# An Array of Disjoint Maximal Constant Weight Codes

Christine Berkesch

April 21, 2004

**Abstract**

We show that when $\gcd(n, w) = 1$, the set of binary words of length $n$ and weight $w$ can be partitioned to give $n$ maximal $w$-weight codes. It follows that under the same hypothesis, the least cardinal of a maximal constant weight code is at most $\frac{1}{n}\binom{n}{w}$.

## 1 Introduction

Communication plays a vital role in the world today. With rapid advances in technology, an increasing amount of information must be exchanged between the individual and machines. When a message is sent through a channel, interference, called noise, can cause errors. The noise includes human error, crosstalk, lightning, thermal noise, or impulse noise. Whether this information is being saved to a computer disk, sent via FAX, or transmitted in radio or mobile communication, the presence of inaccuracies can have massive consequences. Fortunately, these errors can be overcome when an initial language, or code, is constructed with certain error-detecting or correcting features. Coding theory is a relatively new branch of mathematics devoted to discovering languages, called codes, which are being successfully employed to transmit messages accurately and in a most efficient manner.

Coding theory began in 1948 after Claude Shannon, an employee at Bell Laboratories in the USA, showed that it was possible to encode messages so that the number of extra bits was as small as possible. This resulted in the publication of the classic paper *The Mathematical Theory of Communication* in the Bell System Technical Journal. Two years later, Richard Hamming, another employee at Bell Labs, devised a family of single error-correcting codes in order to enhance the performance of a computer [1].

A popular example which demonstrates the value of coding theory is the communication of Mariner 9 with NASA. This was a probe sent to Mars in 1971. In

order to transmit gray-scale pictures of that planet, individual pixels were sent on a scale from 0 to 63. To send these numbers, NASA implemented a Reed-Muller code capable of correcting up to 7 errors in a single 32-bit word (consisting of 6 data bits and 26 check bits). Mariner 9 was able to transmit over 16,000 bits per second back to Earth.

An everyday example of error-correcting codes can be found in the compact disc. On CDs, the signal is encoded digitally. To prevent skips from scratches, two "interleaved" codes are used which can correct up to 4,000 consecutive errors (about 2.5 mm of track) [4].

## 2 Definitions

Let $\mathbb{F}_2$ denote the finite field of order 2, and $\mathbb{F}_2^n$ the $n$-dimensional vector space over $\mathbb{F}_2$. The elements of $\mathbb{F}_2^n$ are called *binary words* of length $n$. The *weight* of a word $\mathbf{x} = (x_1, x_2, \ldots, x_n)$, denoted $||\mathbf{x}||$, is defined to be $|\{i : x_i \neq 0\}|$. The *Hamming distance* $H$ between words $\mathbf{x}$ and $\mathbf{y} \in \mathbb{F}_2^n$ is the number of positions in which they differ. In other words

$$H(\mathbf{x}, \mathbf{y}) := |\{i : 1 \leq i \leq n, x_i \neq y_i\}| = ||\mathbf{x} - \mathbf{y}||.$$

A non-empty subset $C$ of $\mathbb{F}_2^n$ is called a *binary code*, and the elements of $C$ are *binary codewords*. $C$ is called an $(n, d)$ *binary code* if $d \leq H(\mathbf{x}, \mathbf{y})$ for all $\mathbf{x}, \mathbf{y} \in C, \mathbf{x} \neq \mathbf{y}$. The *minimum distance* in a code $C$ is defined to be $\min\{H(\mathbf{x}, \mathbf{y}) : \mathbf{x}, \mathbf{y} \in C, \mathbf{x} \neq \mathbf{y}\}$.

The codewords of length $n$ in a binary code are sometimes referred to as *$n$-bit codewords*. When dealing with such codes, $\mathbf{0}$ and $\mathbf{1}$ denote the $n$-bit words of all 0's and 1's, respectively.

Let $P$ be a property of binary codes. A code $C$ is said to be *maximal* $P$ provided $C$ has $P$ and no code having $P$ properly contains $C$. A code $C$ is *optimal* $P$ provided $C$ has property $P$ and for any code $D$ having $P$, $|D| \leq |C|$.

