

# UNPREDICTABLE BINARY STRINGS

R.M. LOW, M. STAMP, R. CRAIGEN, AND G. FAUCHER

**ABSTRACT.** We examine a class of binary strings arising from considerations about stream cipher encryption: to what degree can one guarantee that the number of pairs of entries distance  $k$  apart that disagree is equal to the number that agree, for all small  $k$ ? In a certain sense, a keystream with such a property achieves a degree of unpredictability. The problem is also restated combinatorially in terms of seating arrangements.

We examine sequences  $s$  of length  $2n$  in which this property holds for all  $k \leq M_n$ , where  $M_n$  is the largest number for which this is possible among strings of length  $2n$ . We give upper and lower bounds for  $M_n$ , and give optimal sequences of all lengths up to  $n = 26$ . We also show how to obtain classes of special orthogonal arrays and balanced sign graphs from such sequences.

## 1. BACKGROUND

A stream cipher cryptosystem is illustrated in Figure 1. The original message, or plaintext, is encrypted by adding (elementwise modulo 2) a pseudo-random sequence of bits to the message. This pseudo-random sequence of bits is known as a keystream. The resulting ciphertext can then be transmitted over insecure lines. The recipient can recover the plaintext by adding, modulo 2, the same keystream to the ciphertext.

Stream ciphers are a natural generalization of the one-time pad, or Vernam cipher. With a one-time pad, a “random” string of bits, or pad, is used to encrypt, and this pad can only be used once. While the one-time pad is provably secure [12], the drawbacks are many. For example, the random pad is the same length as the message, and the pad must be securely transmitted to the recipient before the ciphertext can be decrypted.

Stream ciphers replace the random sequence of the one-time pad with a pseudo-random string of bits that are generated from a short secret key. The result is a more practical cipher since only the short secret key needs to be securely transmitted before using the system. The tradeoff is that the stream cipher does not inherit the provable security of the one-time pad.

---

*Date:* May 26, 2005. Draft version 0.4.

*2000 Mathematics Subject Classification.* 05B15.

The authors wish to thank John F. Dillon for his insightful comments.

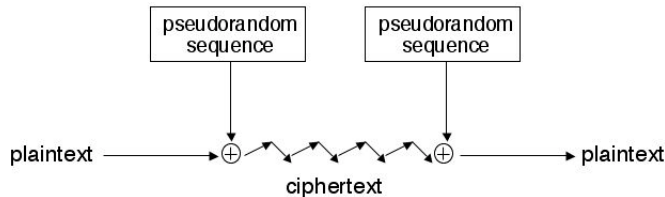


FIGURE 1. Keystream encryption

Suppose that a message is encrypted with a stream cipher. In many situations an attacker will know (or can guess) part of the plaintext message, from which they can then recover part of the keystream. An attacker who can deduce more of the keystream from a short captured segment can then recover more of the plaintext. Therefore, the pseudo-random sequence used in a stream cipher must be “unpredictable” in the sense that it is computationally infeasible to recover more of the sequence from a short captured segment. The term *cryptographically strong* is often used as a synonym for this type of unpredictability.

Several approaches to the topic of cryptographically strong sequences have appeared in the literature. For example, Shamir [11] and Blum and Micali [1] provide a theoretical framework based on one-way functions, while Rueppel [10] and Stamp and Martin [13] present methods for measuring some aspects of the unpredictability of sequences.

Pseudo-random sequences also arise in many non-cryptographic applications. For example, simulations often make use of such sequences. However, in most non-cryptographic applications, the sequences need to be statistically random (i.e., they must satisfy some desired statistical properties), but they do not need to be unpredictable. A particularly dramatic example of the difference between statistical randomness and unpredictability is provided by the so-called  $m$ -sequences [3]. An  $m$ -sequence has many desirable statistical properties and consequently such sequences are often used in simulations and other applications. However,  $m$ -sequences are extremely predictable—from any set of  $2n$  consecutive bits, the entire  $m$ -sequence of length  $2^n - 1$  can be determined using the well-known Berlekamp–Massey algorithm [7]. In a sense,  $m$ -sequences are the most predictable sequences of all [13].

