A. M. Patel

# Maximal Group Codes with Specified Minimum Distance*

**Abstract:** All $n$-digit maximal block codes with a specified minimum distance $d$ such that $2d \geq n$ can be constructed from the Hadamard matrices. These codes meet the Plotkin bound. In this paper we construct all maximal group codes in the region $2d \geq n$, where $d$ is a specified minimum distance and $n$ is the number of digits per code word. Unlike the case of block codes, the Plotkin upper limit, in general, fails to determine the number of code words $B(n, d)$ in a maximal group code in the region $2d \geq n$. We show that the value of $B(n, d)$ largely depends on the binary structure of the number $d$. An alogrithm is developed that determines $B(n, d)$, the maximum number of code words for given $d$ and $n \leq 2d$. The maximal code is, then, given by its modular representation, explicitly in terms of certain binary coefficients and constants related to $n$ and $d$. As a side result, we obtain a new upper bound on the number of code words in the region $2d < n$ which is, in general, stronger than Plotkin's extended bound.

## 1. Introduction

All maximal block codes in the region $2d \geq n$ can be constructed from Hadamard matrices.[1] It has been shown that these codes meet the Plotkin[2] upper limit. In the present paper, we consider maximal group codes in the region $2d \geq n$.

The set $C_n$ of all $n$-digit sequences in ZERO and ONE is an Abelian group. The group operation $\oplus$ between its elements $\alpha$ and $\beta$ is defined as

$$\alpha \oplus \beta = [\alpha(1) +_2 \beta(1), \alpha(2) +_2 \beta(2), \cdots, \alpha(n) +_2 \beta(n)],$$

where $+_2$ denotes addition modulo 2. The unity element of $C_n$ is

$$\phi = [0, 0, \cdots, 0].$$

If $M$ is a subgroup of $C_n$, $M$ forms a group code that possesses the group property that $\phi$ is a code word in $M$ and, if $\omega_i$ and $\omega_j$ are code words in $M$, then $\omega_i \oplus \omega_j$ is also a code word in $M$. It is easy to show that the number of code words in a binary group code is equal to $2^k$, where $k$ is a positive integer.† If $d$ is the minimum Hamming distance[3] between the code words, the group code $M$ may be denoted by $M(n, d; 2^k)$.

Let $B(n, d)$ denote the number of code words in a maximal $n$-digit group code with minimum distance $d$. Since group codes are a subclass of block codes, it is clear that

$$B(n, d) \leq 2^k \leq A(n, d) < 2^{k+1}, \tag{1}$$

where $A(n, d)$ denotes the number of code words in the corresponding maximal block code.

One might be tempted to conjecture[4] that

$$B(n, d) = 2^k \leq A(n, d) < 2^{k+1}.$$

This would be false, in general, because it can be shown by exhaustive search that a 16-word group code does not exist for $n = 19$ and $d = 10$ in spite of the fact that $A(n, d) = 20$. Indeed, there exists no general method of obtaining maximal group codes from the corresponding maximal block codes, or vice versa. The Hamming[3] codes, Golay codes,[5] MacDonald codes[6] and Solomon-Stiffler codes[7] (which are a generalization of the MacDonald codes) are examples of maximal group codes. The attainment of maximal group codes satisfying $2d \geq n$ has been studied before, most notably by Griesmer.[8] He gave a bound on the minimum value of $n$, given $k$ and $d$, while the present paper considers the maximum $k$, given $n$ and $d$. The new results provide necessary *and* sufficient conditions, whereas the Griesmer Bound is a necessary condition. In the present paper we construct *all* the maximal group codes in the region $2d \geq n$, and provide a new bound for maximal codes in the region $2d < n$. In the process, we show that $B(n, d)$ is strongly related to the binary structure of the number $d$.

## 2. Preliminaries

Here, we introduce some notation and pertinent theory regarding group codes. It can be easily verified that the group set $C_n$ of all $n$-digit sequences in ZERO and ONE forms an $n$-dimensional vector space over the field GF(2), and a group code $M$ is a $k$-dimensional subspace of $C_n$. Any set of basis vectors for the subspace $M$ can be con-

The author is located at the IBM Systems Development Division Laboratory in Poughkeepsie, New York 12602.

* The material of this paper is taken from the author's Ph.D. dissertation, University of Colorado, 1969.

† All numbers will be assumed to be positive integers unless otherwise stated.

sidered as rows of a matrix **G**, called the generator of $M$. Clearly, the number of rows in **G** must be $k$. Hence, the columns of **G** are elements of the set of $k$-tuples in ZERO and ONE. The column of all ZEROS may be ruled out as useless since it does not contribute to the distance between code words. Thus, there are $2^k - 1$ different types of columns possible. If rearrangement of columns is unimportant, a group code can be specified by a list of the number of columns of each type in the generator matrix **G**. This is known as the modular representation[5,6,9] of the group code.

In a group code, because of its group property, the distance between any two code words is the weight of some non-$\phi$ code word, and vice versa. A list of weights of the non-$\phi$ code words, called the weight vector, completely specifies the distance properties of the group code. The minimum distance of a group code is the minimum of the nonzero weights of its code words.

The weight vector and the modular representation of a group code are related to each other by a simple matrix equation.[5,6] The results concerning this relationship are stated below.

