ON SUM-DISTINCT ELEMENTS OF A POWER SET

Dušan B. Jevtić
Department of Mathematical Sciences
University of Alaska Fairbanks
Fairbanks, Alaska 99775-1110
email: ffdbj@Alaska

Abstract. We consider a subset-sum problem in $(2^{\mathcal{S}},\cup)$, $(2^{\mathcal{S}},\triangle)$, $(2^{\mathcal{S}},\uplus)$ and $(\mathcal{S}_n,+)$, where $\mathcal{S}$ is an n-element set, $\mathcal{S}_n \triangleq \{0,1,2,\cdots,2^n-1\}$ and $\cup$, $\triangle$, $\uplus$ and $+$ stand for set-union, symmetric set-difference, multiset-union, and real-number addition, respectively. Simple relationships between compatible pairs of sum-distinct sets in these structures are established. The behavior of a sequence $\{n^{-1}|\mathcal{Z}| \mid n = 2,3,\cdots\}$, where $\mathcal{Z}$ is the maximum cardinality sum-distinct subset of $\mathcal{S}$ (or $\mathcal{S}_n$), is described in each of the four structures.

1. Introduction

Let $\mathcal{X}$ be a finite set in which a commutative and associative binary operation is defined. For this operation we choose the additive notation $+$ and call $x_{i_1}+x_{i_2}+\cdots+x_{i_m}$, $x_{i_k}\in\mathcal{X}$, a *sum*. It is assumed that $\mathcal{X}$ has a zero element 0, i.e., $x+0=0+x=x$ for all x from $\mathcal{X}$. The sum of all elements from $\mathcal{Z}$ is called the *sum of* $\mathcal{Z}$. The sum of $\emptyset$ is 0. The number of elements in $\mathcal{Z}\subseteq\mathcal{X}$ is denoted by $|\mathcal{Z}|$. The power set of $\mathcal{Z}$ is denoted by $2^{\mathcal{Z}}$.

A set $\mathcal{Z}\subset\mathcal{X}$ is said to be *sum-distinct* in $(\mathcal{X},+)$, if the $2^{|\mathcal{Z}|}$ sums of subsets of $\mathcal{Z}$ are all distinct.

In the sequel, we will specify $\mathcal{X}$ either as $2^{\mathcal{S}}$, $|\mathcal{S}|=n$, or as $\mathcal{S}_n \triangleq \{0,1,2,\cdots,2^n-1\}$, and consider sum-distinct sets in $(2^{\mathcal{S}},\cup)$, $(2^{\mathcal{S}},\triangle)$, $(2^{\mathcal{S}},\uplus)$ and $(\mathcal{S}_n,+)$, where $\cup$, $\triangle$, $\uplus$ and $+$ stand for set-union, symmetric set-difference, multiset-union, and real-number addition, respectively. The pair $(2^{\mathcal{S}},\cap)$, being the dual of $(2^{\mathcal{S}},\cup)$, is not included. The pair $(\mathcal{S}_n,+)$ is motivated by $|\mathcal{Z}|\in\{1,2,\cdots,2^n-1\}$ if $\mathcal{Z}$ is a proper subset of $2^{\mathcal{S}}$. (The rationale behind $2^n\notin\mathcal{S}_n$ is to make $\mathcal{S}_n$ a decimal equivalent of $\{0,1\}^n$.) By identifying: 1) a subset $\mathcal{A}$ of $\mathcal{S}$ with the characteristic function $I_{\mathcal{A}}$ of $\mathcal{A}$ on $\mathcal{S}$, and 2) $\cup$, $\triangle$, and $\uplus$ with the corresponding bitwise operations on elements from $\{0,1\}^n$, we identify the pairs $(2^{\mathcal{S}},\cup)$, $(2^{\mathcal{S}},\triangle)$ and $(2^{\mathcal{S}},\uplus)$, with

$(\{0,1\}^n,+_\flat)$ where $+_\flat$ is a coordinate-wise Boolean addition,
$(\{0,1\}^n,\oplus)$ where $\oplus$ is a coordinate-wise modulo 2 addition,

$(\{0,1\}^n,+)$ where $+$ is a coordinate-wise real-number addition, [1] respectively. For example, if $\mathcal{S}=\{a,b,c\}$, then, due to the table below

	$I_{\{a,b\}}$	$I_{\{a,c\}}$	$I_{\{b,c\}}$	$I_{\{a\}}$
a	1	1	0	1
b	1	0	1	0
c	0	1	1	0

verifying the sum-distinctness of $\{\{a,b\},\{a,c\},\{b,c\},\{a\}\}$ in $(2^{\{a,b,c\}},\uplus)$ amounts to verifying the sum-distinctness of $\{110,101,011,100\}$ in the corresponding structure $(\{0,1\}^3,+)$. The latter seems to be more appealing (at least to author). Note that $\mathcal{X}$ does not have to be closed with respect to $+$, e.g. $\{a,b\}\uplus\{b,c\}=\{a,2b,c\}\notin 2^{\{a,b,c\}}$.

