Coding Theory

January 20, 2011
INTRODUCTION

LINEAR CODES

DECODING

BOUNDS

CYCLIC CODES

FURTHER EXAMPLES
The channel coding scheme consists of mainly three parts. A message $s$ is encoded by the function $E : Q^k \rightarrow Q^n$ where $Q$ denotes the alphabet which is a finite set of symbols and $k$ is the length of the message. The codeword $c = E(s)$ of length $n$ is sent through a noisy channel and $r$ is the received message. The receiver uses the decoder $D : Q^n \rightarrow Q^k$ to infer the message sent, where $D(r) = D(E(s)) = \hat{s}$. If $\hat{s} = s$, then the message is transmitted successfully.

The design of the encoding and decoding functions is the basic tool of coding theory, where various mathematical tools are used. In particular, algebraic coding theory is said to be the analysis of the linear block codes.
INTRODUCTION

LINEAR CODES
  Generator and Parity Check Matrices
  Graphical Representation

DECODING

BOUNDS

CYCLIC CODES

FURTHER EXAMPLES
Linear Block Codes

Consider the linear map \( E : \mathbb{F}_q^k \to \mathbb{F}_q^n \), where \( \mathbb{F}_q \) is a finite field with \( q \) elements, \( n \geq k \), so that \( C(\lambda s_1 + \mu s_2) = \lambda C(s_1) + \mu C(s_2) \), \( \forall \lambda, \mu \in \mathbb{F}_q \) and \( \forall s_1, s_2 \in \mathbb{F}_q^k \).

Say the message \( s = (s_0, s_1, \ldots, s_{k-1}) \in \mathbb{F}_q^k \) is mapped to \( E(s) = E(s_0, s_1, \ldots, s_{k-1}) = (c_0, c_1, \ldots, c_{n-1}) = c \in \mathbb{F}_q^n \).

The **linear code** \( C \) is all those vectors \( c = E(s) \), which are called **codewords**, i.e. \( C = E(\mathbb{F}_q^k) \).
Generator Matrix

The set of codewords is a subspace of $\mathbb{F}_q^n$ of dimension $k$. So, a linear coding map can be represented by a $k \times n$ matrix $G$, whose rows is a basis for $C$. $G$ is called as the generator matrix for $C$. In systematic form it’s written as $G = [I_k \ | \ P]$, where $I_k$ is the identity matrix of order $k$ and $P$ is a $k \times (n - k)$ matrix.

Example

Hamming Code (7, 4) over $\mathbb{F}_2$

$$G = \begin{pmatrix}
1 & 0 & 0 & 0 & 1 & 0 & 1 \\
0 & 1 & 0 & 0 & 0 & 1 & 1 \\
0 & 0 & 1 & 0 & 1 & 1 & 1 \\
0 & 0 & 0 & 1 & 1 & 1 & 0
\end{pmatrix}$$
The encoding is \( c = s \cdot G \mod q \), which means that a k-bit \( q \)-ary message \( s \) is encoded as the n-bit \( q \)-ary vector \( c \).

In systematic form \( c = s \cdot G = s \cdot [I_k \mid P] = [s \mid s \cdot P] \). Observe that every message \( s \) is encoded as an n-tuple where the first \( k \) components is \( s \) itself and the remaining \( n - k \) symbols come from \( s \cdot P \). These extra symbols actually characterizes the encoding rules and they are called **redundancy** symbols.

If \( C \) has a \( k \times n \) generator matrix \( G \) (i.e. \( E : \mathbb{F}_q^k \to \mathbb{F}_q^n \)), then the **code length** (block length) of \( C \) is \( n \) and the dimension of the code is \( k \).
For a linear code whose generator matrix is given in a systematic form $G = [I_k \mid P]$, we can introduce the corresponding parity check matrix as $H = [-P^T \mid I_{n-k}]$ with $HG^T = 0 \mod q$. Since $\forall c \in C, \ c = s \cdot G \mod q$ for some $s \in \mathbb{F}_q^k$, it follows that $H \cdot c = 0 \mod q$ if and only if $c$ is a codeword.

