Ho Chi Minh University of Science HCMUS

Update: April 18, 2009

Finite fields: some applications

Michel Waldschmidt [4]

Fourth course

April 15, 2009

4 Cyclic codes

4.1 Definitions

A *cyclic code* $\mathcal{C}$ of length n over an alphabet with q elements is a $\mathbf{F}_q$–vector subspace of $\mathbf{F}_q^n$ such that, for any $(a_1, a_2, \dots, a_{n-1}, a_n) \in \mathcal{C}$, the element $(a_n, a_1, a_2, \dots, a_{n-1})$ also belongs to $\mathcal{C}$. We speak of a q-ary code as a reference to the number of elements of the alphabet; it is a binary code for $q = 2$, a ternary code for $q = 3$.

We denote by $T : \mathbf{F}_q^n \longrightarrow \mathbf{F}_q^n$ the linear map (*right shift*)

$$(a_1, a_2, \dots, a_{n-1}, a_n) \longmapsto (a_n, a_1, a_2, \dots, a_{n-1}).$$

In the group of automorphism of the $\mathbf{F}_q$–vector space $\mathbf{F}_q^n$, this element T satisfies $T^n = I$ (the unit of $\mathrm{Aut}_{\mathbf{F}_q}(\mathbf{F}_q^n)$, namely the identity map). This is how the polynomial $X^n - 1$ comes into the picture.

Assume $\gcd(n, q) = 1$. A natural basis of the $\mathbf{F}_q$–space $\mathbf{F}_q[X]/(X^n - 1)$ is given by the classes modulo $X^n - 1$ of $1, X, \dots, X^{n-1}$. This gives a $\mathbf{F}_q$–isomorphism

$$\begin{array}{cccc} \Psi : & \mathbf{F}_q^n & \longrightarrow & \mathbf{F}_q[X]/(X^n - 1) \\ & (a_0, a_1, \dots, a_{n-1}) & \longmapsto & a_0 + a_1 X + \cdots + a_{n-1} X^{n-1}. \end{array}$$

Rewrite the definition of T with the indices $\{0, \dots, n-1\}$ in place of $\{1, \dots, n\}$:

$$T(a_0, a_1, \dots, a_{n-1}) = (a_{n-1}, a_0, \dots, a_{n-2});$$

hence

$$\Psi \circ T(a_0, a_1, \dots, a_{n-1}) = X(a_0 + a_1 X + \cdots + a_{n-1} X^{n-1}) \pmod{X^n - 1}.$$

[4] This text is accessible on the author's web site http://www.math.jussieu.fr/~miw/coursVietnam2009.html

As a consequence, a subset $\mathcal{C}$ of $\mathbf{F}_q^n$ is stable under the shift T if and only if $\Psi(\mathcal{C})$ is stable under multiplication by X in $\mathbf{F}_q[X]/(X^n-1)$.

A vector subspace $\mathcal{I}$ of $\mathbf{F}_q[X]/(X^n-1)$ is stable under multiplication by X if and only if $\mathcal{I}$ is an ideal of the quotient ring $\mathbf{F}_q[X]/(X^n-1)$. Furthermore, there is a bijective map between the ideals of $\mathbf{F}_q[X]/(X^n-1)$ and the ideals of $\mathbf{F}_q[X]$ which contain X^n-1. Since the ring $\mathbf{F}_q[X]$ is principal, the ideals containing X^n-1 are the ideals (Q) generated by a divisor Q of X^n-1. Given such an ideal, there is a single generator Q which is monic. If d is the degree of Q, then the ideal of $\mathbf{F}_q[X]/(X^n-1)$ generated by the class of Q modulo X^n-1 is a $\mathbf{F}_q$–vector space of dimension $r=n-d$: a basis is $Q, XQ, \dots, X^{r-1}Q$. Also the following sequence of $\mathbf{F}_q$–linear maps is exact:

$$0 \longrightarrow \frac{(Q)}{(X^n-1)} \longrightarrow \frac{\mathbf{F}_q[X]}{(X^n-1)} \longrightarrow \frac{\mathbf{F}_q[X]}{(Q)} \longrightarrow 0.$$

The dimensions of these three vector spaces are r, n and d with $n=r+d$, as it should. Combining these results with Proposition 36, we deduce

Proposition 39. *Given a finite field $\mathbf{F}_q$ and an integer n with $\gcd(n,q)=1$, there are bijective maps between the following subsets.*
(i) The codes $\mathcal{C}$ of length n over $\mathbf{F}_q$.
(ii) The ideals $\mathcal{I}$ of $\mathbf{F}_q[X]/(X^n-1)$.
(iii) The monic divisors Q of X^n-1 in $\mathbf{F}_q[X]$.
(iv) The subsets I of $\mathbf{Z}/n\mathbf{Z}$ which are stable under multiplication by q.
Under this correspondence, the dimension d of the code is the dimension of the $\mathbf{F}_q$–vector space $\mathcal{I}$, the degree of Q is $r=n-d$, and the number of elements in I is also r.

