THE RAMANUJAN JOURNAL 2, 271–281 (1998)

Linear Forms in Finite Sets of Integers

SHU-PING HAN
Department of Mathematics, CUNY Graduate School, New York, NY 10036

CHRISTOPH KIRFEL chrisk@lsv.hib.no
Department of Mathematics, Bergen College, N-5030 Landås, Bergen, Norway

MELVYN B. NATHANSON* nathansn@alpha.lehman.cuny.edu
Department of Mathematics, Lehman College (CUNY), Bronx, NY 10468

Dedicated to Paul Erdős

Received January 5, 1998; Accepted February 6, 1998

Abstract. Let $A_1, \ldots, A_r$ be finite, nonempty sets of integers, and let $h_1, \ldots, h_r$ be positive integers. The linear form $h_1 A_1 + \cdots + h_r A_r$ is the set of all integers of the form $b_1 + \cdots + b_r$, where b_i is an integer that can be represented as the sum of h_i elements of the set A_i. In this paper, the structure of the linear form $h_1 A_1 + \cdots + h_r A_r$ is completely determined for all sufficiently large integers h_i.

Key words: sums of sets of integers, sumsets, growth in semigraphs, additive number theory

1991 Mathematics Subject Classification: Primary 11B05; Secondary 11B13, 11B75, 11P99

1. Sums of sets of integers

Let A be a nonempty set of integers. For every positive integer h, the sumset hA is the set of all integers that can be represented as the sum of exactly h not necessarily distinct elements of A. For example,

$$2\{1, 2, 4\} = \{2, 3, 4, 5, 6, 8\}$$

and

$$3\{0, 2, 5\} = \{0, 2, 4, 5, 6, 7, 9, 10, 12, 15\}.$$

We define $hA = \{0\}$ for $h = 0$. For any set A and integers a_0 and δ, we define

$$a_0 + A = \{a_0 + a : a \in A\},$$
$$a_0 - A = \{a_0 - a : a \in A\},$$

*The work of M.B.N. was supported in part by grants from the PSC-CUNY Research Award Program and the National Security Agency Mathematical Sciences Program.

and

$$\delta * A = \{\delta a : a \in A\}.$$

Let $[x, y]$ denote the interval of integers n such that $x \leq n \leq y$.

A finite set A of integers is called *normalized* if it consists of 0 and a nonempty set of relatively prime positive integers. If A is a finite set of integers with $|A| \geq 2$, we can normalize A as follows. Let a_0 be the least element of A, and let δ be the greatest common divisor of the positive integers of the form $a - a_0$ for $a \in A$. The normalized form of A is the set

$$A^{(N)} = \left\{ \frac{a - a_0}{\delta} : a \in A \right\}.$$

Then

$$A = a_0 + \delta * A^{(N)}$$

and

$$hA = ha_0 + \delta * hA^{(N)}. \tag{1}$$

Note that A is normalized if and only if $A = A^{(N)}$.

Nathanson [5, 6] completely determined the structure of the sumset hA for all nonempty, finite sets A and all sufficiently large integers h. By (1), it suffices to describe the structure of sumsets of normalized sets.

Theorem 1 (Nathanson). *Let A be a normalized finite set of integers, and let a^* be the greatest element of A. There exist integers c and d and sets $C \subseteq [0, c-2]$ and $D \subseteq [0, d-2]$ such that, for h sufficiently large,*

$$hA = C \cup [c, ha^* - d] \cup \left(ha^* - D\right).$$

In this paper we generalize this result to linear forms in finite sets of integers. Let $r \geq 1$. If $A_1, \ldots, A_r$ are nonempty sets of integers and $h_1, \ldots, h_r$ are positive integers, then the sumset

$$h_1 A_1 + \cdots + h_r A_r \tag{2}$$

is the set of all integers that can be represented in the form $b_1 + \cdots + b_r$, where $b_i \in h_i A_i$ for $i = 1, \ldots, r$. The sumset (2) is called a *linear form* in the sets $A_1, \ldots, A_r$. We shall describe explicitly the structure of linear forms in finite sets of integers for all sufficiently large values of $h_1, \ldots, h_r$.

