CRD

COMMUNICATIONS RESEARCH DIVISION WORKING PAPER

DIFFERENCE COVERINGS AND A SEQUENCE PROBLEM

W.H. Mills and D.H. Wiedemann

August 1987

S–230,074

IDA-CRD Log No. 81999

DIFFERENCE COVERINGS AND A SEQUENCE PROBLEM

W.H. Mills and D.H. Wiedemann

ABSTRACT

We study the least possible density that a 0-1 sequence can have subject to the property that a coincidence (a pair of corresponding ones) between it and any left shift of itself occurs at least once in any w consecutive positions. A difference covering modulo m is a non-empty set of residues such that every non-zero residue modulo m can be written in at least one way as the difference of two residues in the set. In all cases where the answer to our sequence problem is known, an optimal sequence can be constructed using a minimum difference covering modulo m for some $m \leq w$. Using the computer we list minimum difference coverings modulo m for $m \leq 110$ and prove that for all $w \leq 65$ this list contains optimal solutions to the original problem.

DIFFERENCE COVERINGS AND A SEQUENCE PROBLEM

W.H. Mills and D.H. Wiedemann

1. Introduction and Summary

The Problem. For a given integer $w \geq 1$, what is the least density of an infinite sequence of 0's and 1's, $x_1, x_2, \ldots$ with the property described below?

The Window Property. If the product of the sequence with any left shift of itself does not have w consecutive 0's we will say that it has the w-window property. More precisely, for all $t, j \geq 0$,

$$\sum_{i=t+1}^{i=t+w} x_i x_{i+j} \geq 1.$$

We postpone the actual definition of density until the next section. It will be shown that we can restrict ourselves to periodic sequences, and it is obvious that the density of a periodic sequence with period m should be defined to be the number of 1's in a period divided by m.

Good examples of low density sequences with the window property can be constructed using *simple cyclic difference sets* [1]. A simple cyclic difference set for a modulus $m \geq 1$ is a non-empty set of residues modulo m, $\{r_0, \ldots, r_k\}$ with the property that any non-zero residue modulo m can be uniquely written as $r_i - r_j$ for residues in the set. An easy counting argument shows that $m = k^2 + k + 1$ for $k \geq 0$. Curiously, simple cyclic difference sets are known to exist whenever k is a $0, 1$, or a prime power,

but there is no other case where a simple cyclic difference set is known to exist.

To any simple cyclic difference set of modulus m, associate a sequence of period m with 1's at positions congruent modulo m to some residue in the difference set. It is easy to check that the difference set property implies that the sequence has the window property with $w = m$. In Section 3 we prove that simple cyclic difference sets solve the problem in that the associated sequence minimizes the density of all sequences with window property m.

Example. The set of residues $\{1, 2, 4\}$ forms a simple cyclic difference set modulo 7. The period 7 sequence associated with this is

$$0, 1, 1, 0, 1, 0, 0, 0, 1, 1, 0, 1, 0, 0, \ldots$$

and has the 7-window property.

The definition of simple cyclic difference sets may be relaxed such that each non-zero residue can be written in *at least* one way as the difference of two elements in the set. We call such sets *difference coverings* modulo m. The infinite sequence associated with a difference covering modulo m will have the window property for $w = m$, and therefore for every $w \geq m$.

For all values of $w \leq 65$ we have been able to use the CRAY-1 to solve the problem. In all these cases some sequence derived from a difference covering achieves the minimum density. We do not know if this is always the case, but this prompted our finding the smallest difference covering for all moduli $m \leq 110$. This at least provides a reasonable upper bound for the solution to the problem for all $w \leq 110$.

2. Forcing Periodicity

In this section we want to precisely define the density of a sequence

and then show that the sequences can be assumed to be periodic.

Definition. If $x_1, x_2, \ldots$ is an infinite 0-1 sequence, then its *density* is

$$\liminf_{n\to\infty} n^{-1} \sum_{i=1}^{n} x_i.$$

The problem is then to find the greatest lower bound, say $f(w)$, of the densities of all sequences with the w-window property.