## 2. INTRODUCTION

In this paper, we consider unpredictable binary sequences from a novel perspective. Note that we shall use the terms “strings” and “sequences”, and the corresponding terms “bits” and “entries” interchangeably. Consider a pseudo-random binary sequence

$$s = (s_0, s_1, s_2, s_3, \dots)$$

Ideally, we would like to choose  $s$  so that no function of the preceding  $n$  bits predicts the next bit of  $s$  with a probability different from  $1/2$ . In [5], where “sliding block maps” are discussed, it is shown that it suffices to consider only polynomial functions. In an effort to make this problem more tractable, we might restrict our attention to linear functions of the preceding  $n$  bits. But even this condition is hard to verify, and perhaps too restrictive, so we’ll consider a much more restricted case.

We consider a binary string,  $s = (s_1, \dots, s_n)$ , of fixed length  $n$ , from which a periodic sequence (with period  $n$ ) of arbitrary length can be obtained by repeated use of  $s$ . First, we want  $s$  to contain an equal number of 0’s and 1’s. Suppose that the bit  $s_i$  agrees with bit  $s_{i+1 \pmod n}$  exactly half of the time, for all  $i$ . Then, knowledge of one bit does not provide any information about the next bit; such a string may be considered less predictable than a string that does not satisfy this condition. Now suppose also that bit  $s_i$  agrees with bit  $s_{i+2}$  exactly half of the time—such a string may be considered even more unpredictable than one satisfying only the first condition. Considering strings of a given length satisfying similar conditions for bits at all distances  $1, 2, 3, \dots, k$ , we may consider their unpredictability (in this limited sense) to increase with  $k$ .

The following restatement of our condition in the setting of a more familiar combinatorial problem of arrangements may be helpful.

**Combinatorial Problem:** A group of  $n$  men and  $n$  women are to be seated around a (circular) table. Can this be so arranged that, of the  $2n$  immediate neighbors of the women (counted with repetition) exactly half are men and half are women. Can we simultaneously satisfy the corresponding condition for those seated two seats away from the women? Three seats? For a given  $n$ , how far can we continue this process?

Let us introduce some terms and notations that will be useful. Let  $S^{2n} = \{s \in \{0, 1\}^{2n} : \text{Hamming weight of } s \text{ is } n\}$ . For  $k \in \{1, 2, 3, \dots\}$ , let  $\omega_k(s)$  be the Hamming weight of  $s \oplus \sigma^k(s)$ , where  $\sigma$  is the right cyclic shift operator and  $\oplus$  denotes termwise addition modulo 2. Let  $g(s) = \min\{k : \omega_{k+1}(s) \neq n\}$ . Finally, let  $M_n = \max\{g(s) : s \in S^{2n}\}$ . The combinatorial problem stated above is solved by binary strings  $s$  satisfying  $g(s) = M_n$ .

**Definition.** Let  $s = (s_1, \dots, s_r)$  be any list of  $r$  numbers. For any  $k \in \mathbb{Z}^+$ , the  $k \times r$  circulant matrix generated by  $s$  is the matrix

$$\text{circ}_{k \times r}(s) := [s_{j-i+1 \pmod r}]_{k \times r}.$$

That is,  $\text{circ}_{k \times r}(s)$  is the  $k \times r$  matrix whose  $(i, j)$  entry is equal to  $s_{j-i+1}$ , where the index is reduced modulo  $r$ . If  $k = r$ , we simply write  $\text{circ}(s)$ .

*Example.*  $\text{circ}_{2 \times 3}(a, b, c) = \begin{pmatrix} a & b & c \\ c & a & b \end{pmatrix}$ ;  $\text{circ}(a, b, c) = \begin{pmatrix} a & b & c \\ c & a & b \\ b & c & a \end{pmatrix}$ .

We shall use, without comment, a few well-known and easily demonstrated facts about square circulant matrices  $A, B$ : (i) all row and column sums of matrix  $A$  are equal; (ii)  $A^T$  is circulant; and (iii)  $AB$  is circulant.

For  $m \in \mathbb{Z}^+$ , let us write  $J_m$  (or  $J$ , if  $m$  is understood) for the  $m \times m$  matrix of 1's. Alternately,  $J = \text{circ}(1, 1, \dots, 1)$ . Further, if  $A$  is a square circulant of the same order, then  $AJ = rJ$ , where  $r$  is the row-sum of  $A$ .