2. The structure of linear forms

The system of sets $A_1, \ldots, A_r$ is *normalized*, if each A_i is a finite set of nonnegative integers, if $0 \in A_i$ for $i = 1, \ldots, r$, and if $\bigcup_{i=1}^{r} A_i \backslash \{0\}$ is a nonempty set of relatively prime positive integers. For example, the sets $A_1 = \{0, 6\}$, $A_2 = \{0, 10\}$, and $A_3 = \{0, 15\}$ are a normalized system, since $(A_1 \cup A_2 \cup A_3)\setminus\{0\} = \{6, 10, 15\}$, and $(6, 10, 15) = 1$.

Let $A_1, \ldots, A_r$ be nonempty, finite sets of integers such that $|A_i| \geq 2$ for all i. We shall normalize this system of sets as follows. Let $a_{i,0}$ be the smallest element in A_i. Let δ be the greatest common divisor of the integers in the set

$$\bigcup_{i=1}^{r} \{a_{i,j} - a_{i,0} : a_{i,j} \in A_i\}.$$

Let

$$A_i^{(N)} = \left\{ \frac{a_{i,j} - a_{i,0}}{\delta} : a_{i,j} \in A_i \right\}.$$

The system of sets $A_1^{(N)}, \ldots, A_r^{(N)}$ is normalized, and

$$A_i = a_{i,0} + \delta * A_i^{(N)}$$

for all $i = 1, \ldots, r$. For any positive integers $h_1, \ldots, h_r$, we have

$$\sum_{i=1}^{r} h_i A_i = \left(\sum_{i=1}^{r} h_i a_{i,0} \right) + \delta * \sum_{i=1}^{r} h_i A_i^{(N)}. \tag{3}$$

By (3), it suffices to describe the structure of sums of normalized systems of finite sets of integers.

Let A be a set of nonnegative integers that contains 0, let $\gcd(A)$ denote the greatest common divisor of the elements of A, and let

$$a^* = \max(A).$$

We define the *reflected set*

$$\hat{A} = a^* - A = \{a^* - a : a \in A\}.$$

Then $\hat{A}$ is also a set of nonnegative integers that contains 0,

$$\max(\hat{A}) = \max(A) = a^*,$$
$$\gcd(A) = \gcd(\hat{A}),$$

and

$$\hat{\hat{A}} = A.$$

For any positive integer h, we have $0 \in hA$, and

$$\max(hA) = ha^*,$$

and so

$$\begin{aligned} h\hat{A} &= \left\{ \sum_{j=1}^{h} (a^* - a_j) : a_j \in A \right\} \\ &= ha^* - \left\{ \sum_{j=1}^{h} a_j : a_j \in A \right\} \\ &= ha^* - hA \\ &= \widehat{hA}. \end{aligned}$$

Lemma 1. *Let $A_1, \ldots, A_r$ be a normalized system of finite sets of integers, and let $a_i^* = \max(A_i)$ for $i = 1, \ldots, r$. The reflected sets $\hat{A}_1, \ldots, \hat{A}_r$ also form a normalized system. For any integer x,*

$$x \in \sum_{i=1}^{r} h_i \hat{A}_i$$

if and only if

$$\sum_{i=1}^{r} h_i a_i^* - x \in \sum_{i=1}^{r} h_i A_i.$$

Moreover,

$$\left[d, \sum_{i=1}^{r} h_i a_i^* - d' \right] \subseteq \sum_{i=1}^{r} h_i \hat{A}_i,$$

if and only if

$$\left[d', \sum_{i=1}^{r} h_i a_i^* - d \right] \subseteq \sum_{i=1}^{r} h_i A_i.$$

Proof: For $i = 1, \ldots, r$, let

$$d_i = \gcd(A_i).$$

If the system $A_1, \ldots, A_r$ is normalized, then

$$1 = \gcd\left(\bigcup_{i=1}^{r} A_i \setminus \{0\} \right) = (d_1, \ldots, d_r).$$

Since

$$\gcd(\hat{A}_i) = \gcd(A_i) = d_i,$$

it follows that

$$\gcd\left(\bigcup_{i=1}^{r} \hat{A}_i \setminus \{0\}\right) = (d_1, \ldots, d_r) = 1,$$

and so the system $\hat{A}_1, \ldots, \hat{A}_r$ is also normalized.