Proposition 2.1. If $x_1, x_2, \ldots$ is a 0-1 sequence with the w-window property and if a sequence $y_1, y_2, \ldots$ has the property that every w-tuple of consecutive elements from the second sequence appears somewhere in the first sequence, then the second has the w-window property.

Proof. A 0-1 sequence has the w-window property if and only if the coordinatewise product of any two w-tuples appearing in it is never all zeros. □

Proposition 2.2. The function $f(w)$ remains unchanged if we require the sequences to be periodic.

Proof. Let $x_1, x_2, \ldots$ be a density d sequence with the w-window property. We say the tuple at position t is $(x_{t+1}, x_{t+2}, \ldots, x_{t+w-1})$. Let

$$e(t) = t^{-1} \sum_{i=1}^{t} x_i$$

be the normalized partial density. By assumption there is an infinite sequence of positions $j_1 < j_2 < \cdots$ such that $e(j_a), a = 1, 2, \ldots$ converges to d. There must be a $(w-1)$-tuple $\mathbf{y}$ appearing infinitely often among the tuples at positions $j_1, j_2, \ldots$. In fact, j_a can be replaced by a subsequence such that $\mathbf{y}$ appears at each j_a and $e(j_a)$ converge to d.

For any $\epsilon > 0$ there is a b such that

$$(j_b - j_1)^{-1} \sum_{i=j_1+1}^{j_b} x_i < d + \epsilon.$$

By Proposition 2.1 the sequence with period $j_b - j_1$,

$$x_{j_1+1}, x_{j_1+2}, \ldots, x_{j_b}, x_{j_1+1}, \ldots$$

has the w-window property because all its w-tuples appear in $x_1, x_2, \ldots$.

Thus, there exist periodic sequences with the w-window property and densities arbitrarily close to $f(w)$. □

Proposition 2.3. The function $f(w)$ remains unchanged if we require the sequences to be periodic with period of at most 2^{w-1}.

Proof. By the preceding proposition, for any $\epsilon > 0$, there is a periodic sequence with density $d < f(w) + \epsilon$ and with the w-window property. Let p be the period of such a sequence. If $p > 2^{w-1}$ it will be shown how the sequence can be replaced with one of smaller period.

Assuming $p > 2^{w-1}$, there must be a $(w-1)$-tuple, $\mathbf{y}$, that appears in at least two places in the cycle. If necessary, replace the sequence by a left shift of itself so that the first $(w-1)$-tuple is $\mathbf{y}$. Let i be the next position of $\mathbf{y}$. Since there are dp ones in the entire period, either the first i elements have at most di ones or the last $p-i$ elements have at most $d(p-i)$ ones. In the first case, Proposition 2.1 shows that if the first i elements are formed into a sequence of period i, it will have the w-window property and density at most d. The second case is similar.

The period of the sequence can therefore be reduced until it is no more than 2^{w-1}, without increasing the density. Thus, there is a series of

sequences with periods no more than 2^{w-1}, having the w-window property and having densities converging to $f(w)$. □

Theorem 2.4. For every $w > 0$ there is a sequence with density exactly $f(w)$, the w-window property and periodic with period at most 2^{w-1}.

Proof. By Proposition 2.3 there exists a series of sequences with period at most 2^{w-1}, the w-window property, and densities converging to $f(w)$. But there are only a finite number of sequences with period at most 2^{w-1}, so one member of the series has density exactly $f(w)$. □

The above results are not of much help in finding $f(w)$ because the bound on the period is too large. Curiously, there are no violations known of the following.

Hypothesis. For every $w > 0$ there is a sequence with density $f(w)$, the w-window property and periodic with period at most w.

By computation we have verified this hypothesis for $w \leq 65$, but we have been unable to find an intuitive reason why it should be true or false.

The next proposition shows that given any periodic solution, we may produce a left shift such that *all* partial densities are at most $f(w)$.

