On the distribution of linear codes around the $[7, 4, 3]$-Hamming code

Toshiyuki Katsura*
Graduate School of Mathematical Sciences, University of Tokyo
Tokyo, 153-8914 Japan
E-mail: tkatsura@ms.u-tokyo.ac.jp

1 Introduction

Let $\mathbf{F}_q$ be a finite field with q elements ($q = p^m$, p : a prime number, m : a positive integer), and let $\mathbf{F}_q^*$ be the multiplicative group of non-zero elements of $\mathbf{F}_q$. A linear code $C \subset \mathbf{F}_q^n$ is called an $[n, k, d]$-linear code if $\dim_{\mathbf{F}_q} C = k$ and the minimum distance of C is equal to d. We also denote by $d(C)$ the the minimum distance of the code C. Two linear codes C, C' in $\mathbf{F}_q^n$ are said to be generalized eqivalent to each other if there exist a permutation σ of coodinates of $\mathbf{F}_q^n$ and an element $\lambda = (\lambda_1, \ldots, \lambda_n) \in \mathbf{F}_q^{*n}$ such that $\lambda\sigma(C) = C'$. Here, the multiplication by λ is given by componentwise multiplication. It is clear that if two linear codes $C, C' \subset \mathbf{F}_q^n$ are generalized equivalent to each other, then we have $\dim_{\mathbf{F}_q} C = \dim_{\mathbf{F}_q} C'$ and $d(C) = d(C')$. We regard two generalized equivalent codes as the same codes. We denote by LC_q the set of all linear codes over a fixed finite field $\mathbf{F}_q$. We consider the following mapping:

$$\begin{array}{cccc} \varphi : & LC_q & \longrightarrow & [0,1]^2 \\ & [n,k,d]\text{-code } C & \longmapsto & (d/n, k/n). \end{array}$$

The set U_q of the limit points of Im φ was investigated by Yu. I. Manin, and for the shape of U_q he showed the following:

*Partially supported by Max-Planck-Institut, and Grand-in-Aid, Ministry of Education of Japan

Theorem 1.1 (Manin [3]) *There exists a continuous function*

$$\alpha_q : [0,1] \longrightarrow [0,1]$$

such that

$$U_q = \{(\delta, R) \mid 0 \leq R \leq \alpha_q(\delta)\}$$

and that

$$\alpha_q(0) = 1, \;\; \alpha_q(\delta) \leq \max\{1 - (q/(q-1))\delta, 0\}.$$

As is well-known, the function $\alpha_q(\delta)$ is strictly decreasing for $\delta \in [0, (q-1)/q]$ (cf. Tsfasman and Vladut [4]), and is estimated from below by the Varshamov-Gilbert bound. In this paper, we consider linear codes C such that $\varphi(C) \notin U_q$. In this case, the point $\varphi(C)$ is an isolated point in $\varphi(LC_q)$ and the following two questions arise.

Question 1 Is there any linear code C' such that $\varphi(C') = \varphi(C)$ and that C' is not generalized equivalent to C?

Question 2 What are the linear codes C' such that the points $\varphi(C')$ are nearest to $\varphi(C)$? What is the distance between $\varphi(C')$ and $\varphi(C)$?

The codes C' in **Question 2** are called the nearest codes to C, and the distance between $\varphi(C')$ and $\varphi(C)$ is called the isolated radius of C (for the precise definitions, see Section 2) . We will give answers to these two questions in the case of $[7,4,3]$-Hamming code C (cf. Theorems 2.6). In fact, we will show that there exists no linear code C' such that $\varphi(C') = \varphi(C)$ and $C' \neq C$. We also show that the $[11,6,4]$-linear codes are the nearest codes to C, and so the isolated radius of the $[7,4,3]$-Hamming code is equal to $2\sqrt{2}/7 \cdot 11$. We also have a similar result for the Golay code (cf. Theorem 5.3). The isolated radius is an invariant for linear codes to measure how different the linear code is from other linear codes.

The author thanks Miss M. Suzuki for informing him of the result by C. L. M. van Pul in Section 5. This paper was written while the author was visiting Max-Planck-Institut für Mathematik in Bonn. He heartily thanks Max-Planck-Institut for hospitality.

2 Definitions and the main theorem

To begin with, we give the precise definition of U_q:

$$U_q = \left\{ (\delta, R) \in [0,1] \times [0,1] \;\middle|\; \begin{array}{l} \exists \text{ linear code } C_i \text{ of type } [n_i, k_i, d_i] \\ \text{such that} \\ n_i \longrightarrow \infty \quad \text{as} \quad i \to \infty \\ \text{and} \\ (d_i/n_i, k_i/n_i) \longrightarrow (\delta, R) \quad \text{as} \quad i \to \infty \end{array} \right\}$$

We are mainly interested in points which are not contained in U_q.