Let $\mathbf{G}_0$ be a $k \times (2^k - 1)$ matrix in which the $j$th column is a binary $k$-tuple that represents the binary number $j$, $j = 1, 2, \cdots, 2^k - 1$. Then $\mathbf{G}_0$ has one column of each type other than the column of all ZEROS. The $j$th column of $\mathbf{G}_0$ can be considered as the column of type $j$. The modular representation of a $k$-dimensional group code is, then, a vector

$$\mathbf{N} = [n_1, n_2, \cdots, n_{2^k-1}]$$

in which $n_i$ is the number of columns of type $i$. If **G** is the generator matrix of a group code, then it is clear that the $(2^k - 1) \times n$ matrix

$$\mathbf{B} = \mathbf{G}_0^T\mathbf{G}, \tag{2}$$

where $T$ signifies the transpose of a matrix, has as rows all possible non-$\phi$ linear combinations of rows of **G**, and thus has all non-$\phi$ code words as rows. Similarly, the matrix

$$\mathbf{C} = \mathbf{G}_0^T\mathbf{G}_0 \tag{3}$$

has as rows all non-$\phi$ code words of the code generated by $\mathbf{G}_0$. It is clear that the matrix **C** is symmetric and contains exactly one column of each type that may appear in a $k$-dimensional group code matrix. If **W** is the vector resulting from multiplying, as matrices of real numbers, the modular representation vector of a group code by the matrix **C**, then the $2^k - 1$ components of the vector **W** represent the weights of the $2^k - 1$ non-$\phi$ code words.

$$\mathbf{W} = \mathbf{NC}. \tag{4}$$

It can be shown[4] that each column of the matrix **C** has $(2^{k-1})$ ONES and $(2^{k-1} - 1)$ ZEROS and that the inverse of the matrix **C** exists and is given by

$$\mathbf{C}^{-1} = \frac{2\mathbf{C} - \mathbf{J}}{2^{k-1}}, \tag{5}$$

where **J** denotes a matrix of all ONES.

Any column in a group code matrix, with the row $\phi$ deleted, must be identical to one of the columns of the matrix **C**. Therefore, all columns in a group code matrix have $(2^{k-1})$ ONES and the total number of ONES in a group code matrix is equal to $n(2^{k-1})$.

The above material on the structure of group codes is needed in the proofs of the various theorems in this paper. The following material is a review of some known[2,3,6] results concerning $B(n, d)$ that are used later in the paper.

Let $m$ and $u$ be positive integers such that $u < m$. Then,

$$B(n, d) \leq 2B(n - 1, d); \tag{6}$$

$$B(n, d) \leq 2^m \leq \frac{2d}{2d - n} < 2^{m+1}, \qquad 2d > n; \tag{7}$$

$$B(n, d) \leq 2^m \leq 4d(2^{n-2d}) < 2^{m+1}, \qquad 2d \leq n; \tag{8}$$

$$B(n, 2m) = B(n - 1, 2m - 1); \tag{9}$$

$$B(2^m - 1, 2^{m-1}) = 2^m; \tag{10}$$

$$B(2^m - 2^u, 2^{m-1} - 2^{u-1}) = 2^m. \tag{11}$$

The results given by Eqs. (6), (7) and (8) are derived from Plotkin's paper,[2] the result given by Eq. (9) is due to Hamming[3] and the results given by Eqs. (10) and (11) are due to MacDonald.[6] In addition, we make the following assertions:

$$B(n, d) \geq B(n, d + m); \tag{12}$$

$$B(n, d) \geq B(n - m, d); \tag{13}$$

$$B(n_1 + n_2, d_1 + d_2) \geq \text{minimum of } [B(n_1, d_1), B(n_2, d_2)]. \tag{14}$$

The truth of assertion (12) is clear since a code $M'(n, d; 2^k)$ can be obtained from a code $M(n, d + m; 2^k)$ simply by replacing $m$ columns that contain a ONE in the position corresponding to a row of nonzero minimum weight in $M$, by $m$ columns of all ZEROS. If columns of all ZEROS must be avoided, one may use columns of the type that contains a ZERO in the position corresponding to the row under consideration. Assertion (13) is obvious. The generator matrix of the code $M(n_1 + n_2, d'; 2^k)$, with $d' \geq d_1 + d_2$, may be obtained by concatenating the rows of the generator matrices of the codes $M(n_1, d_1; 2^k)$ and $M(n_2, d_2; 2^k)$. This proves assertion (14).

Here we summarize the results developed in this paper. In Theorem 1 we obtain all of the group codes in which $B(n, d)$ meets the Plotkin upper limit given by Eq. (7) with an equality sign. These codes are called maximum uniform distance codes (MUDC's).[1] In Theorem 2 we

show that an MUDC can be separated from any group code in which $2d > n$. Theorem 3 then establishes a connection between the maximal codes in the region $2d > n$ and the maximal codes on the line $2d = n$. In fact, it shows that one can be obtained from the other by either separating or concatenating an appropriate MUDC. In Theorem 4, we prove an auxiliary result regarding maximal codes on the line $2d = n$, which is used in Theorem 5 where we uncover the strong dependance of $B(2d, d)$ on the binary structure of the number $d$. The main result of the paper is Theorem 6, where we combine the results of Theorems 3 and 5 and develop an algorithm for obtaining $B(n, d)$ for all pairs of $n$ and $d$ such that $2d \geq n$. The modular representation of any maximal code in the region $2d \geq n$ is obtained explicitly in terms of certain binary coefficients and constants related to $n$ and $d$. We show that in the case of group codes the Plotkin upper limit is not met in an infinite number of cases. As a result, Theorem 7 furnishes a new upper bound on $B(n, d)$ in the region $2d < n$, which is, in general, stronger than Plotkin's result, Eq. (8).

### 3. Maximal codes in the region $2d \geq n$

A maximum uniform distance code (MUDC) is defined as a code $M(n, d; g)$ in which $g = 2d/(2d - n)$. The basic property of the MUDC is that the distance between any two of its code words is equal to its minimum distance. In the case of maximal group codes in the region $2d \geq n$, MUDC's play an important role. In the following theorem we construct all group MUDC's.

*Theorem 1:* Given $n$ and $d$ such that $2d > n$ and $d/(2d - n)$ is a positive integral power of 2, then

$$B(n, d) = 2^k = \frac{2d}{2d - n}.$$

*Proof:* Let $t = 2d - n$, where $t$ is a positive integer. Then,

$2d = t(2^k)$, and

$n = t(2^k - 1)$.

From Eq. (10) we have

$B(2^k - 1, 2^{k-1}) = 2^k$.

