SINGLETON CODES

Natalia Dück[1][§], Karl-Heinz Zimmermann[2]

[1,2]Hamburg University of Technology
20173, Hamburg, GERMANY

Abstract: Each linear code can be described by a so-called code ideal. In order to utilize this ideal, Gröbner bases are required. Since many results depend on the chosen term order, knowledge of the universal Gröbner basis is advantageous. Singleton codes have the property that the universal Gröbner basis for their code ideals consists of all binomials associated to a codeword whose Hamming weight satisfies the Singleton bound. In this paper, properties of Singleton codes will be established and it will be examined which classical binary linear codes belong to the class of Singleton codes.

1. Introduction

Reliable transmission of digital data is an important task in many applications. But transmission channels often suffer from noise and so errors can occur. Error-correcting codes are employed to tackle this problem. By adding redundancy such codes allow to detect and correct a certain amount of transmission errors [8].

   Coding theory was founded by Shannon's seminal paper in 1948 [12] in

---

[§]Correspondence author

which it was proved that reliable communication is possible at any rate below the channel capacity. Since then the construction of such codes and the study of their properties is an ongoing task.

Recently a connection between Gröbner bases and linear codes have been established [3, 5, 10, 11]. Gröbner bases are a powerful tool in commutative algebra providing a uniform approach to grasping a wide range of problems such as solving algebraic systems of equations, ideal membership decision, and effective computation in residue class rings modulo polynomial ideals [1, 2, 6, 13].

The first connection between linear codes and Gröbner bases was established in [5] which soon became known as the "Cooper philosophy". This link was based on the description of cyclic codes as ideals in a certain polynomial ring, where entries of a codeword are viewed as coefficients of a polynomial.

In [3], a different connection between linear codes and ideals in polynomial rings was presented, which was followed up in [10, 11]. In this approach, linear codes are described by a binomial ideal in a polynomial ring over an arbitrary field that can be written as the sum of a toric ideal and a non-prime ideal, the so-called code ideal.

The code ideal holds useful information about the code. However, the code ideal can be exploited only if a Gröbner basis is known. Gröbner bases with respect to any monomial order can be computed by Buchberger's algorithm which is implemented in most computer algebra systems. In [10] it has been shown that the reduced Gröbner basis with respect to the lexicographic order can be easily constructed from a standard generator matrix. Unfortunately, it has been proved that many applications require a degree compatible ordering [3, 4]. And as Gröbner bases vary with the term order and the computation of Gröbner bases can be rather costly, it is advantageous to compute the *universal Gröbner basis* instead, i.e., the union of all reduced Gröbner bases [14]. Surprisingly, although infinitely many term orders exist, there is only a finite number of Gröbner bases [9].

For binary linear codes it has been shown that the universal Gröbner basis equals the set of circuits [11]. Hence the computation of the universal Gröbner basis for binary codes amounts to the computation of its circuits. Furthermore, in [7] the universal Gröbner basis for the code ideal of a binary linear code has been completely described: It consists of all binomials associated with the codewords whose Hamming weight is less than or equal to the Singleton bound and which satisfy a certain rank condition. For some codes *all* codewords of Hamming weight less than or equal to the Singleton bound satisfy the mentioned rank condition. Such codes were called Singleton codes in [7].

In this paper, basic properties of Singleton codes will be established and the classical binary linear codes (Hamming, simplex, Golay, Reed-Muller, and cyclic codes) will be examined whether or not they have the Singleton property. For instance, the binary Golay code and its extension by parity check are Singleton, while not all Hamming and Reed-Muller codes are Singleton.

This paper is organized as follows. The next section provides the required background from coding theory and the third section contains the main results.

2. Linear Codes

Let $\mathbb{F}$ be a finite field and let n and k be positive integers with $n \geq k$. A *linear code* $\mathcal{C}$ of length n and dimension k over $\mathbb{F}$ is the image of a one-to-one linear mapping $\phi : \mathbb{F}^k \to \mathbb{F}^n$; that is, $\mathcal{C} = \{\phi(a) \mid a \in \mathbb{F}^k\}$. Such a code is denoted as $[n, k]$ *code* and its elements are called *codewords*. In algebraic coding, the codewords are always written as row vectors. Alternatively, a code $\mathcal{C}$ can be described as the row space of a matrix $G \in \mathbb{F}^{k \times n}$, whose rows form a basis of $\mathcal{C}$, and the matrix G is then called a *generator matrix* for $\mathcal{C}$.

A code $\mathcal{C}$ is *systematic* if it has a generator matrix which is in *standard form*, i.e., $G = (I_k \mid M)$, where I_k is the $k \times k$ identity matrix. Note that a generator matrix for an $[n, k]$ code can contain zero columns. Such a code can be shortened by deleting those columns giving a code of smaller length and equal dimension. All subsequently considered codes are assumed to have no zero columns.

The *support* of a vector $u \in \mathbb{F}^n$, written $\text{supp}(u)$, is the subset of $\underline{n} = \{1, \ldots, n\}$ given by all indices $i \in \underline{n}$ with $u_i \neq 0$, and the *Hamming weight*, denoted by $\text{wt}(u)$, is the number of non-zero components and so equals the cardinality of the codeword's support. Note that for a binary code, each codeword is completely determined by its support. For binary codes, a *circuit* is a codeword whose support is minimal with respect to inclusion. The *weight distribution* of an $[n, k]$ code $\mathcal{C}$ is a finite sequence of integers $A_0, A_1, \ldots, A_n$, where A_i denotes the number of codewords in $\mathcal{C}$ having Hamming weight i, $0 \leq i \leq n$. The *Hamming distance* between two vectors $u, v \in \mathbb{F}^n$ is the number of positions in which they differ and so is given by the Hamming weight $\text{wt}(u - v)$ of the difference vector. The Hamming distance defines a metric on $\mathbb{F}^n$. The minimum Hamming distance between any to distinct codewords in $\mathcal{C}$ is the *minimum distance* of the code $\mathcal{C}$. An $[n, k]$ code having minimum distance d is denoted as $[n, k, d]$ code. The *Singleton bound* for linear codes states that for each $[n, k, d]$ code, $d \leq n - k + 1$.

