Ansaq Journal for arts, literature and humanities
16th edition
Volume (5) Issue (3)
2024 (1-8)

A Study of Relationship between Hamming bound and Perfect Code with new Examples

Hamid Mohammed Khalaf
Department of Mathematics, College of Basic Education, University of Telafer, Tall'Afar, Mosul, Iraq.
Published on: 19 August 2024

Abstract

The aim of this study is to find theories that show the relationship between the perfect linear code and the Hamming codes, as we study the possibility that each Hamming bound is a perfect code, as well as every binary code is a perfect. We will study new examples in coding theory that prove these theories. In these examples, we found the generator matrix G of a $[m, p]$-code C and the codedwords, weight distribution of C and then we found the minimum (Hamming) distance d. Such that we get a code C that achieves the relationship, if $q^{m-p} > \sum_{i=0}^{d-2} \binom{m-1}{i} (q-1)^i$, then there always exists a $[m, p, d]_q$-code.

Keywords: perfect code, Hamming bound, code word.

*** Introduction**

Suppose there are many important messages to be sent through a noisy communication channel, in order to give these messages some protection against error on the channel, they are encoded in to codewords. The set consisting of these codewords is called a code. (Tsfasman et al., 2007)

Let m, p, d be positive integers. A code $C : [m, p, d]$ over Galois Field $(GF(q))$ is a p-dimensional subspace of $GF(q)^n$ with minimum distance d. In a code C of length m let D_i denote the number of codewords with Hamming weight i. The weight enumerator of C is defined by $1 + D_1 z + D_2 z^2 + \cdots + D_m z^m$.

The sequence $(1, D_1, D_2, \cdots, D_m)$ is called the weight distribution of the code C. If the number of nonzero D_i in the sequence $(1, D_1, D_2, \cdots, D_m)$ is equal to d then, a code C is said to be a d-weight code. A code $[m, p, d]$ over $GF(q)$ is called distance-perfect if there is no $[m, p, d+1] -$ code over $GF(q)$ and dimension-perfect if there is no $[m, p+1, d]$-code over $GF(q)$. A code is said to be perfect if it is both distance-perfect and dimension-perfect.

*** Preliminaries**

Definition (2.1): (Tsfasman et al., 2007)

A set of sequences of 0's and 1's called a binary code; each sequence is a codeword.

Definition (2.2): (Hirschfeld, 2014)

A set of sequences where each symbol is from a set

$\mathbb{F}_q = \{\zeta_1, \zeta_2, \ldots, \zeta_q\}$ is called q-ary code C.

Definition (2.3): (Al-Seraji, 2013), (Al-Seraji & AL-Humaidi, 2018)

Let $\mathbb{F}_q^m$ denotes the vector space of m-tuple. A (m, k)-code C over Galois field $\mathbb{F}_q$ is a subset of $\mathbb{F}_q^m$ of size k. A linear $[m, p]_q$-code over $\mathbb{F}_q$ is a $p-$dimensional subspace of $\mathbb{F}_q^m$ and size $k = q^p$. The vectors in the linear code C are called codewords and them by $u = u_1\ u_2\ \ldots\ u_m$ where $u_i \in \mathbb{F}_q$.

In other words, a q-ary linear code of length m and dimension p, or a $[m, p]_q$ -code, is a p-dimensional subspace of $\mathbb{F}_q^m$, where $p = 4$. Every subspace of C is referred to as a sub code of C. The inner product of two vectors $u = (u_1, u_2, \ldots, u_m)$ and $v = (v_1, v_2, \ldots, v_m)$ from $\mathbb{F}_q^m$ is denoted by $uv = u_1v_1 + u_2v_2 + \cdots + u_mv_m$.

Two vectors are said to be orthogonal if their inner product is 0. The set of all vectors of $\mathbb{F}_q^m$ orthogonal to all codewords from C is called the orthogonal code: $C^{\perp} = \{u \in \mathbb{F}_q^m | uv = 0$ for any $v \in C\}$. By a well-known fact from linear algebra, the code $C^{\perp}$ is a linear $[m, m-p]_q$-code. If C is a linear code that, as a vector space over $\mathbb{F}_q$ has dimension p, then we say that C is a $[m, p]$ linear code over $\mathbb{F}_q$.