It is easily recognized that a $t$-fold concatenation of the code $M(2^k - 1, 2^{k-1}; 2^k)$ gives us a code $M(n, d; 2^k)$; see Eq. (14). This, in conjunction with Eq. (7), implies that

$$B(n, d) = 2^k = \frac{2d}{2d - n}.$$

Note that all the codes of Theorem 1 are MUDC's. In view of Eq. (7) it is clear that these are the only MUDC's in the case of group codes.

MacDonald[3,6] has shown that the modular representation of the maximal code $M(2^k - 1, 2^{k-1}; 2^k)$ [see Eq. (10)] is

$$\mathbf{N} = [1, 1, \cdots, 1]. \tag{15}$$

That is, the code matrix $\mathbf{M}$ contains one and only one column of each type. In Theorem 1, we showed that an MUDC with $2d - n = t > 0$ is a $t$-fold concatenation of the code $M(2^k - 1, 2^{k-1}; 2^k)$. Consequently, its code matrix contains exactly $t$ columns of each type. Using the results of Theorem 2, which follows, we can deduce that this is also a necessary condition for the code matrix of an MUDC. In the MUDC the distance between any two of its code words is equal to the code's minimum distance. This implies that the weight of any non-$\phi$ code word in an MUDC is equal to $d$. Note that in a $k$-dimensional $t$-fold MUDC

$n = t(2^k - 1)$, and

$d = t(2^{k-1})$.

It is interesting to note that the Maximum Length Shift Register Codes[5] are MUDC's and that all MUDC's can be constructed to be cyclic by concatenation of the Maximum Length Shift Register Codes.

In the following theorem we prove that every group code with $2d > n$ contains an MUDC.

*Theorem 2:* In a group code if $2d - n = t > 0$, then the code contains a $t$-fold MUDC; that is, the generator matrix contains at least $t$ columns of each type.

*Proof:* The modular representation of a $k$-dimensional code is the vector

$$\mathbf{N} = [n_1, n_2, \cdots, n_{2^k-1}]$$

in which $n_i$ is the number of columns of type $i$. We show that when $2d - n = t > 0$, $n_i$ is greater than or equal to $t$ for all $i$. This is done by using the relationship between the modular representation vector and the weight vector of the code.

One can obtain the weight vector of the code from Eq. (4) using matrix $\mathbf{C}$ of Eq. (3).

$$\mathbf{W} = \mathbf{NC} = [w_1, w_2, \cdots, w_{2^k-1}], \tag{16}$$

in which $w_i$ is the weight of the $i$th code word. In a group code the constraint of minimum distance requires that $w_i \geq d$ for all $i$. Let

$$w_i = d + x_i, \qquad i = 1, 2, \cdots, 2^k - 1. \tag{17}$$

Then the $x_i$'s are nonnegative integers. Another constraint on the values of $x_i$ comes from the fact that the sum of all $w_i$'s is equal to the total number of ONES in a code matrix, which is equal to $n2^{k-1}$. Hence,

$$\sum_{i=1}^{2^k-1} w_i = (2^k - 1)\, d + \sum_{i=1}^{2^k-1} x_i = n(2^{k-1}),$$

which can be rewritten as

$$\sum_{i=1}^{2^k-1} x_i = d - t(2^{k-1}). \tag{18}$$

Now, substituting for the $w_i$'s in Eq. (16) and then postmultiplying both sides by $\mathbf{C}^{-1}$, we get the modular representation vector

$$\mathbf{N} = [d + x_1, d + x_2, \cdots, d + x_{2^k-1}][\mathbf{C}^{-1}]. \tag{19}$$

Substituting for $\mathbf{C}^{-1}$ from Eq. (5) and denoting the matrix $[2\mathbf{C} - \mathbf{J}]$ by $\mathbf{D}$, one can rewrite Eq. (19) as

$$\mathbf{N} = \frac{1}{2^{k-1}} [d + x_1, d + x_2, \cdots, d + x_{2^k-1}][\mathbf{D}]. \tag{20}$$

From the properties of $\mathbf{C}$ it is clear that $\mathbf{D}$ is a symmetric matrix with all its elements either +ONE or −ONE and that $\mathbf{D}$ has exactly $(2^{k-1})$ +ONES and $(2^{k-1} - 1)$ −ONES in each column. Let $h_{ji}$ denote the element in the $j$th row and $i$th column of $\mathbf{D}$. Then, in view of the properties of $\mathbf{D}$, one can deduce from Eq. (20) that the $i$th element of vector $\mathbf{N}$ is

$$n_i = \frac{1}{2^{k-1}} \left(d + \sum_{j=1}^{2^k-1} h_{ji} x_j\right); \qquad i = 1, 2, \cdots, 2^k - 1. \tag{21}$$

Using Eq. (18), one can rewrite this as

$$n_i = t + \frac{1}{2^{k-1}} \sum_{j=1}^{2^k-1} (1 + h_{ji}) x_j; \qquad i = 1, 2, \cdots, 2^k - 1. \tag{22}$$

Since $x_j$ is nonnegative and $h_{ji}$ is either +ONE or −ONE for all $i$ and $j$, it is clear that all the terms in the summation in Eq. (22) are nonnegative. Thus, $n_i \geq t$, which completes the proof.

Theorem 2 shows that an MUDC can always be separated from a group code in which $2d > n$. In Theorem 3, this fact is used to relate maximal codes in the region $2d > n$, with the maximal codes on the line $2d = n$.