Let $X$ denote the set $\{1, 2, 3, \ldots, n\}$, and $P(X)$ the power set of $X$. The function $r : \mathbb{F}_2^n \to P(X)$, defined by $r(\mathbf{x}) = \{i : x_i = 1\}$ is a one-to-one correspondence between $\mathbb{F}_2^n$ and $P(X)$. Thus, each binary word of length $n$ will be identified with a subset of $X$. Let $\binom{X}{w}$ denote the set of subsets associated with binary words of weight $w$.

For two sets $A$ and $B$, their *symmetric difference* is $A \triangle B := (A - B) \cup (B - A)$. Observe that $H(\mathbf{x}, \mathbf{y}) = |r(\mathbf{x}) \triangle r(\mathbf{y})|$.

## 3 Constant Weight Binary Codes

An $(n, d)$ code $C$ is said to have *constant weight* $w$, and is called an $(n, d, w)$ code, if every word in $C$ has weight $w$. $A(n, d)$ and $A(n, d, w)$ respectively denote the cardinality of an optimal $(n, d)$ and an optimal $(n, d, w)$ code. Similarly, $a(n, d)$ and $a(n, d, w)$ respectively denote the least possible cardinality of a maximal $(n, d)$ code and a maximal $(n, d, w)$ code. In general, the values of all four of these functions are unknown.

In an exhaustive study of the lower bounds for $A(n, d, w)$, A. E. Brouwer, et al. [2] found an elegant method to partition the set of binary words of length $n$ and weight $w$ into $n$ $(n, 4, w)$ codes, thus proving that $A(n, 4, w) \geq \frac{1}{n}\binom{n}{w}$. The purpose of this paper is to show that if $\gcd(n, w) = 1$, then each of the $n$ codes defined in [2] is, in fact, maximal $(n, 4, w)$. Before proving this result, we first briefly sketch the construction given in [2].

**Theorem 1.** $A(n, 4, w) \geq \frac{1}{n}\binom{n}{w}$

Proof. Let $Z_n$ be the group of integers modulo $n$, and consider the function $f : \binom{X}{w} \to Z_n$, defined by $f(A) = (\sum_{a \in A} a) \pmod n$. For an $i$, $0 \leq i \leq n-1$, let $C_i = f^{-1}(i)$. Let $A, B \in C_i$, $A \neq B$. We will show that $|A \triangle B| \geq 4$. Clearly $|A \triangle B| = 2w - 2|A \cap B|$, a positive even integer. Now if $|A \triangle B| = 2$, then there exist $a_0 \in A$ and $b_0 \in B$ such that $A \triangle B = \{a_0, b_0\}$. Since $\sum_{a \in A} a = \sum_{b \in B} b = i \pmod n$, we have that $a_0 = b_0 \pmod n$. But that is impossible because $a_0 \neq b_0$ and both $a_0$ and $b_0$ belong to the set $X = \{1, 2, \ldots, n\}$. It follows that $|A \triangle B| \geq 4$; thus, each $C_i$ is an $(n, 4, w)$ code. As the collection $\{C_0, C_1, \ldots, C_{n-1}\}$ is obviously a partition of $\binom{X}{w}$ and $|\binom{X}{w}| = \binom{n}{w}$, therefore some $C_i$ must have at least $\frac{1}{n}\binom{n}{w}$ elements. □

**Theorem 2.** If $2 \leq w < n$ and $\gcd(n, w) = 1$, then $\binom{X}{w}$ can be partitioned into $n$ maximal $(n, 4, w)$ codes.

Proof. We will prove that under the given hypothesis, each of the sets $C_i$, $0 \leq i \leq n-1$, defined in the proof of Theorem 1 is a maximal $(n, 4, w)$ code. We do this by showing that for each $i$ and for any $A \in \binom{X}{w}$, if $A \notin C_i$, then there is a $B \in C_i$ such that $H(A, B) = 2$. Fix an $i$ and an $A$, as stated. Determine $j$, $0 \leq j \leq n-1$, such that $A \in C_j$ and then let $k = (i - j) \pmod n$. As $A \notin C_i$, $j \neq i$, and thus $0 < k < n$. For each $a \in A$, let $a^* = (a + k) \pmod n$. If for some $a_0 \in A$, $a_0^* \notin A$, then by deleting $a_0$ from $A$ and inserting $a_0^*$ in its place, we obtain

a set $B$ as desired. To see why, note that $A \triangle B = \{a_0, a_0^*\}$, and so $|A \triangle B| = 2$. Also, summing modulo $n$, we have $\sum_{b\in B} b = \sum a \in A + a_0^* - a_0 = j + k = i$. Thus, $B \in C_i$.

