ARTICLE IN PRESS

Available online at www.sciencedirect.com

SCIENCE DIRECT®

ELSEVIER

Discrete Applied Mathematics ▮▮▮ (▮▮▮▮) ▮▮▮–▮▮▮

DISCRETE APPLIED MATHEMATICS

www.elsevier.com/locate/dam

# Sets in $\mathbb{Z}_n$ with distinct sums of pairs☆

Harri Haanpää[a], Antti Huima[a], Patric Östergård[b]

[a] *Department of Computer Science and Engineering, Helsinki University of Technology, P.O. Box 5400, FIN-02015-HUT, Finland*

[b] *Department of Electrical and Communications Engineering, Helsinki University of Technology, P.O. Box 3000, FIN-02015-HUT, Finland*

Received 8 November 2000; received in revised form 2 April 2002; accepted 27 February 2003

**Abstract**

A subset $S = \{s_1, \ldots, s_k\}$ of an abelian group $G$ is called an $S_t$-set of size $k$ if all sums of $t$ different elements in $S$ are distinct. A function with applications in coding theory, $v_\gamma(k)$ denotes the order of the smallest cyclic group in which an $S_2$-set of size $k$ exists. A lower bound for $v_\gamma(k)$ is given in this study, and exact values of $v_\gamma(k)$ are obtained for $k \leqslant 15$. For the related problem in which all sums of any two, not necessarily distinct, elements in $S$ are required to be different, values of the corresponding function $v_\delta(k)$ for each $k \leqslant 14$ are given.


*Keywords:* Additive base; Backtrack search; Difference set; Packing

## 1. Introduction

This work considers packing problems in cyclic groups. A subset $S$ of an abelian group where $|S| = k$ is an $S_t$-set of size $k$ if all sums of $t$ different elements in $S$ are distinct in the group. For the group operations, we use additive notation throughout the paper. See [6,7] for open problems in additive number theory related to $S_t$-sets and similar configurations.

Two central functions in the study of $S_t$-sets are $v(k)$ and $v_\gamma(k)$, which give the order of the smallest abelian group and cyclic group, respectively, in which an $S_2$-set of size $k$ exists. Since cyclic groups are abelian, clearly $v(k) \leqslant v_\gamma(k)$.

☆ Supported by the Academy of Finland under grants 44517 and 100500.

*E-mail addresses:* harri.haanpaa@hut.fi (H. Haanpää), antti.huima@hut.fi (A. Huima), patric.ostergard@hut.fi (P. Östergård).

0166-218X/03/$ - see front matter © 2003 Elsevier B.V. All rights reserved.
doi:10.1016/S0166-218X(03)00273-7

One motivation for studying $v(k)$, $v_\gamma(k)$, and $S_t$-sets is that they have applications in coding theory [3,5,6]. A constant weight error-correcting code is a set of binary vectors of length $k$ and weight $w$ such that the Hamming distance between any two vectors is at least $d$. Given $k$, $w$, and $d$, the maximum size of such a code is denoted by $A(k,d,w)$. In [5, Theorem 16] it is shown that $A(k,6,w) \geqslant \binom{k}{w} / v(k)$.

In this paper we consider $v_\gamma(k)$, which is called $v_{\mathbb{Z}}(k)$ in [5]. In [6], values of $v_\gamma(k)$ and corresponding $S_2$-sets are presented for $k \leqslant 10$. The (computer-aided) methods for obtaining these results are, however, not discussed in [6]. We present an exhaustive search method for determining $v_\gamma(k)$ and finding corresponding $S_2$-sets. We determine $v_\gamma(k)$ for $k \leqslant 15$.

With a slight modification, our method allows us to also consider $v_\delta(k)$, the order of the smallest cyclic group in which there exists a $k$-subset such that the sums of any two elements of the subset, not necessarily different, are distinct. The values of $v_\delta(k)$ and corresponding sets are given for $k \leqslant 10$ and $k=12$ in [6]. In [12], constructions are developed to determine lower bounds for the maximum cardinality of such a subset in a given cyclic group $\mathbb{Z}_n$; also, the maximum cardinality is determined by a computer search for $n \leqslant 134$, whereby $v_\delta(k)$ is determined for $k \leqslant 12$. Using our method we determine $v_\delta(k)$ and the corresponding subsets for $k \leqslant 14$.

