Behavioral Patterns of the Index of Non-Repetitive Sequences

Chanae Ottley
Mentor: Christopher Plyley, Ph.D.
University of the Virgin Islands

Grant Number:
NSF HBCU-UP Scholars SURE: 1137472

An Unsolved Problem in Combinatorial Number Theory

In 2011 the following question was asked in [G]:

> *What is the smallest integer value k such that every sequence of k elements from Z_n contains a zero-sum subsequence with index 1?*

This constant is denoted by T(n). Its determination and general behavior is currently unknown. This open problem is an example of a *zero-sum* problem in combinatorial number theory.

Zero-sum Sequences

Our object of study is a sequence of integers in Z_n (the group of integers modulo n), where the order of the sequence is disregarded. If a sequence of (not necessarily distinct) integers sums to 0 (mod n), then the sequence is called a **zero-sum sequence**.

Example

Let n=7, so that Z_7= {0,1,2,3,4,5,6}. Then the sequence S=(2,3,4,5) is a zero-sum sequence since 2+3+4+5 = 0 (mod 7).

Zero-sum problems study conditions which ensure that a given sequence will have a zero-sum subsequence with additional prescribed properties. More generally, one considers sequences of elements from an arbitrary finite abelian group. For example, the determination of the **Davenport constant** has become one of the most important unsolved problems in group theory since first being posed by Baayen, Erdos and Davenport in 1967.

Definition

For any group G, the smallest integer value k such that every sequence of k elements from G has a zero-sum subsequence is called the Davenport Constant of G.

If we disallow the repetition of elements in the sequence, the study of sequences in Z_n is equivalent to the study of the subsets of Z_n. In this case, the **Olson Constant** is the analog of the Davenport Constant.

Definition

For any group G, the smallest integer value k such that every sequence of k distinct elements from G has a zero-sum subsequence is called the Olson Constant of G (denoted Ol(G)).

Determining the value of Ol(G) is currently an open problem, and the subject of much research. Precise values are known only for very special examples of G.

Index of a Sequence

In 1988, Daniel Kleitman (MIT) and Mark Lemke (Minnesota) introduced the invariant known as the **index** [KL]. If S is a zero-sum sequence in Z_n, then elements of S sum to some multiple of n. The index of S is defined as the smallest integer multiple of n that S can be made to sum to after transformation by group automorphism (such transformed sequences are called **equivalent**). In Z_n a group automorphism is equivalent to multiplication by an integer relatively prime to n.

Definition

Let $S=(a_1,\dots, a_r)$ be a zero-sum sequence in Z_n. The index of S is the minimal possible integer k such that $ma_1 \pmod n + \dots + ma_r \pmod n = kn$, where m is any integer relatively prime to n.

Most research on index centers around finding conditions which ensure that a sequence will contain a zero-sum subsequence with index 1.

Example

In Z_8, the sequence S=(4,5,7) has index 1, since it is equivalent to the sequence 5S = (4,1,3), which has sum 8. The sequence T=(1,4,5,6), has index 2, since each of the equivalent sequences 3T = (3,4,7,5) 5T=(5,4,1,6) and 7T=(7,4,3,2) have sum 16.

Bounds on T(n)

We can obtain upper and lower bounds for T(n) with some observations.

Upper Bound

$T(n) \le \frac{n}{2} + 1$ for every value of n.

Proof: We partition the integers of Z_n into $\frac{n}{2}$ sets of elements each with sum equal to n in the obvious way: (1,n-1), (2,n-2),….($\frac{n}{2}$). If a sequence in Z_n has more than $\frac{n}{2}$ elements, then by the pigeonhole principle it must contain at least two elements from one of these sets, and hence a zero-subsequence of index 1.

Lower Bound

$T(n) \ge Ol(Z_n)$, for every value of n.

Proof: By the minimality of the Olson constant, we know that there exists a sequence of length $Ol(Z_n)$-1 which has no zero-sum subsequence, hence no subsequence of index 1.