Let $\mathcal{Z}$ be the maximum cardinality sum-distinct set in $(\mathcal{X},+)$, $|\mathcal{X}|=2^n$, and let $\mathcal{N}_n\triangleq\{n,n+1,n+2,\cdots\}$. A relative measure of the size of $\mathcal{Z}$ in $\mathcal{X}$, defined by $d_n\triangleq n^{-1}|\mathcal{Z}|$, will be called the *density of $\mathcal{Z}$ in $\mathcal{X}$*. Our criterion in comparing sum-distinct sets in the above structures will be the density sequence $\{d_n\mid n\in\mathcal{N}_2\}$ and, in particular, $d^*\triangleq\max\{d_n\mid n\in\mathcal{N}_2\}$. The definition of density is motivated by an information-theoretic measure in multiple-access coding, [4].

Note that $\omega:\{0,1\}^n\rightarrow\mathcal{S}_n$, defined by $\omega(\bar{z})=\sum_{i=0}^{n-1}z_i2^i$, $\bar{z}=(z_0,\cdots,z_{n-1})$ and $z_i\in\{0,1\}$, is an injection. (The algorithm $\omega^{-1}(\cdot)$ is simple and well known.) The binary equivalent of $x\in\mathcal{S}_n$ will be denoted by $b(x)$, e.g. $b(0)=\bar{0}$. The binary equivalent of a set $\mathcal{Z}\subseteq\mathcal{S}_n$ will be denoted by $b(\mathcal{Z})$; that is, $b(\mathcal{Z})=\{b(z)\mid z\in\mathcal{Z}\}$.

In the sequel, we discuss sum-distinct sets in the above listed structures. We relate them to each other (when possible), describe the behavior of their density sequences and determine d^*. It turns out that, with respect to the above points, problems related to sum-distinct sets in $(\{0,1\}^n,+_\flat)$ and $(\{0,1\}^n,\oplus)$ are simpler than the ones in $(\mathcal{S}_n,+)$ and $(\{0,1\}^n,+)$. We address the former in the next section. In Sections 3 and 4 we discuss sum-distinct sets in $(\mathcal{S}_n,+)$ and $(\{0,1\}^n,+)$, respectively.

2. Sum-distinct sets in $(\{0,1\}^n,+_\flat)$ and $(\{0,1\}^n,\oplus)$

[1] Since the binary operation $+$ is always written as a part of the pair $(\mathcal{X},+)$, the same notation for addition (as in $(\mathcal{S}_n,+)$) should not cause confusion.

Sum-distinct sets in $(\{0,1\}^n, +_\flat)$ and $(\{0,1\}^n, \oplus)$ are related by the following observation.

Remark 1. If $\mathcal{Z}$ is a sum-distinct set in $(\{0,1\}^n, +_\flat)$, then $\mathcal{Z}$ is a sum-distinct set in $(\{0,1\}^n, \oplus)$.

Proof.[2] With little effort one can show that a set $\{\bar{z}_1, \bar{z}_2, \cdots, \bar{z}_T\}$, $\bar{z}_i \neq \bar{0}$, is sum-distinct in $(\{0,1\}^n, \oplus)$ if and only if none of its $2^T - 1$ nonempty subsets has sum equal to zero. Hence, if $\mathcal{Z}$ is not sum-distinct in $(\{0,1\}^n, \oplus)$, then there exists at least one subset of $\mathcal{Z}$, say $\{\bar{z}_{i_1}, \bar{z}_{i_2}, \cdots, \bar{z}_{i_m}\}$, such that

$$\bar{z}_{i_1} \oplus \bar{z}_{i_2} \oplus \cdots \oplus \bar{z}_{i_m} = \bar{0}. \tag{2.1}$$