Proposition 2.5. If $x_1, x_2, \ldots$ is a 0-1 sequence with density d and periodic with period p, there is an $i \geq 0$ such that for all $t > 0$,

$$\sum_{j=1}^{t} x_{i+j} \leq td.$$

Proof. Let

$$q(t) = \sum_{j=1}^{t} x_j - td.$$

Note $q(t)$ has period p. Let i be such that $q(t)$ is maximized at $q(i)$. Then for $t \geq 0$, $q(t+i) - q(i) \leq 0$ and this is equivalent to

$$\sum_{j=1}^{t} x_{i+j} \leq td. \ \square$$

3. Lower Bounds on the Density.

Proposition 3.1. Let $x_1, x_2, \ldots$ be a sequence with density d and the w-window property. If for some $i \geq 0$, $x_{i+1} + \cdots + x_{i+w} = k+1$, then

$$d \geq \frac{w+k}{w(k+1)}.$$

Proof. In the statement of the proposition we may take $i = 0$, since otherwise the sequence can be left shifted i positions. For the remainder of the proof we will use the first $n = mw$ elements of the sequence, where m is some large integer. We set

$$D = \sum_{i=1}^{n} x_i.$$

Let $0 < \alpha_0 < \alpha_1 < \cdots < \alpha_k \leq w$ be the indices of the first $k+1$ ones in $x_1, \ldots, x_w$.

For any s, $1 \leq s \leq n - w$ let

$$E(s) = \sum_{i=1}^{w} x_i x_{i+s}.$$

Note that $E(s) \geq 1$ by the w-window property. Also,

$$\sum_{s=1}^{n-w} E(s) = \sum_{i=1}^{w} x_i \sum_{s=1}^{n-w} x_{i+s} \leq (k+1)D.$$

We note in passing that this implies $n - w \leq (k+1)D$. Since $n = mw$ this gives us $1 - 1/m \leq (k+1)D/n$ for all m. Letting $m \to \infty$ we get $d \geq 1/(k+1)$. However, this inequality is weaker than the one to be proved.

For any j and s, such that $1 \leq j \leq k$ and $1 \leq s \leq n - w$ let

$$e_j(s) = x_{\alpha_0+s} x_{\alpha_j+s}.$$

Because the sequence has the w-window property for the left shift by $\alpha_j - \alpha_0$, we have

$$\sum_{s=1}^{n-w} e_j(s) \geq m - 1.$$

If $x_{\alpha_0+s} = 0$, then

$$E(s) \geq 1 = 1 + \sum_{j=1}^{k} e_j(s).$$

If $x_{\alpha_0+s} = 1$, then

$$E(s) = 1 + \sum_{j=1}^{k} e_j(s).$$

Hence,

$$(k+1)D \geq \sum_{s=1}^{n-w} E(s) \geq n - w + \sum_{j=1}^{k} \sum_{s=1}^{n-w} e_j(s) \geq n - w + k(m-1).$$

Dividing by n and letting $m \to \infty$, we get $(k+1)d \geq 1 + k/w$, which is equivalent to the statement of the proposition. □

Proposition 3.2. Suppose that k and w are positive integers. If $w \geq k^2+1$, then $f(w) \geq (k+1)/w$. If $w \leq k^2 + 1$, then $f(w) \geq (w + k - 1)/(kw)$.

Proof. Let $x_1, x_2, \ldots$ have the w-window property and density $d = f(w)$. Notice that if $w \geq k^2 + 1$, then $(w + k - 1)/(wk) \geq (k + 1)/w$ and if $w \leq k^2 + 1$, then $(k + 1)/w \geq (w + k - 1)/(wk)$. Hence if $d \geq (k + 1)/w$ there is nothing to prove. Now suppose that $d < (k + 1)/w$. Then there is some interval of length w in which the sequence has at most k ones. By Proposition 3.1 we have $f(w) = d \geq (w + k - 1)/(wk)$, and our result follows. □

Corollary 3.3. For $w \geq 1$, $f(w) \geq \frac{1}{\sqrt{w}}$.