The trivial code $\{0\}$ of length n and dimension 0 corresponds to the ideal (0) of $\mathbf{F}_q[X]/(X^n-1)$, to the divisor X^n-1 of X^n-1 and to the empty subset of $\mathbf{Z}/n\mathbf{Z}$.

The full code $\mathbf{F}_q^n$ of length n and dimension n corresponds to the ideal (1) of $\mathbf{F}_q[X]/(X^n-1)$, to the divisor 1 of X^n-1 and to the set $I=\mathbf{Z}/n\mathbf{Z}$ itself.

The repetition code $\{(a,a,\dots,a)\ ;\ a\in\mathbf{F}_q\}\subset\mathbf{F}_q^n$ of length n and dimension 1 corresponds to the ideal $(1+X+\cdots+X^{n-1})$ of $\mathbf{F}_q[X]/(X^n-1)$, to the divisor $(X^n-1)/(X-1)$ of X^n-1 and to the set $I=(\mathbf{Z}/n\mathbf{Z})\setminus\{0\}$.

The hyperplane of equation $x_1+\cdots+x_n=0$ in $\mathbf{F}_q^n$ is a parity bit check code of length n and dimension $n-1$. It corresponds to the ideal $(X-1)$ of $\mathbf{F}_q[X]/(X^n-1)$, to the divisor $X-1$ of X^n-1 and to the subset $I=\{0\}$ of $\mathbf{Z}/n\mathbf{Z}$.

4.2 Hamming codes

From http://en.wikipedia.org/wiki/Hamming_code

> *In telecommunication, a Hamming code is a linear error-correcting code named after its inventor, Richard Hamming. Hamming codes can detect up to two simultaneous bit errors, and correct single-bit errors; thus, reliable communication is possible when the Hamming distance between the transmitted and received bit patterns is less than or equal to one. By contrast, the simple parity code cannot correct errors, and can only detect an odd number of errors.*
>
> *Hamming worked at Bell Labs in the 1940s on the Bell Model V computer, an electromechanical relay-based machine with cycle times in seconds. Input was fed in on punch cards, which would invariably have read errors. During weekdays, special code would find errors and flash lights so the operators could correct the problem. During after-hours periods and on weekends, when there were no operators, the machine simply moved on to the next job.*
>
> *Hamming worked on weekends, and grew increasingly frustrated with having to restart his programs from scratch due to the unreliability of the card reader. Over the next few years he worked on the problem of error-correction, developing an increasingly powerful array of algorithms. In 1950 he published what is now known as Hamming Code, which remains in use in some applications today.*

Let $\mathbf{F}_q$ be a finite field with q elements and let r be a positive integer. Define

$$n = \frac{q^r - 1}{q - 1} = 1 + q + q^2 + \cdots + q^{r-1}.$$

Therefore q is prime to n and the class of q in $(\mathbf{Z}/n\mathbf{Z})^\times$ has order r. The subset $I = \{1, q, q^2, \ldots, q^{r-1}\}$ of $\mathbf{Z}/n\mathbf{Z}$ is stable under multiplication by q. This defines a code of length n and dimension $d = n - r$ over $\mathbf{F}_q$.

We first develop the special case already considered in example 27, where $r = 3$, $q = 2$, hence $n = 7$ and $d = 4$. We have seen in example 29 that the decomposition of Φ_7 over $\mathbf{F}_2$ is

$$\Phi_7(X) = (X^3 + X + 1)(X^3 + X^2 + 1).$$

We choose $Q(X) = 1 + X + X^3$. The vector of its coordinates in the basis $1, X, X^2, X^3, X^4, X^5, X^6$ is $e_0 = (1,1,0,1,0,0,0) \in \mathbf{F}_2^7$. Next define e_1, e_2 and e_3 by taking the coordinates in the same basis of XQ, X^2Q, X^3Q:

$$\begin{array}{rl}
Q(X) = 1 + X + X^3 & e_0 = (1,1,0,1,0,0,0),\\
XQ(X) = X + X^2 + X^4, & e_1 = (0,1,1,0,1,0,0) = Te_0,\\
X^2Q(X) = X^2 + X^3 + X^5, & e_2 = (0,0,1,1,0,1,0) = Te_1,\\
X^3Q(X) = X^3 + X^4 + X^6, & e_3 = (0,0,0,1,1,0,1) = Te_2.
\end{array}$$