*Remark.* Using this notation, a string  $s \in S^{2n}$  satisfies  $g(s) \geq k$  if and only if  $AA^T = \frac{n}{2}(I_{k+1} + J_{k+1})$ , where  $A = \text{circ}_{(k+1) \times 2n}(s)$ .

We denote the concatenation of strings  $s$  and  $t$  by  $st$ . Note that if  $s \in S^{2m}$  and  $t \in S^{2n}$ , then  $st \in S^{2(m+n)}$ .

### 3. BOUNDS ON $M_n$

Let us begin with a few technical lemmas.

**Lemma 1.** *Let  $s \in S^{2n}$ ,  $A = \text{circ}(s)$ . Then,  $g(s) \geq k$  if and only if  $AA^T = \text{circ}(t)$ , where the first  $k+1$  elements of sequence  $t$  are  $(n, \frac{n}{2}, \frac{n}{2}, \dots, \frac{n}{2})$ .*

*Proof.* Observe that  $g(s) \geq k$  if and only if rows  $2, 3, \dots, k+1$  match row 1 in exactly  $n$  positions. Since each row contains exactly  $n$  1's it is equivalent to say that the dot products of these rows with row 1 is  $\frac{n}{2}$ . The  $i$ th entry of  $t$  is the dot product of rows 1 and  $i$  of  $A$ ; the dot product of row 1 with itself is clearly  $n$ . The result follows.  $\square$

One consequence of Lemma 1 is that  $M_n = 0$  when  $n$  is odd.

**Lemma 2.** *Let  $s \in S^{2m}$  and  $t \in S^{2n}$ , such that  $g(s), g(t) \geq k$ , and suppose that the first  $k+1$  elements of  $s$  and  $t$  are equal. Then  $g(st) \geq k$ .*

*Proof.* This follows from Lemma 1 and the observation that, under these conditions,  $\text{circ}(st)$  can be partitioned in the form  $\left( \begin{array}{c|c} \text{circ}(s) & \text{circ}(t) \\ \hline * & * \end{array} \right)$ .  $\square$

**Lemma 3.** *Suppose  $s = (s_1, \dots, s_{2n}) \in S^{2n}$ ,  $g(s) = k$ . Then:*

1. For each  $i = 1, \dots, 2n$ ,  $g(\sigma^i(s)) = k$ ;
2.  $g(s_{2n}, s_{2n-1}, \dots, s_1) = k$ ;
3.  $g(1 - s_1, 1 - s_2, \dots, 1 - s_{2n}) = k$ .

*Proof.* This follows immediately from the combinatorial interpretation of the binary string problem.  $\square$

We now establish upper and lower bounds for  $M_n$ , where  $n$  is even.

**Theorem 1.** *For even  $n \geq 4$ ,  $M_n \leq n - 2$ .*

*Proof.* Let  $s \in S^{2n}$ ,  $g(s) \geq n - 1$ , and  $A = \text{circ}(s)$ . By Lemma 1 and the fact that  $AA^T$  is symmetric,

$$AA^T = \text{circ}\left(n, \frac{n}{2}, \dots, \frac{n}{2}, a, \frac{n}{2}, \dots, \frac{n}{2}\right),$$

where  $a$  occurs in the  $(n + 1)$ -th position. Now,  $AJ = A^T J = nJ$ , so  $AA^T J = n^2 J = (n + (2n - 2)\frac{n}{2} + a) J$ . Accordingly,  $a = 0$ .

Let  $s = pq$ , where  $p, q \in \{0, 1\}^n$ . Then the  $(n + 1)$ -th row of  $A$  is  $qp$ , so  $A = \begin{pmatrix} B & C \\ C & B \end{pmatrix}$ , where  $B$  and  $C$  are  $n \times n$  matrices. Since  $a = 0$ , the dot product of  $p$  and  $q$  is 0; it follows that  $p + q = (1, 1, \dots, 1)$ , and so  $B + C = J$ .

By matrix algebra, we see that every two rows of  $\text{circ}_{n \times 2n}(pq) = (B|C)$  differ in exactly  $n$  positions. Since  $C = J - B$ , two rows of  $B$  differ in exactly the same positions as the corresponding rows of  $C$ . Therefore, every two rows of  $B$  must differ in exactly  $\frac{n}{2}$  positions.