If

$$x \in \sum_{i=1}^{r} h_i \hat{A}_i = \sum_{i=1}^{r} \widehat{h_i A_i} = \sum_{i=1}^{r} (h_i a_i^* - h_i A_i),$$

then there exist integers $b_i \in h_i A_i$ such that

$$x = \sum_{i=1}^{r} (h_i a_i^* - b_i),$$

and so

$$\sum_{i=1}^{r} h_i a_i^* - x = \sum_{i=1}^{r} b_i \in \sum_{i=1}^{r} h_i A_i.$$

The proof in the opposite direction is similar.

We observe that

$$x \in \left[d', \sum_{i=1}^{r} h_i a_i^* - d\right],$$

if and only if

$$\sum_{i=1}^{r} h_i a_i^* - x \in \left[d, \sum_{i=1}^{r} h_i a_i^* - d'\right],$$

and this suffices to prove the last part of the Lemma. □

Theorem 2. *Let $A_1, \ldots, A_r$ be a normalized system of finite sets of integers. Let $a_i^* = \max(A_i)$ for $i = 1, \ldots, r$. There exist integers c and d and finite sets*

$$C \subseteq [0, c-2]$$

and

$$D \subseteq [0, d-2]$$

and there exist integers $h_1^*, \ldots, h_r^*$ such that, if $h_i \geq h_i^*$ for all $i = 1, \ldots, r$, then

$$h_1 A_1 + \cdots + h_r A_r = C \cup \left[c, \sum_{i=1}^{r} h_i a_i^* - d\right] \cup \left(\sum_{i=1}^{r} h_i a_i^* - D\right).$$

Proof: For $i = 1, \ldots, r$, let

$$A_i = \left\{a_{i,0}, a_{i,1}, \ldots a_{i,k_i-1}\right\},$$

where

$$k_i = |A_i|$$

and

$$0 = a_{i,0} < a_{i,1} < \cdots < a_{i,k_i-1}.$$

Renumbering the sets A_i, we can assume that

$$a^* = \max\{a_i^* : i = 1, \ldots, r\} = a_r^* = a_{r,k_r-1}.$$

For any integers c and m^* with $m^* \geq a^*$, we have

$$[c, c + m^* - 1] + A_j = [c, c + m^* - 1 + a_j^*] \tag{4}$$

for $j = 1, \ldots, r$.

Since $\bigcup_{i=1}^r A_i \setminus \{0\}$ is a nonempty set of relatively prime positive integers, it follows that for every integer n there exist integers $x'_{i,j}$ such that

$$n = \sum_{i=1}^{r} \sum_{j=1}^{k_i-1} x'_{i,j} a_{i,j}.$$

For each pair $(i, j) \neq (r, k_r - 1)$, we can choose an integer $x_{i,j}$ such that

$$x'_{i,j} \equiv x_{i,j} \pmod{a^*}$$

and

$$0 \leq x_{i,j} \leq a^* - 1.$$

There exist integers $t_{i,j}$ such that

$$x'_{i,j} = x_{i,j} + t_{i,j} a^*.$$

Then

$$\begin{aligned} n &= \sum_{i=1}^{r} \sum_{j=1}^{k_i-1} x'_{i,j} a_{i,j} \\ &= \sum_{i=1}^{r} \sum_{\substack{j=1 \\ (i,j)\neq(r,k_r-1)}}^{k_i-1} (x_{i,j} + t_{i,j} a^*) a_{i,j} + x'_{r,k_r-1} a_{r,k_r-1} \\ &= \sum_{i=1}^{r} \sum_{j=1}^{k_i-1} x_{i,j} a_{i,j}, \end{aligned}$$

where

$$x_{r,k_r-1} = x'_{r,k_r-1} + \sum_{i=1}^{r} \sum_{\substack{j=1 \\ (i,j)\neq(r,k_r-1)}}^{k_i-1} t_{i,j} a_{i,j}.$$

If

$$n \geq (a^* - 1) \sum_{i=1}^{r} \sum_{\substack{j=1 \\ (i,j)\neq(r,k_r-1)}}^{k_i-1} a_{i,j},$$

then

$$x_{r,k_r-1} \geq 0.$$

Therefore, every sufficiently large integer is a nonnegative integer linear combination of the elements of $\bigcup_{i=1}^{r} A_i$. Let c be the smallest integer such that every integer $n \geq c$ can be represented in the form

$$n = \sum_{i=1}^{r} \sum_{j=1}^{k_i-1} x_{i,j}(n) a_{i,j},$$

where the coefficients $x_{i,j}(n)$ are nonnegative integers. Then

$$c - 1 \notin \sum_{i=1}^{r} h_i A_i$$

for all nonnegative integers $h_1, \ldots, h_r$. For each $i = 1, \ldots, r$, we define

$$h_i^{(1)} = \max \left\{ \sum_{j=1}^{k_i-1} x_{i,j}(n) : n = c, c+1, \ldots, c + a^* - 1 \right\}.$$