*Theorem 3:* Given $n$ and $d$ such that $2d - n = t > 0$, then for any positive integer $k$,

$$B(n, d) \geq 2^k \quad \text{if and only if} \quad B\left(n', \frac{n'}{2}\right) \geq 2^k,$$

where

$$n' = 2d - t2^k.$$

*Proof:* Suppose $B(n, d) \geq 2^k$. Then there exists a group code $M(n, d; 2^k)$. Since $2d - n = t > 0$, Theorem 2 implies that the code matrix $\mathbf{M}$ can be separated into two parts $\mathbf{X}$ and $\mathbf{Y}$, where $\mathbf{X}$ consists of $t$ columns of each type and $\mathbf{Y}$ consists of the remaining columns in $\mathbf{M}$. Clearly, $\mathbf{X}$ is a $k$-dimensional $t$-fold MUDC. Consequently, the distance between any two code words in $\mathbf{X}$ is equal to $t(2^{k-1})$. Since the minimum distance in $\mathbf{M}$ is $d$, it is clear that the distance between any two rows in $\mathbf{Y}$ is at least $d - t(2^{k-1})$. This means that $2^k$ rows in $\mathbf{Y}$ are distinct and the minimum distance in $\mathbf{Y}$ is $d - t(2^{k-1})$. The length of rows in $\mathbf{X}$ is equal to $t(2^k - 1)$. The length of rows in $\mathbf{Y}$ is the difference $n - t(2^k - 1)$, which is $2d - t(2^k)$. Thus, $\mathbf{Y}$ is the group code $Y(n', n'/2; 2^k)$. This proves that $B(n', n'/2) \geq 2^k$.

In order to prove the converse, suppose $B(n', n'/2) \geq 2^k$. Then there exists a group code $Y(n', n'/2; 2^k)$. We can obtain a group code $M(n, d; 2^k)$ by concatenating $\mathbf{Y}$ to a $k$-dimensional $t$-fold MUDC. This implies that $B(n, d) \geq 2^k$, which completes the proof.

*Corollary 3-1:* Given $n$ and $d$ such that $2d - n = t > 0$, then

$$B(n, d) = B\left(n', \frac{n'}{2}\right),$$

where $n' = 2d - tB(n, d)$.
The proof follows directly from Theorem 3.

Corollary 3-1 establishes a connection between maximal codes in the region $2d > n$ and the maximal codes on the line $2d = n$. In fact, one may be obtained from another by either separating or concatenating an appropriate MUDC.

Some known maximal codes on the line $2d = n$ are the codes $M(2^k - 2^u, 2^{k-1} - 2^{u-1}; 2^k)$ given by MacDonald[6] [see Eq. (11)]. These codes can be constructed by taking one column of each type $j$ such that $j \geq 2^u$. Thus the modular representation vector of these codes is given by the vector

$$\mathbf{N} = [0, 0, \cdots, 0, 1, 1, \cdots, 1] \tag{23}$$

composed of $(2^u - 1)$ ZEROS followed by $(2^k - 2^u)$ ONES.

In the following two theorems we deal with $B(n, d)$ on the line $2d = n$ and show that all maximal codes are substantially related to the codes $M(2^k - 2^u, 2^{k-1} - 2^{u-1}; 2^k)$ and the MUDC's, which we call the parent codes.

*Theorem 4:* If $n$ is a positive integer divisible by 4, then

$$B\left(\frac{n}{2}, \frac{n}{4}\right) = \tfrac{1}{2} B\left(n, \frac{n}{2}\right).$$

*Proof:* Suppose $B(n, n/2) = 2^k$ and $\mathbf{M}$ is the code matrix for the corresponding code $M(n, n/2; 2^k)$. Let $\beta_m$ be a row in $\mathbf{M}$ with weight equal to $n/2$. There is at least one such row in $\mathbf{M}$ since the minimum distance is

$n/2$. Clearly, $\beta_m$ consists of $n/2$ ONES and $n/2$ ZEROS. Rearrange the columns of the matrix $\mathbf{M}$ so that the first $n/2$ columns have a ZERO in the place corresponding to the row $\beta_m$. The new matrix, denoted by $\mathbf{M}_p$, is a code $M_p(n, n/2; 2^k)$. Let $\omega_m$ be the row in $\mathbf{M}_p$ corresponding to the row $\beta_m$ in $\mathbf{M}$. Select $k$ rows, one of which is $\omega_m$, from $\mathbf{M}_p$ to form an independent set $\mathbf{G}$. This is possible because $\mathbf{M}_p$ is a $k$-dimensional code. Let $\mathbf{M}_q$ be the code generated by the $k-1$ rows of $\mathbf{G}$, other than $\omega_m$. Clearly, $\mathbf{M}_q$ is a subspace of $\mathbf{M}_p$, and, hence, it is a code $M_q(n, n/2; 2^{k-1})$. Partition the matrix $\mathbf{M}_q$ into two parts $\mathbf{X}$ and $\mathbf{Y}$, where $\mathbf{X}$ consists of first $n/2$ columns of $\mathbf{M}_q$ and $\mathbf{Y}$ consists of the remaining $n/2$ columns of $\mathbf{M}_q$. We show next that $\mathbf{X}$ has minimum distance of at least $n/4$.

If $\alpha_i$, $i = 1, 2, \cdots, 2^k - 1$, is a non-$\phi$ row in $\mathbf{M}_q$, then $\alpha_i \oplus \omega_m$ is a non-$\phi$ row in $\mathbf{M}_p$. Since the minimum distance in $\mathbf{M}_p$ and $\mathbf{M}_q$ is $n/2$, we have

$$|\alpha_i| \geq \frac{n}{2} \quad \text{and} \tag{24}$$

$$|\alpha_i \oplus \omega_m| \geq \frac{n}{2}. \tag{25}$$

Denote the partitions of $\alpha_i$ in $\mathbf{X}$ and $\mathbf{Y}$ by $\alpha_i'$ and $\alpha_i''$, respectively. Similar partitions of $\omega_m$ are denoted by $\omega_m'$ and $\omega_m''$. Then,

$$|\alpha_i| = |\alpha_i'| + |\alpha_i''| \text{ and} \tag{26}$$

$$|\alpha_i \oplus \omega_m| = |\alpha_i' \oplus \omega_m'| + |\alpha_i'' \oplus \omega_m''|. \tag{27}$$