Some theoretical results on $v_\gamma(k)$ and $v_\delta(k)$ are presented in Section 2, and a computational method for finding their values is discussed in Section 3. The search results are documented in Section 4. We list the values of $v_\gamma(k)$ for $k \leqslant 15$ and $v_\delta(k)$ for $k \leqslant 14$ along with the corresponding $k$-element sets.

## 2. Theoretical results

In this section, we review some lower bounds for $v_\delta(k)$ and present a lower bound for $v_\gamma(k)$. These bounds can be used to limit the scope of the computer search.

An $(n,k,\lambda)$-difference set is a $k$-subset of a group of order $n$ such that the list of non-zero differences contains each non-zero group element exactly $\lambda$ times. When the group is cyclic, we have a cyclic difference set [9]. The problems we are considering are closely connected to the theory of difference sets with $\lambda=1$, since $x_1+x_2=x_3+x_4 \Leftrightarrow x_1-x_4=x_3-x_2$. When, in calculating $v_\delta(k)$, we allow taking sums of equal elements, the relation is straightforward and the theory of cyclic difference sets is directly applicable.

By a simple volume argument, we know that $v_\delta(k) \geqslant k(k-1)+1$: from a $k$-subset, $k(k-1)$ differences can be formed, none of which is zero and no two of which may be equal. When $k-1$ is a prime power, a cyclic difference set construction by Singer shows that the bound is sharp [11]. The series of values of $k$ that are not settled by this result is thus $7, 11, 13, 15, 16, \ldots$. Also, when $k$ is a prime power, another cyclic difference set construction by Bose [2] shows that $v_\delta(k) \leqslant k^2 - 1$. In [6] it is shown that $v_\delta(7) = 48$, and our calculations show that the bound by Bose is sharp also for $k = 11$ and 13.

We next present a volume argument for $v_\gamma(k)$. In that case some values of differences may occur more than once, since if $x_3 - x_2 = x_2 - x_1$, then the right-hand side of

$x_1 + x_3 = x_2 + x_2$ is not the sum of two distinct elements, and thus does not contradict the assumption of distinct sums of pairs. Our idea for adapting the volume argument for $v_\gamma(k)$ involves finding upper bounds for the number of difference values that may occur more than once.

Let $S = \{s_1, \ldots, s_k\}$ be a set with distinct sums of pairs in a cyclic group $\mathbb{Z}_n$. For each $d \neq 0$ in $\mathbb{Z}_n$, we construct the undirected graph $G_d$ by taking as the edge set $E(G_d)$ the set of those pairs of elements in $S$ whose difference equals $\pm d$, and as the vertex set $V(G_d)$ the endpoints of the edges.

Following the usual convention, let $P_i$ denote a path of length $i$, and $C_i$ a cycle of length $i$.

**Lemma 1.** *Every non-empty graph $G_d$ is isomorphic to $P_1$, $P_2$, or $C_3$.*

**Proof.** By definition, each vertex of $G_d$ has degree at least one. Each vertex of $G_d$ has degree at most two: no vertex $s$ can be adjacent to a vertex other than $s + d$ or $s - d$. Every pair of edges $\{s_{i_1}, s_{i_2}\}$, $\{s_{i_3}, s_{i_4}\}$ in $G_d$, ordered so that $s_{i_1} - s_{i_2} = s_{i_3} - s_{i_4}$, must have a common endpoint, or $s_{i_1} + s_{i_4} = s_{i_2} + s_{i_3}$ will contradict the assumption of distinct sums of pairs. Since every vertex has degree at least one, and every pair of edges must have an endpoint in common, the graph must be connected. A connected graph with highest degree at most two can only be a path or a cycle. Paths of more than two edges and cycles with more than three edges contain a pair of edges without a common endpoint. Hence, if $G_d$ is non-empty, it must be isomorphic to $P_1$, $P_2$, or $C_3$. □

Let $p$ be the number of graphs $G_d$ isomorphic to $P_2$, and $c$ the number of graphs $G_d$ isomorphic to $C_3$.