Definition 2.1 *For $r \in \mathbf{R}$ and $(a,b) \in [0,1]^2$, we set*

$$B_r((a,b)) = \{[n,k,d]\text{-}linear\ code \mid (d/n - a)^2 + (k/n - b)^2 < r^2\}.$$

Definition 2.2 *Let C be an $[n,k,d]$-linear code. If there exists a positive real number r such that $B_r((d/n, k/n))$ contains only C, then C is said to be completely isolated.*

Example 2.3 *Let C be a linear code such that $\varphi(C) \in U_q$. Then, by the definition of U_q, C is not completely isolated.*

Example 2.4 *Let $\tilde{C}$ be the $[8,4,4]$ extended Hamming code. Then, $\varphi(\tilde{C})$ is not contained in U_2. We set*

$$C' = \mathbf{F}_2(1,0) \subset \mathbf{F}_2^2.$$

Then, it is clear that C' is a $[2,1,1]$-linear code. Since we have

$$\varphi(\tilde{C}) = \varphi(C') = (1/2, 1/2),$$

the $[8,4,4]$ extended Hamming code is not completely isolated.

Definition 2.5 *Let C be an $[n,k,d]$-linear code. If there exists the greatest real number r such that $B_r((d/n, k/n))$ contains only C, then r is called the isolated radius of C. The linear codes C' such that the points $\varphi(C')$ are nearest to $\varphi(C) = (d/n, k/n)$ are called the nearest codes to C.*

Using these notions, we have the following result.

Theorem 2.6 *The $[7,4,3]$-Hamming code is completely isolated. The nearest linear codes to the $[7,4,3]$-Hamming code are the $[11,6,4]$-linear codes. The isolated radius is equal to $2\sqrt{2}/7 \cdot 11 = 0.0367\ldots$*

3 Bounds

¿From here on, we assume $q = 2$, unless otherwise mentioned. To prove Theorem 2.6, we need the following two well-known results (cf. Manin [3], Tsfasman and Vladut [4]).

Lemma 3.1 (Spoiling lemma) *If there exists an $[n, k, d]$-linear code, then an $[n-1, k-1, d]$-linear code and an $[n-1, k, d-1]$-linear code also exist.*

Lemma 3.2 (Plotkin bound for $q = 2$) *If there exists an $[n, k, d]$-linear code, then $d/n \leq 2^{k-1}/(2^k - 1)$ holds.*

Using these two lemmas, we have the following bound.

Theorem 3.3 *Assume that there exists an $[n, k, d]$-linear code with $k \geq 4$. Then, the following bound holds:*

$$k/n + (15/8)(d/n) \leq 1 + 4/n.$$

Proof Put $\ell = k - 4$. Using Lemma 3.1 ℓ times, we have an $[n-k+4, 4, d]$-linear code. Therefore, using the Plotkin bound, we have

$$d/(n-k+4) \leq 2^3/(2^4 - 1),$$

which gives our inequality. ■

Remark 3.4 *We fix a positive integer α. Then, we can generalize Theorem 3.3 as follows:*

Assume that there exists an $[n, k, d]$-linear code with $k \geq \alpha$. Then, the following bound holds:

$$k/n + (2 - 1/2^{\alpha-1})(d/n) \leq 1 + \alpha/n.$$

For our purpose, the case $\alpha = 4$ is most efficient. Note that this inequality contains only linear terms for k, d and n like the Singleton bound.

We also have the following two famous bounds (cf. Tsfasman and Vladut [4], for instance).

Theorem 3.5 *If there exists an* $[n,k,d]$*-linear code, then the following bounds hold:*

$$n \geq \sum_{i=0}^{k-1} \rceil d/2^i \lceil \quad (\textit{Griesmer bound}),$$
$$2^n \geq \sum_{i=0}^{[d/2]} \binom{n}{i} \quad (\textit{Hamming bound}).$$

Here, for a real number r, $\rceil r \lceil$ *means the least integer which is greater than or equal to* r, *and* $[r]$ *means the greatest integer which is less than or equal to* r.

4 Proof of Theorem 2.6

The following lemma is known. I'll give a proof for the readers' convenience.