Since, by construction, $\omega_m'$ is a sequence of all ZEROS and $\omega_m''$ is a sequence of all ONES, we have

$$|\alpha_i' \oplus \omega_m'| = |\alpha_i'|, \tag{28}$$

$$|\alpha_i'' \oplus \omega_m''| = \frac{n}{2} - |\alpha_i''|. \tag{29}$$

From Eqs. (24) through (29) it can be shown that

$$|\alpha_i'| \geq \frac{n}{4}, \qquad i = 1, 2, \cdots, 2^k - 1.$$

This implies that $\mathbf{X}$ represents a code $X(n/2, d_x; 2^{k-1})$ in which

$$d_x \geq \frac{n}{4}.$$

Then from Eq. (12) we obtain

$$B\left(\frac{n}{2}, \frac{n}{4}\right) \geq B\left(\frac{n}{2}, d_x\right) \geq 2^{k-1} = \tfrac{1}{2}B\left(n, \frac{n}{2}\right).$$

Now we give the converse in order to complete the proof of the theorem. Suppose $B(n/2, n/4) = 2^k$. Let $\mathbf{G}$ be the generator matrix of the group code $M(n/2, n/4; 2^k)$. Let $\mathbf{G}_1$ be the matrix obtained by concatenating each row of $\mathbf{G}$ to itself. Clearly, the code $\mathbf{M}_1$ generated by $\mathbf{G}_1$ is $M_1(n, n/2; 2^k)$. Let $\mathbf{G}_2$ be the matrix obtained by adding a row of $n/2$ ZEROS and $n/2$ ONES to the matrix $\mathbf{G}_1$. Obviously the $k+1$ rows of $\mathbf{G}_2$ form an independent set. Let $\mathbf{M}_2$ be the code generated by $\mathbf{G}_2$. Then $\mathbf{M}_1$ is a subspace of $\mathbf{M}_2$. We show next that the minimum distance in $\mathbf{M}_2$ is equal to $n/2$.

If $\omega$ is any non-$\phi$ row in $\mathbf{M}_2$, then either it is in $\mathbf{M}_1$, and is thus a concatenation of some non-$\phi$ row $\alpha$ in $\mathbf{M}$ to itself, or it is a concatenation of some non-$\phi$ row $\beta$ in $\mathbf{M}$ to $\beta \oplus \Psi$, where $\Psi$ is a sequence of $n/2$ ONES. We observe that in one case

$$|\omega| = |\alpha, \alpha| = |\alpha| + |\alpha| \geq \frac{n}{2}$$

and in the other case

$$|\omega| = |\beta, \beta \oplus \Psi| = |\beta| + \frac{n}{2} - |\beta| = \frac{n}{2}.$$

This implies that $\mathbf{M}_2$ is a code $M_2(n, n/2; 2^{k+1})$. That is,

$$B\left(n, \frac{n}{2}\right) \geq 2^{k+1} = 2B\left(\frac{n}{2}, \frac{n}{4}\right).$$

This completes the proof of the theorem.

*Theorem 5:* If $n$ is a positive integer divisible by 4, and $k$ is any positive integer, $n$ may be expressed as

$$n = x2^k - \sum_{m=2}^{k-1} a_m 2^m,$$

where $x$ is a positive integer dependent on $k$ and $a_m = 0$ or 1 for all $m$. Then,

$$B\left(n, \frac{n}{2}\right) \geq 2^k \quad \text{if and only if} \quad x \geq \sum_{m=2}^{k-1} a_m.$$

*Proof:* Suppose

$$x \geq \sum_{m=2}^{k-1} a_m.$$

Let

$$t = x - \sum_{m=2}^{k-1} a_m;$$

$t$ is a nonnegative integer. Then $n$ can be partitioned as $n_1 + n_2$, where

$$n_1 = t2^k,$$

$$n_2 = \sum_{m=2}^{k-1} a_m(2^k - 2^m).$$

From Eq. (13) and Theorem 1, it is clear that

$$B\left(n_1, \frac{n_1}{2}\right) \geq B\left(n_1 - t, \frac{n_1}{2}\right) = 2^k.$$

Also, by applying Eq. (14) repeatedly to Eq. (11), we obtain

$$B\left(n_2, \frac{n_2}{2}\right) \geq 2^k.$$

Applying Eq. (14) once again, we obtain

$$B\left(n, \frac{n}{2}\right) \geq \text{ minimum of } B\left(n_1, \frac{n_1}{2}\right) \text{ and}$$

$$B\left(n_2, \frac{n_2}{2}\right) \geq 2^k.$$

We prove the remaining part of the theorem by proving the contrapositive. That is, we show that if

$$B\left(n, \frac{n}{2}\right) \geq 2^k \quad \text{when} \quad x < \sum_{m=2}^{k-1} a_m,$$

then a known result is contradicted. We assume that $k > 2$, since the other cases are not applicable here. For the proof we need the following two lemmas.

*Lemma 1:* If $n < 2^{k-1}$, then $B(n, n/2) < 2^k$.

*Proof:* If $n < 2^{k-1}$, from Eq. (7) it is clear that

$$B\left(n - 1, \frac{n}{2}\right) \leq n < 2^{k-1}.$$

Then from Eq. (6) we get

$$B\left(n, \frac{n}{2}\right) \leq 2B\left(n - 1, \frac{n}{2}\right) < 2^k.$$

*Lemma 2:* Given

$$n = 2^k - \sum_{m=2}^{k-1} a_m 2^m.$$

If

$$\sum_{m=2}^{k-1} a_m > 1, \quad \text{then} \quad B\left(n, \frac{n}{2}\right) < 2^k.$$

*Proof:* If $a_{k-1} = 1$, clearly $n < 2^{k-1}$ and Lemma 1 completes the proof. In general, say

$$a_{k-1} = a_{k-2} = \cdots = a_{k-j} = 0 \text{ and } a_{k-j-1} = 1.$$

Then

$$n = 2^k - \sum_{m=2}^{k-j-1} a_m 2^m \quad \text{and} \quad \sum_{m=2}^{k-j-1} a_m = \sum_{m=2}^{k-1} a_m > 1.$$