Example

Hamming Code $(7, 4)$ over $\mathbb{F}_2$

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

We have $C = \text{span}(G) = \text{ker}(H)$. 
Examples

Example
Repetition Code
In this code every codeword is of the form \((a, a, ..., a)\) where \(a \in \mathbb{F}_q\). Thus, we have a \([n, 1]\)-code with parity-check matrix 
\[ H = [-P^T | I_{n-k}] \]

Example
Parity Check Code
Let \(q = 2\) and the encoding as \((c_1, ..., c_n) \mapsto (c_1, ..., c_n, \sum_{i=1}^n c_i)\). Then this code is a binary linear \([n + 1, n]\)-code with parity-check matrix 
\[
H = \begin{pmatrix}
1 & 1 & 1 & \ldots & 1
\end{pmatrix}
\]
Minimal Distance

The *Hamming distance* between two codewords $c_1 = (c_{11}, \ldots, c_{1n})$, $c_2 = (c_{21}, \ldots, c_{2n})$ is $d(c_1, c_2) = \{ \#i \mid c_{1i} \neq c_{2i}\}$. The **minimal distance** of a code $C$ is $d(C) = \min \{ d(c_i, c_j) \mid i \neq j \}$ for all $c_i, c_j \in C$.

The *Hamming weight* $w(c)$ of the codeword $c$ is the number its nonzero coordinates, i.e. $w(c) = \{ \#i \mid c_i \neq 0, c = (c_1, \ldots, c_n) \}$

A linear code $C$ with a $k \times n$ generator matrix $G$ and minimal distance $d(C) = d$ is called an $[n, k, d]$-Code.
**Dual Code**

For a linear code $C \in \mathbb{F}_q^n$ we define its **dual code** as

$$C^\perp = \{ c' \in \mathbb{F}_q^n : c \cdot c' = 0 \text{ for all } c \in C \}.$$

If $C$ is an $[n, k, d]$-code, then $C^\perp$ is an $[n, n-k, d']$-code. Moreover, $C \oplus C^\perp = \mathbb{F}_q^n$

If $G$ is a generator matrix of $C$, then $G$ is a parity check matrix for $C^\perp$. Conversely, if $H$ is a parity check matrix of $C$, then $H$ is a generator matrix for $C^\perp$.

$C$ is called self-dual, if $C = C^\perp$. Observe that in this case $C$ is an $[n, \frac{n}{2}, d]$-code, where $n$ is even.
The **code rate** of $C$ is the ratio of the dimension of the code to the length of the codeword, namely $\log_q |C|/n$. For the linear binary codes we define the rate as $R = k/n$. So, it states what portion of the total amount of information that is useful. In general, $R \leq 1$. 
Tanner Graph

A graph $G = (U, V, E)$ is defined to be a bipartite graph where the vertex set consists of two disjoint sets $U$ and $V$ such that every edge connects a vertex in $U$ to another in $V$.

**Definition**

A *Tanner graph* of a code with parity check matrix $H$ is a bipartite graph such that there exists one vertex in $U$ for each row of $H$ and one vertex in $V$ for each column of $H$ and there is an edge between two vertices $i$ and $j$ exactly when $H_{ij} = 1$. 
Example

The creation of such a graph is straightforward, consider Hamming Code (7, 4)

\[
\begin{pmatrix}
1 & 1 & 1 & 0 & 1 & 0 & 0 \\
0 & 1 & 1 & 1 & 0 & 1 & 0 \\
1 & 0 & 1 & 1 & 0 & 0 & 1
\end{pmatrix}
\]

Its Tanner graph is given as follows
INTRODUCTION

LINEAR CODES

DECODING

Discrete Memoryless Channels
Minimum Distance Decoding
Hard Decision

BOUNDS

CYCLIC CODES

FURTHER EXAMPLES
Definition

We define a *discrete channel* to be a system consisting of two finite sets, namely input alphabet $\Sigma$ and output alphabet $\Theta$ and a probability transition matrix $P$ expressing the probability of observing the output $Y$ given that the input $X$ is sent. The channel is *memoryless* if the probability distribution of the output depends only on the input at that time and is conditionally independent of previous channel inputs or outputs.