Proof. Choose k so that $k^2 + 1 \leq w \leq (k + 1)^2$. Then Proposition 3.2 gives us

$$f(w) \geq \frac{k+1}{w} \geq \frac{1}{\sqrt{w}}. \square$$

Proposition 3.2 shows that when $k^2 + 1 \leq w \leq k^2 + k + 1$ and there is a difference covering consisting of $k + 1$ elements modulo w, then $f(w) = (k + 1)/w$. This is because the difference covering construction gives a sequence with density equal to the bound in the proposition. For $w > k^2 + k + 1$ there is no such difference covering.

Suppose $k^2 + 1 < w \leq k^2 + k + 1$ for some k and there is a periodic sequence with the w-window property and density equal to the lower bound $(k + 1)/w$ of the proposition. If the sequence has an interval of length w with at most k ones, then Proposition 3.1 gives us

$$d \geq \frac{w + k - 1}{wk} > \frac{k + 1}{w}.$$

Thus every interval of length w has at least $k + 1$ ones. Furthermore, if there were ever more than $k+1$ ones in an interval of length w the sequence would have density greater than $(k+1)/w$. Thus every interval of length w

has exactly $k+1$ ones. This implies that the sequence has period w and it follows that it must correspond to a cyclic difference covering. Therefore, for $w > k^2+1$, when a cyclic difference covering with $k+1$ elements modulo w exists, all optimal periodic solutions for that value of w come from such cyclic difference coverings.

We sum up our main results in the following theorem.

Theorem 3.4. Let k and w be positive integers. If $k^2+1 \leq w \leq k^2+k+1$, then $f(w) \geq (k+1)/w$. We have equality if a difference covering of $k+1$ elements modulo w exists. For $w > k^2+1$, the only periodic sequences with the w-window property and density $(k+1)/w$ are those that correspond to such a difference covering. If $k^2+k+1 < w \leq (k+1)^2$, then

$$f(w) \geq (k+w)/(kw+w).$$

The last assertion of Theorem 3.4 comes directly from Proposition 3.2.

4. Computational Results

Two backtrack programs were run on the CRAY-1. One of these found the smallest difference covering for a given modulus m. It begins with 0 and 1 already assumed to be residues in the covering. It then branches on the next residue to place in the covering and backtracks when too many differences are repeated. As a parameter the program is given the cardinality of the difference covering. To speed the program up, for all but the last four levels of the backtrack tree a test was inserted to see if the covering as constructed so far was minimum under the automorphism group consisting of all translations and multiplications by units modulo m.

Using this program we found the lexicographically first covering of minimum cardinality for each modulus $m \leq 110$. The results are given in

the table at the end of this section. This table also gives the density of the associated periodic sequence. A star next to the density means that the density is smaller than the density for any preceding moduli. A second star next to the density means that the covering is also a difference set.

Let $f^*(w)$ be the smallest density for any difference covering modulo $m \leq w$. We have $f(w) \leq f^*(w)$, so that $f^*(w)$ is an upper bound for $f(w)$. The hypothesis of Section 2 asserts that $f(w) = f^*(w)$. For $w \leq 110$ the value of $f^*(w)$ can be read off of the table. Indeed it is the density for the largest modulus $m \leq w$ for which the density is starred. Thus our program constructively provides the upper bound $f^*(w)$ for the function $f(w)$.

A second program was written to search exhaustively for an infinite sequence $x_1, x_2, \ldots$ with the w-window property. At stage l of the branching it tests that the sequence $x_1, x_2, \ldots, x_l, 1, 1, 1, \ldots$ has the w-window property. Also, the program is given an upper bound b on $f(w)$ as an input parameter. If $x_1 + \cdots + x_l > lb$ then the program backtracks, in light of Proposition 2.5. This is a slower running program than the first, but it verified that $f(w) = f^*(w)$ for all $w \leq 65$. Thus, no counterexamples to the hypothesis of Section 2 were found.