Let $\bar{u} = \bar{z}_{i_1} +_\flat \bar{z}_{i_2} +_\flat \cdots +_\flat \bar{z}_{i_{m-1}}$ and let $\vee$ denote logical or. For any given coordinate $l \in \{0, 1, \cdots, n-1\}$, either $\bar{u}[l] = 1$ or $\bar{u}[l] = 0$. If $\bar{u}[l] = 1$, then $(\bar{u} +_\flat \bar{z}_{i_m})[l] = \bar{u}[l] \vee \bar{z}_{i_m}[l] = 1 \vee \bar{z}_{i_m}[l] = 1$. If $\bar{u}[l] = 0$, then, by the nature of $+_\flat$ in $\{0,1\}^n$, $\bar{z}_{i_j}[l] = 0$ for $j = 1, 2, \cdots, m-1$ and hence, by (2.1), $\bar{z}_{i_m}[l] = 0$. Thus, $\bar{u}[l] = 0$ and (2.1) imply $(\bar{u} +_\flat z_{i_m})[l] = 0$. So, (2.1) implies $\bar{u} +_\flat \bar{z}_{i_m} = \bar{u}$ and thus $\mathcal{Z}$ is not sum-distinct in $(\{0,1\}^n, +_\flat)$. ∎

The converse of the above remark is not true. The 'smallest' example are the sets $\{11, 01\}$ and $\{11, 10\}$.

Since $\{0,1\}^n$ is closed with respect to $+_\flat$ and $\oplus$, it follows that $d_n \leq 1$ for all $n \in \mathcal{N}_2$ in both $(\{0,1\}^n, +_\flat)$ and $(\{0,1\}^n, \oplus)$. Clearly, $b(\mathcal{I}_n)$, where $\mathcal{I}_n \triangleq \{2^i \mid i = 0, 1, \cdots, n-1\}$, is the only sum-distinct set in $(\{0,1\}^n, +_\flat)$ for which $d_n = 1$. There are, however, many n-element sum-distinct sets in $(\{0,1\}^n, \oplus)$. One can generate them easily by noticing that if $\{\bar{z}_{i_1}, \bar{z}_{i_2}, \cdots, \bar{z}_{i_m}\}$, $m \leq n$, is sum-distinct in $(\{0,1\}^n, \oplus)$ and $\bar{v} = \bar{z}_{i_1} \oplus \bar{z}_{i_2} \oplus \cdots \oplus \bar{z}_{i_m}$, then $\{\bar{v}, \bar{z}_{i_1}, \bar{z}_{i_2}, \cdots, \bar{z}_{i_{m-1}}\}$ is sum-distinct in $(\{0,1\}^n, \oplus)$. Hence, $d^* = 1$ in both $(\{0,1\}^n, +_\flat)$ and $(\{0,1\}^n, \oplus)$.

3. Sum-distinct sets in $(\mathcal{S}_n, +)$

Let $c \in \mathcal{N}_0$ and let $\mathcal{B}_c$ be a class of sum-distinct sets from $(\mathcal{S}_n, +)$ such that $\mathcal{Z} \in \mathcal{B}_c$ if and only if $|\mathcal{Z}| = n + c$. That is,

$$\mathcal{B}_c \triangleq \{\mathcal{Z} \subset \mathcal{S}_n \mid c \in \mathcal{N}_0,\ |\mathcal{Z}| = n + c,\ \mathcal{Z} \text{ is sum} - \text{distinct}\}. \tag{3.1}$$

If $n < 3$, then $\mathcal{B}_1 = \emptyset$. If $n = 3$, only $\{3, 5, 6, 7\}$ is in $\mathcal{B}_1$. If $n > 3$, then $\mathcal{B}_1$ contains at least $(32\sqrt{2})^{-1} 2^n$ elements, [6]. Hence, if $n \in \mathcal{N}_3$, then $d_n \geq 1 + n^{-1}$ in $(\mathcal{S}_n, +)$. According to [2] and [5], there is a 23-element sum-distinct set in $(\mathcal{S}_{21}, +)$.

---

[2] The proof given below was suggested to author by Professor R.W. Gatterdam.

Lemma 1. Let $G(n) = 1 + n^{-1}\log_2(n + \log_2(1 + n + \log_2 n))$, $n \in \mathcal{N}_2$. Then, $d_n < G(n)$ in $(\mathcal{S}_n, +)$.

Proof. Let $\mathcal{Z}$ be a sum-distinct set from $(\mathcal{S}_n, +)$. Lemma follows from $|\mathcal{Z}| < 2^n$ by a repeated substitution of the upper bound for $|\mathcal{Z}|$ into the right-hand side of $2^{|\mathcal{Z}|} < |\mathcal{Z}|2^n$. That is, $2^{|\mathcal{Z}|} < |\mathcal{Z}|2^n$ and $|\mathcal{Z}| < 2^n$ yield $|\mathcal{Z}| < 2n$ which, together with $2^{|\mathcal{Z}|} < |\mathcal{Z}|2^n$, yields $|\mathcal{Z}| < 1 + n + \log_2 n$ and so on. ∎

Note that $G(n)$ monotonically decreases with $n \in \mathcal{N}_2$ and $G(2) = 2$. Hence, for any given $\alpha \in (1, 2]$, there exists $n_\alpha \in \mathcal{N}_2$ such that $G(n) < \alpha$ for $n > n_\alpha$. In particular, $n_{\frac{4}{3}} = 12$.