Let $X,Y$ be a discrete random variables with alphabets $\Sigma, \Theta$ and probability mass functions $p(x) = Pr(X = x), \ x \in \Sigma$, $p(y) = Pr(Y = y), \ y \in \Theta$ respectively. Then the probability transition matrix $P$ consists of the values $Pr(Y = y_j|X = x_i)$ where $x_i \in \Sigma, \ y_j \in \Theta$. 
The most famous examples are the binary symmetric channel (BSC) and the binary erasure channel (BEC). Across the BSC, the input symbols are flipped with probability $p$. The received message bits through BEC are erased with probability $\epsilon$.

**Example**

A binary symmetric channel with parameter $p$ is a binary DMC with $Pr(Y = 0|X = 1) = Pr(Y = 1|X = 0) = p$ and $Pr(Y = 0|X = 0) = Pr(Y = 1|X = 1) = 1 - p$ for all $i = 1, \ldots, n$. 

\[
\begin{pmatrix}
0 & 1-p \\
p & p \\
1-p & 1-p
\end{pmatrix}
\]

\[
P = 
\begin{pmatrix}
p & 1-p \\
1-p & p
\end{pmatrix}
\]
Example

A binary erasure channel with parameter $\epsilon$ is a binary DMC with $Pr(Y_i = \text{erased}|X_i = 1) = Pr(Y_i = \text{erased}|X_i = 0) = \epsilon$ and $Pr(Y_i = 1|X_i = 1) = Pr(Y_i = 0|X_i = 0) = 1 - \epsilon$ for all $i = 1, \ldots, n$.

\[
\begin{pmatrix}
0 & 1 - \epsilon & 0 & \epsilon \\
\epsilon & 0 & 1 - \epsilon & \epsilon \\
\end{pmatrix}
\]
If the codewords are uniformly distributed and sent through a BSC with crossover probability $p$, then the **block error probability** $P_e$ of the code is estimated as:

$$P_e(C) = \sum_{c \in C} \Pr(c) \cdot \Pr\{MDD(c + e) \neq c\}$$

$$\leq 1 - \Pr\{e \in MDD_c\}$$

$$\leq 1 - \Pr\{|e| \leq t\}$$

$$= \sum_{i=t+1}^{n} \Pr\{|e| = i\} \quad \text{where } t = \left\lfloor \frac{d - 1}{2} \right\rfloor$$

$$= \sum_{i=t+1}^{n} \binom{n}{i} p^i (1 - p)^{n-i}$$
This description is enough to set up our decoding problem: By considering the linear codes, $\Sigma = \Theta = \mathbb{F}_q^n$, where $\tilde{c}$ is the input and $\tilde{r}$ is the output.

The probability transition matrix have the entries $Pr(r_i|c_i)$, $i = 1, \ldots, n$. The receiver makes an error, if $D(\tilde{r}) = \hat{s} \neq \tilde{s}$.

Instead of tracing back to $\tilde{s}$, the receiver can reduce the decoding scheme to find $\tilde{c} = C(\tilde{s})$, in that case, the difference $\tilde{e} = \tilde{r} - \tilde{c}$ is said to be the error.

The error correction strategy tries to find the most probable codeword $\tilde{c}$ by fixing the corrupted bits in $\tilde{r}$.
For the received message $r$, **minimum error decoding (MED)** is the strategy to find $\tilde{c} \in C$ minimizing the probability of error, or equivalently to maximize $\Pr(\tilde{c} \text{ sent} \mid \tilde{r} \text{ received})$ i.e. pick the codeword $\tilde{c}$ that is most likely to be received as $\tilde{r}$.