The components of e_0, e_1, e_2, e_3 in $\mathbf{F}_2^7$ are the rows of the following matrix

$$G = \begin{pmatrix} 1 & 1 & 0 & 1 & 0 & 0 & 0\\ 0 & 1 & 1 & 0 & 1 & 0 & 0\\ 0 & 0 & 1 & 1 & 0 & 1 & 0\\ 0 & 0 & 0 & 1 & 1 & 0 & 1 \end{pmatrix}.$$

The elements in the code $\mathcal{C}$ are the 16 elements

$$m_0e_0 + m_1e_1 + m_2e_2 + m_3e_3$$

with $(m_0, m_1, m_2, m_3) \in \mathbf{F}_2^4$. This subspace $\mathcal{C}$ of $\mathbf{F}_2^7$ has dimension 4, hence it is an intersection of 3 hyperplanes. Let us recall how to find a basis of the $\mathbf{F}_q$–vector space of linear forms vanishing on a subspace V of F^n given by a basis with d elements. We write the $d \times n$ matrix whose rows are the coordinates of the given basis. We add one further row with the variables $x_1, \dots, x_n$. By elementary columns operations (replacing a column by its sum with a linear combination of the other columns, which corresponds to the multiplication on the right by a regular $n \times n$ matrix), we get a matrix of the form

$$\begin{pmatrix} & I_d & & & 0 & \dots & 0\\ y_1 & y_2 & \dots & y_d & y_{d+1} & \cdots & y_n \end{pmatrix}$$

where I_d is the identity $d \times d$ matrix and $y_1, \dots, y_n$ are linearly independent linear forms in $x_1, \dots, x_n$. Then the $(n-d)$–tuple $y_{d+1}, \dots, y_n$ is a basis of the space of linear forms vanishing on V. This can be checked by reducing to the simple case of a hyperplane $x_n = t_1x_1 + \cdots + t_{n-1}x_{n-1}$ with $d = n-1$ and the matrix

$$\begin{pmatrix} & & & & t_1\\ & & I_{n-1} & & \vdots\\ & & & & t_{n-1}\\ x_1 & x_2 & \dots & x_{n-1} & x_n \end{pmatrix}$$

We perform this process with the matrix G: therefore we introduce

$$\begin{pmatrix} 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 1 \\ x_0 & x_1 & x_2 & x_3 & x_4 & x_5 & x_6 \end{pmatrix}.$$

Here is the last row of the successive matrices obtained by the triangulation process (we work over $\mathbf{F}_2$)

x_0	x_1	x_2	x_3	x_4	x_5	x_6
x_0	$x_1 + x_0$	x_2	$x_3 + x_0$	$x_4 + x_0 + x_1$	x_5	x_6
x_0	$x_1 + x_0$	$x_2 + x_0 + x_1$	$x_3 + x_0$	$x_4 + x_0 + x_1$	x_5	x_6
x_0	$x_1 + x_0$	$x_2 + x_0 + x_1$	$x_3 + x_1 + x_2$	$x_4 + x_0 + x_1$	$x_5 + x_0 + x_1 + x_2$	x_6
x_0	$x_1 + x_0$	$x_2 + x_0 + x_1$	$x_3 + x_1 + x_2$	$x_4 + x_0 + x_2 + x_3$	$x_5 + x_0 + x_1 + x_2$	x_6
x_0	$x_1 + x_0$	$x_2 + x_0 + x_1$	$x_3 + x_1 + x_2$	$x_4 + x_0 + x_2 + x_3$	$x_5 + x_0 + x_1 + x_2$	$x_6 + x_1 + x_2 + x_3$

Therefore we introduce the three linear forms

$$\begin{aligned} L_0(\underline{x}) &= x_0 + x_2 + x_3 + x_4 \\ L_1(\underline{x}) &= x_0 + x_1 + x_2 + x_5 \\ L_2(\underline{x}) &= x_1 + x_2 + x_3 + x_6. \end{aligned}$$

$$H = \begin{pmatrix} 1 & 0 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 1 \end{pmatrix} \tag{40}$$

The 7 column vectors are all the non–zero elements in $\mathbf{F}_2^3$. The product $G \cdot {}^tH$ of G with the transpose of H is the zero 4×3 matrix.