For any matrix $G \in \mathbb{F}^{k \times n}$ and any subset $J \subseteq \underline{n}$ of indices, let G_J denote the $k \times |J|$ submatrix of G consisting of the columns in G with indices from J. Similarly, let c_J be the vector of length $|J|$ consisting of the coordinates in c with indices from J. A subset $J \subseteq \underline{n}$ of cardinality k is called an *information set* of an $[n, k]$ code with generator matrix G if the $k \times k$ submatrix G_J has rank k. The following are equivalent: (1) The set of indices J is an information set. (2) For each $m \in \mathbb{F}^k$ there is a unique $c \in \mathcal{C}$ with $c_J = m$. (3) For any generator matrix G of $\mathcal{C}$, G_J has rank k. By the second assertion, a code cannot contain an information set $J \subseteq \underline{n} \setminus \mathrm{supp}(c)$ at the zero positions of a non-zero codeword c.

The dual code $\mathcal{C}^\perp$ of an $[n, k]$ code $\mathcal{C}$ over $\mathbb{F}$ is an $[n, n - k]$ code consisting of all words $u \in \mathbb{F}^n$ such that $u \cdot c = uc^T = 0$ for each $c \in \mathcal{C}$, where c^T denotes the transposed of c. If $G = (I_k \mid M)$ is a generator matrix for $\mathcal{C}$, then $H = \left(-M^T \mid I_{n-k}\right)$ is a generator matrix for $\mathcal{C}^\perp$. For each word $c \in \mathbb{F}^n$, we have $c \in \mathcal{C}$ if and only if $Hc^T = \mathbf{0}$. The matrix H is a *parity check matrix* for $\mathcal{C}$.

3. Singleton Codes

This section is devoted to the question which of the classical binary codes are Singleton. The results require only linear algebra and therefore, basics about Gröbner bases and code ideals can be omitted.

3.1. General Considerations

A binary linear code is called a *Singleton code* (or simply *Singleton*) if every codeword satisfying the Singleton bound is a circuit [7].

The well-known *maximum distance separable* (MDS) codes attain the Singleton bound with equality and so are Singleton codes. In particular, the trivial codes are MDS and so are Singleton.

Proposition 1 ([7]). *A binary $[n, k]$ code $\mathcal{C}$ is Singleton if and only if for every generator matrix G of $\mathcal{C}$ each codeword c in $\mathcal{C}$ with $\mathrm{wt}(c) \leq n - k + 1$ satisfies*

$$rk\left(G_{\underline{n} \setminus \{supp(c)\}}\right) = k - 1. \tag{1}$$

Proof. Let $\mathcal{C}$ be a Singleton code and $c \in \mathcal{C}$ a circuit. Claim that c satisfies (1). Indeed, there is a non-zero information word $x \in \mathbb{F}_2^k$ with $x \cdot G = c$ for any generator matrix G. But $x \cdot G_{\underline{n} \setminus \mathrm{supp}(c)} = \mathbf{0}$ and so $G_{\underline{n} \setminus \mathrm{supp}(c)}$ cannot have

full row rank k. Suppose $G_{\underline{n}\backslash\mathrm{supp}(c)}$ has a smaller rank than $k-1$. Then by the dimension formula for linear mappings,

$$k = \dim \ker G_{\underline{n}\backslash\mathrm{supp}(c)} + \dim \mathrm{im}\, G_{\underline{n}\backslash\mathrm{supp}(c)} < \dim \ker G_{\underline{n}\backslash\mathrm{supp}(c)} + (k-1)$$

and so $\dim \ker G_{\underline{n}\backslash\mathrm{supp}(c)} > 1$. Thus there must be another information word $x' \in \mathbb{F}_2^k$ with $x' \cdot G_{\underline{n}\backslash\mathrm{supp}(c)} = \mathbf{0}$. Put $c' = x' \cdot G$. So for each index i in $\underline{n}\backslash\mathrm{supp}(c)$, $c_i' = x' \cdot G_{\{i\}} = 0$ and thus $\mathrm{supp}(c') \subseteq \mathrm{supp}(c)$. But the encoding is one-to-one and so the codeword c' is distinct from c. It follows that $\mathrm{supp}(c') \subsetneq \mathrm{supp}(c)$ contradicting the hypothesis that c is a circuit. Hence, the rank of $G_{\underline{n}\backslash\mathrm{supp}(c)}$ is equal to $k-1$.

Conversely, let c be a codeword with $\mathrm{wt}(c) \le n-k+1$ satisfying (1). Then c has at least $k-1$ entries that are 0 and so by hypothesis, among those one can find exactly $k-1$ coordinates $J \subseteq \underline{n} \setminus \mathrm{supp}(c)$ such that G_J has rank $k-1$. But as the generator matrix G has rank k there must be another column in G, say indexed by i, with $c_i = 1$, such that $G_{J\cup\{i\}}$ is a $k \times k$ matrix of rank k; that is, $J \cup \{i\}$ is an information set. Let $c' \in \mathcal{C}$ be a non-zero codeword such that $\mathrm{supp}(c') \subseteq \mathrm{supp}(c)$. Clearly, $c_\ell = 0$ implies $c_\ell' = 0$ and thus $c_J' = \mathbf{0} = c_J$. Moreover, $c_i' = 1 = c_i$ and so $c_{J\cup\{i\}}' = c_{J\cup\{i\}}$ because $J \cup \{i\}$ is an information set and c' is not the zero codeword. This implies that $c = c'$. Hence, c is a circuit. $\square$

Proposition 2. *A binary $[n,k]$ code $\mathcal{C}$ is a Singleton code if and only if there are no non-zero codewords c_1, c_2 in $\mathcal{C}$ such that $\mathrm{supp}(c_1) \cap \mathrm{supp}(c_2) = \emptyset$ and $\mathrm{wt}(c_1) + \mathrm{wt}(c_2) \le n-k+1$.*

Proof. Suppose there are non-zero words $c_1, c_2 \in \mathcal{C}$ such that $\mathrm{supp}(c_1) \cap \mathrm{supp}(c_2) = \emptyset$ and $\mathrm{wt}(c_1) + \mathrm{wt}(c_2) \le n-k+1$. Then $c = c_1 + c_2 \in \mathcal{C}$ is a codeword with Hamming weight $\mathrm{wt}(c) = \mathrm{wt}(c_1) + \mathrm{wt}(c_2)$ satisfying the Singleton bound. But the codeword c is not a circuit since $\mathrm{supp}(c_i) \subsetneq \mathrm{supp}(c)$ for $i = 1, 2$ and so $\mathcal{C}$ is not a Singleton code.