There are 24 values of $w \leq 65$ such that $f(w)$ is equal to the lower bound given by Theorem 3.4. These are all values for which $k^2 + 1 \leq w \leq k^2 + k + 1$ for some k, and $f(w) = (k+1)/w$. These values of w are 1, 2, 3, 5, 6, 7, 10, 11, 12, 13, 17, 18, 19, 21 26, 27, 28, 31, 37, 39, 50, 51, 57, 65.

The graph gives the value of $1/f^*(w)$, for $w \leq 110$.

GRAPH

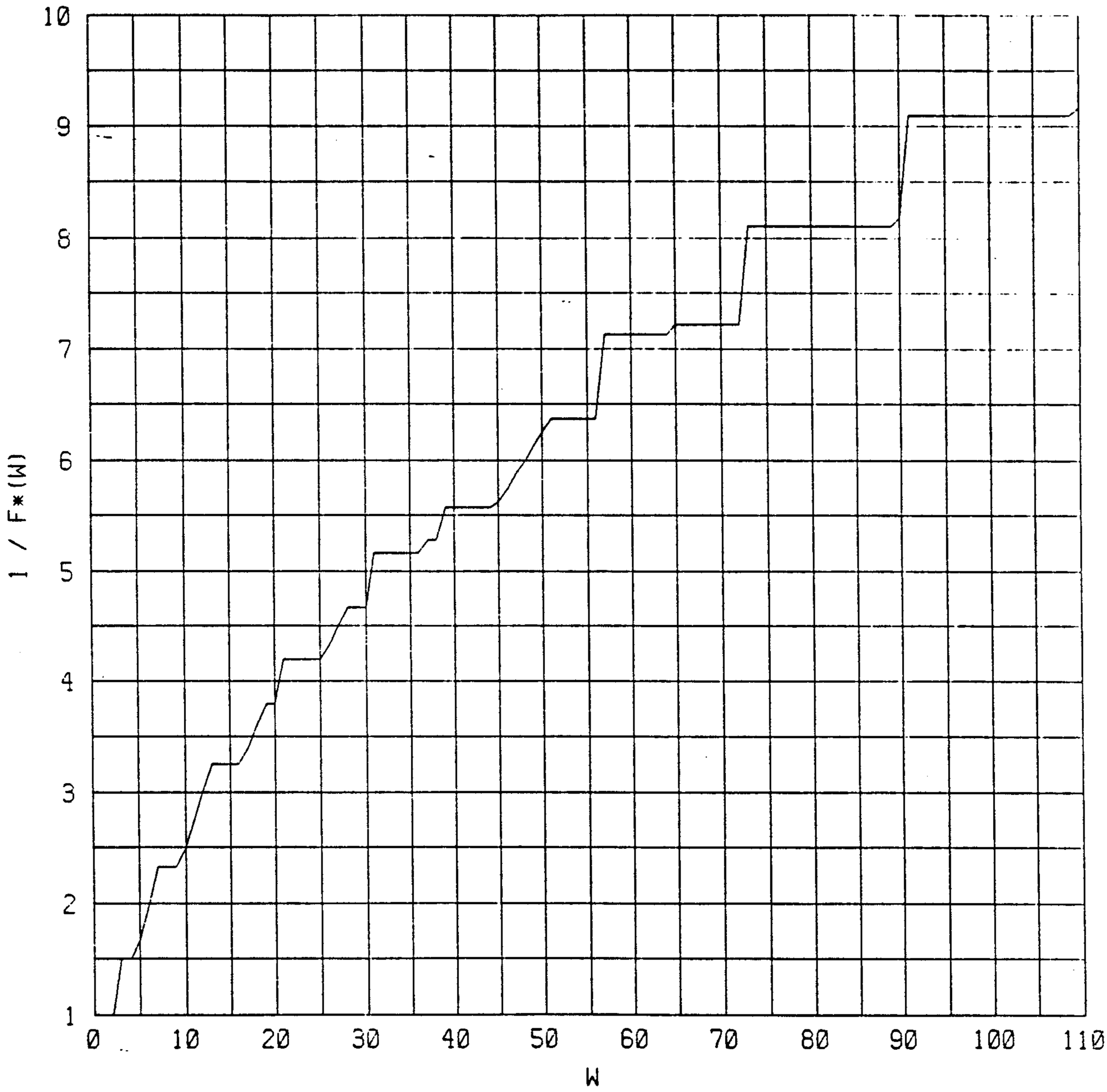

TABLE
Difference coverings of minimal size for a given modulus

m	Size	Covering	Density
2	2	0,1	*1.00000
3	2	0,1	**0.66667
4	3	0,1,2	0.75000
5	3	0,1,2	*0.60000
6	3	0,1,3	*0.50000
7	3	0,1,3	**0.42857
8	4	0,1,2,4	0.50000
9	4	0,1,2,4	0.44444
10	4	0,1,2,5	*0.40000
11	4	0,1,2,5	*0.36364
12	4	0,1,3,7	*0.33333
13	4	0,1,3,9	**0.30769
14	5	0,1,2,3,7	0.35714
15	5	0,1,2,3,7	0.33333
16	5	0,1,2,5,8	0.31250
17	5	0,1,2,4,12	*0.29412
18	5	0,1,2,5,11	*0.27778
19	5	0,1,2,6,9	*0.26316
20	6	0,1,2,3,6,10	0.30000
21	5	0,1,4,14,16	**0.23810
22	6	0,1,2,3,7,11	0.27273
23	6	0,1,2,3,7,11	0.26087
24	6	0,1,2,3,7,15	0.25000
25	6	0,1,2,3,8,12	0.24000
26	6	0,1,2,5,9,15	*0.23077
27	6	0,1,2,5,13,22	*0.22222