The same construction can be performed in the general case of $\mathbf{F}_q^n$ with $n = (q^r - 1)/(q - 1)$. In $\mathbf{F}_q^r$, there are $q^r - 1$ non–zero elements, each of them defines a line ($\mathbf{F}_q$–subspace of dimension 1) having $q - 1$ non–zero elements, and therefore there are n lines: on each of them we select one element. We take for H the $r \times n$ matrix whose columns are the coordinates of these elements. Any two rows of H are linearly independent over $\mathbf{F}_q$. The intersection of the r hyperplanes $\mathbf{F}_q^n$ defined by the rows of H is a code which is *Hamming code of length n and dimension $d = n - r$ over $\mathbf{F}_q$.* The corresponding subset I of $\mathbf{Z}/n\mathbf{Z}$ is $\{1, q, q^2, \ldots, q^{r-1}\}$. Let ζ be a primitive n–th root of unity. Given a message $(m_r, \ldots, m_{n-1}) \in \mathbf{F}_q^d$, one computes $(m_0, \ldots, m_{r-1}) \in \mathbf{F}_q^r$, so that

$$m_0 + m_1\zeta + \cdots + m_{r-1}\zeta^{r-1} = -m_r\zeta^r - \cdots - m_{n-1}\zeta^{n-1}$$

and the associated codeword is $\underline{c} = (m_0, \dots, m_{n-1}) \in \mathbf{F}_q^n$. For $\underline{x} \in \mathbf{F}_q^n$, we have

$$\underline{x} = (x_0, \dots, x_n) \in \mathcal{C} \quad \text{if and only if} \quad \sum_{i=0}^{n-1} x_i \zeta^i = 0.$$

If this sum is nonzero and if there exists $\underline{c} \in \mathcal{C}$ with $d(\underline{x}, \underline{c}) \leq 1$, then the error $\epsilon = \underline{x} - \underline{c} = (0, \dots, 0, \epsilon_k, 0, \dots, 0) \in \mathbf{F}_q^n$ has its nonzero component in position k with

$$\epsilon_k \zeta^k = -\sum_{i=0}^{n-1} x_i \zeta^i.$$

4.3 Generator matrix and check matrix

Among many others, a reference for this section is [2], Chapter 3.

Given a linear code $\mathcal{C}$ of dimension d and length n over $\mathbf{F}_q$, a *generator matrix* is a $d \times n$ matrix G with coefficients in $\mathbf{F}_q$, the rows of which are the components of a basis of $\mathcal{C}$. The code is the set of elements $\underline{m}G$ where $\underline{m}$ ranges over $\mathbf{F}_q^d$ (viewed as a $1 \times d$ row vector). From the definition it follows that G has rank d.

A *check matrix* is a $(n-d) \times n$ matrix H with coefficients in $\mathbf{F}_q$, the rows of which are the components of a basis of the space of linear forms vanishing on $\mathcal{C}$. The code $\mathcal{C}$ is the set of elements $\underline{c}$ in $\mathbf{F}_q^n$ such that $H \cdot {}^t\underline{c} = 0$, where ^{t} denotes the transposition, so that ${}^t\underline{c}$ is a $n \times 1$ column vector in $\mathbf{F}_q^n$. Therefore

$$G \cdot {}^tH = 0$$

where G is a $d \times n$ matrix of rank d and H a $r \times n$ matrix of rank $r = n - d$.

The code is said to be *in systematic form* if $H = \begin{pmatrix} A & I_r \end{pmatrix}$, where I_r is the identity $r \times r$ matrix and A is a $r \times d$ matrix. .

Two codes are *isomorphic* if they have the same check matrix in suitable bases - for instance the two descriptions that we gave of the Hamming code of length 7 and dimension 4 in example 27 and § 4.2 are isomorphic but not identical.

5 Error correcting codes: further definitions.

From http://en.wikipedia.org/wiki/Coding_theory

> *Coding theory is an approach to various science disciplines – such as information theory, electrical engineering, digital communication, mathematics, and computer science – which helps*

> *design efficient and reliable data transmission methods so that redundancy can be removed and errors corrected.*
>
> *Channel encoding adds extra data bits to make the transmission of data more robust to disturbances present on the transmission channel.*

Definitions of error detection and error correction:
Error detection is the ability to detect the presence of errors caused by noise or other impairments during transmission from the transmitter to the receiver.
Error correction is the additional ability to reconstruct the original, error–free data.

