Reconstructing integer sets from their representation functions

Vsevolod F. Lev

Department of Mathematics
University of Haifa at Oranim
Tivon 36006, Israel
seva@math.haifa.ac.il

Submitted: Oct 5, 2004; Accepted: Oct 27, 2004; Published: Nov 3, 2004
Mathematics Subject Classifications: 11B34, 05A17, 11B13

Abstract

We give a simple common proof to recent results by Dombi and by Chen and Wang concerning the number of representations of an integer in the form $a_1 + a_2$, where a_1 and a_2 are elements of a given infinite set of integers. Considering the similar problem for differences, we show that there exists a partition $\mathbb{N} = \cup_{k=1}^{\infty} A_k$ of the set of positive integers such that each A_k is a perfect difference set (meaning that any non-zero integer has a unique representation as $a_1 - a_2$ with $a_1, a_2 \in A_k$). A number of open problems are presented.

1 Introduction and summary

For a set $A \subseteq \mathbb{Z}$ and an integer $n \in \mathbb{Z}$ consider the representation functions

$$R_A^{(1)}(n) = \{(a_1, a_2) \in A \times A\colon a_1 + a_2 = n\},$$
$$R_A^{(2)}(n) = \{(a_1, a_2) \in A \times A\colon a_1 + a_2 = n,\ a_1 < a_2\},$$

and

$$R_A^{(3)}(n) = \{(a_1, a_2) \in A \times A\colon a_1 + a_2 = n,\ a_1 \le a_2\}.$$

To what extent do $R_A^{(j)}(n)$ determine the set A? Problems of this sort were, to our knowledge, first studied by Nathanson in [N78]. Let $\mathbb{N}$ denote the set of all positive integers. In his research talks and private communications, Sárközy has raised the following question: do there exist $A, B \subseteq \mathbb{N}$ with the infinite symmetric difference such that $R_A^{(j)}(n) = R_B^{(j)}(n)$ for all, but finitely many $n \in \mathbb{N}$?

Dombi noticed in [D02] that the answer is negative for $j = 1$, by the simple observation that $R_A^{(1)}(n)$ is odd if and only if $n = 2a$ for some $a \in A$. On the other hand, he has shown that for $j = 2$ the answer is positive and indeed, there is a partition $\mathbb{N} = A \cup B$ such that $R_A^{(2)}(n) = R_B^{(2)}(n)$ for all $n \in \mathbb{N}$.

Theorem 1 (Dombi [D02]) *Define the mapping* $T\colon \mathbb{N} \to \{1,-1\}$ *by*

$$T(1) = 1,\ \ T(2n) = -T(2n-1),\ T(2n+1) = T(n+1); \quad n \in \mathbb{N}$$

and let

$$A = \{n \in \mathbb{N}\colon T(n) = 1\},\ B = \{n \in \mathbb{N}\colon T(n) = -1\}.$$

Then $R_A^{(2)}(n) = R_B^{(2)}(n)$ *for all* $n \in \mathbb{N}$.

For the function $R_A^{(3)}(n)$ the problem was solved by Chen and Wang in [CW03].

Theorem 2 (Chen and Wang [CW03]) [1] *Define the mapping* $T\colon \mathbb{N} \to \{1,-1\}$ *by*

$$T(1) = 1,\ T(2n) = -T(2n-1),\ T(2n+1) = -T(n+1); \quad n \in \mathbb{N}$$

and let

$$A = \{n \in \mathbb{N}\colon T(n) = 1\},\ B = \{n \in \mathbb{N}\colon T(n) = -1\}.$$

Then $R_A^{(3)}(n) = R_B^{(3)}(n)$ *for all integer* $n \geq 3$.

Below we give Theorems 1 and 2 a new simple proof, establishing both results through one common argument which also shows that the constructions of Dombi and Chen-Wang are, essentially, unique. We then proceed to investigate the parallel problem for differences.