28	6	0,1,4,15,20,22	*0.21429
29	7	0,1,2,3,4,9,14	0.24138
30	7	0,1,2,3,4,9,19	0.23333

TABLE (continued)

m	Size	Covering	Density
31	6	0,1,3,8,12,18	**0.19355
32	7	0,1,2,3,7,11,19	0.21875
33	7	0,1,2,3,6,16,27	0.21212
34	7	0,1,2,3,7,12,20	0.20588
35	7	0,1,2,3,8,12,21	0.20000
36	7	0,1,2,5,12,14,20	0.19444
37	7	0,1,2,4,10,15,22	*0.18919
38	8	0,1,2,3,4,8,14,23	0.21053
39	7	0,1,2,4,13,18,33	*0.17949
40	8	0,1,2,3,4,9,14,24	0.20000
41	8	0,1,2,3,4,9,15,25	0.19512
42	8	0,1,2,3,4,9,15,25	0.19048
43	8	0,1,2,3,4,10,15,26	0.18605
44	8	0,1,2,3,6,16,27,38	0.18182
45	8	0,1,2,3,5,12,18,26	*0.17778
46	8	0,1,2,3,6,18,25,38	*0.17391
47	8	0,1,2,3,5,16,22,40	*0.17021
48	8	0,1,2,5,9,20,26,36	*0.16667
49	8	0,1,2,5,24,33,36,44	*0.16327
50	8	0,1,3,8,17,28,32,38	*0.16000
51	8	0,1,2,5,11,18,30,38	*0.15686
52	9	0,1,2,3,4,6,14,21,30	0.17308
53	9	0,1,2,3,4,7,21,29,44	0.16981
54	9	0,1,2,3,4,9,15,21,31	0.16667
55	9	0,1,2,3,4,6,19,26,47	0.16364
56	9	0,1,2,3,4,11,16,33,39	0.16071
57	8	0,1,3,13,32,36,43,52	**0.14035
58	9	0,1,2,3,7,21,33,37,50	0.15517
59	9	0,1,2,3,6,13,21,35,44	0.15254
60	9	0,1,2,4,9,15,25,30,42	0.15000