Since  $q = (1, \dots, 1) - p$ , each row of  $B$  is obtained by circulating the previous row and then interchanging 0 and 1 in the first position. Thus, the difference of the row sums of consecutive rows is odd. Consequently, the difference of row sums of rows two positions apart is even.

Let  $a, b, c, d$  be the number of positions in which two rows of  $B$ , made into a  $2 \times n$  array, have columns of the forms  $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ ,  $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ ,  $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$  and  $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ , respectively. From the above discussion, we have that  $a + d = b + c = \frac{n}{2}$ . Further, if the rows are consecutive, then  $(a + b) - (a + c) = b - c = \frac{n}{2} - 2b$  is odd; thus  $\frac{n}{2}$  is odd. On the other hand, if the rows differ in position by 2, this number must be even—a contradiction (for even  $n \geq 4$ ).  $\square$

**Theorem 2.** *For even  $n \geq 6$ ,  $M_n \geq 4$ .*

*Proof.* Observe that  $g(s) = 4$  and  $g(t) = 6$ , where  $s = (101100001011) \in S^{12}$  and  $t = (1011011110001000) \in S^{16}$ . Also, the first five entries of  $s$  and  $t$  match. By repeated application of Lemma 2, if  $u$  is any string obtained by concatenating multiple copies of  $s$  and  $t$ , then  $g(u) \geq 4$ . Since 12, 16 and any multiple of 4 greater than 20 are nonnegative integer combinations of 12 and 16,  $u$  can be taken to have any such length. A string  $u'$  of length 20 with  $g(u') = 6$  is given in Table 1. The result follows.  $\square$

The lower bound of Theorem 2 can be improved for binary strings of length a power of two.

**Theorem 3.**  *$M_{2^k} \geq k$ , for all  $k \geq 1$ .*

*Proof.* Let  $s$  be a de Bruijn sequence [14] of length  $2 \cdot 2^k = 2^{k+1}$ . Clearly,  $s \in S^{2^{n+1}}$ . Let  $A = \text{circ}_{(n+1) \times 2^{n+1}}(s)$ .

Since  $s$  is a de Bruijn sequence, every binary  $(n + 1)$ -tuple appears exactly once as a column of  $A$ , and the Hamming distance between any two rows of  $A$  is  $2^n$ . Therefore  $g(s) \geq n$ ; the result follows.  $\square$

Figure 2 illustrates Theorem 3, with  $n = 2$ .

$$A = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 & 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 & 1 & 0 & 1 \end{bmatrix}.$$

FIGURE 2. Matrix formed from the de Bruijn sequence  $s = 00010111$ .

#### 4. GENERATION OF SEQUENCES $s$ WITH $g(s) = M_n$ , FOR SMALL $n$

An obvious way to compute  $M_n$  is to generate all  $\binom{2n}{n}$  balanced binary strings of length  $2n$  and for each such sequence  $s$ , shift and add to find  $g(s)$ . In this approach,  $\binom{2n}{n}$  strings must be tested and, asymptotically, little economy is achieved over trying all  $2^{2n}$  binary strings. In particular, the work grows by a factor of 16 for each subsequent even value of  $n$ .

A slight improvement can be obtained by only testing those strings  $s$  for which  $g(s) \geq 1$ . Any such string  $s$  must contain  $n/2$  blocks of 0's interlaced with  $n/2$  blocks of 1's. The sum of the lengths of the  $n/2$  blocks of 0's is  $n$  and similarly for the blocks of 1's. Consequently, we can generate all partitions of the integer  $n$  into  $n/2$  parts [6], with each part greater than zero, and interlace these to generate the candidate strings  $s$  that must be tested. Each such string will satisfy  $g(s) \geq 1$ . The number of strings that must be tested in this case also grows exponentially and, apparently, by the same factor as in the naive case. For the small cases under consideration, the initial savings are sufficient to allow additional cases to be computed using this approach.