Let $r_A(n)$ denote the number of representations of the integer n as a difference of two elements of the set $A \subseteq \mathbb{Z}$:

$$r_A(n) = \{(a', a'') \in A \times A\colon a'' - a' = n\}.$$

It is not difficult to see that for any finite partition of $\mathbb{N}$ one can find a partition set, say A, such that there are arbitrarily large integer n with $r_A(n) = \infty$. On the other hand, we were able to partition $\mathbb{N}$ into the *infinite* number of subsets with identical finite difference representation functions. Indeed, our subsets are perfect difference sets. (Recall, that $A \subseteq \mathbb{Z}$ is a *perfect difference set* if any non-zero integer has a unique representation as a difference of two elements of A; in our terms, $r_A(n) = 1$ for any $n \in \mathbb{N}$.) Moreover, one can arrange it so that the subsets in question have completely different structure.

Theorem 3 *There is a partition* $\mathbb{N} = \cup_{k=1}^{\infty} A_k$ *of the set of all positive integers such that each* A_k *is a perfect difference set and* $|A_i \cap (A_j + z)| \leq 2$ *for any* $i, j, z \in \mathbb{N}$.

[1] The way we present Theorems 1 and 2 emphasizes the striking similarity between the partitions $\mathbb{N} = A \cup B$ considered in these theorems. Ironically, Dombi conjectured that sets $A, B \subseteq \mathbb{N}$ with the infinite symmetric difference satisfying $R_A^{(3)}(n) = R_B^{(3)}(n)$ (for n large enough) do *not* exist. The result of Chen and Wang shows, however, that such sets do exist and can be obtained by a very minor modification of Dombi's original construction.

2 The proofs

Proof of Theorems 1 and 2. Let A_2, B_2, and T_2 denote the two sets and the mapping of Theorem 1, and let A_3, B_3, and T_3 denote the two sets and the mapping of Theorem 2. For $j \in \{2, 3\}$ define

$$\alpha_j(x) = \sum_{a \in A_j} x^a,\ \beta_j(x) = \sum_{b \in B_j} x^b,\ \text{and } \tau_j(x) = \sum_{n \in \mathbb{N}} T_j(n) x^n.$$

Thus $\alpha_j(x), \beta_j(x)$, and $\tau_j(x)$ converge absolutely for $|x| < 1$ and satisfy

$$\alpha_j(x) + \beta_j(x) = \frac{x}{1-x},\ \alpha_j(x) - \beta_j(x) = \tau_j(x). \tag{1}$$

Moreover, it is easily seen that

$$2 \sum_{n \in \mathbb{N}} R_A^{(j)}(n) x^n = (\alpha_j(x))^2 + (-1)^{j+1} \alpha_j(x^2) \tag{2}$$

and similar identity holds with B and β substituted for A and α, respectively. Taking into account that the sum $T_j(1) + \dots + T_j(n-1)$ vanishes for n odd and equals $-T_j(n)$ for n even, we derive from (1) and (2) that

$$\begin{aligned}
2 \sum_{n \in \mathbb{N}} \big(R_A^{(j)}(n) - R_B^{(j)}(n)\big) x^n &= \big((\alpha_j(x))^2 - (\beta_j(x))^2\big) + (-1)^{j+1}\big(\alpha_j(x^2) - \beta_j(x^2)\big) \\
&= \frac{x}{1-x}\, \tau_j(x) + (-1)^{j+1} \tau_j(x^2) \\
&= \sum_{n \in \mathbb{N}} \Big(\sum_{1 \le i \le n-1} T_j(i) \Big) x^n + (-1)^{j+1} \sum_{n \in \mathbb{N}} T_j(n) x^{2n} \\
&= \sum_{n \in \mathbb{N}} \big(-T_j(2n) + (-1)^{j+1} T_j(n) \big) x^{2n}
\end{aligned}$$

for $|x| < 1$ and $j \in \{2, 3\}$. It remains to observe that $T_j(2n) = (-1)^{j+1} T_j(n)$, except if $j = 3$ and $n = 1$. $\square$

Suppose that $\mathbb{N} = A \cup B$ is a partition of the set of positive integers and let $T(n) = 1$ if $n \in A$ and $T(n) = -1$ if $n \in B$. Our proof of Theorems 1 and 2 shows that then $R_A^{(j)}(n) = R_B^{(j)}(n)$ for all sufficiently large n if and only if $T(1) + \dots + T(2n) = 0$ and $T(2n) = (-1)^{j+1} T(n)$, for all but finitely many $n \in \mathbb{N}$. The reader will easily check that this is equivalent to the assertion that there exists $n_0 \in \mathbb{N}$ such that $T(2n) = -T(2n-1)$ and $T(2n-1) = (-1)^j T(n)$ for $n \ge n_0$, and $T(1) + \dots + T(2n_0) = 0$. That is, any partition $\mathbb{N} = A \cup B$ satisfying $R_A^{(j)}(n) = R_B^{(j)}(n)$ for all sufficiently large n is obtained essentially as in Theorems 1 and 2.