The probability of error is said to be the expected probability that a codeword is decoded or corrected wrongly. In general, it is not easy to derive this. MED is an optimal correction which maximizes $\Pr(\tilde{s} \mid \tilde{r}, C)$ but without knowledge of the probability distribution of codewords we are not able to implement this method. A clearer setup could be achieved by the following:
For the received message $r$, maximum likelihood decoding (MLD) is the strategy to find $\bar{c} \in C$ maximizing $Pr(r \text{ received} \mid c \text{ sent})$

Note that if all codewords are equally likely to be sent, then this scheme is equivalent to minimum error decoding:

$$Pr(r \text{ received} \mid c \text{ sent}) = \frac{Pr(r \text{ received}, c \text{ sent})}{Pr(c \text{ sent})} = Pr(c \text{ sent} \mid r \text{ received}) \cdot \frac{Pr(c \text{ received})}{Pr(c \text{ sent})} = Pr(c \text{ sent} \mid r \text{ received})$$
Notice that $Pr(r \text{ received } | c \text{ sent}) \approx Pr(\text{Error } e = r - c \text{ is met}).$
This allows us to derive a more reasonable and natural model:

For the received message $r$, \textbf{minimal distance decoding (MDD)} finds $c \in C$ to minimize the Hamming distance:

$$d(r, c) = \#\{ i : c_i \neq r_i \} = w(e)$$

So, for every bit of the codeword $c$ we say an error is occurred if there are bits in the received message $r$ which differ from the original codeword. The probability of such a corruption could be easily observed given that a bit, say 0, is sent and 1 is received with probability $p$. 
Across the binary symmetric channel, MDD is equivalent to MLD.

The message $r$ is received when the codeword $c$ is sent only if the error $e = r - c$ occurs. Let $w(e) = t$, then the probability of error $e$ is $p_t = p^t(1 - p)^{n-t}$ across the BSC. So we have the following order: $p_0 > p_1 > \ldots > p_t > \ldots > p_n$ which implies $Pr(r \text{ received} \mid c \text{ sent})$ is maximized when $d(r, c) = w(e)$ is minimized.

So we have the following relation: $\text{ME} \xrightarrow{i.i.d.} \text{ML} \xrightarrow{\text{BSC}} \text{MD}$
In case of linear codes, the binary message $s$ is encoded as $s \cdot G = c$ and the message $r$ is received. If $s$ is uniformly distributed and $e$ is independent of $s$, then the optimal decoder is equivalent to the MLD and we have $r = s \cdot G + e$. Furthermore, by using the parity check matrix we get $Hr = Hc + He = He = z$. The decoding is therefore reduced to find the most probable $e$ satisfying this equality. The MLD for the problem is to find $e$ which maximizes $Pr(e|H, z)$. The vector $z$ is said to be the syndrome of the received message and this decoding task is known as syndrome decoding.
**Theorem**

A linear \([n, k, d]\)-code \(C\) can **detect** up to \(d - 1\) errors and **correct** up to \(t = \left\lfloor \frac{d-1}{2} \right\rfloor\) errors made by the channel, under the MDD.

**Proof.**

Let \(c, c' \in C\) with \(d(C) = d\). For \(w(e) \leq \frac{d-1}{2}\) we have
\[
d \leq d(c, c') \leq d(c, c + e) + d(c + e, c') \Rightarrow \\
d(c + e, c') \geq d - d(c, c + e) = \frac{d+1}{2} \Rightarrow c\text{ is the closest codeword to } c + e.\]

Hence MDD can correct an error \(w(e) \leq \frac{d-1}{2}\).

