theorems 2.3 and 2.4 to the sets

$$B'_j := \frac{1}{N}\, B_j, \qquad C'_j := \frac{1}{N}\,(C_j - t).$$

We have

$$|A_j| \ge \frac{1}{(2\sigma - 2 - \varepsilon)N}\,(b_j + c_j).$$

We notice that, for ε sufficiently small, since $\sigma < 5/3$,

$$\frac{1}{2\sigma - 2 - \varepsilon} > \frac{3}{4}.$$

Fix $\delta > 0$ such that

$$|A_j| \ge \left(\frac{3}{4} + \delta\right)\frac{b_j + c_j}{N}.$$

Using Lemma 3.5, we obtain

$$|A_j| \ge \left(\frac{3}{4} + \delta\right)\frac{(2-\varepsilon')y_j}{N}$$

for ε' arbitrarily small. Therefore, there exists a positive constant δ' such that, for j sufficiently large,

$$|A_j| \ge \left(\frac{3}{2} + \delta'\right)\frac{y_j}{N}.$$

It follows that

$$|B'_j| = |B_j| = |A_j| - |C_j| \ge |A_j| - \frac{y_j}{N} \ge \left(\frac{1}{2} + \delta'\right)\frac{y_j}{N} \ge \left(\frac{1}{2} + \delta'\right)\max(B'_j).$$

Thus we can apply Theorem 2.3 to B'_j. We can do the same for C'_j. Moreover, we have

$$|B'_j| + |C'_j| = |A_j| \ge \left(\frac{3}{2} + \delta'\right)\frac{y_j}{N},$$

so we can also apply Theorem 2.4.

For j sufficiently large, we then have

$$(9)\qquad \begin{aligned} |A_j+A_j| &= |B_j+B_j|+|B_j+C_j|+|C_j+C_j| \\ &= |B_j'+B_j'|+|B_j'+C_j'|+|C_j'+C_j'| \\ &\ge |B_j'|+\frac{y_j}{N}+(1-\varepsilon')\,\frac{y_j}{N}+|B_j'|+(1-\varepsilon')\,\frac{y_j}{N}+|C_j'| \\ &= 2|B_j|+|C_j|+(3-2\varepsilon')\,\frac{y_j}{N}, \end{aligned}$$

assuming, without loss of generality, that $|B_j|\ge|C_j|$. Here, ε' is arbitrarily small, by Lemma 3.5.

Now, we also have, for j sufficiently large,

$$(10)\qquad |A_j+A_j|\le(2\sigma+\varepsilon')|A_j|.$$

Since $|B_j|\ge|C_j|$, inequality (9) implies that

$$|A_j+A_j|\ge\frac{3}{2}\,|B_j|+\frac{3}{2}\,|C_j|+(3-2\varepsilon')\,\frac{y_j}{N}=\frac{3}{2}\,|A_j|+(3-2\varepsilon')\,\frac{y_j}{N}.$$

Combining this with (10), we obtain

$$|A_j|\ge\frac{6-4\varepsilon'}{4\sigma-3+2\varepsilon'}\,\frac{y_j}{N}.$$

Now, dividing by y_j and sending j to infinity, we obtain

$$\alpha\ge\frac{6-4\varepsilon'}{4\sigma-3+2\varepsilon'}\,\frac{1}{N}.$$

Since ε' is arbitrary, this proves the required inequality of Theorem 1.2:

$$\overline{d}(A)\ge\frac{6}{(4\sigma-3)N}.$$

3.2. *The archimedean case.* Now we assume that $N=1$. We show that, in this case, the sequence $(d_{2,j})_{j\ge1}$ cannot be bounded. Suppose the contrary; then we could extract a subsequence of $(y_j)_{j\ge1}$ such that $d_{2,j}=t$ for all j and act as in the non-archimedean case, i.e. find a common isomorphism between every A_j and a part of two lines of $\mathbb{Z}^2$. This isomorphism could be written

$$\varphi:\mathbb{Z}\times\{0,1\}\to\mathbb{N},\qquad (x,y)\mapsto x+ty.$$

This is impossible because, for j sufficiently large, we would have an element of $A\cap\varphi(\{y=0\})$ greater than t (remember that there are infinitely many elements of $\varphi^{-1}(A)$ on each line) so that t would have two inverse images under φ (one on each line), which contradicts the definition of an F_2-isomorphism.

Therefore, we can choose $(y_j)_{j\ge1}$ and consequently $t_j:=d_{2,j}$ such that t_j is a strictly increasing sequence. Thus, as in the non-archimedean case, we can have $\varphi_j((0,0)) = 0$, $\varphi_j((1,0)) = 1$, $\varphi_j((1,0)) = t_j$ and

$$\varphi_j : \mathbb{Z} \times \{0,1\} \to \mathbb{N}, \qquad (x,y) \mapsto x + yt_j.$$

We shall apply Theorem 2.11 to the sets A_j. Then, we can include A_j in a set A_j^0 which is the union of two arithmetical progressions B_j^0 and C_j^0 (of difference $N = 1$ here). We denote as usual by $b_j := l_{2,j} = |B_j^0|$ and $c_j := l_{2,j} = |C_j^0|$ the respective lengths, where $0 \in B_j^0$ and $y_j \in C_j^0$. Indeed, those two elements cannot be in the same progression: in this case, A would be in an arithmetical progression of dimension 1, say B_j^0. This case, which is the *single line case*, is already excluded by $\alpha < \alpha_0$. Those lengths being supposed minimal, we have $y_j - t_j = l_{2,j}$ and $\max(B_j^0) = b_j$.