The *Hamming distance* on the set $\mathbf{F}_q^n$ is

$$d(\underline{x}, \underline{y}) = \#\{i\,; 1 \le i \le n,\ x_i \ne y_i\}$$

for $\underline{x} = (x_1, \dots, x_n)$ and $\underline{y} = (y_1, \dots, y_n)$. It satisfies, as it should with the name *distance* (see for instance [1], Prop. 10.D),

$$d(\underline{x}, \underline{y}) = 0 \Longleftrightarrow \underline{x} = \underline{y}$$

and

$$d(\underline{y}, \underline{x}) = d(\underline{x}, \underline{y})$$

for $\underline{x}$ and $\underline{y}$ in $\mathbf{F}_q^n$, as well as the triangle inequality for $\underline{x}$, $\underline{y}$ and $\underline{z}$ in $\mathbf{F}_q^n$,

$$d(\underline{x}, \underline{z}) \le d(\underline{x}, \underline{y}) + d(\underline{y}, \underline{z}).$$

We define the *minimum distance* $d(\mathcal{C})$ of a code $\mathcal{C} \subset \mathbf{F}_q^n$ by

$$d(\mathcal{C}) = \min\{d(\underline{x}, \underline{y})\ ;\ \underline{x}, \underline{y} \in \mathcal{C},\ \underline{x} \ne \underline{y}\}.$$

The *Hamming weight* $w(\underline{x})$ of an element of $\mathbf{F}_q^n$ is its Hamming distance with 0: for $\underline{x} = (x_1, \dots, x_n)$:

$$w(\underline{x}) = \#\{i\ ;\ 1 \le i \le n,\ x_i \ne 0\}.$$

Hence, for $\underline{x}$ and $\underline{y}$ in $\mathbf{F}_q^n$,

$$d(\underline{x}, \underline{y}) = w(\underline{x} - \underline{y}).$$

For a linear code, $d(\mathcal{C})$ is the minimal weight of a non–zero element in $\mathcal{C}$.

For t a non–negative integer, the *Hamming ball* $B(\underline{c},t)$ of center $\underline{c} \in \mathbf{F}_q^n$ and radius t is the set of elements of $\mathbf{F}_q^n$ having Hamming distance to $\underline{c}$ at most t:

$$B(\underline{c},t) = \{\underline{x} \in \mathbf{F}_q^n \ ; \ d(\underline{x},\underline{c}) \le t\}.$$

The number of elements in $B(\underline{c},t)$ is 1 for $t=0$, it is $1+n(q-1)$ for $t=1$, and more generally

$$\#B(\underline{c},t) = 1 + \binom{n}{1}(q-1) + \cdots + \binom{n}{t}(q-1)^t \qquad \text{for } t \ge 0. \tag{41}$$

A *transmission with at most* t *errors* is a mapping $f : \mathcal{C} \longrightarrow \mathbf{F}_q^n$ such that for all $\underline{c} \in \mathcal{C}$,

$$d\bigl(f(\underline{c}),\underline{c}\bigr) \le t.$$

The *error* is $\epsilon(\underline{c}) = f(\underline{c}) - \underline{c}$. The message which is sent is $\underline{c}$, a codeword, the message which is received is $f(\underline{c})$.

The first question is to detect if an error occurred, that means to detect whether $\epsilon(\underline{c})$ is zero ot not. A code $\mathcal{C}$ *can detect* t *errors* if for all $\underline{c} \in \mathcal{C}$,

$$B(\underline{c},t) \cap \mathcal{C} = \{\underline{c}\}.$$

This means that for a transmission $f : \mathcal{C} \longrightarrow \mathbf{F}_q^n$ with at most t errors, $f(\underline{c}) \in \mathcal{C}$ if and only if $\epsilon(\underline{c}) = 0$. The receiver checks whether $f(\underline{c})$ is in $\mathcal{C}$ or not (for instance by using a check matrix H). If $f(\underline{c}) \in \mathcal{C}$, if the code is t–error detecting and if the transmission had at most t errors, then $\epsilon(\underline{c}) = 0$: there was no error.

A code $\mathcal{C}$ of length n over $\mathbf{F}_q$ *can correct* t *errors* (one also says that it is t*–error correcting*) if for all $\underline{x} \in \mathbf{F}_q^n$,

$$\#B(\underline{x},t) \cap \mathcal{C} \le 1.$$

This means that any transmission $f : \mathcal{C} \longrightarrow \mathbf{F}_q^n$ with at most t errors is injective: for all $\underline{y} \in f(\mathcal{C})$ there is a single $\underline{c}$ such that $\underline{y} = f(\underline{c})$. After receiving $\underline{y} = f(\underline{c})$, knowing that the transmission had at most t errors, the receiver computes the unique $\underline{c}$ for which $d(y;\underline{c}) \le t$. Then he knows that $f(\underline{c}) = \underline{y}$ and he also knows the error $\epsilon(\underline{c}) = f(\underline{c}) - \underline{y}$.