Then

$$[c, c + a^* - 1] \subseteq \sum_{i=1}^{r} h_i^{(1)} A_i.$$

It follows that

$$c + a^* - 1 \leq \max \left(\sum_{i=1}^{r} h_i^{(1)} A_i \right) = \sum_{i=1}^{r} h_i^{(1)} a_i^*,$$

and so

$$c' = \sum_{i=1}^{r} h_i^{(1)} a_i^* - (c + a^* - 1) \geq 0.$$

We shall prove that if $h_i \geq h_i^{(1)}$ for all $i = 1, \ldots, r$, then the sumset $\sum_{i=1}^{r} h_i A_i$ contains the interval of integers

$$\left[c, \sum_{i=1}^{r} \left(h_i - h_i^{(1)}\right) a_i^* + c + a^* - 1\right] = \left[c, \sum_{i=1}^{r} h_i a_i^* - c'\right].$$

The proof is by induction on

$$\ell = \sum_{i=1}^{r} \left(h_i - h_i^{(1)}\right).$$

If $\ell = 0$, then $h_i = h_i^{(1)}$ for all $i = 1, \ldots, r$, and the assertion is true.

Let $\ell \geq 1$, and assume that the statement holds for $\ell - 1$. Then $h_j \geq h_j^{(1)} + 1$ for some j. By the induction assumption, we have

$$\begin{aligned}\sum_{\substack{i=1\\ i\neq j}}^{r} h_i A_i + (h_j - 1) A_j &\supseteq \left[c, \sum_{\substack{i=1\\ i\neq j}}^{r} h_i a_i^* + (h_j - 1) a_j^* - c'\right] \\ &= \left[c, \sum_{i=1}^{r} \left(h_i - h_i^{(1)}\right) a_i^* - a_j^* + c + a^* - 1\right].\end{aligned}$$

Applying (4) with

$$m^* = a^* + \sum_{i=1}^{r} \left(h_i - h_i^{(1)}\right) a_i^* - a_j^* \geq a^*,$$

we obtain

$$\begin{aligned}\sum_{i=1}^{r} h_i A_i &= \left(\sum_{\substack{i=1\\ i\neq j}}^{r} h_i A_i + (h_j - 1) A_j\right) + A_j \\ &\supseteq \left[c, \sum_{i=1}^{r} \left(h_i - h_i^{(1)}\right) a_i^* - a_j^* + c + a^* - 1\right] + A_j \\ &= \left[c, \sum_{i=1}^{r} \left(h_i - h_i^{(1)}\right) a_i^* + c + a^* - 1\right] \\ &= \left[c, \sum_{i=1}^{r} h_i a_i^* - c'\right].\end{aligned}$$

This completes the induction.

If the system of sets $A_1, \ldots, A_r$ is normalized, then the system of reflected sets $\hat{A}_1, \ldots, \hat{A}_r$ is also normalized. Applying the previous argument to the reflected system, we obtain integers $d, d', h_1^{(2)}, \ldots, h_r^{(2)}$ such that d is the largest integer with the property that $d - 1$ cannot be written as a nonnegative integral linear combination of the elements of $\bigcup_{i=1}^r \hat{A}_i$, and

$$\left[d, \sum_{i=1}^r h_i a_i^* - d'\right] \subseteq \sum_{i=1}^r h_i \hat{A}_i$$

if $h_i \geq h_i^{(2)}$ for $i = 1, \ldots, r$. By Lemma 1,

$$\left[d', \sum_{i=1}^r h_i a_i^* - d\right] \subseteq \sum_{i=1}^r h_i A_i$$

and

$$\sum_{i=1}^r h_i a_i^* - d + 1 \notin \sum_{i=1}^r h_i A_i$$

for all nonnegative integers $h_1, \ldots, h_r$.