To complete the proof, we show that under the given hypothesis, there exists an $a \in A$ for which $a^* \notin A$. By way of contradiction, suppose such an $a$ does not exist. Let $\langle k \rangle$ be the subgroup of $Z_n$ generated by $k$. Then for each $a \in A$, the coset $\langle k \rangle + a$ is contained in $A$. Therefore, $A$ is a union of a set of cosets of $\langle k \rangle$ in $Z_n$. As the order of the subgroup $\langle k \rangle$ is $\frac{n}{g}$ where $g = \gcd(k, n)$, it follows that the cardinality of $A$, namely $w$, is a multiple of $\frac{n}{g}$. This, in view of the hypothesis that $\gcd(n, w) = 1$ implies that $\frac{n}{g} = 1$, so $n = \gcd(n, k)$. But then $k \geq n$, a contradiction. $\square$

**Corollary.** If $\gcd(n, w) = 1$, then $a(n, d, w) \leq \frac{1}{n}\binom{n}{w}$.

The proof of Theorem 2 implies that if $\gcd(n, w) = g > 1$, then at least one of the $C_i$ is not a maximal $(n, 4, w)$ code. To see this, let $w = gw_1$ and $n = gn_1$, and note that the subgroup $\langle n_1 \rangle$ of $Z_n$ is of order $g$. Choose any $w_1$ of the cosets of $\langle n_1 \rangle$ in $Z_n$, and let $A$ be their union, so that $A \in \binom{X}{w}$. Then, for each $a \in A$, $(a + n_1) \pmod n \in A$. Thus, if $A \in C_j$ and if $i = (j + n_1) \pmod n$, then $A \notin C_i$ even though $A$ is at a distance at least 4 from each member of $C_i$, as shown in the proof of Theorem 1. Thus, $C_i$ is not a maximal $(n, 4, w)$ code, and we have shown the following.

**Theorem 3.** If $\gcd(n, w) > 1$, then at least one $C_i$ is not a maximal $(n, 4, w)$ code.

An interesting interpretation of Theorem 2 in terms of "graph-coloring" is as follows: If $\gcd(n, w) = 1$, it is possible to color each binary word of length $n$ and weight $w$ in one of $n$ given colors so that each of the monochromatic classes is a maximal $(n, 4, w)$ code. It is noteworthy that similar "color" partitions into optimal codes do not always exist. This is due to the fact that $A(n, d)$ and $A(n, d, w)$ generally fail to divide $2^n$ and $\binom{n}{w}$, respectively. For example, $A(9, 4) = 40$ which does not divide $2^9$, so a monochromatic partition of $\mathbb{F}_2^9$ into optimal $(9, 4)$ codes does not exist. Similarly, $A(11, 4, 4) = 35$, which does not divide $\binom{11}{4}$, so the set of binary words with length 11 and weight 4 cannot be partitioned into optimal $(11, 4, 4)$ codes.

**Acknowledgements.** This research was done during Summer 2003 at Butler University under the guidance of Dr. Prem Sharma as a part of the Butler Summer Institute, funded by a Holcomb Undergraduate Grant.

## References


[1] A History of Fundamental Mathematics Research at Bell Labs (1998). (http://cm.bell-labs.com/cm/ms/departments/fm/history.html)

[2] A. E. Brouwer, James B. Shearer, N. J. A. Sloane, and Warren D. Smith, A New Table of Constant Weight Codes, IEEE Transactions on Information Theory, Vol. 36, No. 6, November 1990.

[3] J. H. van Lint, Introduction to Coding Theory, 3rd Edition, Springer-Verlag, Berlin 1999.

[4] Richard Pinch, Coding theory: the first 50 years, PASS Maths, Issue No. 3: September 1997. (http://plus.maths.org/issue3/codes/index.html)

[5] E. M. Rains and N. J. A. Sloane, Table of Constant Weight Binary Codes. (http://www.research.att.com/ njas/codes/Andw/)