Lemma 42. *A code* $\mathcal{C}$ *of length* n *over* $\mathbf{F}_q$ *can detect* t *errors if and only if* $d(\mathcal{C}) \ge t+1$. *The code* $\mathcal{C}$ *can correct* t *errors if and only if* $d(\mathcal{C}) \ge 2t+1$.

Proof. The condition $d(\mathcal{C}) \ge t+1$ means that a message at Hamming distance at most t from an element $\underline{c}$ of $\mathcal{C}$ and distinct from $\underline{c}$ does not belong to $\mathcal{C}$. This is equivalent to saying that $\mathcal{C}$ can detect t errors.

For the second part of the lemma, assume first that $d(\mathcal{C}) \geq 2t+1$. Let $\underline{x} \in \mathbf{F}_q^n$ and let $\underline{c}_1$ and $\underline{c}_2$ in $\mathcal{C}$ satisfy $d(\underline{x}_1, \underline{c}_1) \leq t$ and $d(\underline{x}_2, \underline{c}_2) \leq t$. Then by the triangle inequality

$$d(\underline{c}_1, \underline{c}_2) \leq 2t < d(\mathcal{C}).$$

Therefore $\underline{c}_1 = \underline{c}_2$.
Conversely, assume $d(\mathcal{C}) \leq 2t$: there is a non–zero element $\underline{c}$ in $\mathcal{C}$ with $w(\underline{c}) \leq 2t$, hence $\underline{c}$ has at most $2t$ non–zero components. Split the set of indices of the non–zero components of $\underline{c}$ into two disjoint subsets I_1 and I_2 having each at most t elements. Next define $\underline{x} \in \mathbf{F}_q^n$ as the point having the same components x_i as $\underline{c}$ for $i \in I_1$ and 0 for i not in I_1. Then in the Hamming ball of center $\underline{x}$ and radius t there are at least two points of $\mathcal{C}$, namely 0 and $\underline{c}$. Hence $\mathcal{C}$ is not t–error correcting.

□

Proposition 43. *For a code $\mathcal{C}$ of length n and dimension d, the minimum distance is bounded by*

$$d(\mathcal{C}) \leq n+1-d.$$

Proof. The subspace

$$V = \{(x_1, \dots, x_{n+1-d}, 0, \dots, 0)\ ;\ (x_1, \dots, x_{n+1-d}) \in \mathbf{F}_q^{n+1-d}\}$$

of $\mathbf{F}_q^n$ has dimension $n+1-d$, the sum of this dimension with the dimension d of $\mathcal{C}$ exceeds the dimension n of the ambient space $\mathbf{F}_q^n$, hence there is a non–zero element in the intersection. This is a non–zero element of $\mathcal{C}$ with weight $\leq n+1-d$. □

A code $\mathcal{C}$ of length n and dimension d for which $d(\mathcal{C}) = n+1-d$ is called MDS (*Maximal Distance Separable*).

Hamming code of length 7 and dimension 4 has minimum distance 3, hence is not MDS.

From (41) we deduce Hamming's bound on the error correcting capacity of a code of length n and dimension r over $\mathbf{F}_q$ (see [2] Theorem 3.3.1).

Theorem 44. *For a linear code $\mathcal{C}$ in $\mathbf{F}_q^n$ of dimension r which is t–error correcting,*

$$1 + \binom{n}{1}(q-1) + \cdots + \binom{n}{t}(q-1)^t \leq q^{n-r}.$$

A t–error correcting code over $\mathbf{F}_q$ of length n is *perfect* if $\mathbf{F}_q^n$ is the disjoint union of the balls of radius t around the codewords in $\mathcal{C}$.

For a perfect 1–error correcting code over $\mathbf{F}_q$ of length n and dimension d, the union of the q^d Hamming balls of radius 1 gives a packing of the set $\mathbf{F}_q^n$ with q^n elements, hence

$$q^d\bigl(1+n(q-1)\bigr)=q^n.$$

We set $d=n-r$, so that $n=(q^r-1)/(q-1)$. As we have observed, for these parameters the polynomial Φ_n splits into irreducible factors of degree r. Each of these factors gives a cyclic code which is Hamming q-ary code of length n and dimension d.

For instance take $q=2$. For $r=2$ we have $n=3$, $d=1$ and this is the repetition code $\{(0,0,0)\ ,\ (1,1,1)\}$. For $r=3$ we have $n=7$, $d=4$ which are the parameters of Hamming code considered in example 27 and § 4.2.