**Lemma 2.** $p + 3c \leqslant 2k$.

**Proof.** The vertices that correspond to a given element of $S$ can have degree two in at most two graphs $G_d$. Otherwise, among the (at least three) graphs there would have to be two graphs, say $G_d$ and $G_{d'}$, where $d \neq -d'$. Then $\{s - d', s - d, s + d, s + d'\}$ would be a set of four distinct elements in $S$, and the assumption of distinct sums of pairs would be contradicted by $(s - d') + (s + d') = (s - d) + (s + d)$. From this, and noting that the graphs $G_d$ isomorphic to $P_2$ have one vertex of degree two whereas those isomorphic to $C_3$ have three, the lemma follows. □

**Theorem 3.** $v_\gamma(k) \geqslant k(k - 3)$.

**Proof.** Let us calculate an upper bound for the total number of times the different values may occur as the difference of two elements of $S$. The difference value $d$ occurs once if $d \neq -d$ and $G_d$ is isomorphic to $P_1$; twice if either $G_d$ is isomorphic to $P_1$ and $d = -d$, which is possible for at most one value in $\mathbb{Z}_n$, or if $G_d$ is isomorphic to $P_2$; and three times if $G_d$ is isomorphic to $C_3$. The number of difference values that occur at least once is then at most $n - 1$, the number of difference values occurring at

least twice is at most $1 + p + c$, and the number of difference values occurring three times is $c$. The total, then, is at most $n + p + 2c$. From a $k$-element subset, $k(k-1)$ differences can be formed; thus $k(k-1) \leqslant n + p + 2c = n - c + p + 3c \leqslant n + 2k$ (by Lemma 2), and the theorem follows. □

This bound has the same asymptotic behavior as a similar bound in [6, Lemma 6], but it is slightly stronger.

## 3. Computer search

### 3.1. A backtrack algorithm

The natural way to approach this problem is to consider backtrack algorithms. In a backtrack search, there are two main ways of pruning the search tree. First, a search branch may be pruned as soon as it is clear that a configuration cannot lead to a desired solution, and, second, of search branches leading to equivalent configurations one may prune all but one.

Before the search, we fix $n$, the order of the cyclic group. For equivalence of sets, we use the following definition, which is used in [1] for difference sets. Two sets $S_1, S_2 \subseteq \mathbb{Z}_n$ are *equivalent* if $S_2 = S_1 a + b := \{sa + b \,|\, s \in S_1\}$, where $b$ is any group element and $a$ is co-prime with $n$. For the purposes of canonicity testing, $\mathbb{Z}_n$ is to be interpreted as a ring, i.e., addition and multiplication are carried out modulo $n$. It is not difficult to see that if all sums of pairs in $S_1$ are different, then this also holds for all sums of pairs in $S_2$. Hence, only one set in each equivalence class needs to be considered in the search. Since a given set may have as many as $n\phi(n)$ different equivalent sets, where $\phi$ is the Euler totient function, a considerable speedup may be achieved if equivalence testing can be carried out fast.

All one-element sets are clearly equivalent, so we may start from $\{0\}$. Throughout the search, we only need examine the search branches that correspond to the set $S'$ if no set equivalent to $S'$ has been considered earlier in the search. Since the algorithm searches all possible sets in lexicographical order, this is the case precisely when $S'$ is the lexicographically first member of its equivalence class. Such a set is called the canonical representative of its equivalence class. Pruning the search when sums of pairs occur more than once or when $S'$ is not in canonical form we get the following algorithm. At line 2, if a bound such as Theorem 3 shows that no $S_2$-set with $|S| + 1$ elements can exist, one may stop the search.

```
function recsearch(S, p)
if |S| > maxk then
   larger distinct-sum set found; store; maxk ← |S|
end if
for i = p to n − 1 do
   S′ ← S ∪ {i}
```

```
    if sums of pairs occur at most once then
        if S' is a canonical representative then
            recsearch(S', i + 1)
        end if
    end if
end for
return

function search
maxk ← 1
recsearch({0}, 1)
return
```

Our backtrack search method with isomorph pruning is an *orderly algorithm* [10]. Any set $S$ with at least two elements that is in canonical form may be reached in the search from another set in canonical form. Thus, the algorithm searches all sets with distinct sums of pairs that are in canonical form.