Lemma 4.1 *There exists an* $[11,6,4]$*-code.*

Proof We take the projective plane $\mathbf{P}^3(\mathbf{F}_2)$ of dimension 3 over $\mathbf{F}_2$. We denote by (x_0, x_1, x_2, x_3) the homogeneous coodinates of $\mathbf{P}^3(\mathbf{F}_2)$. Since $\mathbf{P}^3(\mathbf{F}_2)$ has fifteen $\mathbf{F}_2$-rational points, we denote them by P_i $(i = 1, \ldots, 15)$. We can assume $P_{12} = (1,0,0,0)$, $P_{13} = (0,1,0,0)$, $P_{14} = (0,0,1,0)$, $P_{15} = (0,0,0,1)$. We consider the vector space V over $\mathbf{F}_2$ defined by

$$V = \oplus_{i<j} \mathbf{F}_2 x_i x_j.$$

Then, we have $\dim_{\mathbf{F}_2} V = 6$ and for $V \ni f$ we have $f(P_k) = 0$ $(k = 12, 13, 14, 15)$. We consider the linear mapping defined by

$$\begin{array}{cccc} \phi: & V & \longrightarrow & \mathbf{F}_2^{11} \\ & f & \mapsto & (f(P_1), \ldots, f(P_{11})). \end{array}$$

Then, it is easy to prove that the mapping ϕ is injective and that Im ϕ is an $[11,6,4]$-linear code. ■

Now, we prove Theorem 2.6. We prove first that the $[7,4,3]$-Hamming code C is completely isolated.

By Manin's theorem, it is clear that the point $\varphi(C) = (3/7, 4/7)$ is not contained in the set U_2. Therefore, the point $(3/7, 4/7)$ is isolated. Suppose there exists an $[n,k,d]$-linear code C' such that $\varphi(C') = \varphi(C) = (3/7, 4/7)$. By $3/7 = d/n$ and $4/7 = k/n$, there exists a positive integer ℓ such that

$$n = 7\ell, \ k = 4\ell, \ d = 3\ell.$$

Using the spoiling lemma 3ℓ times, we have a $[4\ell, \ell, 3\ell]$-linear code. Applying the Plotkin bound to this code, we have an inequality

$$3\ell/4\ell \leq 2^{\ell-1}/(2^\ell - 1).$$

Suppose $\ell \geq 2$. Then, we have

$$3/4 \leq 2/(4 - (1/2^{\ell-2})) \leq 2/3,$$

which is a contradiction. Therefore, we have $\ell = 1$ and the code C' must be a $[7,4,3]$-linear code. As is well-known, a [7, 4, 3]-linear code is a perfect code and unique up to generalized equivalence. Hence, the $[7,4,3]$-Hamming code C is completely isolated.

Secondly, we look for the nearest linear codes to C and calculate the isolated radius of C. For this purpose, we consider the extended linear code $\tilde{C}$ of C. As is well-known (or by an easy calculation), $\tilde{C}$ is an $[8,4,4]$-linear code. Since $\varphi(\tilde{C}) = (1/2, 1/2)$, the distance between $\varphi(C)$ and $\varphi(\tilde{C})$ is equal to $1/\sqrt{98}$. We examine the codes which are contained in $B_{1/\sqrt{98}}((3/7, 4/7))$.

Suppose that an $[n, k, d]$-code C' is contained in $B_{1/\sqrt{98}}((3/7, 4/7))$. Then, we have

$$(\delta/n - 3/7)^2 + (k/n - 4/7)^2 < 1/98.$$

If $k \leq 3$, then we have $4/7 - 1/\sqrt{98} \leq k/n \leq 3/n$. Therefore, we have $n \leq 7$. Now, we assume $k \geq 4$. Then, by Theorem 3.3 we have an inequality $k/n + (15/8)(d/n) \leq 1 + 4/n$. We consider the line L and the circle Z defined as follows:

$$\begin{array}{ll} L: & R + (15/8)\delta = 1 + 4/n, \\ Z: & (\delta - 3/7)^2 + (R - 4/7)^2 = 1/98. \end{array}$$

The point $\varphi(C') = (d/n, k/n)$ exists in the circle Z and below the line L. Therefore, for the existence of such a point, it is necessary that part of the circle exists below the line L. Considering the condition that L and Z intersect each other, we conclude $n \leq 24$, which holds also in the case of $k \leq 3$. Thus, we have only finitely many cases.