Now let \(w(e') = \frac{d+1}{2}\), then
\[
d \leq d(c, c') \leq d(c, c + e') + d(c + e', c') \Rightarrow \\
d(c + e', c') \geq d - d(c, c + e') = \frac{d-1}{2}.\]

Since every codeword is packed by disjoint spheres of diameter \(d - 1\), any error having at most \(d - 1\) distance to a codeword can be detected. \(\square\)
The algorithm which returns the unique closest codeword $c$ which lies in the $\left\lfloor \frac{d-1}{2} \right\rfloor$ neighborhood of the received message $r$ is called bounded minimum distance decoding or in short, bounded distance decoding (BDD). So, linear codes are able to correct an error $|e| \leq t$. 
Theorem

A linear code $C$ of length $n$ with parity-check matrix $H$ has a minimum distance $d(C) \geq d + 1$ if and only if any $d$ columns of $H$ are linearly independent.

Proof.

Suppose there are $d$ linearly dependent columns of $H$, then $Hc = 0$ implies $w(c) \leq d$ hence $d(C) \leq d$. Conversely, if any $d$ columns are linearly independent, then there is no codeword $c$ with $w(c) \leq d$, hence $d(C) \geq d + 1$. 

□
Message Passing Algorithms

We can describe a general class of decoding algorithms, which are called **message passing algorithms** (MPA), and are iterative. The reason for their name is that at each round of the iteration, messages are passed from message nodes to check nodes, and from check nodes back to message nodes. The messages from message nodes to check nodes are computed based on the observed value of the message node and some of the messages passed from the neighboring check nodes to that message node.
Hard Decision Algorithm

Consider the code with the parity check matrix

\[
\begin{pmatrix}
1 & 1 & 0 & 1 & 0 & 0 \\
0 & 1 & 1 & 0 & 1 & 0 \\
1 & 0 & 0 & 0 & 1 & 1 \\
0 & 0 & 1 & 1 & 0 & 1
\end{pmatrix}
\]

Its corresponding Tanner graph is given by:
Example

Initialization

Set $v_i = y_i$, $i = 1, \ldots, 6$
Check Messages

Now every check node calculates a value to every single bit due to the other received bits of the equation.
**Example**

**Bit update**

The message bits are updated so that the decision is to be made due to the majority of the check messages and its initial value.

![Diagram showing bit update process](image-url)
Example

Test

After the bit update the signal values are sent again to be checked if they satisfy the parity equations.

In our case, all the parity check equations are valid after one iteration, so the algorithm decodes \( c = [001011] \) where the received message is corrected without requiring a consideration of all possible codewords.
Bad Example

As a second decoding scheme, let the parity check matrix is given as

\[
\begin{pmatrix}
1 & 1 & 0 & 1 & 0 & 0 \\
1 & 1 & 0 & 0 & 1 & 0 \\
0 & 0 & 1 & 0 & 1 & 1 \\
0 & 0 & 1 & 1 & 0 & 1
\end{pmatrix}
\]

and suppose \( y = [101001] \) is sent through the same BSC channel. A valid codeword is \( c = [001001] \)
First Iteration

Initialization
Set $v_i = y_i$, $i = 1, ..., 6$

Check Messages 1
Check nodes calculate a value to every single bit due to the other received bits.
First Iteration

Bit update 1

The message bits are updated due to the majority decision.

Test

The new signal values are sent again to be checked if they satisfy the parity equations.
Check Messages 2
Now the check equations try the updated bits

Bit update 2
So, the two check equations are still unsatisfied and the algorithm returns again the initial codeword, which would continue as a loop.

At this moment, the second test would yield the same result as in the initial one and we see that the codeword is not corrected with hard-decision algorithm.
A 4-cycle in the Tanner graph corresponds a 2x2-submatrix of 1’s in the parity check equation.

\[
\begin{pmatrix}
1 & 1 & 0 & 1 & 0 & 0 \\
1 & 1 & 0 & 0 & 1 & 0 \\
0 & 0 & 1 & 0 & 1 & 1 \\
0 & 0 & 1 & 1 & 0 & 1 \\
\end{pmatrix}
\]

**Note:** A bipartite graph does not contain any odd-length cycles.
The Channel Coding Theorem

**Shannon 1948**

All rates below capacity $C$ are achievable. Specifically, for every rate $R \leq C$, there exists a sequence of codes with maximum probability of error 0. Conversely, any sequence of codes with $p \to 0$ must have $R \leq C$. 
**Good Codes vs. Very Good Codes**

**Very good codes**

Given a channel, a family of block codes that achieve arbitrarily small probability of error at any communication rate up to the capacity of the channel are called ‘very good’ codes for that channel.