Conversely, suppose $\mathcal{C}$ is not a Singleton code. Then there exists a codeword c in $\mathcal{C}$ of Hamming weight $\mathrm{wt}(c) \le n-k+1$ which is not a circuit. This implies that there is another codeword c_1 with the property that $\mathrm{supp}(c_1) \subsetneq \mathrm{supp}(c)$. Put $c_2 = c + c_1$. Then $\mathrm{supp}(c_2) = \mathrm{supp}(c)\backslash\mathrm{supp}(c_1)$ and so $\mathrm{supp}(c_1) \cap \mathrm{supp}(c_2) = \emptyset$. Moreover, $n-k+1 \ge \mathrm{wt}(c) = \mathrm{wt}(c_1) + \mathrm{wt}(c_2)$. $\square$

Proposition 3. *Let $\mathcal{C}$ be a binary $[n,k,d]$ code. If the minimum distance of $\mathcal{C}$ satisfies $d > \frac{1}{2}(n-k+1)$, then $\mathcal{C}$ is a Singleton code.*

Proof. Suppose $\mathcal{C}$ is not a Singleton code. Then by Prop. 2, there are non-zero codewords $c_1, c_2 \in \mathcal{C}$ such that $\mathrm{supp}(c_1) \cap \mathrm{supp}(c_2) = \emptyset$ and $\mathrm{wt}(c_1) + \mathrm{wt}(c_2) \leq n - k + 1$. But if $\mathrm{wt}(c_2) > \frac{1}{2}(n - k + 1)$, then $\mathrm{wt}(c_1) \leq \frac{1}{2}(n - k + 1)$ because the codewords have disjoint support. This contradicts the hypothesis. $\square$

3.2. Code Modifications and Constructions

The closure properties of the class of Singleton codes under prominent code modifications and constructions are studied. First, three basic code modifications are considered. For this, let $\mathcal{C}$ be a binary $[n, k]$ code.

- Puncturing: The code punctured at a parity-check position $i \in \underline{n}$ is an $[n - 1, k]$ code given as

$$\dot{\mathcal{C}}_i = \{(c_1, \ldots, c_{i-1}, c_{i+1}, \ldots, c_n) \in \mathbb{F}_2^{n-1} \mid$$
$$(c_1, \ldots, c_{i-1}, c_i, c_{i+1}, \ldots, c_n) \in \mathcal{C} \text{ for some } c_i \in \mathbb{F}_2\}.$$

- Shortening: The code shortened at an information position $i \in \underline{n}$ is an $[n - 1, k - 1]$ code defined as

$$\check{\mathcal{C}}_i = \{(c_1, \ldots, c_{i-1}, c_{i+1}, \ldots, c_n) \in \mathbb{F}_2^{n-1} \mid$$
$$(c_1, \ldots, c_{i-1}, 0, c_{i+1}, \ldots, c_n) \in \mathcal{C}\}.$$

- Extension: The code extended by an overall parity check is an $[n + 1, k]$ code represented as

$$\hat{\mathcal{C}} = \{(c_1, \ldots, c_n, c_{n+1}) \in \mathbb{F}_2^{n+1} \mid (c_1, \ldots, c_n) \in \mathcal{C}, \ c_{n+1} = \sum_{i=1}^{n} c_i\}.$$

Punturing a Singleton code may not result in another Singleton code as the following example will show.

Example 1. Consider the binary $[8, 4]$ code $\mathcal{C}$ with generator matrix

$$G = (I_4 \mid M), \quad M = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \end{pmatrix}.$$

Puncturing at the last coordinate gives a $[7, 4]$ code $\dot{\mathcal{C}}$ with generator matrix $\dot{G} = G_{\{1, \ldots, 7\}}$. The codeword $c = (1, 1, 1, 1, 0, 0, 0) \in \dot{\mathcal{C}}$ attains the Singleton bound but the rank condition in Eq. (1) is not fulfilled:

$$\mathrm{rk}\left(\dot{G}_{\underline{7} \backslash \mathrm{supp}(c)}\right) = \mathrm{rk} \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix} = 2 < 3 = 4 - 1.$$

Hence, the punctured code $\dot{C}$ is not Singleton although the code $\mathcal{C}$ is.

On the other hand, the class of Singleton codes is closed under shortening.

Proposition 4. *The shortening of a binary Singleton code gives another binary Singleton code.*

Proof. Let $\mathcal{C}$ be an $[n, k]$ Singleton code. Without restriction, the shortened code $\check{\mathcal{C}}_n$ may be taken. Let $\check{G} \in \mathbb{F}_2^{k-1 \times n-1}$ be a generator matrix for $\check{\mathcal{C}}$ and let c be a codeword in $\mathcal{C}$ with $c_n = 1$. Note that such a codeword must exist because only codes without zero columns are considered. Then the following matrix extended by the word c is a generator matrix for the code $\mathcal{C}$,

$$G = \begin{pmatrix} \check{G} & \mathbf{0} \\ & c \end{pmatrix} = \begin{pmatrix} & \check{G} & & \begin{matrix} 0 \\ \vdots \\ 0 \end{matrix} \\ * & \cdots & * & 1 \end{pmatrix} \in \mathbb{F}_2^{k \times n}. \tag{2}$$