Definition (2.4): (Assmus & Mattson, 1978), (Al-Seraji, 2012)

The Hamming weight $w(u)$ of a vector u in $\mathbb{F}_q^m$ is the number of its nonzero coordinates, such that $\mathbb{F}_q^m = \{(u_1, u_2, \ldots, u_m) | u_1, \ldots, u_m \in \mathbb{F}_q\}$. For $u = (u_1, \ldots, u_m), v = (v_1, \ldots, v_m) \in \mathbb{F}_q^m$, the Hamming distance between u and v is

$d(u,v) = |\{i|u_i \neq v_i\}|$. The weight of u is $w(u) = |\{i|u_i \neq 0\}| = d(u,0)$.

Since a $[m,p,d]_q$-code C means a p-dimensional subspace of $\mathbb{F}_q^m$ with minimum distance d, then $d = \min\{d(u,v)|u \neq v,\ u,v \in C\} = \min\{w(u)|w(u) \neq 0,\ u \in C\}$.

It is important to see that the minimum weight is the same as the minimum distance if C is a linear code. This comes from the fact that the zero vector is always in a linear code.

In other words, if $C \subseteq \mathbb{F}_q^m$ is a linear code and $u,v \in C$ with $u \neq v$, then
$u - v \in C \setminus \{0\}$ and therefore, $w(C) \leq d(C)$.

Conversely, since the vector $0 \in C$ and $w(u) = d(v,0)$, then $w(C) \geq d(C)$.

Hence, the minimum distance for linear codes, and the minimum weight are coincide.

Definition (2.5): (El-atrash & Al-Ashker, 2003), (Ma et al., 2021)

A code C is called a $[m,p,d]$-code if d is the minimum nonzero weight in C. The weight enumerator of the code C is defined by:
$1 + D_1u + D_2u^2 + \cdots + D_mu^m$

The list D_i for $0 \leq i \leq m$ is called the weight distribution of C.

In general, it is difficult to determine the weight distribution of a given linear code C. If the number of nonzero D_i in $(1, D_1, D_2, \ldots, D_m)$ is equal to d, then C is called a d-weight code.

Theorem (2.6): (El-atrash & Al-Ashker, 2003)

Let C be a $[m,p,d]$-code over $\mathbb{F}_q$ then
$1 - D_0(C) + D_1(C) + \cdots + D_m(C) = q^p$
$2 - D_0(C) = 1$ and $D_1(C) = D_2(C) = \cdots = D_{d-1}(C) = 0$.

Theorem: (2.7): (Tsfasman et al., 2007)

For every linear $[m,p,d]$-code C,$0 < D_1(C) < D_2(C) < \cdots < D_p(C) \leq m$.

Definition (2.8): (Tsfasman et al., 2007)

To define the Hamming codes Ham (h,q) over $\mathbb{F}_q$, where $m = \frac{q^h-1}{q-1}$, $h = m - p$, a parity-check matrix Ħ is specified. First, consider the case $q = 2$. For any positive integer h, let Ħ be an $h \times m$ matrix, $m = 2h - 1$, whose columns are the elements of $V(h,2)\backslash\{0\}$.

The parameters of Hamming codes: $[\frac{q^h-1}{q-1}, \frac{q^h-1}{q-1} - h, 3]$.

Theorem (2.9): (Hamming bound) (Kageyama & Maruta, 2016)

Let C be a $[m, p, d]_q$-code, then $\sum_{i=0}^{\left\lfloor \frac{d-1}{2} \right\rfloor} \binom{m}{i} (q-1)^i \leq q^{m-p}$, where denotes $\lfloor u \rfloor$ the greatest integer less than or equals to u.

Theorem (2.10): (Griesmer Bound) (Klein & Storme, 2011), (Ma & Luo, 2020)

Let $m_q(p, d)$ denote the minimal m for which a $[m, p, d]_q$-code exists, then $m \geq m_q(p, d) = \sum_{i=0}^{p-1} \left\lceil \frac{d}{q^i} \right\rceil$, where denotes $\lceil u \rceil$ the smallest integer larger than or equals to u.