Theorem 1. $d^* = \frac{4}{3}$ in $(\mathcal{S}_n, +)$. The corresponding sum-distinct set is $\{3, 5, 6, 7\}$.

Proof. By using repeated substitution described in the above lemma, it is easy to show that $|\mathcal{Z}| < n + \log_2(1 + n + \log_2 n)$ for a sum-distinct set $\mathcal{Z} \in \mathcal{S}_n$. Hence, $|\mathcal{Z}| \leq 16$ for $n = 12$ and thus $d_{12} \leq \frac{4}{3}$. Let $\mathcal{B}_c \neq \emptyset$ for $n \leq 12$. Then $n^{-1}|\mathcal{Z}| = 1 + n^{-1}c \geq \frac{4}{3}$ implies $n \leq 3c$. Inspection shows that $d_3 = \frac{4}{3}$ for $\{3, 5, 6, 7\} \in \mathcal{B}_1$ and $\mathcal{B}_c = \emptyset$ for $c = 2, 3, 4$ and $n = 3c$. ∎

Inequality $2^{|\mathcal{Z}|-1} < 2^{n+1}\sqrt{|\mathcal{Z}|}$, derived from the result in [3, p. 137], provides an alternative way of proving the above theorem. The proof starts with the obvious $|\mathcal{Z}| < 2^n$ and uses a repeated substitution of the upper bound for $|\mathcal{Z}|$ in $\sqrt{|\mathcal{Z}|}$ to obtain a suitable upper bound for $|\mathcal{Z}|$.

Theorem 1 utilizes $|\mathcal{Z}| < n + k + k_1 \log|\mathcal{Z}|$, where $k \in \mathcal{N}_1$ and $k_1 \in R^+$, for a sum-distinct set $\mathcal{Z}$ from $(\mathcal{S}_n, +)$. The author believes this to be a poor estimate of $|\mathcal{Z}|$. A futile computer-search for sum-distinct sets whose cardinality meets the upper bound stated in Lemma 1 indicates that, perhaps, $k_1 = 0$. As poor as it is, this argument prompts the following conjecture.

Conjecture 1. There exist a finite positive integer k and a fixed positive integer n_0 such that $\mathcal{B}_c = \emptyset$ for all $c > k$ and all $n \geq n_0$.

4. Sum-distinct sets in $(\{0,1\}^n, +)$

A sum-distinct set in $(\{0,1\}^n, +)$ is also called a detecting set of vectors. This term comes from the 'coin weighing' problem, [7]. Sum-distinct sets in $(\mathcal{S}_n, +)$ and $(\{0,1\}^n, +)$ are related by the following observation.

Remark 2. A binary equivalent of a sum-distinct set from $(\mathcal{S}_n, +)$ is a sum-distinct set in $(\{0,1\}^n, +)$.

Proof. If $b(\mathcal{X}) = \{\bar{x}_1, \bar{x}_2, \cdots, \bar{x}_T\}$ is not sum-distinct in $(\{0,1\}^n, +)$, then there exist at least two subsets of $b(\mathcal{X})$, say $\{\bar{x}_{i_1}, \bar{x}_{i_2}, \cdots, \bar{x}_{i_m}\}$ and $\{\bar{x}_{j_1}, \bar{x}_{j_2}, \cdots, \bar{x}_{j_k}\}$ such that $(\bar{x}_{i_1} + \bar{x}_{i_2} + \cdots + \bar{x}_{i_m})[l] = (\bar{x}_{j_1} + \bar{x}_{j_2} + \cdots + \bar{x}_{j_k})[l]$ for $l = 0, 1, \cdots, n-1$. After multiplying this by 2^l and summing up we have

$\sum_{r=1}^{m} x_{i_r} = \sum_{r=1}^{k} x_{j_r}$ and thus $\mathcal{X} = \{x_i \mid b(x_i) = \bar{x}_i\}$ is not sum distinct in $(\mathcal{S}_n, +)$. ∎

The converse of the above remark is not true. For example, the set $\{b(2), b(3), b(5)\}$ has eight distinct sums, that is, 000, 010, 110, 101, 120, 111, 211, 221. By the above remark, a lower bound, $1 + n^{-1}$, for d_n in $(\{0,1\}^n, +)$ follows from $\mathcal{B}_1 \neq \emptyset$ for $n \in \mathcal{N}_3$. However, a much stronger result holds in $(\{0,1\}^n, +)$.