For the sake of completeness, we give, in Table 1, all currently known values of  $M_n$  and, for each, a sequence attaining this value. These values of  $M_n$  were obtained and confirmed by exhaustive computer programs based on two different algorithms by the authors and also by Ritter (for  $n \leq 16$ , [9]) and Newton (for  $n \leq 22$ , [8])—but, so far, only one of our algorithms has established  $M_{26}$ .

Table 1 uses the following, more compact, representation for these sequences.

*Notation.* Let  $s \in S^{2n}$  begin with 0 and end with 1. A *0-block* in such a string is a maximal nonempty substring of 0's; *1-blocks* are defined similarly. By Lemma 3, each sequence  $s \in S^{2n}$  can be transformed, by a cyclic

permutation, into an equivalent sequence  $\hat{s}$ , where the largest 0-block is the first block. The *block-representation* of  $s$  (denoted by  $s'$ ) is

$$s' = (s_1, s_2, \dots, s_l) = (|A_1|, |B_1|, |A_2|, |B_2|, \dots, |A_l|, |B_l|),$$

where  $A_1, \dots, A_l$  are the 0-blocks of  $\hat{s}$  in order,  $B_1, \dots, B_l$  are the 1-blocks,  $|A_i|$  and  $|B_i|$  denote the lengths of the blocks, and  $|A_1| \geq |A_i|$  for all  $i$ .

*Example.* If  $s = 01110001$ , then  $\hat{s} = 00010111$  and  $s' = (3,1,1,3)$ .

The following result gives some basic properties of block representations.

**Lemma 4.** *Let  $s \in S^{2n}$  begin with 0 and end with 1,  $g(s) > 0$ , and let  $s'$  be its block-representation. Then:*

1. *The sum of the entries of  $s'$  in even positions, and the sum of the entries of  $s'$  in odd positions, are both equal to  $n$ .*
2.  *$s'$  has length  $n$ .*
3.  *$n$  is even.*

*Proof.* Part 1 is an immediate consequence of the definition of  $S^{2n}$ .

The length of  $s'$  is the number of blocks in  $s$ . This is equal to the number of cyclically adjacent pairs of entries in  $s$  that differ. If  $g(s) > 0$ , then rows 1 and 2 of  $\text{circ}(s)$  differ in  $n$  positions. Part 2 follows.

Finally, if  $n$  is odd,  $s$  has an odd number of blocks. Because they alternate between 0-blocks and 1-blocks, it follows that the last block of  $s$  is a 0-block, contradicting the condition that  $s_{2n} = 1$ .  $\square$

$n$	$M_n$	Block-representation $s' = (s_1, \dots, s_n)$
2	1	(2,2)
4	2	(3,3,1,1)
6	4	(4,2,1,3,1,1)
8	6	(4,3,1,3,1,1,2,1)
10	6	(4,3,1,3,1,2,1,1,3,1)
12	8	(6,2,1,1,2,3,1,3,1,1,1,2)
14	8	(5,2,4,2,1,3,1,3,1,2,1,1,1,1)
16	11	(5,1,1,2,1,2,3,4,1,3,1,2,3,1,1,1)
18	13	(6,3,1,3,2,2,1,1,1,1,1,3,2,2,3,2,1,1)
20	16	(5,3,4,1,2,3,1,2,1,1,1,1,1,3,1,4,2,1,2,1)
22	15	(5,3,2,1,2,1,5,1,1,1,1,4,2,3,1,2,1,1,1,3,1,2)
24	16	(7,4,1,2,1,1,1,1,2,2,1,3,1,1,1,2,1,3,1,1,3,2,4,2)
26	19	(7,1,1,2,2,1,1,4,2,5,2,2,2,4,2,1,3,1,1,1,1,2,1,1,1,1)

TABLE 1. Known  $M_n$  values and sequences that attain them.

In general, for a given  $n$ , many  $s \in S^{2n}$  (i.e., not a single sequence and those obtained from it by Lemma 3) may attain  $M_n$ . For some  $n$ , there are surprisingly few such strings; but there is no obvious pattern among these occurrences.

## 5. ASSOCIATED COMBINATORIAL STRUCTURES

In studying the unpredictable binary string problem, we note a couple of other combinatorial objects which arise naturally, namely orthogonal arrays and balanced signed graphs.

Recall the following definition [4].