Let $\check{c} \in \check{\mathcal{C}}$ be a non-zero codeword of Hamming weight $\mathrm{wt}(\check{c}) \leq (n-1) - (k-1) + 1 = n - k + 1$. Then by definition $c' = (\check{c}, 0)$ is a codeword in $\mathcal{C}$ of the same Hamming weight. But $\mathcal{C}$ is a Singleton code and so by Prop. 1,

$$\mathrm{rk}\left(G_{\underline{n} \backslash \mathrm{supp}(c')}\right) = k - 1, \tag{3}$$

where the submatrix $G_{\underline{n} \backslash \mathrm{supp}(c')}$ is composed of

$$\begin{pmatrix} & \check{G} & \\ * & \cdots & * \end{pmatrix}_{\underline{n-1} \backslash \mathrm{supp}(\check{c})} \quad \text{and} \quad \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}.$$

However, the last row of this submatrix does not belong to the span of the first $k - 1$ rows. Thus Eq. (3) yields

$$\mathrm{rk}\left(\check{G}_{\underline{n-1} \backslash \mathrm{supp}(\check{c})}\right) = k - 2.$$

It follows that each codeword in $\check{\mathcal{C}}$ whose Hamming weight satisfies the Singleton bound also satisfies the rank condition in Eq. (1). Hence, $\check{\mathcal{C}}$ is a Singleton code. $\qquad\square$

The extension of a Singleton code by adding an overall parity check yields a Singleton code under certain conditions.

Proposition 5. *Let $\mathcal{C}$ be a binary $[n, k]$ code. If $\mathcal{C}$ is a Singleton code which contains no codeword of Hamming weight $n - k + 2$, then the extended code $\hat{\mathcal{C}}$ obtained by adding an overall parity check is also a Singleton code.*

Proof. Let $\mathcal{C}$ be a binary $[n, k]$ code and let G be a generator matrix for $\mathcal{C}$. Then the extended code $\hat{\mathcal{C}}$ has the generator matrix $\hat{G} = (G \mid v)$, where $v \in \mathbb{F}_2^k$ is the sum of all column vectors in G.

Let c be a codeword in $\hat{\mathcal{C}}$ of Hamming weight $\mathrm{wt}(c) \leq n - k + 2$ and denote by c' the codeword obtained from c by deleting the last coordinate. Two cases can occur.

First, the last coordinate of c is 1. Then the codeword c' has Hamming weight $\mathrm{wt}(c') = \mathrm{wt}(c) - 1 \leq n - k + 1$ and so is a circuit in $\mathcal{C}$ satisfying the rank condition $\mathrm{rk}\left(G_{\underline{n}\setminus\mathrm{supp}(c')}\right) = k - 1$. Moreover,

$$\mathrm{rk}\left(\hat{G}_{\underline{n+1}\setminus\mathrm{supp}(c)}\right) = \mathrm{rk}\left(G_{\underline{n}\setminus\mathrm{supp}(c')}\right) = k - 1$$

shows that c is a circuit in $\hat{\mathcal{C}}$ (see Prop. 1).

Second, the last coordinate of c is 0. Then the codeword c must be of Hamming weight $\mathrm{wt}(c) \leq n - k + 1$ because otherwise $\mathrm{wt}(c') = n - k + 2$, which is excluded by the hypothesis. A similar argument as in the first case then exhibits that c is a circuit in $\hat{\mathcal{C}}$. $\qquad\square$

Second, two basic code constructions are considered.

- Direct sum: If $\mathcal{C}_1$ is a binary $[n_1, k_1]$ code and $\mathcal{C}_2$ is a binary $[n_2, k_2]$ code, then the direct sum is a binary $[n_1 + n_2, k_1 + k_2]$ code given by

$$\mathcal{C}_1 \oplus \mathcal{C}_2 = \left\{(c_1, c_2) \in \mathbb{F}_2^{n_1+n_2} \mid c_1 \in \mathcal{C}_1, c_2 \in \mathcal{C}_2\right\}.$$

- $(u \mid u + v)$ construction: If $\mathcal{C}_1$ is a binary $[n, k_1]$ code and $\mathcal{C}_2$ is a binary $[n, k_2]$ code, then the $(u \mid u + v)$ construction yields an $[2n, k_1 + k_2]$ code defined by

$$\mathcal{C} = \left\{(c_1, c_1 + c_2) \in \mathbb{F}_2^{2n} \mid c_1 \in \mathcal{C}_1, c_2 \in \mathcal{C}_2\right\}.$$

The direct sum $\mathcal{C}_1 \oplus \mathcal{C}_2$ of any two Singleton codes $\mathcal{C}_1$ and $\mathcal{C}_2$ cannot be a Singleton code. To see this, let $\mathcal{C}_i$ be a binary $[n_i, k_i]$ Singleton code with generator matrix G_i, $i \in \{1, 2\}$. Then the direct sum $\mathcal{C} = \mathcal{C}_1 \oplus \mathcal{C}_2$ has the generator matrix

$$G = \begin{pmatrix} G_1 & \mathbf{0} \\ \mathbf{0} & G_2 \end{pmatrix}.$$

So for any codeword $c = (c_1, c_2) \in \mathcal{C}$, where $c_i \in \mathcal{C}_i$, $i = 1, 2$,

$$\mathrm{rk}\left(G_{\underline{2n}\backslash\mathrm{supp}(c)}\right) = \mathrm{rk}\left((G_1)_{\underline{n}\backslash\mathrm{supp}(c_1)}\right) + \mathrm{rk}\left((G_2)_{\underline{n}\backslash\mathrm{supp}(c_2)}\right)$$
$$\leq (k_1 - 1) + (k_2 - 1) = k_1 + k_2 - 2 < k_1 + k_2 - 1.$$

Fortunately, the direct sum of two codes is rather uninteresting in applications because the minimum distance of such composite codes does not exceed the minimum distance of the component codes.

The $(u \mid u+v)$-construction will be later discussed in connection with Reed-Muller codes.

3.3. Hamming Codes and Simplex Codes

For any number $r \geq 3$, the rth binary Hamming code is a linear code of length $n = 2^r - 1$, dimension $k = n - r$, and minimum distance $d = 3$. This code can be defined as the kernel of a matrix whose columns are exactly the vectors in $\mathbb{F}_2^r$. Different arrangements of the columns give equivalent codes.