Theorem 2. Let $A(n)$, $n \in \mathcal{N}_1$, denote the number of 1's in the binary representation of the first n positive integers. Then, $n^{-1}A(n) \leq d_n < \log_2(1 + n + n^2)$, $n \in \mathcal{N}_2$, in $(\{0,1\}^n, +)$.

Proof. It has been shown (by effective design), [5], that for every $n \in \mathcal{N}_3$ there is a set $\mathcal{Z}$ which is sum-distinct in $(\{0,1\}^n, +)$ and $|\mathcal{Z}| \geq A(n)$. To establish the upper bound, note that $|\mathcal{Z}| < 2^n$ and $2^{|\mathcal{Z}|} < (1 + |\mathcal{Z}|)^n$. Hence, $|\mathcal{Z}| < n^2 + n$ and $2^{|\mathcal{Z}|} < (1 + n + n^2)^n$. This upper bound can be slightly improved by a repeated substitution of the upper bound for $|\mathcal{Z}|$ into $2^{|\mathcal{Z}|} < (1 + |\mathcal{Z}|)^n$. ∎

From [5, p. 479] one obtains $2^{|\mathcal{Z}|-n} \leq c^n |\mathcal{Z}|^{\frac{n}{2}}$ where $c < 4e$. That yields asymptotically stronger upper bound in Theorem 2, namely $d_n \leq k + \frac{1}{2}\log_2(\frac{1}{2}n^2 + kn)$, where $k < 3 + \log_2 e$. The estimate $A(n) = \frac{1}{2}n \log_2 n + 0(n)$ was announced in [1]. A more precise result may be found in [8].

Call a sum-distinct set $\mathcal{Z} = \{z_1, z_2, \cdots, z_T\}$, $z_i \in \mathcal{X}$, *full* in $(\mathcal{X}, +)$ if $\{z\} \cup \mathcal{Z}$ is not sum-distinct for any z from $\mathcal{X}$.

A full sum-distinct set does not necessarily have to be the maximum cardinality sum-distinct set, e.g. $\mathcal{I}_n$. However, the binary equivalent of a full sum-distinct set from $(\mathcal{S}_n, +)$ may or may not be a full sum-distinct set in $(\{0,1\}^n, +)$. For example, $\mathcal{I}_n$ and $b(\mathcal{I}_n)$ are full sum-distinct sets in $(\mathcal{S}_n, +)$ and $(\{0,1\}^n, +)$, respectively. However, $\{1, 3, 5\}$ is full in $(\mathcal{S}_3, +)$ while $\{b(1), b(3), b(5)\}$ is not in $(\{0,1\}^3, +)$. [3)]

Problem 1. Let $\mathcal{Z}$ be the maximum cardinality sum-distinct set from $(\mathcal{S}_n, +)$. Under what conditions is $b(\mathcal{Z})$ a full sum-distinct set in $(\{0,1\}^n, +)$?

Acknowledgment

The author wishes to thank the anonymous referee for several useful remarks.

References

[3)] One can verify that $\{b(1), b(3), b(5), b(6)\}$ is sum-distinct in $(\{0,1\}^3, +)$.

[1] R. Bellman and H.N. Shapiro, "On a problem in additive number theory," *Annals of Mathematics*, vol. 49, no. 2, pp. 333-340, April 1948.

[2] J.H. Conway and R.K. Guy, "Solution of a problem of P. Erdös", *Colloquium Mathematicum*, vol. 20, fasc. 2, p. 307, 1969.

[3] P. Erdös, "Problems and results in additive number theory", *Colloque sur la Theorie des Nombres*, Bruxelles, 1955, Liege and Paris, 1956, pp. 127-137, esp. p. 137.

[4] D.B. Jevtić, "Optimum multiple access coding, Problem 91-2," *SIAM Review*, vol. 33, no. 1, pp. 114-115, March 1991.

[5] B. Lindström, "On a combinatorial problem in number theory," *Canad. Math. Bull.*, vol. 8, no. 4, pp. 477-490, June 1965.

[6] P. Smith, "Problem E 2536", *Amer. Math. Monthly*, vol. 82, no. 3, p. 300, 1975. Solutions and comments, vol. 83, no. 6, p. 484, 1976.

[7] H.S. Shapiro, "Problem E 1399," *Amer. Math. Monthly*, vol. 67, no. 82, p. 697, Sept. 1960.

[8] M.D. McIlroy, "The number of 1's in binary integers: bounds and extremal properties," *SIAM J. on Computing*, vol. 3, no. 4, pp. 255-261, December 1974.