**Definition.** Let  $S$  be a set of  $s$  symbols, denoted by  $0, 1, 2, \dots, s - 1$ . An  $N \times k$  array  $A$  with entries from  $S$  is said to be an *orthogonal array* with  $N$  runs,  $k$  factors,  $s$  levels, *strength*  $t$  (for some  $t$  in the range  $0 \leq t \leq k$ ) and *index*  $\lambda$  if every  $N \times t$  subarray of  $A$  contains each  $t$ -tuple of symbols from  $S$  exactly  $\lambda$  times as a row. An orthogonal array with these parameters is denoted by  $OA(N, k, s, t)$ .

We now construct a particular class of  $OA(4n, k, 2, 2)$ , for  $n \geq 1$ . These orthogonal arrays have some nice properties. For  $n \geq 1$ , let  $s \in S^{4n}$  where  $g(s) = M_{2n}$ , and take  $A = \text{circ}(s)^T$ .

*Remarks.*

1.  $A$  is a  $4n \times (M_{2n} + 1)$  matrix whose transpose is circulant.
2. The Hamming weight of each column of  $A$  is  $2n$ .
3. The Hamming distance between any two columns of  $A$  is  $2n$ .

**Theorem 4.** For  $n \geq 1$ ,  $A$  is an  $OA(4n, M_{2n} + 1, 2, 2)$ .

*Proof.* We need only show that each 2-tuple,  $(0\ 0)$ ,  $(0\ 1)$ ,  $(1\ 0)$ ,  $(1\ 1)$ , appears exactly  $n$  times as a row in every  $4n \times 2$  submatrix of  $A$ . Let  $a, b, c$  and  $d$  be the number of times each appears in the submatrix defined by any pair of columns of  $A$ . By Remark 1 above,  $a + b + c + d = 4n$ ; by Remark 2,  $b + d = c + d = 2n$ ; by Remark 3,  $a + d = 2n$ . Solving this system gives  $a = b = c = d = n$ , as required.  $\square$

One can also view the binary string problem in a graph-theoretic context by using our relatively unpredictable strings to construct balanced signed graphs with nice additional properties. Recall the following definitions and theorem from [2].

**Definition.** A *signed graph*  $G$  is a graph, where each edge has been assigned the symbol  $+$  or  $-$ .

A signed graph  $G$  is *balanced* if its vertex set can be partitioned into two subsets (one of which may be empty) so that each edge joining two vertices in the same subset is positive, while each edge joining vertices in different subsets is negative.



**Theorem 5.** A signed graph  $G$  is balanced if and only if every cycle of  $G$  has an even number of edges labeled  $-$ .

Let  $s \in S^{2n}$ , with  $n$  even. We associate to  $s$  a signed graph  $G_s$  by labeling the vertices of the  $2n$ -cycle (counter-clockwise) with the digits of  $s$ . The function  $l : E(G_s) \rightarrow \{+, -\}$ , defined by

$$l(uv) = \begin{cases} + & \text{if } u + v \equiv 0 \pmod{2} \\ - & \text{otherwise,} \end{cases}$$

induces a labeling of the edges  $uv$  in this cycle.

*Remarks.*

1. The signed graph  $G_s$  is a balanced.
2.  $\omega_1(s)$  is equal to the number of edges of  $G_s$  labeled  $-$ .
3. If  $s \in S^{2n}$ ,  $g(s) > 0$ , then  $G_s$  has  $n$  edges labeled  $+$  and  $n$  labeled  $-$ .
4. Balanced graphs satisfying Remark 3 above can also be obtained by this procedure from sequences  $s \notin S^{2n}$ . (See Figure 3.)

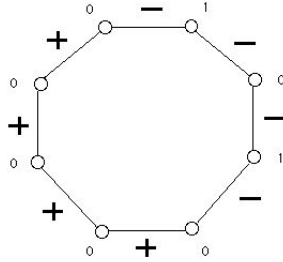


FIGURE 3.  $s = 00000101 \notin S^{2n}$ .

But we can do more than this.

*Notation.* Let  $G_s(i)$  be the signed graph (possibly disconnected) having the same vertices as  $G_s$ , but whose edges correspond to pairs of entries of  $s$  distance  $i$  apart, and edge labels are induced by function  $l$  above.