Simplex codes are the duals of the Hamming codes. The rth binary simplex code is a binary $[n = 2^r - 1, r]$ code with minimum distance $d = 2^{r-1}$ and weight distribution $A_0 = 1$, $A_d = 2^r - 1$, and $A_i = 0$ for $i \in \underline{n} \setminus \{0, d\}$ [8]. Therefore, all non-zero codewords are circuits and so the following holds.

Proposition 6. *Every simplex code is Singleton.*

Proposition 7. *The third and fourth binary Hamming codes are Singleton, and for any integer $r \geq 5$ the rth binary Hamming code is not Singleton.*

Proof. The Singleton bound for the rth binary Hamming code is $S = r + 1$.

By Prop. 3, the rth binary Hamming code is Singleton if the minimum distance satisfies $3 = d > \frac{1}{2}(r + 1)$ which only holds if $r = 3$ or $r = 4$.

Let $r \geq 5$ and let $\mathcal{C}$ denote the rth binary Hamming code with generator matrix $G = (I_k \mid M)$. The $k \times r$ matrix M consists of all row vectors in $\mathbb{F}_2^r$

except for the zero vector and the unit vectors. Thus it can be assumed that the first two rows are

$$m_1 = (0, 1, \ldots, 1, 0, \ldots, 0) \quad \text{and} \quad m_2 = (0, 0, \ldots, 0, 1, \ldots, 1),$$

where m_1 has Hamming weight $\left\lfloor \frac{r-1}{2} \right\rfloor$ and m_2 has Hamming weight $\left\lceil \frac{r-1}{2} \right\rceil$ and their supports are disjoint. Since $r \geq 5$, the weights satisfy $\left\lfloor \frac{r-1}{2} \right\rfloor \geq 2$ and $\left\lceil \frac{r-1}{2} \right\rceil \geq 2$ and so both m_1 and m_2 are not unit vectors.

Adding the first two row vectors in G gives a codeword c with support $\{1, 2, k + 2, \ldots, n\}$ and Hamming weight $\mathrm{wt}(c) = 2 + r - 1 = r + 1 = S$. However, the first two rows of the submatrix $G_{\underline{n} \backslash \mathrm{supp}(c)}$ are zero and thus its rank is $\leq k - 2$. In other words, c violates the rank condition in Eq. (1) and so is not a circuit. $\qquad\square$

3.4. Binary Golay Code

The binary Golay code is a $[23, 12, 7]$ code with weight distribution [8]

i	0	7	8	11	12	15	16	23
A_i	1	253	506	1288	1288	506	253	1

The minimum distance satisfies $7 = d > \frac{1}{2}(n - k + 1) = 6$ and so by Prop. 3 gives rise to the following result.

Proposition 8. *The binary Golay code is Singleton.*

The parity check extension of the binary Golay coded is the extended binary $[24, 12, 8]$ Golay code. Note that this code contains no codeword of Hamming weight $23 - 12 + 2 = 13$ and so Prop. 5 yields the following result.

Proposition 9. *The extended binary Golay code is Singleton.*

3.5. Reed-Muller Codes

Reed-Muller codes can be introduced by using the $(u \mid u + v)$-construction [8].

Let $m \geq 2$ and r be integers with $0 \leq r \leq m$. The rth order Reed-Muller code of length 2^m denoted by $\mathcal{R}(r, m)$ is for $r = 0$ defined as the binary repetition code, i.e., $\mathcal{R}(0, m) = \{\mathbf{0}, \mathbf{1}\} \subset \mathbb{F}_2^{2^m}$, where $\mathbf{1}$ is the all-1 word, for $r = m$ as the ambient space, i.e., $\mathcal{R}(m, m) = \mathbb{F}_2^{2^m}$, and for $1 \leq r < m$ as

$$\mathcal{R}(r, m) = \{(u, u + v) \mid u \in \mathcal{R}(r, m - 1), v \in \mathcal{R}(r - 1, m - 1)\}.$$

The dimension of the code $\mathcal{R}(r, m)$ is $k = \sum_{i=0}^{r} \binom{m}{i}$ and the minimum distance is $d = 2^{m-r}$. Furthermore, the dual of the rth order Reed Muller code is the

Reed Muller code $\mathcal{R}^\perp(r, m) = \mathcal{R}(m - r - 1, m)$ for $0 \le r < m$ and $\mathcal{R}^\perp(m, m) = \{\mathbf{0}\}$.

Let $G(r, m)$ denote a generator matrix for $\mathcal{R}(r, m)$. Clearly, $G(0, m)$ is the all-1 row vector and $G(m, m)$ is the identity matrix I_{2^m}. By the $(u \mid u + v)$-construction, a generator matrix for the other Reed-Muller codes is given as follows,

$$G(r, m) = \begin{pmatrix} G(r, m - 1) & G(r, m - 1) \\ \mathbf{0} & G(r - 1, m - 1) \end{pmatrix}. \qquad (4)$$

As already pointed out, the trivial codes $\mathcal{R}(0, m)$ and $\mathcal{R}(m, m)$ are Singleton for all integers $m \ge 2$.

Proposition 10. *The first order Reed-Muller codes are Singleton.*

Proof. Let $m \ge 2$. The code $\mathcal{R}(1, m)$ is a $[2^m, m + 1]$ code with weight distribution $A_0 = 1$, $A_d = 2^{m+1} - 2$, and $A_n = 1$, where $d = 2^{m-1}$ is the minimum distance of the code. The minimum-weight codewords satisfy the Singleton bound $2^{m-1} = d \le n - k + 1 = 2^m - m$ and so the code is Singleton. $\qquad \square$

Proposition 11. *For any integer $m \ge 2$, the $(m - 1)$th order Reed-Muller code is Singleton.*

Proof. The code $\mathcal{R}(m - 1, m)$ is an even-weight code consisting of all vectors in $\mathbb{F}_2^{2^m}$ with even Hamming weight. Its Singleton bound is 2 and so it is an MDS code. $\qquad \square$

Proposition 12. *The second order Reed-Muller codes $\mathcal{R}(2, 3)$ and $\mathcal{R}(2, 4)$ are Singleton, and for any integer $m \ge 5$ the second-order Muller code $\mathcal{R}(2, m)$ is not Singleton.*

Proof. The code $\mathcal{R}(2, m)$ is an $\left[n = 2^m, k = 1 + m + \binom{m}{2}\right]$ code with minimum distance $d_m = 2^{m-2}$ and the Singleton bound is $S_m = 2^m - \frac{m(m+1)}{2}$.