Proof of Theorem 3. Fix a function $f \colon \mathbb{N} \to \mathbb{N}$ satisfying

$$f(2m-1) \le m,\ f(2m) = m+1; \quad m = 1, 2, \dots \tag{3}$$

and such that for any $k \in \mathbb{N}$ the inverse image $f^{-1}(k) = \{n \in \mathbb{N}\colon k = f(n)\}$ is infinite. Set $A_k = \varnothing$ for all $k \in \mathbb{N}$. Our construction involves infinitely many steps which we enumerate by positive integers. At the nth step we add one or two elements to the set $A_{f(n)}$ so that (i) every positive integer is added to some A_k at certain step; (ii) no positive integer is added to several distinct A_k at different steps; (iii) for any $d, k \in \mathbb{N}$ there is a step such that the element(s) added at this step to A_k produce(s) a pair $(a_1, a_2) \in A_k \times A_k$ with $a_2 - a_1 = d$; (iv) the element(s) added to A_k at any step produce(s) no non-trivial equality of the form $a_1 - a_2 = a_3 - a_4$ with $a_1, a_2, a_3, a_4 \in A_k$; (v) the element(s) added to A_k at any step produce(s) no triple $(a_1, a_2, a_3) \in A_k \times A_k \times A_k$ which is a shift of another triple $(b_1, b_2, b_3) \in A_l \times A_l \times A_l$ with some $l \neq k$. Once we manage to satisfy (i)–(v), our proof is over; we now proceed to describe exactly how the elements to be added to $A_{f(n)}$ at the nth step are chosen.

If $n = 2m$ is even then it follows from (3) that the set $A_{f(n)} = A_{m+1}$ was not affected by steps $1, \dots, n-1$. This set, therefore, remains empty by the beginning of the nth step, and we initialize it inserting to it the smallest positive integer not contained in $\cup_{l=1}^{m} A_l$.

Suppose now that n is odd and write for brevity $k = f(n)$. Let d be the smallest positive integer, not representable as $a_1 - a_2$ with $a_1, a_2 \in A_k$. We insert to A_k two numbers z and $z + d$, where z is to satisfy the following conditions:

(a) $\{z, z+d\} \cap \left(\cup_{l=1}^{\infty} A_l \right) = \varnothing$;

(b) equality $a_1 - a_2 = a_3 - a_4$ with $a_1, a_2, a_3, a_4 \in A_k \cup \{z, z+d\}$ holds only trivially; that is, if and only if either $a_1 = a_3$ and $a_2 = a_4$, or $a_1 = a_2$ and $a_3 = a_4$;

(c) none of the triples $(a_1, z, z+d)$, (a_1, a_2, z), $(a_1, a_2, z+d)$ with $a_1, a_2 \in A_k$ are translates of a triple (b_1, b_2, b_3) with $b_1, b_2, b_3 \in A_l,\ l \neq k$.

Clearly, condition (a) excludes only a finite number of possible values of z, and a little meditation shows that this is the case also with condition (b). Concentrating on condition (c), we notice that the actual number of values of l to be taken into account is finite, as all but $(n+1)/2$ sets A_l are empty by the beginning of the nth step. Furthermore, for any fixed l the number of triples (b_1, b_2, b_3) with $b_1, b_2, b_3 \in A_l$ is finite, and the number of possible values of $a_1 \in A_k$ is finite, too. It follows that condition (c) also excludes only finite number of z. Thus choosing z is always possible, and this concludes the proof. □

3 Open problems

We list below some related problems.