Suppose $B(n, n/2) \geq 2^k$; then, applying Eq. (6), one obtains

$$B\left(n - 1, \frac{n}{2}\right) \geq \tfrac{1}{2} B\left(n, \frac{n}{2}\right) \geq 2^{k-1}.$$

Note that Theorem 3 may be applied. Separating an MUDC, we get

$$B\left(n_1, \frac{n_1}{2}\right) \geq 2^{k-1},$$

where

$$n_1 = n - 2^{k-1} = 2^{k-1} - \sum_{m=2}^{k-j-1} a_m 2^m.$$

Repeating the above procedure an appropriate number of times, we can deduce that

$$B\left(n_2, \frac{n_2}{2}\right) \geq 2^{k-i},$$

where

$$n_2 = 2^{k-i} - \sum_{m=2}^{k-j-1} a_m 2^m.$$

Since

$$a_{m-j-1} = 1 \quad \text{and} \quad \sum_{m=2}^{k-j-1} a_m > 1,$$

we have $n_2 < 2^{k-i-1}$. This contradicts the result of Lemma 1.

*Proof of Theorem 5 (cont.):* In the case where

$$x < \sum_{m=2}^{k-1} a_m,$$

suppose $B(n, n/2) \geq 2^k$. If $x = 1$, Lemma 2 is contradicted. Suppose $x > 1$. In general, if $a_m = 0$ for $m = 2, 3, \cdots, j$ and $a_{j+1} = 1$, then applying Theorem 4 $j$ times, we obtain

$$B\left(n_1, \frac{n_1}{2}\right) \geq 2^{k-i},$$

where

$$n_1 = \frac{n}{2^i} = x2^{k-i} - \sum_{m=j+1}^{k-1} a_m 2^{m-i}.$$

Note that

$$\sum_{m=j+1}^{k-1} a_m = \sum_{m=2}^{k-1} a_m > x.$$

Since $a_{j+1} = 1$, it is clear that $n_1/2$ is an odd integer. Hence, using Eq. (9) we deduce that

$$B\left(n_1 + 1, \frac{n_1}{2} + 1\right) = B\left(n_1, \frac{n_1}{2}\right) \geq 2^{k-i}.$$

Note that $2(n_1/2 + 1) - (n_1 + 1) = 1$. Hence, according to Theorem 3, we can separate an MUDC and obtain

$$B\left(n_2, \frac{n_2}{2}\right) \geq 2^{k-i},$$

where

$$n_2 = 2\left(\frac{n_1}{2} + 1\right) - 2^{k-i}$$

$$= (x - 1)2^{k-i} - \sum_{m=i+1}^{k-1} a_m 2^{m-i} + 2.$$

Substituting $x_1$ for $x - 1$, $k_1$ for $k - j$ and $a'_m$ for $a_{m-i}$ and rearranging the expression, we can rewrite $n_2$ as

$$n_2 = x_1 2^{k_1} - \sum_{m=2}^{k_1-1} a'_m 2^m.$$

It can be verified that

$$\left[\sum_{m=2}^{k_1-1} a'_m\right] - x_1 = \left[\sum_{m=2}^{k-1} a_m\right] - x.$$

By repeating the above procedure an appropriate number of times we can show that

$$B\left(n_3, \frac{n_3}{2}\right) \geq 2^{k_2},$$

where

$$n_3 = 2^{k_2} - \sum_{m=2}^{k_2-1} a''_m 2^m$$

and $k_2$ and $a''_m$ are the numbers obtained in the process such that

$$\left[\sum_{m=2}^{k_2-1} a''_m\right] - 1 = \left[\sum_{m=2}^{k-1} a_m\right] - x > 0.$$

But this contradicts Lemma 2, which completes the proof of the theorem.

*Corollary 5-1:* If $n$ is an even integer and $k$ is any positive integer, $n$ may be expressed as

$$n = x2^k - \sum_{m=1}^{k-1} a_m 2^m,$$

where $x$ is a positive integer dependent on $k$ and $a_m = 0$ or 1 for all $m$. Then

$$B\left(n, \frac{n}{2}\right) \geq 2^k \quad \text{if and only if} \quad x \geq \sum_{m=1}^{k-1} a_m.$$

*Proof:* Note that $2n$ is divisible by 4 and

$$2n = x2^{k+1} - \sum_{m=1}^{k-1} a_m 2^{m+2}.$$

According to Theorems 4 and 5,

$$B\left(n, \frac{n}{2}\right) = \tfrac{1}{2}B(2n, n) \geq 2^k$$

if and only if $x \geq \sum_{m=1}^{k-1} a_m$.

In the following theorem we extend the results of Corollary 5-1 to the entire region $2d \geq n$ by using the relation obtained in Theorem 3 between maximal codes in the region $2d > n$ and maximal codes on the line $2d = n$.

*Theorem 6:* Given $n$ and $d$ such that $2d - n = t > 0$, for any positive integer $k$, $2d$ can be expressed as

$$2d = x2^k - \sum_{m=1}^{k-1} a_m 2^m,$$

where $x$ is a positive integer dependent on $k$ and $a_m = 0$ or 1 for all $m$. Then

$$B(n, d) \geq 2^k \quad \text{if and only if} \quad x \geq t + \sum_{m=1}^{k-1} a_m.$$

*Proof:* The case $t = 0$ is proved in Theorem 5. We assume here that $t > 0$. Suppose

$$x \geq t + \sum_{m=1}^{k-1} a_m.$$

Let

$$n_1 = 2d_1 = (x - t)2^k - \sum_{m=1}^{k-1} a_m 2^m.$$

Then, applying Corollary 5-1, we obtain

$$B(n_1, d_1) \geq 2^k.$$

Let

$$n_2 = t(2^k - 1) \text{ and } 2d_2 = t2^k.$$

Then, according to Theorem 1,

$$B(n_2, d_2) = 2^k.$$

Note that $n_1 + n_2 = n$ and $d_1 + d_2 = d$. Thus, applying Eq. (14) we obtain

$$B(n, d) \geq 2^k.$$

In order to prove the converse, suppose $B(n, d) \geq 2^k$. Since $2d - n = t > 0$, Theorem 3 implies that

$$B\left(n_1, \frac{n_1}{2}\right) \geq 2^k,$$

where

$$n_1 = 2d - t2^k = (x - t)2^k - \sum_{m=1}^{k-1} a_m 2^m.$$

Then Corollary 5-1 implies that

$$(x - t) \geq \sum_{m=1}^{k-1} a_m.$$

This completes the proof.