We have $S_3 = 2$ and $d_3 = 2$ as well as $S_4 = 6$ and $d_4 = 4$. In each case, $d_m > \frac{1}{2} \cdot S_m$ and thus by Prop. 3, the codes $\mathcal{R}(2, 3)$ and $\mathcal{R}(2, 4)$ are Singleton.

Claim that for any $m \ge 5$, the code $\mathcal{R}(2, m)$ is not Singleton. Indeed, let $m \ge 5$. A generator matrix for $\mathcal{R}(2, m)$ is

$$G(2, m) = \begin{pmatrix} G(2, m - 1) & G(2, m - 1) \\ \mathbf{0} & G(1, m - 1) \end{pmatrix}.$$

Since $G(1, m - 1)$ is a generator matrix for the first-order Reed-Muller code of length 2^{m-1}, the vector $c = (0, \dots, 0, 1, \dots, 1)$ of length 2^m consisting of 2^{m-1}

entries 0 and 2^{m-1} entries 1 belongs to the code $\mathcal{R}(2, m)$. The codeword c satisfies $\text{wt}(c) = 2^{m-1} \leq 2^m - \frac{m(m+1)}{2} = S_m$ for any $m \geq 5$. Moreover, the submatrix of $G(2, m)$ corresponding to the zero entries of c is

$$G(2, m)_{\underline{n}\backslash\text{supp}(c)} = \begin{pmatrix} G(2, m-1) \\ \mathbf{0} \end{pmatrix}$$

and has rank

$$\text{rk}\,(G(2, m-1)) = \sum_{i=0}^{2} \binom{m-1}{i} = m + \binom{m-1}{2}.$$

But $k - 1 = m + \binom{m}{2}$ and so $\text{rk}\,\big(G(2, m)_{\underline{n}\backslash\text{supp}(c)}\big) < k - 1$. Thus the codeword c violates rank condition in Eq. (1) and hence the code $\mathcal{R}(2, m)$ is not Singleton. ☐

Lemma 13. *For any integers $m \geq 2$ and r with $1 \leq r \leq m$, the code $\mathcal{R}(r, m)$ contains the codeword $(0, \ldots, 0, 1, \ldots, 1)$ of Hamming weight 2^{m-r+1}.*

Proof. Note that the code $\mathcal{R}(1, m)$ contains the all-1 word of Hamming weight 2^m.

Let $r > 1$. By definition, the codewords of $\mathcal{R}(r, m)$ are of the form $(u, u+v)$, where $u \in \mathcal{R}(r, m-1)$ and $v \in \mathcal{R}(r-1, m-1)$. By induction, the codeword $c = (0, \ldots, 0, 1, \ldots, 1)$ of length 2^{m-1} and Hamming weight $2^{(m-1)-(r-1)+1} = 2^{m-r+1}$ belongs to $\mathcal{R}(r-1, m-1)$. Therefore, the codeword $(\mathbf{0} \mid \mathbf{0}+c) = (\mathbf{0} \mid c)$ has the required property. ☐

Proposition 14. *Let $m \geq 2$ and r be integers with $1 \leq r \leq m$. If*

$$2^m - \sum_{i=0}^{r} \binom{m}{i} + 1 \geq 2^{m-r+1}, \tag{5}$$

then the rth order Reed-Muller code of length 2^m is not Singleton.

Proof. The code $\mathcal{R}(r, m)$ has length $n = 2^m$, dimension $k = \sum_{i=0}^{r} \binom{m}{i}$, and by Eq. (4), the following generator matrix in upper-triangular block-diagonal form,

$$G(r, m) = \begin{pmatrix} G(r, m{-}1) & * & & * & * \\ \mathbf{0} & G(r{-}1, m{-}2) & & * & * \\ \vdots & \mathbf{0} & \ddots & \vdots & \vdots \\ & & & G(2, m{-}r{+}1) & * \\ \mathbf{0} & & & \mathbf{0} & G(1, m{-}r{+}1) \end{pmatrix}.$$

By Lemma 13, the vector $c = (0, \ldots, 0, 1, \ldots, 1)$ with Hamming weight 2^{m-r+1} lies in $\mathcal{R}(r, m)$. Moreover by hypothesis, Eq. (5), the Hammig weight of this codeword satisfies the Singleton bound. The submatrix of $G(r, m)$ corresponding to the zero entries of c is obtained from $G(r, m)$ by removing the last 2^{m-r+1} columns, i.e.,

$$G(r, m)_{\underline{n} \backslash \text{supp}(c)} = \begin{pmatrix} G(r, m-1) & * & & * \\ \mathbf{0} & G(r-1, m-2) & & * \\ \vdots & \mathbf{0} & \ddots & \vdots \\ & & & G(2, m-r+1) \\ \mathbf{0} & & & \mathbf{0} \end{pmatrix}.$$

Since the last $m-r+2$ rows of this submatrix are zero, its rank is $k-(m-r+2) < k-1$. Hence, the code $\mathcal{R}(r, m)$ is not Singleton. $\qquad\square$

The above result is in accordance with Cor. 10 and Prop. 11, 12 showing that the codes $\mathcal{R}(r, m)$ are Singleton for $r = 0, 1, m-1, m$.

The next result shows that the number of Reed-Muller codes which are Singleton is finite.