The implementation of the two tests in lines 7 and 8 of our algorithm is crucial for its performance. These will be discussed in the following subsections.

### 3.2. Testing occurrences of sums

In order to check whether an element $c$ can be added to $S$ without violating the distinct sums condition, we need to calculate the new sums that are possible to obtain after adding the element $c$ to the set $S$ and verify that none of them are obtainable from the elements of $S$ prior to adding $c$ to the set. When calculating $v_\gamma(k)$, we restrict ourselves to sums of two distinct elements; in that case the new sums are obtained by adding $c$ to each element of $S$. In calculating $v_\delta(k)$ we also have sums of two identical elements, and we need to take also the additional new sum $c + c$ into account.

As for testing occurrences of sums, we obtained the best performance with bitmask representations. These were also used recently in a search for Golomb rulers [4], another similar problem in additive number theory.

In the bitmask representation, the search set $S$ is represented as a bit sequence that contains $n$ bits. The $i$th bit is 1 if and only if $i \in S$ (the 0th bit is leftmost). Similarly, the set of sums that can be composed from the elements in $S$ is represented as $a$ $v$-bit sequence with 1s in the corresponding positions.

The crucial point of this representation is that if $S$ is represented as a bit sequence $s$, the set $S + c$ is represented as $s \ggg c$, where $\ggg$ denotes right rotation by $c$ bits, allowing us to calculate very quickly the new sums introduced by adding the element $c$ to the set $S$.

Assume that all sums of pairs in $S$ are different on line 6 of our algorithm. Let $s$ be the bit sequence representing $S$ and $u$ the sequence representing the sums that can be composed from the elements of $S$. To perform the test in line 7, it is enough to calculate

$$(s \ggg i) \text{ AND } u, \tag{1}$$

where the AND operation is carried out bitwise. This can be done very fast. If the value of (1) is 0, then the test is passed. In calculating $v_\delta(k)$, the new sums are most conveniently calculated by substituting the bitmap $s'$, corresponding to the set $S'$, for $s$ in (1).

Another benefit of the bitmask representation is that it can be used to compare the lexicographical ordering of two sets very efficiently, also when several machine words are used for their representation in the actual implementation.

### 3.3. Testing canonicity

The efficiency of our algorithm very much depends on efficient testing of canonicity of $S'$ in line 8. The most straightforward, but also the most time-consuming way to do this is to compute all $n\phi(n)$ equivalent sets corresponding to the allowable $\phi(n)$ different values of $a$ and $n$ different values of $b$, find the lexicographically first one of them and compare it to $S'$. To speedup this direct approach, we may use some additional information that we have about the canonical representative.

We know that 0 is always in the canonical representative of any non-empty set. For any set with at least two elements, the second smallest element in the canonical representative is $c = \min_{x_1,x_2 \in S'} \gcd(x_2 - x_1, n)$. It is therefore sufficient to only check such choices of $a$ and $b$ that $f(x) = (ax + b) \bmod n$ maps a pair $(x_1, x_2)$ in $S'$ to $(0, c)$. The canonicity test is then performed by finding all such choices of $a$ and $b$, applying the corresponding mapping to $S'$, and checking whether the result lexicographically precedes $S'$. In this manner it is only necessary to compare $\mathcal{O}(k^2)$ sets. Of course, if one of the mappings gives a set that lexicographically precedes $S'$, we may stop testing right away—we already know that $S'$ is not in canonical form.

## 4. Computational results

We implemented the algorithm described in Section 3 in C. For best possible performance, canonicity testing was not performed on all levels in the search tree. For example, with $k = 12$, the test should be done with at most about 7 (depending somewhat on the value of $n$) elements in $S'$.