## 4. An algorithm for constructing maximal codes

Theorem 6 suggests the following algorithm to obtain $B(n, d)$ for given $n$ and $d$ in the region $2d \geq n$:

Given $n$ and $d$, $2d - n = t \geq 0$, $2d$ may be expressed as

$$2d = x_j 2^j - \sum_{m=1}^{j-1} a_m 2^m, \qquad j = 1, 2, \cdots, \tag{30}$$

where $x_j$ is a positive integer dependent on $j$ and $a_m = 0$ or 1 for all $m$. If

$$y_j = \sum_{m=1}^{j-1} a_m, \tag{31}$$

then $x_j$ and $y_j$ may be obtained from the iterative formulas

$$x_{j+1} = \begin{cases} \dfrac{x_j}{2} & \text{if } x_j \text{ is even} \\ \dfrac{x_j + 1}{2} & \text{if } x_j \text{ is odd, and} \end{cases} \tag{32}$$

$$y_{j+1} = \begin{cases} y_j & \text{if } x_j \text{ is even} \\ y_j + 1 & \text{if } x_j \text{ is odd,} \end{cases} \tag{33}$$

where $x_1 = d$ and $y_1 = 0$.

Note that in all codes $n$ is greater than or equal to $d$. Hence $d \geq 2d - n = t$; that is, $x_1 \geq y_1 + t$. Moreover, $x_j$ is a strictly decreasing function of $j$ and $y_j$ is a non-decreasing function of $j$. Therefore, it is clear that there exists a unique positive integer $k$ that satisfies the inequalities

$$x_k \geq y_k + t \text{ and } x_{k+1} < y_{k+1} + t. \tag{34}$$

Then, Theorem 6 implies that

$$B(n, d) = 2^k. \tag{35}$$

*Example 1:* Suppose $n = 762$ and $d = 386$. Then $2d - n = t = 10$. We tabulate $x_j$ and $y_j + t$;

| $j$ | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| $x_j$ | 386 | 193 | 97 | 49 | 25 | 13 |
| $y_j + t$ | 10 | 10 | 11 | 12 | 13 | 14 |

Thus, $j = 5$ is the largest value of $j$ for which $x_j \geq y_j + t$. This implies that $B(762, 386) = 2^5$.

The algorithm can also be used to find minimum $n$ for given $d$ and $k = j$ such that $x_j \geq y_j$. In this case one finds $t_j = x_j - y_j$ from $d$ and $k = j$. Then the minimum $n$ for a $2^k$-word group code with minimum distance $d$ is equal to $n_{\min} = 2d - t_j$. With specified $d = 386$ we find $n_{\min}$ for each $k = j$ as follows:

| $k = j$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|
| $x_j$ | 386 | 193 | 97 | 49 | 25 | 13 | 7 | 4 |
| $y_j$ | 0 | 0 | 1 | 2 | 3 | 4 | 5 | 6 |
| $t_j$ | 386 | 193 | 96 | 47 | 22 | 9 | 2 | — |
| $n_{\min}$ | 386 | 579 | 676 | 725 | 750 | 763 | 770 | — |

Now we are in a position to give the modular representations of the maximal codes in the region $2d \geq n$. The reader may note that the modular representation of a code is, in general, not unique, since permutations of rows in a generator matrix give rise to a different modular representation for the same code. Using Eqs. (30) to (35) of the algorithm and denoting $(x_k - y_k)$ by $t_k$, we may express $n$ and $d$ as

$$2d = t_k(2^k) + \sum_{m=1}^{k-1} a_m(2^k - 2^m) \quad \text{and} \tag{36}$$

$$n = t_k(2^k - 1) + \sum_{m=1}^{k-1} a_m(2^k - 2^m) + (t_k - t). \tag{37}$$

Note that $(t_k - t)$ is a positive integer. If $n'$ denotes $n - (t_k - t)$, then

$$n' = t_k(2^k - 1) + \sum_{m=1}^{k-1} a_m(2^k - 2^m) \quad \text{and} \tag{38}$$

$$B(n', d) = B(n, d) = 2^k.$$

It is easy to recognize from Eqs. (36) and (38) that the code $M(n', d; 2^k)$ can be obtained by concatenating the parent codes given by the modular representations of Eqs. (15) and (23). Thus the modular representation of the code $M(n', d; 2^k)$ is given by the vector

$$\mathbf{N} = [n_1, n_2, n_3, \cdots, n_{2^k-1}],$$

where $n_i$ may be given by

$$\begin{aligned} n_1 &= t_k \\ n_2 = n_3 &= t_k + a_1 \\ n_4 = n_5 = n_6 = n_7 &= t_k + a_1 + a_2 \\ n_8 = n_9 \cdots = n_{15} &= t_k + a_1 + a_2 + a_3 \\ &\vdots \\ n_{2^s} = \cdots = n_{2^{s+1}-1} &= t_k + \sum_{m=1}^{s} a_m, \qquad s \leq k - 1, \end{aligned}$$

or, written differently,

$$n_{2^s} = \cdots = n_{2^{s+1}-1} = x_k - \sum_{m=s}^{k-1} a_m.$$

The code $M'(n, d; 2^k)$ may be obtained from $M(n', d; 2^k)$ trivially by adding $(t_k - t)$ columns of all ZEROS. It can

easily be shown that adding any number of columns of even type, say type 2, to any of the parent codes does not increase the minimum distance. In fact, if $d$ is even, adding any number of columns of any one type does not increase the minimum distance of the parent codes. So if columns of all ZEROS must be avoided, they may be replaced by columns of any even type.