Proposition 15. *For any integer $r \geq 2$, there exists a positive number M_0 such that for all integers $m \geq M_0$ the code $\mathcal{R}(r, m)$ is not Singleton.*

Proof. Claim that

$$\lim_{m \to \infty} \frac{2^m - \sum_{i=0}^{r} \binom{m}{i} + 1}{2^{m-r+1}} = 2^{r-1} > 0. \tag{6}$$

Indeed,

$$\lim_{m \to \infty} \frac{2^m - \sum_{i=0}^{r} \binom{m}{i} + 1}{2^{m-r+1}} = 2^{r-1} - \frac{\lim\limits_{m \to \infty} \sum_{i=1}^{r} \binom{m}{i}}{\lim\limits_{m \to \infty} 2^{m-r+1}}. \tag{7}$$

In the quotient on the right-hand side, both denominator and numerator converge to infinity for $m \to \infty$. Thus L'Hpital's rule can be used to compute the limit. To this end, note that $\sum_{i=0}^{r} \binom{m}{i}$ as a polynomial in m is of degree r and 2^{m-r+1} as a function in m differentiated a few times yields $C \cdot 2^{m-r+1}$, where C is a constant. Therefore, if L'Hpital's rule is applied r times, then

$$\frac{\lim\limits_{m \to \infty} \sum_{i=1}^{r} \binom{m}{i}}{\lim\limits_{m \to \infty} 2^{m-r+1}} = \lim_{m \to \infty} \frac{\tilde{C}}{C \cdot 2^{m-r+1}} = 0,$$

where $\tilde{C}$ is a constant. Inserting this into Eq. (7) yields Eq. (6) proving the claim.

By Eq. (6), for any $\epsilon > 0$ there exists a number M_0 such that for all integers $m \geq M_0$,

$$\left| 2^{r-1} - \frac{2^m - \sum_{i=0}^{r} \binom{m}{i} + 1}{2^{m-r+1}} \right| < \epsilon. \tag{8}$$

Note that taking the absolute value on the left-hand side of Eq. (8) is not necessary since the value is already non-negative. Setting $\epsilon = 2^{r-1} - 1$ gives

$$2^{m-r+1} < 2^m - \sum_{i=0}^{r} \binom{m}{i} + 1.$$

Hence by Prop. 14, the result follows.   □

3.6. Cyclic Codes

Binary cyclic codes form a useful class of codes [8]. They contain the binary Hamming codes and so by Prop. 7 not all binary cyclic codes can be Singleton.

A binary linear code $\mathcal{C}$ of length n is *cyclic* if the cyclic shift of coordinates $i \mapsto i+1$ modulo n of any codeword also yields a codeword. Each binary vector $c = (c_0, c_1, \ldots, c_{n-1})$ can be associated with a polynomial $c(x) = c_0 + c_1 x + \cdots + c_{n-1} x^{n-1}$ in $\mathbb{F}_2[x]$. In this case, the cyclic shift of a codeword c corresponds to the codeword obtained by multiplying the polynomial $c(x)$ by x modulo $x^n - 1$. It follows that the binary cyclic codes of length n are precisely the ideals in the quotient ring $\mathbb{F}_2[x]/\langle x^n - 1 \rangle$. Furthermore, for each non-zero cyclic code $\mathcal{C} \subset \mathbb{F}_2^n$ there exists a polynomial $g(x)$ in $\mathbb{F}_2[x]$ called *generator polynomial* of $\mathcal{C}$ with the following properties:

- $g(x)$ is the unique monic polynomial in $\mathcal{C}$ of minimal degree and $g(x)$ divides $x^n - 1$,

- $g(x)$ generates $\mathcal{C}$ as an ideal in $\mathbb{F}_2[x]/\langle x^n - 1 \rangle$, written $\mathcal{C} = \langle g(x) \rangle$,

- $k = n - \deg g(x)$ is the dimension and $\{g(x), xg(x), \ldots, x^{k-1}g(x)\}$ is a basis for $\mathcal{C}$.

These properties imply that the binary cyclic codes of length n are in one-to-one correspondence with the factors of the polynomial $x^n - 1$ in $\mathbb{F}_2[x]$.

In the following, the length n is assumed to be odd (in this case, the polynomial $x^n - 1$ has no repeated irreducible factors).

Example 2. Consider the binary cyclic codes of length 7. The factorization of $x^7 - 1$ into irreducible components gives

$$x^7 - 1 = (1 + x)(1 + x + x^3)(1 + x^2 + x^3).$$

Thus there are eight binary cyclic codes of length 7. Computing the code parameters and applying Prop. 3 reveals that all these cyclic codes are Singleton (Table 1).

k	generator polynomial	d	S	Singleton?
1	$1 + x + x^2 + x^3 + x^4 + x^5 + x^6$	7	7	$\checkmark$
3	$1 + x^2 + x^3 + x^4$	4	5	$\checkmark$
3	$1 + x + x^3 + x^4$	4	5	$\checkmark$
4	$1 + x^2 + x^3$	3	4	$\checkmark$
4	$1 + x + x^3$	3	4	$\checkmark$
6	$1 + x$	2	2	$\checkmark$

Table 1: Parameters of the binary cyclic codes of length 7.

Example 3. Consider the binary cyclic codes of length 9. The factorization of $x^9 - 1$ into irreducible polynomials yields

$$x^9 - 1 = (1 + x)(1 + x + x^2)(1 + x^3 + x^6).$$

It follows that there are eight binary cyclic codes of length 9. By Prop. 3, the cyclic codes of dimension 2 and 7 are Singleton as are the trivial ones (Table 2).

It remains to inspect the codes $\mathcal{C}_1 = \langle 1 + x^3 \rangle$ and $\mathcal{C}_2 = \langle 1 + x^3 + x^6 \rangle$. The codeword $(1, 0, 1, 0, 0, 0, 0, 0, 0)$ and its cyclic shift $(0, 1, 0, 1, 0, 0, 0, 0, 0)$ belong to $\mathcal{C}_1$. Their supports are disjoint and their Hamming weights sum up to 4 which is the Singleton bound. Thus by Prop. 2, the code $\mathcal{C}_1$ is not Singleton. A similar argument exhibits that the code $\mathcal{C}_2$ is not Singleton.

In Prop. 3 it has been shown that $d > \frac{1}{2}(n - k + 1)$ is a sufficient though not necessary condition for a code to be Singleton. The next result deals with the case $d \leq \frac{1}{2}(n - k + 1)$.