Generally, the program finds the maximum subset with distinct sums fairly quickly. Verifying that no larger subset with distinct sums exists takes significantly longer. To determine the value of $v_\gamma(k)$, and analogously $v_\delta(k)$, we have to calculate a lower bound for $v_\gamma(k)$ (e.g., from Theorem 3), and run our algorithm for all consecutive values of $n$ starting from the lower bound until we find a packing that gives the value of $v_\gamma(k)$. This holds only for the modular versions of packing and covering problems, whereas, for example, the Golomb ruler problem can be solved in a more direct way (see [6, Lemma 4] and the comments thereafter).

We used the program to compute $v_\gamma(k)$ for $k \leqslant 15$. Our new results on $v_\gamma(k)$ for $11 \leqslant k \leqslant 15$ improve on the bounds given for $v(k)$ in [3, Table 5]. Our results on $v_\gamma(k)$ for $11 \leqslant k \leqslant 15$ are thus also the best-known upper bounds on $v(k)$. The results are presented in Table 1. We similarly calculated the values of $v_\delta(k)$ and the corresponding

Table 1
Values of $v_\gamma(k)$ and the corresponding sets for $k \leqslant 15$

| $k$ | $v_\gamma(k)$ | The corresponding sets |
|---|---|---|
| 1 | 1 | $\{0\}$ |
| 2 | 2 | $\{0,1\}$ |
| 3 | 3 | $\{0,1,2\}$ |
| 4 | 6 | $\{0,1,2,4\}$ |
| 5 | 11 | $\{0,1,2,4,7\}$ |
| 6 | 19 | $\{0,1,2,4,7,12\}$ |
| 7 | 28 | $\{0,1,2,4,8,15,20\}$, |
| | | $\{0,1,2,5,9,17,23\}$ |
| 8 | 40 | $\{0,1,5,7,9,20,23,35\}$ |
| 9 | 56 | $\{0,1,2,4,7,13,24,32,42\}$ |
| 10 | 72 | $\{0,1,2,4,7,13,23,31,39,59\}$ |
| 11 | 96 | $\{0,1,2,4,10,16,30,37,50,55,74\}$, |
| | | $\{0,1,2,4,11,21,40,52,70,75,83\}$, |
| | | $\{0,1,2,4,13,26,34,40,50,55,78\}$, |
| | | $\{0,1,2,4,16,22,27,35,52,59,69\}$ |
| 12 | 114 | $\{0,1,4,14,22,34,39,66,68,77,92,108\}$ |
| 13 | 147 | $\{0,1,2,4,7,29,40,54,75,88,107,131,139\}$ |
| 14 | 178 | $\{0,1,2,4,16,51,80,98,105,111,137,142,159,170\}$ |
| 15 | 183 | $\{0,1,2,14,18,21,27,52,81,86,91,128,139,161,169\}$ |

Table 2
Values of $v_\delta(k)$ and corresponding sets for $k \leqslant 15$

| $k$ | $v_\delta(k)$ | Lexicographically first set |
|---|---|---|
| 1 | 1 | $\{0\}$ |
| 2 | 3 | $\{0,1\}$ |
| 3 | 7 | $\{0,1,3\}$ |
| 4 | 13 | $\{0,1,3,9\}$ |
| 5 | 21 | $\{0,1,4,14,16\}$ |
| 6 | 31 | $\{0,1,3,8,12,18\}$ |
| 7 | 48 | $\{0,1,3,15,20,38,42\}$ |
| 8 | 57 | $\{0,1,3,13,32,36,43,52\}$ |
| 9 | 73 | $\{0,1,3,7,15,31,36,54,63\}$ |
| 10 | 91 | $\{0,1,3,9,27,49,56,61,77,81\}$ |
| 11 | 120 | $\{0,1,3,20,31,35,45,53,58,74,114\}$ |
| 12 | 133 | $\{0,1,3,12,20,34,38,81,88,94,104,109\}$ |
| 13 | 168 | $\{0,1,3,11,30,34,46,83,103,108,121,147,162\}$ |
| 14 | 183 | $\{0,1,3,16,23,28,42,76,82,86,119,137,154,175\}$ |

lexicographically first maximum sets for $k \leqslant 14$. Our results are summarized in Table 2. The value of $v_\delta(13)$ is new. The $k$-element subsets given in Tables 1 and 2 are unique up to equivalence, with the exceptions of $v_\gamma(7)$ and $v_\gamma(11)$, for which the lexicographically first $k$-subsets of each equivalence class is given.