Using a computer, we can choose the cases which satisfy both the Griesmer bound and the Hamming bound. (The author used Mathematica to choose the cases.) Then, we have the possibilities of codes of the following types:

$$[16,8,6],\ [11,6,4],\ [10,5,4],\ [9,5,3],\ [8,4,3],\ [7,4,3],\ [5,3,2].$$

If a $[16,8,6]$-linear code exists, then using the spoiling lemma, we have a $[13,6,5]$-linear code. However, it is known that there exist no such linear codes (see van Lint [2], p.63 Example (a)). Therefore, $[16,8,6]$-linear codes are excluded. By Lemma 4.1, we have an $[11,6,4]$-linear code. So, using the spoiling lemma, we have linear codes of the following types :

$$[10,5,4],\ [9,5,3],\ [8,4,3].$$

Since we have the $[7,4,3]$-Hamming code, by the spoiling lemma we have a $[5,3,2]$-linear code. It is a simple calculation to pick up the nearest linear codes to the $[7,4,3]$-Hamming code C among these linear codes. Hence, we conclude that the nearest linear codes to C are the $[11,6,4]$-linear codes, and that the isolated radius is equal to $2\sqrt{2}/7 \cdot 11 = 0.0367\ldots$.

Corollary 4.2 $B_{1/\sqrt{98}}((3/7,4/7))$ *consists of the linear codes of the following types:*

$$[11,6,4],\ [10,5,4],\ [9,5,3],\ [8,4,3],\ [7,4,3],\ [5,3,2].$$

5 Some comments

For a point $(a,b) \in [0,1]^2$, it will also be interesting to investigate the set $S_{(a,b)}$ defined by

$$S_{(a,b)} = \{[n,k,d]\text{-linear code} \mid a < d/n,\ b < k/n\}.$$

For the $[7,4,3]$-Hamming code C, we have $\varphi(C) = (3/7,4/7)$. For this point, we have the following:

Theorem 5.1 *The set* $S_{(3/7,4/7)}$ *consists of the following four linear codes:*

$$[1,1,1]\text{-code},\ [2,2,1]\text{-code},\ [3,2,2]\text{-code},\ [4,3,2]\text{-code}.$$

Proof Let C' be an $[n,k,d]$-linear code in $S_{(3/7,4/7)}$. If $k \leq 3$, then $4/7 \leq k/n \leq 3/n$. Therefore, we have $n \leq 5$. If $k \geq 4$, then by Theorem 3.3 we have

$$4/7 + (15/8)(3/7) \leq k/n + (15/8)(d/n) \leq 1 + 4/n.$$

Therefore, we have $n \leq 10$, which holds also in the case of $k \leq 3$.

Assume $n = 10$. Then, by $3/7 \leq d/10$ and $4/7 \leq k/10$, we have $5 \leq d$ and $6 \leq k$. Using the Singleton bound, we have

$$5 \leq d \leq n - k + 1 \leq 10 - 6 + 1 = 5.$$

Therefore, we have $d = 5, k = 6$, which contradicts the Griesmer bound. Applying the same method to the cases of $1 \le n \le 9$, we get our list. It is easy to show the existence of these codes (also well-known) . ∎

Remark 5.2 *To prove Theorem* 5.1*, we can use the inequality in Theorem* 3.3 *instead of the Singleton bound. Then, the proof becomes shorter, but the calculation becomes harder. In any case, whichever we use, we have not much difference.*

It will be interesting to calculate the isolated radius for various codes. For example, in the case of Golay code (the [23, 12, 7]-linear code), we have a similar result to Theorem 2.6.

Theorem 5.3 (T. Katsura and J. Katsuta) *The* $[23, 12, 7]$*-Golay code is completely isolated. The nearest linear codes to the* $[23, 12, 7]$*-Golay code are the* $[17, 9, 5]$*-linear codes. The isolated radius is equal to* $5/11 \cdot 23 = 0.0127\ldots$

The existence of a $[17, 9, 5]$-linear code is due to Karlin [1]. To prove this theorem, we need to show the non-existence of $[27, 14, 8]$-linear codes. This follows from the spoiling lemma and the result by C. L. M. van Pul. In fact, he showed in 1985 that there exist no $[26, 13, 8]$-linear codes (cf. http://www.win.tue.nl/win/math/dw/voorlincod.html). If there exists a $[27, 14, 8]$-linear code, then using the spoiling lemma, we can construct a $[26, 13, 8]$-linear code, which is a contradiction. The details on the Golay code will be discussed elsewhere.

References

[1] L. M. Karlin, New binary coding results by circulants, IEEE Trans. Inform. Theory, IT-15 (1969), 81-92.

[2] J. H. van Lint, *Introduction to Coding Theory,* Graduate Texts in Mathematics 86, Springer-Verlag, 1991.

[3] Yu. I. Manin, What is the maximal number of points on a curve over $\mathbf{F}_2$, J. Fac. Sci. Univ. Sect. Tokyo IA Math., 28 (1982), 715-720.

[4] M. A. Tsfasman and S. G. Vladut, *Algebraic-Geometric Codes,* Kluwer Academic Publishers, 1991.