The binary Golay code of length 23, dimension 12

A perfect code with $q=2$, $n=23$, $d=12$ and minimal distance 7 (hence it is 3–error correcting but not MDS) has been constructed by Golay as follows.

We have $2^{11}-1=23\times 89=2047$, which is the smallest integer of the form $M_p=2^p-1$ with p prime but which is not itself a prime (primes of the form $M_p=2^p-1$ are called *Mersenne primes*). We take the subset I of $(\mathbf{Z}/23\mathbf{Z})^\times$ generated by 2, which is

$$I=\{1,2,3,4,6,8,9,12,13,16,18\}.$$

The decomposition of Φ_{23} over $\mathbf{F}_2$ has been given in exercise 32.

There are 2^{12} codewords, for each of them the Hamming ball of radius 3 has

$$\binom{23}{0}+\binom{23}{1}+\binom{23}{2}+\binom{23}{3}=1+23+253+1771=2048=2^{11}$$

elements, these balls are disjoint and the total number of elements in their union is $2^{21}2^{12}=2^{23}$.

The ternary Golay code of length 11, dimension 6

An other perfect code constructed by Golay has the parameters $q=3$, $n=11$, $d=6$ and minimal distance 5 (it is 2–error correcting not MDS). We have $3^5-1=11\times 23$. We take the subset I of $(\mathbf{Z}/11\mathbf{Z})^\times$ generated by 3, which is $I=\{1,3,4,5,9\}$. The decomposition of Φ_{11} over $\mathbf{F}_3$ has been given in exercise 31.

There are 3^6 codewords, for each of them the Hamming ball of radius 2 has

$$\binom{11}{0} + 2\binom{11}{1} + 2^2\binom{11}{2} = 1 + 22 + 220 = 243 = 3^5$$

elements, they are disjoint the total number of elements in $\mathbf{F}_3^{11}$ is $3^6 3^5 = 3^{11}$.

BCH (Bose–Chaudhuri–Hocenghem) codes

Given a finite field $\mathbf{F}_q$ and an integer r, let n be a divisor of $q^r - 1$. Hence the order of q modulo n divides r. Let $\zeta \in \mathbf{F}_q^r$ be a primitive n-th root of unity and let $\delta \geq 2$ be an integer. Consider the morphism of rings

$$\begin{array}{ccc} \mathbf{F}_q[X]/(X^n - 1) & \longrightarrow & \mathbf{F}_q^{\delta-1} \\ P & \longmapsto & \bigl(P(\zeta^j)\bigr)_{1 \leq j \leq \delta-1} \end{array}$$

The kernel is a cyclic q–ary code of length n and minimal distance δ, the generating polynomial is the lcm of the minimal polynomials over $\mathbf{F}_q$ of the elements ζ^j, $1 \leq j \leq \delta - 1$: the subset I of $\mathbf{Z}/n\mathbf{Z}$ is the smallest subset containing $\{1, \ldots, q\}$ and stable under multiplication by q.

Reed–Solomon code

The Reed–Solomon codes are special cases of BCH codes. Let $q = 2^m$, $n = q - 1$ and let ζ be a primitive n–th–th root of unity, that means a generator of $\mathbf{F}_q^\times$. For $1 \leq d \leq n$ the code associated with the subset $I = \{1, 2, 3, \ldots, n - d\}$ of $\mathbf{Z}/n\mathbf{Z}$ and to the polynomial

$$\prod_{i=1}^{n-d} (X - \zeta^i)$$

has dimension d and minimal distance $q - d$. This code is MDS; it is used in CD's.

It is known that the only perfect codes are

- *The trivial code with a single element* 0.
- *The full code* $\mathbf{F}_q^n$.
- *A binary repetition code with odd length (see [2] Exercise 3.12).*
- *For* $r \geq 2$*, the* q*-ary Hamming code of length* $n = (q^r - 1)/(q - 1)$*, dimension* $n - r$*, and minimal distance* 3.
- *The ternary Golay code over* $\mathbf{F}_3$ *of length* 11*, dimension* 6 *and minimal distance* 5.