A Linear codes attaining the Griesmer bound, i.e. linear codes with parameters $[m_q(p, d), p, d]_q$, are called Griesmer codes.

Definition (2.11): (Al-Seraji, 2013), (Ma & Luo, 2020)

A generator matrix G of a $[m, p]$-code C is a $(p \times m)$ matrix whose rows form a basis for C.

$$G = \begin{bmatrix} g_0 \\ g_1 \\ . \\ . \\ . \\ g_{p-1} \end{bmatrix} =$$

$$\begin{bmatrix} g_{0,0} & g_{0,1} & & g_{0,m-1} \\ . & . & & . \\ . & . & & . \\ . & . & & . \\ . & . & & . \\ g_{p-1,0} & g_{p-1,1} & & g_{p-1,m-1} \end{bmatrix}$$

, Such that $V = U_{1\times p}\, G_{p\times m}$,

(U is a different row vectors)

Codewords of C is a linear combinations of row of G.

Generator matrix G not uniqe-elementary row operations gives the same code.

Definition (2.12): (Hirschfeld, 2014)

A code is e-error correcting if it can correct e errors.

3- The Perfect Code: (MacWilliams and Sloane, 1977)

If each codeword in the code C consists of m letters taken from an alphabet A of length q, where C consisting of k codewords and every two distinct codewords differ in at least $d = 2e + 1$ places. Then C is said to be perfect if for every possible word w_0 of length m with letters in A, there is a unique codeword w in C in which at most e letters of w differ from the corresponding letters of w_0.

Definition (3.1): (Bonisoli et al., 1996)

An e-error-correcting code C in $\mathbb{F}_q^m$ is perfect if any vector in $\mathbb{F}_q^m$ is at distance at most e from exactly one codeword; that is, every received message is corrected.

Theorem (3.2): (Bonisoli et al., 1996)

Let C be a code in $\mathbb{F}_q^m$

1- If $d \geq e + 1$, then C can detect up to e errors.

2- If $d \geq 2e + 1$, then C can correct up to e errors.

Corollary(3.3): (Bonisoli et al., 1996), (Hirschfeld, 2014)

If C has minimum distance d, then it can detect $d-1$ errors and correct $e=\lfloor (d-1)/2 \rfloor$ errors.

Theorem (3.4): (Al-Seraji & Ajaj, 2019)

A q-ary $[m,p,2e+1]$-code C satisfies: $p\{\binom{m}{0}+\binom{m}{1}(q-1)+\cdots+\binom{m}{e}(q-1)^e\} \leq q^m$.

Corollary (3.5): (Al-Seraji & Ajaj, 2019)

A q-ary $[m,p,2e+1]$-code C is perfect if and only if equality holds in Theorem (3.4).

Theorem (3.6): Every binary code is a perfect code

Proof: It is straightforward to show that a code C is perfect if $\sum_{i=0}^{\lfloor \frac{(d-1)}{2} \rfloor}\binom{m}{i}(q-1)^i=\frac{q^m}{k}$ …. (1).

If we take $k=q^p$ then, from the Hamming bound we have: $\sum_{i=0}^{e=\lfloor \frac{(d-1)}{2} \rfloor}\binom{m}{i}(q-1)^i \leq q^{m-p}$ … (2). That means the eq.(1) satisfies.

Then, eq.(1) satisfy the condition of Hamming bound.

Now, since the Hamming bound definition as in eq.(2), and the condition of a binary code to be perfect code is $\sum_{i=0}^{\lfloor \frac{(d-1)}{2} \rfloor}\binom{m}{i}=2^{m-p}$ … (3).

From eq.(2) in eq.(3) we get : $2^{m-p}(q-1)^i \leq q^{m-p}$. Since it's a binary code; $q=2$ then, $2^{m-p}(1)^i \leq 2^{m-p}$. That's means every binary code is a perfect code.