We pointed out previously that in case of $n = 19$, $d = 10$ the Plotkin limit Eq. (7) is not met. Using the algorithm and the Plotkin limit, one can easily verify that

$$B(19, 10) = 8$$

$$A(19, 10) = 20.$$

Here we show that indeed there exists an infinite number of pairs $(n, d)$, even when $d$ is even, for which $B(n, d)$ does not meet the Plotkin upper limit. Consider the following example: Let

$$2d = 2^p + 4 = 2(2^p) - \sum_{m=2}^{p-1} a_m 2^m,$$

where $a_m = 1$ for $m = 2, 3, \cdots, p-1$ and $n = 2^p + 3$. Then $t = 1$, $x_p = 2$ and $y_p = p - 2$.

Clearly, $B(n, d) < 2^p$ for all $p > 3$, whereas according to Eq. (7) the Plotkin upper limit is $2^p$. In the following theorem we obtain a new bound on $B(n, d)$ in the region $2d < n$, which is, in general, stronger than Plotkin's extended bound, Eq. (8).

*Theorem 7:* Given $n$ and $d$ such that $n > 2d$. For any positive integer $j$, $2d$ may be expressed as

$$2d = x_j 2^j - \sum_{m=1}^{j-1} a_m 2^m,$$

where $x_j$ is a positive integer dependent on $j$ and $a_m = 0$ or 1 for all $m$. If $k$ is the largest value of $j$ such that

$$x_k \geq \sum_{m=1}^{k-1} a_m,$$

then $B(n, d)$ is bounded above by

$$B(n, d) \leq 2^{k+n-2d}.$$

*Proof:* $B(2d, d) = 2^k$ by the algorithm of Theorem 6. Applying Eq. (1) $n - 2d$ times, we obtain

$$B(n, d) \leq 2^{k+n-2d}.$$

To illustrate the bound of Theorem 7, consider $B(n, d)$ for $d = 66$. According to Plotkin's extended bound, Eq. (8),

$$B(n, d) \leq 256(2^{n-2d}).$$

**Table 1** Dependence of $A(2d, d)$ and $B(2d, d)$ on binary value of $d$.

| $d$ | $A(2d, d)$ | $B(2d, d)$ | $d$ | $A(2d, d)$ | $B(2d, d)$ |
|---|---|---|---|---|---|
| 32 | 128 | 128 | 49 | 100 | 32 |
| 33 | 68 | 16 | 50 | 200 | 32 |
| 34 | 136 | 32 | 51 | 104 | 32 |
| 35 | 72 | 32 | 52 | 208 | 64 |
| 36 | 144 | 32 | 53 | 108 | 32 |
| 37 | 76 | 32 | 54 | 216 | 64 |
| 38 | 152 | 32 | 55 | 112 | 64 |
| 39 | 80 | 32 | 56 | 224 | 128 |
| 40 | 160 | 64 | 57 | 116 | 32 |
| 41 | 84 | 32 | 58 | 232 | 64 |
| 42 | 168 | 32 | 59 | 120 | 64 |
| 43 | 88 | 32 | 60 | 240 | 128 |
| 44 | 176 | 64 | 61 | 124 | 64 |
| 45 | 92 | 32 | 62 | 248 | 128 |
| 46 | 184 | 64 | 63 | 128 | 128 |
| 47 | 96 | 64 | 64 | 256 | 256 |
| 48 | 192 | 128 | | | |

However, $2d = 5(32) - 16 - 8 - 4$ and $B(2d, d) = 32$. Hence, Theorem 7 gives a closer bound

$$B(n, d) \leq 32(2^{n-2d}).$$

## 5. Conclusions

The main contribution of this paper is that the problem of obtaining maximal group codes in the region $2d \geq n$ is completely solved. Furthermore, a strengthened bound is obtained for $B(n, d)$ in the region $2d < n$. However, the strong dependence of the value of $B(2d, d)$, and hence of $B(n, d)$ on the binary structure of the number $d$ should not go unnoticed. To illustrate the effect of this, we give the values of $A(2d, d)$ and $B(2d, d)$ for $2^5 \leq d \leq 2^6$ in Table 1.

As shown in Theorems 3 and 7, $B(n, d)$ substantially depends on $B(2d, d)$. Hence, if one has a choice, $d$ should be chosen to have the simplest binary structure possible; that is, to have a minimal number of terms in $y_k$ of Eq. (33). The simple binary structure for $d$ means the $B(n, d)$ is large and hence the resulting code yields a relatively high rate of transmission of information.


## Acknowledgments

The author expresses his gratitude to R. A. Roberts and L. J. Griffiths of the University of Colorado for their interest, encouragement and comments.


## References


1. A. M. Patel, "Maximal Codes with Specified Minimum Distance," IBM Technical Report TR 44.0085, Nov. 1969.
2. M. Plotkin, "Binary Codes with Specified Minimum Distance," *IRE Trans. Info. Theory* **IT-6,** 445 (1960).

3. R. W. Hamming, "Error Detecting and Error Correcting Codes," *Bell System Tech. J.* **29,** 147 (1950).
4. M. Nadler, "A 32-Point $n = 12$, $d = 5$ Code," *IRE Trans. Info. Theory* **IT-8,** 58 (1962).
5. W. W. Peterson, *Error Correcting Codes,* MIT Press, Cambridge, Mass., 1961.
6. J. E. MacDonald, "Design Methods for Maximum Minimum-Distance Error-Correcting Codes," *IBM J. Res. Develop.* **4,** 43 (1960).
7. G. Solomon and J. J. Stiffler, "Algebraically Punctured Cyclic Codes," *Info. and Control* **8,** 170 (1965).
8. J. H. Griesmer, "A Bound for Error Correcting Codes," *IBM J. Res. Develop.* **4,** 532 (1960).
9. D. Slepian, "A Class of Binary Signaling Alphabets," *Bell System Tech. J.* **35,** 203 (1956).

*Received September 10, 1969*