LEMMA 3.6. *There exists a sequence $(y_j)_{j\geq 1}$ such that, for all $\varepsilon > 0$, there exists $j_0 \geq 1$ such that for all $j \geq j_0$,*

$$|A_j| \geq \left(\frac{1}{2\sigma - 2} - \varepsilon\right)(l_{1,j} + l_{2,j}).$$

Proof. It suffices to apply Theorem 2.11 for j sufficiently large:

$$\begin{aligned} l_{1,j} + l_{2,j} &\leq |A_j + A_j| - 2|A_j| + 3 \leq (2\sigma - 2 + \varepsilon')|A_j| + 3 \\ &\leq (2\sigma - 2 + \varepsilon'')|A_j|, \end{aligned}$$

where ε' is arbitrarily small and $\varepsilon'' = O(\varepsilon')$. ∎

From now on, $(y_j)_{j\geq 1}$ is a sequence of integers as in the last lemma.

If $b_j \geq t_j$, then $l_{1,j} + l_{2,j} \geq y_j$ and, by Lemma 3.6 and the range of values of σ, we have $|A_j| \geq \frac{3}{4} y_j$, which is incompatible with $\alpha < 1/2$. Therefore, $b_j < t_j$, and thus

$$A(b_j, t_j) = 0. \tag{11}$$

Now we define $B_j := A \cap [0; b_j]$ and $C_j := A \cap [t_j; y_j]$ with $b_j < t_j$.

The quotient $X_j := |B_j|/b_j$ cannot be too large, otherwise we would obtain, considering the sets $A(b_j)$, a too large value for α. Clearly, we have

$$X := \limsup_{j\to\infty} X_j \leq \alpha. \tag{12}$$

Let us show in which sense b_j is necessarily small compared with $l_{2,j}$.

LEMMA 3.7. *Define $\lambda_j := b_j / l_{2,j}$. Then*

$$\lambda := \limsup_{j\to\infty} \lambda_j \leq \frac{2\sigma - 3}{2\sigma - 2}\left(\frac{1}{2\sigma - 2} - X\right)^{-1}. \tag{13}$$

Proof. We use Lemma 3.6, noting that

$$|A_j| = |B_j| + |C_j| = X_j \lambda_j l_{2,j} + |C_j|.$$

For all $\varepsilon > 0$, and j sufficiently large, we obtain

$$X_j \lambda_j l_{2,j} + |C_j| \geq \left(\frac{1}{2\sigma - 2} - \varepsilon\right)(\lambda_j + 1) l_{2,j},$$

and so,

$$|C_j| \geq l_{2,j}\left(\frac{1}{2\sigma - 2} - \varepsilon + \lambda_j\left(\frac{1}{2\sigma - 2} - \varepsilon - X_j\right)\right). \tag{14}$$

Now, we know that $|C_j| \leq l_{2,j}$, and therefore we obtain the upper bound

$$\lambda_j \leq \left(\frac{2\sigma - 3}{2\sigma - 2} + \varepsilon\right)\left(\frac{1}{2\sigma - 2} - \varepsilon - X_j\right)^{-1}.$$

It remains to recall that $X \leq \alpha$ to obtain

$$\lambda \leq \frac{2\sigma - 3}{2\sigma - 2}\left(\frac{1}{2\sigma - 2} - \alpha\right)^{-1}. \ \blacksquare$$

Let us take as a new sequence $(y_j)_{j \geq 1}$ a subsequence such that $\lim_{j \to \infty} \lambda_j = \lambda$. It suffices to look again at the relation (14) to obtain

$$r_j = \frac{|C_j|}{l_{2,j}} \geq \frac{1}{2\sigma - 2} + \lambda\left(\frac{1}{2\sigma - 2} - X\right)$$

for infinitely many values of j.

Then, a last extraction of a subsequence allows us to suppose that the bounded sequence $(r_j)_{j \geq 1}$ has a limit r such that

$$r \geq \frac{1}{2\sigma - 2} + \lambda\left(\frac{1}{2\sigma - 2} - \alpha\right). \tag{15}$$

Hence, putting together (11), (12), Lemma 3.7 and (15) we conclude the proof of the archimedean case of Theorem 1.2.

References

[1] Y. Bilu, *Addition of sets of integers of positive density*, J. Number Theory 64 (1997), 233–275.

[2] —, *Structure of sets with small sumset*, Astérisque 258 (1999), 77–108.

[3] G. A. Freiman, *Foundations of a Structural Theory of Set Addition*, Translat. Math. Monographs 37, Amer. Math. Soc., 1973.

[4] H. Halberstam and K. F. Roth, *Sequences,* 2nd ed., Springer, New York, 1983.

[5] R. Jin, *Inverse problem for upper asymptotic density*, Trans. Amer. Math. Soc. 355 (2002), 57–78.

[6] —, *Solution to the inverse problem for upper asymptotic density*, to appear.

[7] M. Kneser, *Abschätzung der asymptotischen Dichte von Summenmengen*, Math. Z. 58 (1953), 459–484.

[8] V. F. Lev and P. Y. Smeliansky, *On addition of two distinct sets of integers*, Acta Arith. 70 (1995), 85–91.

[9] M. B. Nathanson, *Additive Number Theory. Inverse Problems and the Geometry of Sumsets*, Grad. Texts in Math. 165, Springer, 1996.

[10] Y. Stanchescu, *On addition of two distinct sets of integers*, Acta Arith. 75 (1996), 191–194.

Département de Mathématiques
Université de Bordeaux 1
Talence Cedex, France
E-mail: bordes@math.u-bordeaux1.fr

Received on 30.7.2004
and in revised form on 7.6.2005 (4814)