The maximum subsets of cyclic groups of small order may be obtained electronically from the WWW page ⟨http://www.tcs.hut.fi/∼haha/Zn/⟩.

The computations were run in a heterogeneous, non-dedicated network of PCs, using the *autoson* distributed batch system [8]. For determining $v_\gamma(15)$, Theorem 3 gives us the lower bound $v_\gamma(15) \geqslant 180$. For each $180 \leqslant n \leqslant 182$ an exhaustive search shows that $v_\gamma(15) \neq n$, as $\mathbb{Z}_n$ contains no 15-element subset. Each of these searches takes an estimated two to three days on a 1.4 GHz AMD Athlon PC. It takes 15 min to find the 15-element subset of $\mathbb{Z}_{183}$ listed in Table 1, which shows that $v_\gamma(15) \leqslant 183$, and by combining these results we get $v_\gamma(15) = 183$.

Verifying that the values given in Tables 1 and 2 are upper bounds of $v_\gamma(k)$ and $v_\delta(k)$ is straightforward. Verifying that they are minimal is much harder. However, for $v_\gamma(k)$, two independent implementations gave the same results for $k \leqslant 13$. Additionally, our computations confirm the computational results in [12] and the computational results on $v_\gamma(k)$ and $v_\delta(k)$ in [5].

While for most values of $k$ the best-known upper bound for $v(k)$ is derived from a subset in a cyclic group, for $k \in \{6, 7, 9\}$ the bounds in [3] show that for those values $v(k) < v_\gamma(k)$. Exhaustively searching abelian groups of small order for maximum subsets with distinct sums of pairs would be a natural way to continue this research.

## Acknowledgements

The authors wish to thank professor Alexander Pott for his most helpful comments.

## References

[1] L.D. Baumert, Cyclic Difference Sets, in: Lecture Notes in Mathematics, Vol. 182, Springer, Berlin, 1971.
[2] R.C. Bose, An affine analogue of Springer's theorem, J. Indian Math. Soc. 6 (1942) 1–15.
[3] A.E. Brouwer, J.B. Shearer, N.J.A. Sloane, W.D. Smith, A new table of constant weight codes, IEEE Trans. Inform. Theory 36 (1990) 1334–1380.
[4] A. Dollas, W.T. Rankin, D. McCracken, A new algorithm for Golomb ruler derivation and proof of the 19 mark ruler, IEEE Trans. Inform. Theory 44 (1998) 379–382.
[5] R.L. Graham, N.J.A. Sloane, Lower bounds for constant weight codes, IEEE Trans. Inform. Theory 26 (1980) 37–43.
[6] R.L. Graham, N.J.A. Sloane, On additive bases and harmonious graphs, SIAM J. Algebraic Discrete Methods 1 (1980) 382–404.
[7] R.K. Guy, Unsolved Problems in Number Theory, 2nd Edition, Springer, New York, 1994.
[8] B.D. McKay, Autoson—a distributed batch system for UNIX workstation networks (version 1.3), Technical Report TR-CS-96-03, Department of Computer Science, Australian National University, 1996.
[9] A. Pott, Finite Geometry and Character Theory, in: Lecture Notes in Mathematics, Vol. 1601, Springer, Berlin, 1995.
[10] R.C. Read, Every one a winner *or* How to avoid isomorphism search when cataloguing combinatorial configurations, Ann. Discrete Math. 2 (1978) 107–120.
[11] J. Singer, A theorem in finite projective geometry and some applications to number theory, Trans. Amer. Math. Soc. 43 (1938) 377–385.
[12] C.N. Swanson, Planar cyclic difference packings, J. Combin. Des. 8 (2000) 426–434.