Methodology

To determine T(n) for each Z_n we select a length k and search every sequence of length k for a subsequence of index 1. If we are able to find a sequence with no subsequence of index 1, we proceed to the next value of k. Thus, our general procedure is to, for a given value of n, to select a length of sequence within our bounds, and then take the following steps:

STEP 1 – Utilise a computer program to generate a list of all sequences.

- Sage Math Cloud (https://cloud.sagemath.com).
- Determines the number of sequences needed and generates a list of all such sequences.

STEP 2 - Eliminate as many sequences as possible which contain a subsequence of index 1.

- Eliminate every sequence with a subsequence that sums to exactly n;
- Eliminate every sequence not containing 1:
  - If n is prime, every sequence is equivalent to a sequence containing 1;
  - If n is composite, handle the scenario where no element in the sequence is coprime to n, then suppose that our sequence contains a 1.

STEP 3 - Check remaining sequences by hand to see if they have subsequences of index 1.

- Use some basic ideas and known results to identify sequences of index 1.
  - Any zero-sum sequence of length 3 must have index 1.
  - If n is prime, then any zero-sequence of length 4 has index 1 [L].

Note that the number of sequences increase rapidly as n increases; for example, there are over 125,000 sequences of length 8 in Z_{20} and over 12 billion sequences of length 13 in Z_{40}.

Results

We determined the minimal length k so that every sequence in Z_n of length k contains a subsequence of index 1 for all n<42. These values are shown below and represent the highest known calculated values of the constant T(n).

In order to demonstrate that T(n)=k we had to demonstrate the following:

- Every sequence in Z_n of length k contains a subsequence of index 1;
- There exists at least one sequence in Z_n of length k-1 with no subsequence of index 1.

We call the latter sequences "counterexamples" and they represent the maximal length sequences of index 2 or higher for each n, and are of potential interest. The counterexamples can be used to search for patterns to improve bounds, or studied in their own right. For example, we can demonstrate the following fact using our counterexample from n=20 to show the following:

If n>16 is even, then the sequence $(1,3,\frac{n}{2},\frac{n}{2}+1,\frac{n}{2}+3,n-6,n-2)$ has no zero-subsequence of index 1.

The list of all values of T(n) with our counterexample sequences are given. When viewed in conjunction with our predetermined upper and lower bounds, we can observe that the value of T(n) appears to remain closer to the value of $Ol(Z_n)$ than our upper bound.

n	T(n)	Counterexample
2	2	(1)
3	2	(1)
4	3	(1,2)
5	3	(1,2)
6	4	(1,3,4)
7	4	(1,2,3)
8	5	(1,4,5,6)
9	5	(2,3,5,8)
10	5	(1,5,6,8)
11	6	(1,3,4,5,9)
12	6	(1,2,6,7,8)
13	6	(1,2,6,8,9)
14	7	(1,2,6,7,9,10)
15	6	(1,2,3,4,5)
16	7	(1,8,9,12,13,14)
17	7	(1,2,5,6,7,13)
18	8	(1,2,9,10,11,12,13)
19	7	(1,2,3,9,11,12)
20	8	(1,2,3,9,11,12)
21	8	(1,2,7,8,9,15,16)
22	8	(1,2,3,10,13,14,15)
23	8	(1,2,14,16,17,18,19)
24	10	(1,4,5,9,12,13,16,17,21)
25	8	(1,2,5,6,7,8,21)

n	T(n)	Counterexample
26	9	(1,2,3,13,14,15,16,17)
27	9	(1,2,9,10,11,12,19,20)
28	10	(1,2,7,8,9,14,15,22,23)
29	9	(1,3,4,13,14,17,19,20)
30	10	(1,2,3,4,15,16,17,18,19)
31	9	(1,3,4,9,12,13,25,29)
32	11	(1,2,8,9,10,16,17,18,25,26)
33	10	(1,2,3,11,12,14,23,24,25)

n	T(n)	Counterexample
34	11	(1,2,3,4,17,18,19,20,21, 22)
35	11	(1,5,6,10,11,15,16,21,26,31)
36	11	(1,2,3,4,18,19,20,21,22,24)
37	9	(1,2,3,4,5,6,7,8)
38	11	(1, 2, 3, 4, 19, 20, 21, 22, 23, 24)
39	12	(1,2,3,13,14,15,16,27,28,29,30)
40	13	(1, 5, 6, 11, 16, 20, 21, 25, 26, 30, 31, 36)
41	10	(1, 2, 3, 13, 14, 15, 16, 29, 30)