- *The binary Golay code over $\mathbf{F}_2$ of length 23, dimension 12 and minimal distance 7.*

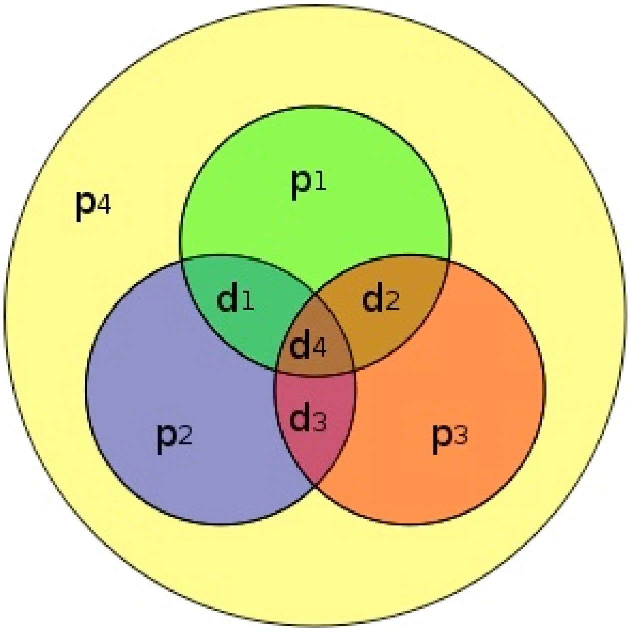
We state two results which are useful tools to compute the minimum distance of a code. For the first one, see [1], Prop. 11C.

Proposition 45. *Let $\mathcal{C}$ be a linear code over $\mathbf{F}_q$ of length n with check matrix H and let s be a positive integer. Then $\mathcal{C}$ has minimum distance $\geq s+1$ if and only if any s columns of H are linearly independent over $\mathbf{F}_q$.*

As a consequence, if any s columns of H are linearly independent over $\mathbf{F}_q$, and if further there exists $s+1$ columns of H which are linearly dependent over $\mathbf{F}_q$, then $d(\mathcal{C}) = s+1$. This enables one to check that Hamming code has minimum distance 3. Indeed in the matrix (40) all rows are non–zero and distinct (hence any two rows are linearly independent over $\mathbf{F}_2$), but there are sets of three rows which are linearly dependent. If we add a row with 1's, then for the new matrix any sum of an odd number of rows is non–zero, hence any three rows are linearly independent. This means that we extend the code of Hamming of lenth 7 to a code of length 8 by adding a parity check bit.

$$G = \begin{pmatrix} 1 & 1 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 0 & 1 & 1 \end{pmatrix}. \qquad H = \begin{pmatrix} 1 & 0 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}$$

This code has therefore minimum distance 4, it cannot correct more than one error, but it can detect up to 3 errors.



Hamming extended (8,4) code

To any code $\mathcal{C} \subset \mathbf{F}_q^n$ we can associate an *extended code* $\widetilde{\mathcal{C}} \subset \mathbf{F}_q^{n+1}$ by adding a parity bit:

$$\widetilde{\mathcal{C}} = \{(x_1, \dots, x_{n+1}) \in \mathcal{C} \times \mathbf{F}_q \; ; \; (x_1, \dots, x_n) \in \mathcal{C}, \; x_1 + \cdots + x_{n+1} = 0\} \subset \mathbf{F}_q^{n+1}.$$

One can check $d(\mathcal{C}) \leq d(\widetilde{\mathcal{C}}) \leq d(\mathcal{C})$.

A variant is to take the *even subcode*

$$\mathcal{C}' = \{(x_1, \dots, x_n) \in \mathcal{C} \; ; \; x_1 + \cdots + x_n = 0\} \subset \mathbf{F}_q^n.$$

Then $d(\mathcal{C}) \leq d(\mathcal{C}')$.

Proposition 46. *Let $\mathcal{C}$ be a cyclic linear code of length n over $\mathbf{F}_q$ associated with a subset I of $\mathbf{Z}/n\mathbf{Z}$ stable under multiplication by q. Assume that there exist i and s such that $\{i+1, i+2, \dots, i+s\} \subset I$. Then $d(\mathcal{C}) \geq s+1$.*

For instance Hamming code is associated with the subset $I = \{1, 2, 4, \dots, 2^{r-1}\}$ of $\mathbf{Z}/n\mathbf{Z}$, with two consecutive elements, hence its distance is at least 3 (and here it is just 3).

5.1 Some historical dates

Among important dates are the following

• 1949: Marcel Golay (specialist of radars): produced two remarkably efficient codes.

• 1950: Richard W. Hamming, *Error detecting and error correcting codes*, The Bell System Technical Journal **26** (April 1950), N° 2, 147–160.

• 1955: Convolutional codes.

• 1959: Bose Chaudhuri Hocquenghem codes (BCH codes).

• 1960: Reed Solomon codes.

• 1963 John Leech uses Golay's ideas for sphere packing in dimension 24 - classification of finite simple groups

• 1971: no other perfect code than the two found by Golay.

• 1970: Goppa codes.

• 1981: Algebraic geometry codes.