Example(3.6): If we take $m=5$ such that, $m \geq \sum_{i=0}^{p-1}\left\lceil \frac{d}{q^i} \right\rceil$, with Generator matrix:

$G=\begin{bmatrix}1&0&0&0&1\\0&1&0&0&1\\0&0&1&0&1\\0&0&0&1&1\end{bmatrix}$, in this construction $PG(p-1,2)=\{u_1,u_2,\dots,u_m\}$ where $m=q^p=(2)^4=16$ and u_i is defined as follows $\forall\, i$, $PG(p-1,2)=\{[1,0,0,0],\dots,[1,1,1,1]\}$, therefore $G=[u_1^T,\dots,u_m^T]$. The codewords are generated by multiplying each point in $PG(p-1,2)$ to G. Any permutation of the rows of C or multiplication of a row of C by an element of $\mathbb{F}_2$ gives another matrix of G, therefor the codewords in $[5,4]_2$ linear code is defined as follows: $C.W.=\{[u_i*G]_{q^p*m} \setminus u_i \in PG(p-1,q)\}=$

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

With weight distribution: 0^1,2^{10},4^5. We note that the minimum distance d is 2. Since, if $q^{m-p} > \sum_{i=0}^{d-2}\binom{m-1}{i}(q-1)^i$, then there always exists a $[m,p,d]_q$-code, so we get C is a $[5,4,2]_2$-code.

Since, $m \geq \sum_{i=0}^{p-1}\left\lceil\frac{d}{q^i}\right\rceil = \left\lceil\frac{2}{2^0}\right\rceil + \left\lceil\frac{2}{2^1}\right\rceil + \left\lceil\frac{2}{2^2}\right\rceil + \left\lceil\frac{2}{2^3}\right\rceil = 5$, then C is Griesmer code with $e = \left\lfloor\frac{d-1}{2}\right\rfloor = 0$. Now since, $p\sum_{i=0}^{e}\binom{m}{i}(q-1)^i = 4\left[\binom{5}{0}\right] = 4 \leq 2^5$ and $2^{m-p}(q-1)^i \leq 2^{m-p}$ then, it's attaining the Hamming bound, this means that the binary code C: $[5,4,2]_2$ is perfect code.

Theorem (3.7): Every Hamming codes of order h over $GF(q)$ are perfect codes.

Proof: From Hamming bound: $\sum_{i=0}^{e}\binom{m}{i}(q-1)^i \leq q^{m-p}$ multiplication by q^p we get: $q^p\sum_{i=0}^{e}\binom{m}{i}(q-1)^i \leq q^m$, when $q^p = k$ then it's satisfies the condition of perfect code. Now, we take $d = 3$ then $2e+1 = 3 \Rightarrow e = 1$ then,$k\sum_{i=0}^{e}\binom{m}{i}(q-1)^i = q^p\sum_{i=0}^{1}\binom{m}{i}(q-1)^i = q^p\left[\binom{m}{0}(q-1)^0 + \binom{m}{1}(q-1)^1\right] = q^p\left[\frac{m!}{m!0!} + \frac{m!}{(m-1)!1!}(q-1)\right] = q^p[1+m(q-1)] = q^p\left[1+\frac{q^h-1}{q-1}(q-1)\right] = q^p[1+q^h-1] = q^{p+h} \leq q^m$.

And this leads to, $k\{\binom{m}{0} + \binom{m}{1}(q-1) + \cdots + \binom{m}{e}(q-1)^e\} \leq q^m$.

Then, the Ham(h,q) codes are perfect codes.

Example (3.8): Consider $m = 7, p = 4, h = 3$, The code C : $[7,4]_2$ -code, generated by the rows of G:

$$G = \begin{bmatrix}1&0&0&0&1&1&0\\0&1&0&0&0&1&1\\0&0&1&0&1&1&1\\0&0&0&1&1&0&1\end{bmatrix}$$

By multiplying G on the left by the 16 different binary row vectors of

length 4, we get the 16 codewords. So for instance we get codewords:

(0,0,0,0) . G = (0,0,0,0,0,0,0)

(1,0,1,1) . G = (1,0,1,1,1,0,0)

: : : : : :

(1,1,0,0) . G = (1,1,0,0,1,0,1)

(1,1,0,0) . G = (1,1,0,0,1,0,1).