The proof of Theorem 3 can be simplified if we wish to construct just one perfect difference set $A \subseteq \mathbb{N}$. In this case we can start with the empty set $A^{(0)} = \varnothing$ and define at the nth step $A^{(n)} = A^{(n-1)} \cup \{z_n, z_n + d_n\}$, where d_n is the smallest non-negative integer not representable as $a_1 - a_2$ with $a_1, a_2 \in A^{(n-1)}$, and z_n is to be so chosen that $z_n, z_n + d_n \notin A^{(n-1)}$, and no non-trivial equality $a_1 - a_2 = a_3 - a_4$ with $a_1, a_2, a_3, a_4 \in$

$A^{(n-1)} \cup \{z_n, z_n + d_n\}$ is created. The number of choices of z_n excluded by these conditions is $O(n^3)$ and since $d_n = O(n^2)$, the nth element of the resulting set A is $O(n^3)$. It follows that the counting function $A(x) = |A \cap [1, x]|$ satisfies $A(x) \gg x^{1/3}$. On the other hand, it is easily seen that for any perfect difference set $A \subseteq \mathbb{N}$ we have $A(x) \ll x^{1/2}$.

Problem 1 Does there exist a perfect difference set $A \subseteq \mathbb{N}$ with the counting function $A(x) \gg x^{1/2}$? If not, is it true that for any $\varepsilon > 0$ there exists a perfect difference set $A \subseteq \mathbb{N}$ with $A(x) \gg x^{1/2-\varepsilon}$? If not, how large can $\liminf_{x\to\infty} \ln A(x)/\ln x$ for a perfect difference set $A \subseteq \mathbb{N}$ be?

The definition of $R_A^{(1)}(n)$ extends readily onto the case where A is a subset of an arbitrary abelian group G and n is an element of the group. Suppose that G is finite and for a group character χ let $\widehat{A}(\chi) = |G|^{-1} \sum_{a\in A} \overline{\chi}(a)$, the Fourier coefficient of the indicator function of A. The identity $R_A^{(1)} = R_B^{(1)}$ translates easily into the requirement that either $\widehat{A}(\chi) = \widehat{B}(\chi)$ or $\widehat{A}(\chi) = -\widehat{B}(\chi)$ hold for any character χ. Though this seems to be a rather strong condition, numerical computations show that for certain groups pairs (A, B) such that $R_A^{(1)} = R_B^{(1)}$ are not that rare as one could expect. Quite likely, these pairs are not limited to simple special cases as for instance $|A| = |G|/2,\ B = G \setminus A$, or $B = \{a + d \colon a \in A\}$ with a fixed element $d \in G$ of order two. Nevertheless we state our next problem in the most general form.

Problem 2 For any finite abelian group G determine all pairs of subsets $A, B \subseteq G$ such that $R_A^{(1)} = R_B^{(1)}$.

We note that if G is of odd order then no non-trivial pairs exist, as in this case the values of n for which $R_A^{(1)}(n)$ is odd determine the set A uniquely. On the other hand, if G is an elementary 2-group then any two perfect difference sets $A, B \subseteq G$ satisfy $R_A^{(1)} = R_B^{(1)}$.

As a common generalization of the representation functions $R_A^{(1)}$ and r_A, one can consider two potentially different sets $A, B \subseteq \mathbb{Z}$ and for $n \in \mathbb{Z}$ define

$$r_{A,B}(n) = \#\{(a, b) \in A \times B \colon a + b = n\}.$$

An unpublished observation due to Freiman, Yudin, and the present author is as follows. Suppose that A and B are finite and non-empty, and for $k \in \mathbb{N}$ let ν_k denote the kth largest value attained by $r_{A,B}$. Thus $\{\nu_k\}$ is the spectrum of the function $r_{A,B}$ and we have $\nu_1 \geq \nu_2 \geq \cdots,\ \nu_1 + \nu_2 + \cdots = |A||B|$, and $\nu_k = 0$ for all k large enough. Then

$$\nu_k^2 \leq \nu_k + \nu_{k+1} + \nu_{k+2} + \cdots \tag{4}$$

for any $k \in \mathbb{N}$.