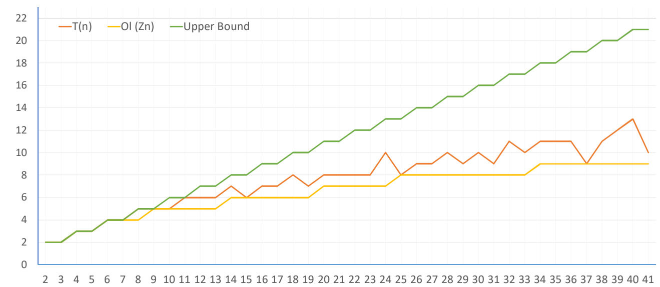

Figure 1. Values of T(n) along with the bounds derived above.

Conclusions

Our data suggest that the value of T(n) is likely to remain relatively close to the value of $Ol(Z_n)$ as n increases, however the value of T(n) appears to behave more sporadically than that of $Ol(Z_n)$. This supports the idea that sequences of index higher than 1 are relatively sparse in sequences which contain multiple zero-sums. It suggests that once a sequence is long enough to ensure zero-sum sequences, sequences without a subsequence of index 1 are relatively rare.

Further, it appears that the prime factors of n influence the value of T(n). When n is sufficiently large, we see general increases in the value of T(n) when n has a high number of prime factors (i.e. n=24, n=32) and decreases in the value of T(n) when n is prime (n=31, 37). This is likely related to the relative size of the automorphism group of Z_n, which is largest when n is prime.

Future Work

1. Determine the value of T(n) for all n less than 100 and compare with the known values of the $Ol(Z_n)$. This would likely require high powered computing and more advanced results on the index of sequences.
2. Improve the bounds on the value of T(n). For example, find a constant $c<\frac{1}{2}$ so that for sufficiently large n, T(n) <cn or (analogous to Olson constant) a constant d so that $T(n) < d\sqrt{n}$.
3. Derive a function which expresses/bounds T(n) in terms of $Ol(Z_n)$.
4. Allow for repetition in the sequence. In other words, let t(n) be the smallest integer k such that every sequence of length k contains a subsequence a zero-sum subsequence of index 1, and find t(n).

References

[G] W. Gao, Y. Li, J. Peng, C. Plyley, and G. Wang, *On the index sequences in cyclic groups*. Acta Arithmetica, 148(2), (2011) 119-134.
[KL] D.J. Kleitman and P. Lemke, *An addition theorem on the integers modulo n*. J. Number Theory, 31 (1989) 335-345.
[L] Y. Li, C. Plyley, P. Yuan and X. Zeng, *Minimal zero sum sequences of length four over finite cyclic groups*. Journal of Number Theory, 130 (2010) 2033 – 2048
[O] Olson, J. E. *An addition theorem modulo p*, J. Combin. Theory, 5 (1968) 45–52.
[S] Subocz G., Julio C., *Some Values of Olson's Constant*. Divulgaciones Mathemáticas, 8(2), (2000) 121-128.

Acknowledgements

This project was completed with the support of mentor Chris Plyley. I would like to thank the University of the Virgin Islands, the Emerging Caribbean Scientists Program, the National Science Foundation and the Summer Undergraduate Research Experience Program for the opportunity to be a mathematician and for funding me over this summer. Also, special thanks to Israel Dennie for assisting with coding of the computer program that allowed us to reduce the lists of sequences.