So, the list of all the codewords is:

0 0 0 0 0 0 0	1 1 0 1 0 0 0
0 1 1 0 1 0 0	0 0 1 1 0 1 0
0 0 0 1 1 0 1	1 0 0 0 1 1 0
0 1 0 0 0 1 1	1 0 1 0 0 0 1
1 1 1 1 1 1 1	0 0 1 0 1 1 1
1 0 0 1 0 1 1	1 1 0 0 1 0 1
1 1 1 0 0 1 0	0 1 1 1 0 0 1
1 0 1 1 1 0 0	0 1 0 1 1 1 0

We assume G has no all-zero column. The weight distribution $(w.d.)$ of C is the list of numbers $D_i = |\{u \in \mathsf{C} \mid w(u) = i\}|$.

The weight distribution with $(D_0, D_d, \dots, D_i, \dots) = (1, \alpha, \dots, w, \dots)$ is also expressed as $0^1, d^\alpha, \cdots, i^w \cdots$. Thus, we get:

weight distribution: $0^1, 3^7, 4^7, 7^1$. Hence, Ham(3,2) is equivalent to the perfect $[7,4,3]_2$ – code. We note that $m = \frac{q^h-1}{q-h} = 7, p = \frac{q^h-1}{q-1} - h$ and $d = 3$ then, $q^{m-h}\big(1 + m(q-1)\big) = q^{m-h}\left\{1 + \frac{q^h-1}{q-1}(q-1)\right\} = q^{m-h}(1 + q^h - 1) = q^m$.Hence, the code C is perfect.

*** Conclusions**

From the above results, we found the relationship between the perfect linear code and the Hamming codes, as we study the possibility that each Hamming bound is a perfect code, as well as every binary code is a perfect. We found a new examples in coding. In these examples, we found the generator matrix G of a [m,p]-code C and the codedwords, weight distribution of C and then we found the minimum (Hamming) distance d.

*** Conflicts of Interest**

The author declare that there are no conflicts of interest regarding this work

*** Acknowledgment**

The research is supported by the College of Computer Sciences and Mathematics, University of Mosul, and College of Basic Education, University of Telafer, Mosul, Republic of Iraq.

*** REFERENCES**

Al-Seraji, N. A. M. & Ajaj, H. L. (2019). Some Applications of Coding Theory in the Projective Plane of Order Four. Al-Mustansiriyah Journal of Science, 30(1), 152-157.

Al-Seraji, N. A. M. & AL-Humaidi, R. I. (2018). Some application of coding theory in the

projective plane of order three. Iraqi Journal of Science, 1947-1951.

Al-Seraji, N. A. M. (2012). On Optimal Codes. Mustansiriyah Journal of Science, 23(6).

Al-Seraji, N. A. M. (2013). Generalized of Optimal Codes. Al-Mustansiriyah Journal of Science, 24(6), 101-108.

Assmus, E. & Mattson, H. (1978). The weight-distribution of a coset of a linear code (Corresp.). IEEE Transactions on Information Theory, 24(4), 497.

Bonisoli, A., Cossidente, A. & Saeli, D. (1996). Partitioning projective geometries into Segre varieties. Journal of Geometry, 57(1), 58-62.

El-atrash, M. S. & Al-Ashker, M. M. (2003). Linear codes over F_2+ uF_2. Journal of The Islamic University of Gaza, 11(2).

Hirschfeld, J. W. P. (2014). Coding Theory. Lectures, Sussex University, UK.

Kageyama, Y. & Maruta, T. (2016). On the geometric constructions of optimal linear codes. Designs, Codes and Cryptography, 81(3), 469-480.

Klein, A. & Storme, L. (2011). Applications of finite geometry in coding theory and cryptography. Information Security, Coding Theory and Related Combinatorics, 29, 38-58.

Ma, L., Li, G. & Liu, F. (2021). More Constructions of 3-Weight Linear Codes. Journal of Mathematics, 2021.

Ma, W. & Luo, J. (2020). Construction on Griesmer Codes with Dimension Not Less than Five. arXiv preprint arXiv:2009.11998.

MacWilliams, F. J. and Sloane, N. J. A. (1977). The theory of error correcting codes (Vol. 16). Elsevier.

Tsfasman, M. A., Vlădu, S. G. and Nogin, D. (2007). Algebraic Geometric Codes: Basic Notions: Basic Notions (Vol. 1). American Mathematical Soc.