Choose $h_i^{(3)} \geq \max\{h_i^{(1)}, h_i^{(2)}\}$ such that

$$c' + d' \leq \sum_{i=1}^r h_i^{(3)} a_i^*.$$

If $h_i \geq h_i^{(3)}$, then

$$\left[c, \sum_{i=1}^r h_i a_i^* - d\right] \subseteq \sum_{i=1}^r h_i A_i.$$

Since

$$c - 1 \notin \sum_{i=1}^r h_i A_i$$

and

$$\sum_{i=1}^r h_i a_i^* - d + 1 \notin \sum_{i=1}^r h_i A_i$$

for all nonnegative integers $h_1, \ldots, h_r$, it follows that if

$$h_i \geq h_i^* = \max\left\{h_i^{(3)}, c, d\right\},$$

then there exist sets $C \subseteq [0, c-2]$ and $D \subseteq [0, d-2]$ such that

$$\sum_{i=1}^{r} h_i A_i = C \cup \left[c, \sum_{i=1}^{r} h_i a_i^* - d \right] \cup \left(\sum_{i=1}^{r} h_i a_i^* - D \right).$$

This completes the proof. □

Theorem 3. *Let $A_1, \ldots, A_r$ be a normalized system of finite sets of integers, and let $a_i^* = \max(A_i)$ for $i = 1, \ldots, r$. Let B be a finite set of nonnegative integers with $0 \in B$ and $b^* = \max(B)$. There exist integers c and d and finite sets*

$$C \subseteq [0,\ c-2]$$

and

$$D \subseteq [0,\ d-2]$$

such that

$$B + h_1 A_1 + \cdots + h_r A_r = C \cup \left[c, b^* + \sum_{i=1}^{r} h_i a_i^* - d \right] \cup \left(b^* + \sum_{i=1}^{r} h_i a_i^* - D \right)$$

for all sufficiently large intergers h_i.

Proof: This is a simple consequence of Theorem 2. □

3. The cardinality of linear forms

Theorem 3 immediately implies the following estimate for the size of a sumset of integers.

Theorem 4. *Let $A_1, \ldots, A_r$ be a normalized system of finite sets of integers, and let B be a nonempty, finite set of nonnegative integers. There exist positive integers $a_1^*, \ldots, a_r^*$ and nonnegative integers b^* and Δ such that*

$$|B + h_1 A_1 + \cdots + h_r A_r| = \sum_{i=1}^{r} a_i^* h_i + b^* + 1 - \Delta$$

for all sufficiently large integers h_i.

Theorem 4 shows that the cardinality of the sumset $B + h_1 A_1 + \cdots + h_r A_r$ is a linear polynomial in the variables $h_1, \ldots, h_r$. This is a special case of the following very general result. Let S be an arbitrary abelian semigroup, written additively, and let $B, A_1, \ldots, A_r$ be finite, nonempty subsets of S. We can define the sumset $B + h_1 A_1 + \cdots + h_r A_r$ in S exactly as we defined sumsets in the semigroup of integers. Extending results of Khovanskii [1, 2] for the case $r = 1$, Nathanson [7] proved that there exists a polynomial $p(x_1, \ldots, x_r)$ such that

$$|B + h_1 A_1 + \cdots + h_r A_r| = p(h_1, \ldots, h_r)$$

for all sufficiently large integers h_i. For an arbitrary semigroup S, it is not known how to compute this polynomial, nor even to determine its degree.

References

1. A.G. Khovanskii, "Newton polyhedron, Hilbert polynomial, and sums of finite sets," *Funktsional. Anal. i Prilozhen*. **26** (1992), 276–281.
2. A.G. Khovanskii, "Sums of finite sets, orbits of commutative semigroups, and Hilbert functions," *Funktsional. Anal. i Prilozhen*. **29** (1995), 102–112.
3. V.F. Lev, "Structure theorem for multiple addition and the Frobenius problem," *J. Number Theory* **58** (1996), 79–88.
4. V.F. Lev, "Addendum to 'Structure theorem for multiple addition'," *J. Number Theory* **65** (1997), 96–100.
5. M.B. Nathanson, "Sums of finite sets of integers," *Amer. Math. Monthly* **79** (1972), 1010–1012.
6. M.B. Nathanson, *Additive Number Theory: Inverse Problems and the Geometry of Sumsets*, volume 165 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1996.
7. M.B. Nathanson, "Growth of sumsets in abelian semigroups," Preprint, 1997.
8. Ö.J. Rödseth, "On *h*-bases for *n*," *Math. Scand*. **48** (1981), 165–183.