Proposition 16. *Let $\mathcal{C}$ be a binary cyclic $[n, k, d]$ code. If*

$$d \leq \frac{1}{2}(n - k + 1) \quad \text{and} \quad d - 1 < \frac{n - d}{k - 1}, \tag{9}$$

k	generator polynomial	d	S	Singleton?
1	$1 + x + x^2 + x^3 + x^4 + x^5 + x^6 + x^7 + x^8$	9	9	✓
2	$1 + x + x^3 + x^4 + x^6 + x^7$	6	8	✓
3	$1 + x^3 + x^6$	3	7	✗
6	$1 + x^3$	2	4	✗
7	$1 + x + x^2$	2	3	✓
8	$1 + x$	2	2	✓

Table 2: Parameters of the binary cyclic codes of length 9.

then $\mathcal{C}$ is not a Singleton code.

Proof. Suppose both inequalities hold.

Let $c \in \mathcal{C}$ be a codeword of minimal Hamming weight d. Since $\mathcal{C}$ is cyclic, any number of cyclic shifts of c yields another codeword in $\mathcal{C}$. Assume that c can be shifted by one position such that the resulting codeword c' has support disjoint from that of c. Then the codeword $c + c'$ would have Hamming weight $2d \leq n - k + 1$ and by Prop. 2 it would follow that the code $\mathcal{C}$ is not Singleton.

Claim that one cyclic shift of c yields a codeword c' whose support is disjoint from supp(c). To this end, denote by s the number of blocks of consecutive 1's, which is also the number of blocks of consecutive 0's, because every block of consecutive ones is followed by a block of consecutive zeros and vice versa. Here a number of consecutive ones or zeros at the beginning and at the end counts as a single block. The number s can be bounded as follows,

$$\frac{n - d}{k - 1} \leq s \leq d. \tag{10}$$

The upper bound is obvious. In order to obtain the lower bound, note that any block of consecutive zeros is of length at most $k - 1$. Because otherwise there would be a codeword corresponding to a polynomial of degree less than $n - k$, which would be smaller than the degree $n - k$ of the generator polynomial. It follows that $s \cdot (k - 1) \geq n - d$, since a codeword of minimum Hamming weight has $n - d$ zeros.

Combining the inequalities (9) and (10) gives $d - 1 < s \leq d$ and hence $s = d$. Thus the codeword c has d blocks of consecutive ones which are all of length 1.

Since c has Hamming weight d, it follows that each one-entry is followed by at least one zero-entry. This proves the claim and the result is established. $\square$

Example 4. Consider the binary cyclic codes of length 15 generated by $g(x) = 1 + x^5 + x^{10}$. This is an $[15, 5, 3]$ code and so, $\frac{n-d}{k-1} = \frac{12}{4} = 3 > 2 = d - 1$ and $d = 3 \le \frac{1}{2} 11 = \frac{1}{2}(15 - 5 + 1)$. Thus by Prop. 16 this code is not Singleton.

The previous examples and others exhibit that the binary cyclic codes up to length 19 are Singleton if the condition $d > \frac{1}{2}(n - k + 1)$ is satisfied. However, this condition is not necessary as will be shown next.

Example 5. Consider the binary cyclic code of length 21 generated by the polynomial
$$g(x) = 1 + x + x^2 + x^4 + x^5 + x^8 + x^9.$$

The dimension of this code is $k = 12$ and so the Singleton bound is $S = 10$. Since the polynomial
$$c(x) = 1 + x^{10} + x^{12} + x^{17} + x^{18} = g(x) \cdot (1 + x + x^3 + x^5 + x^9)$$

belongs to the code, the minimum distance is $d \le 5 = \frac{1}{2} \cdot S$. In fact, the weight distribution of the code is

i	0	5	6	7	8	9	10
A_i	1	21	168	360	210	280	1008

Note that the 21 cyclic shifts of the polynomial $c(x)$ provide all codewords of minimal Hamming weight. It follows that there are no codewords as in Prop. 2 and hence this code is Singleton.

References

[1] W. Adams, P. Loustaunau, *An Introduction to Gröbner Bases*, American Mathematical Society, 1994.

[2] T. Becker, V. Weispfenning, *Gröbner Bases – A Computational Approach to Commutative Algebra*, Springer, 1998.

[3] M. Borges-Quintana, M.A. Borges-Trenard, P. Fitzpatrick, E. Martinez-Moro, Gröbner bases and combinatorics for binary codes, *AAECC*, **19**, No. 5 (2008), 393-411.

[4] M. Borges-Quintana, M.A. Borges-Trenard, I. Marquez-Corbella, E. Martinez-Moro, An algebraic view to gradient descent decoding, *IEEE information theory workshop (itw)* (2010), 1-4.

[5] A.B. Cooper, Towards a new method of decoding algebraic codes using Gröbner bases, In: *Transactions 10-th Army Conf. Appl. Math. Comp.*, **93** (1992), 293-297.

[6] D. Cox, J. Little, D. O'Shea, *Ideals, Varieties, and Algorithms*, Springer, 1996.

[7] N. Dück, K.-H. Zimmermann, Universal Gröbner bases for binary linear codes, *International Journal of Pure and Applied Mathematics*, **86**, No. 2 (2013).

[8] F.J. MacWilliams, N.J.A. Sloane, *The Theory of Error-Correcting Codes*, North Holland, New York, 1977.

[9] T. Mora, L. Robbiano, The Gröbner Fan of an Ideal, *Journal of Symbolic Computation*, **6** (1988), 183-208.

[10] M. Saleemi, K.-H. Zimmermann, Gröbner bases for linear codes, *International Journal of Pure and Applied Mathematics*, **62** (2010), 481-491.

[11] M. Saleemi, K.-H. Zimmermann, Linear codes as binomial ideals, *International Journal of Pure and Applied Mathematics*, **61** (2010), 147-156.

[12] C. Shannon, A mathematical theory of communication, *Bell System Technical Journal*, **27** (1948), 379-423.

[13] B. Sturmfels, *Gröbner Bases and Convex Polytopes*, American Mathematical Society, 1996.

[14] V. Weispfenning, Constructing universal Gröbner bases, In: *Applied Algebra, Algebraic Algorithms and Error-Correcting Codes (AAECC-5)*, Volume 356 of *Lecture Notes in Computer Science* (1987), 408-417.