*Remark.*

5. Let  $s \in S^{2n}$ ,  $g(s) = M_n$ . Then,  $G_s, G_s(2), \dots, G_s(M_n)$  are all balanced signed graphs, and each has  $n$   $+$ 's and  $n$   $-$ 's. Further, since they are edge-disjoint and  $G = G_s \cup G_s(2) \cup \dots \cup G_s(M_n)$  is a balanced signed graph with an equal number ( $nM_n$ ) of edges labeled  $+$  and  $-$ . (See Figure 4.)

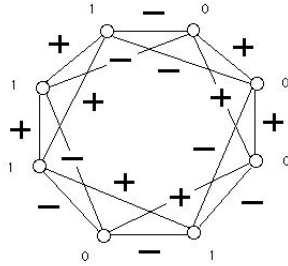


FIGURE 4.  $s = 00010111$ , attaining  $M_4 = 2$ , and the corresponding balanced sign graph.

## 6. CONJECTURES AND OPEN PROBLEMS

**Conjecture 1.** For even  $n \geq 2$ ,  $M_n \geq \frac{n}{2}$ .

**Conjecture 2.** For all  $n > 8$ ,  $M_n < n - 2$ .

**Open Problem 1.** Find a generating function or recurrence relation for the sequence  $\{M_{2n}\} = \{1, 2, 4, 6, 6, 8, 8, 11, 13, 16, 15, 16, 19, \dots\}$

**Open Problem 2.** What is the limiting behavior of  $f(n) = \frac{M_n}{n}$ ?

**Open Problem 3.** Aside from  $M_{22}$ ,  $M_n$  appears to be monotonically increasing as a function of  $n$ . Prove this or find another counterexample.

## REFERENCES

- [1] Blum, M. and Micali, S., *How to generate cryptographically strong pseudorandom sequences*, SIAM Journal of Computing, **13**, no. 4 (1984), 850–864.
- [2] Chartrand, G., *Introductory Graph Theory*, Dover Publications, (1985).
- [3] Golomb, S.W., *Shift Register Sequences*, Aegean Park Press, (1987).
- [4] Hedayat, A.S., Sloane, N.J.A., and Stufken, John, *Orthogonal Arrays: Theory and Applications*. New York: Springer-Verlag (1999).
- [5] Hedlund, G., *Endomorphisms and automorphisms of the shift dynamical system*, Mathematical Systems Theory, **3**, no. 4 (1969), 320–375.
- [6] Kreher, D.L. and Stinson, D.R., *Combinatorial Algorithms*, CRC Press, (1999).
- [7] Massey, J.L., *Shift-register synthesis and BCH decoding*, IEEE Transactions on Information Theory, **IT-15** (1969), 122–127.
- [8] Newton, D., personal communication.
- [9] Ritter, J., personal communication.
- [10] Rueppel, R.A., *Analysis and Design of Stream Ciphers*, Springer, (1986).
- [11] Shamir, A., *On the generation of cryptographically strong pseudorandom sequences*, 8th International Colloquium on Automata, Languages and Programming (Lecture Notes in Computer Science 62), Springer (1981), 544–550.
- [12] Shannon, C.E., *Communication theory of secrecy systems*, Bell System Technical Journal, **28-4** (1949), 656–715.

- [13] Stamp, M. and Martin, C.F., *An algorithm for the  $k$ -error linear complexity of binary sequences with period  $2^n$* , IEEE Transactions on Information Theory, **IT-39**, no. 4 (1993), 1398–1401.
- [14] West, D., *Introduction to Graph Theory*, Second Edition, Prentice Hall, (2000).

DEPARTMENT OF MATHEMATICS, SAN JOSE STATE UNIVERSITY, SAN JOSE, CA 95192, USA

*E-mail address:* low@math.sjsu.edu

DEPARTMENT OF COMPUTER SCIENCE, SAN JOSE STATE UNIVERSITY, SAN JOSE, CA 95192, USA

*E-mail address:* stamp@cs.sjsu.edu

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF MANITOBA, WINNIPEG, MANITOBA R3T 2N2, CANADA

*E-mail address:* craigenr@cc.umanitoba.ca

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF MANITOBA, WINNIPEG, MANITOBA R3T 2N2, CANADA

*E-mail address:* umfauch2@cc.umanitoba.ca