For the proof we write $A = \{a_1, \dots, a_l\}$ and $B = \{b_1, \dots, b_m\}$ where the elements are numbered in the increasing order, and notice first that

$$r_{A,B}(a_i + b_j) \leq \min\{i + j - 1, l + m - (i + j - 1)\}; \quad 1 \leq i \leq l,\ 1 \leq j \leq m.$$

For if $a_i + b_j = a_u + b_v$ then either $u \leq i$, or $v \leq j$; since there are at most i such representations with $u \leq i$ and at most j representations with $v \leq j$, and one representation satisfies both $u \le i$ and $v \le j$, we conclude that $r_{A,B}(a_i + b_j) \le i + j - 1$. The proof of the estimate $r_{A,B}(a_i + b_j) \le l + m - (i + j - 1)$ is almost identical; just notice that if $a_i + b_j = a_u + b_v$ then either $u \ge i$ or $v \ge j$. Now we have

$$\begin{aligned}
\nu_1 + \cdots + \nu_k &\le \#\{(i,j)\colon r_{A,B}(a_i + b_j) \ge \nu_k\} \\
&\le \#\{(i,j)\colon \min\{i + j - 1, l + m - (i + j - 1)\} \ge \nu_k\} \\
&= \#\{(i,j)\colon \nu_k \le i + j - 1 \le l + m - \nu_k\} \\
&= lm - 2\,\#\{(i,j)\colon i + j \le \nu_k\} \\
&= lm - \nu_k(\nu_k - 1)
\end{aligned}$$

and (4) follows from $lm = \nu_1 + \cdots + \nu_k + \nu_{k+1} + \cdots$.

Problem 3 What are the general properties shared by the functions $r_{A,B}(n)$ (for all finite non-empty $A, B \subseteq \mathbb{Z}$), other than that reflected by (4)?

Since the spectrum $\{\nu_k\}$ defines a partition of the integer $|A||B|$, it can be visualized with a Ferrers diagram corresponding to this partition; that is, an arrangement of $|A||B|$ square boxes in bottom-aligned columns such that the leftmost column is of height ν_1, the next column is of height ν_2, and so on. For any $t \in \mathbb{N}$, the length of the tth row of this diagram (counting the rows from the bottom) is then $N_t = \#\{n\colon r_{A,B}(n) \ge t\}$. We notice that from a well-known result of Pollard [P75] it follows that $N_1 + \cdots + N_t \ge t(|A| + |B| - t)$ for any $t \le \min\{|A|, |B|\}$; one can derive this inequality as a corollary of (4), too.

We conclude our note with two problems due to Gowers and Konyagin, presented here from their kind permission. Both problems pertain to the group $\mathbb{Z}/p\mathbb{Z}$ of residue classes modulo a prime p.

Problem 4 (Gowers, personal communication) For a prime p, let $A \subseteq \mathbb{Z}/p\mathbb{Z}$ be a subset of cardinality $|A| = (p+1)/2$. The average value of $R_A^{(1)}$ is then $(p+1)^2/(4p) = p/4 + O(1)$. Is it true that for any positive constant ε and any sufficiently large p, there exists $n \in \mathbb{Z}/p\mathbb{Z}$ satisfying $|R_A^{(1)}(n) - p/4| < \varepsilon p$?

Problem 5 (Konyagin, personal communication) Do there exist positive constants ε and C such that for any sufficiently large prime p and any non-empty subset $A \subseteq \mathbb{Z}/p\mathbb{Z}$ of cardinality $|A| < \sqrt{p}$, there is $n \in \mathbb{Z}/p\mathbb{Z}$ satisfying $1 \le R_A^{(1)}(n) \le C|A|^{1-\varepsilon}$?

References

[CW03] Y.-G. Chen and B. Wang, On the additive properties of two special sequences, *Acta Arithmetica* **110** (2003), *no.* 3, 299–303.

[D02] G. Dombi, Additive properties of certain sets, *Acta Arithmetica* **103** (2002), *no.* 2, 137–146.

[N78] M.B. Nathanson, Representation functions of sequences in additive number theory, *Proc. Amer. Math. Soc.* **72** (1978), *no.* 1, 16–20.

[P75] J.M. Pollard, Addition properties of residue classes, *J. London Math. Soc.* **2** (11) (1975), 147–152, 460–462.