TABLE (continued)

m	Size	Covering	Density
61	9	0,1,2,3,7,15,25,36,45	0.14754
62	9	0,1,2,4,10,32,39,46,51	0.14516
63	9	0,1,2,6,8,20,38,41,54	0.14286
64	9	0,1,2,5,14,16,34,42,59	0.14062
65	9	0,1,2,6,10,28,35,51,54	*0.13846
66	10	0,1,2,3,4,5,13,19,39,46	0.15152
67	10	0,1,2,3,4,5,12,20,26,39	0.14925
68	10	0,1,2,3,4,10,16,21,38,45	0.14706
69	10	0,1,2,3,4,10,17,22,33,45	0.14493
70	10	0,1,2,3,4,9,20,35,49,62	0.14286
71	10	0,1,2,3,4,10,18,23,34,46	0.14085
72	10	0,1,2,3,6,11,18,31,37,51	0.13889
73	9	0,1,3,7,15,31,36,54,63	**0.12329
74	10	0,1,2,3,7,28,30,43,57,65	0.13514
75	10	0,1,2,5,8,18,30,32,41,56	0.13333
76	10	0,1,2,6,9,25,35,46,58,63	0.13158
77	10	0,1,2,4,10,15,37,49,56,61	0.12987
78	10	0,1,2,7,13,16,33,51,55,70	0.12821
79	10	0,1,2,6,13,28,31,47,48,71	0.12658
80	11	0,1,2,3,4,5,10,23,40,56,71	0.13750
81	11	0,1,2,3,4,5,12,20,26,39,53	0.13580
82	11	0,1,2,3,4,5,12,20,26,40,53	0.13415
83	11	0,1,2,3,4,5,12,21,27,40,54	0.13253
84	11	0,1,2,3,4,7,18,26,46,54,75	0.13095
85	11	0,1,2,3,4,9,13,25,40,54,68	0.12941
86	11	0,1,2,3,4,11,17,24,29,48,54	0.12791
87	11	0,1,2,3,4,10,42,54,62,67,73	0.12644
88	11	0,1,2,3,5,11,24,29,36,43,73	0.12500
89	11	0,1,2,3,5,12,18,43,57,65,71	0.12360
90	11	0,1,2,3,6,33,46,54,67,74,81	*0.12222

TABLE (continued)

m	Size	Covering	Density
91	10	0,1,3,9,27,49,56,61,77,81	**0.10989
92	11	0,1,2,4,40,50,51,59,64,71,77	0.11957
93	11	0,1,2,5,14,20,24,31,52,60,68	0.11828
94	12	0,1,2,3,4,5,6,14,23,30,46,61	0.12766
95	11	0,1,2,5,8,17,28,39,53,63,82	0.11579
96	12	0,1,2,3,4,5,8,21,30,53,62,86	0.12500
97	12	0,1,2,3,4,5,9,17,33,43,54,79	0.12371
98	12	0,1,2,3,4,5,11,27,40,54,69,81	0.12245
99	12	0,1,2,3,4,5,12,21,27,34,48,62	0.12121
100	12	0,1,2,3,4,5,13,20,28,34,56,63	0.12000
101	12	0,1,2,3,4,5,12,49,63,72,78,85	0.11881
102	12	0,1,2,3,4,6,13,28,34,42,50,85	0.11765
103	12	0,1,2,3,4,7,38,53,62,77,85,93	0.11650
104	12	0,1,2,3,4,9,19,32,46,57,72,84	0.11538
105	12	0,1,2,3,4,10,15,36,39,61,66,89	0.11429
106	12	0,1,2,3,5,48,53,69,76,82,89,97	0.11321
107	12	0,1,2,3,5,20,27,35,42,48,58,98	0.11215
108	12	0,1,2,3,7,12,20,34,41,49,57,85	0.11111
109	12	0,1,2,3,7,15,39,49,58,83,89,94	0.11009
110	12	0,1,2,6,17,25,39,43,46,52,80,100	*0.10909

REFERENCE

[1] Leonard D. Baumert, Cyclic Difference Sets, Springer-Verlag Lecture Notes in Mathematics 182, New York, 1971.