**Good codes**

Code families that achieve arbitrarily small probability of error at non-zero communication rates up to some maximum rate that may be less than the *capacity* of the given channel.

**Bad codes**

Families that cannot achieve arbitrarily small probability of error, or that can achieve arbitrarily small probability of error only by decreasing the information rate to zero. Repetition codes are an example of a bad code family. (Bad codes are not necessarily useless for practical purposes.)

**Practical codes**

Code families that can be encoded and decoded in time and space polynomial in the block length.
Gilbert-Varshamov Bound

There exists a linear $[n, k]$-code over $\mathbb{F}_q$ with minimum distance $\geq d$ whenever
$$q^{n-k} \geq \sum_{i=0}^{d-2} \binom{n}{i} (q-1)^i.$$ 

Let $A_q(n, d)$ be the maximum possible size of a code $C$ over $\mathbb{F}_q$ with length $n$ and with minimum distance $d$, then:
$$A_q(n, d) \geq \frac{q^n}{\sum_{i=0}^{d-1} \binom{n}{i} (q-1)^i}.$$ 

For any $x \in \mathbb{F}_q^n$ we have $d(c, x) \leq d - 1 \ \forall c \in C$, since otherwise we would contradict the maximality of $A_q(n, d)$. Hence
$$\mathbb{F}_q^n = \bigcup_{c \in C} B_q(c, d - 1)$$
$$\Rightarrow |\mathbb{F}_q^n| \leq \sum_{c \in C} |B_q(c, d - 1)|,$$ 

where there are
$$\sum_{i=0}^{d-1} \binom{n}{i} (q-1)^i$$

vectors in each ball.
Gilbert-Varshamov conjecture

This distance, $d_{GV}$, is known as the *Gilbert-Varshamov distance* for rate $R$ and block length $n$.

The **Gilbert-Varshamov conjecture**, widely believed, asserts that (for large $n$) it is not possible to create binary codes with minimum distance significantly greater than $d_{GV}$.

The *Gilbert-Varshamov rate* $R_{GV}$ is the maximum rate at which you can reliably communicate with a bounded-distance decoder, assuming that the Gilbert-Varshamov conjecture is true.

The codes which attain GV-bound are called **optimal**.
Sphere Packing Bound/ Hamming Bound

For a linear \([n, k]\)-code over \(\mathbb{F}_q\) we have

\[
\sum_{i=0}^{t} \binom{n}{i} (q - 1)^i \leq q^{n-k}
\]

where \(t = \left\lfloor \frac{d-1}{2} \right\rfloor\).

The proof is trivial, the number of vectors of length \(n\) with weight \(m\) is \(\binom{n}{m}(q - 1)^m\). The balls of radius \(t\) centered at the codewords are disjoint and contain \(\sum_{i=0}^{t} \binom{n}{i} (q - 1)^i\) vectors and there are \(q^k\) such balls. The codes achieving Sphere Packing Bound are called perfect.
Singleton Bound

For a linear \([n, k, d]\)-code over \(\mathbb{F}_q\) we have \(d \leq n - k + 1\).

Consider the case that we take first \(k - 1\) components from every codeword. Since the dimension of the code is \(k\), there are \(q^k\) distinct codewords. Therefore by pigeonhole principle at least two codewords should meet in their first \(k - 1\) components. But then they should differ in at most the remaining \(n - (k - 1)\) components. The code had minimum distance \(d\), hence we have \(d \leq n - k + 1\). The codes achieving Singleton Bound are called Maximum Distance Separable (MDS) codes.
Plotkin Bound

For a linear \([n, k, d]\)-code over \(\mathbb{F}_q\) we have \(d \leq \frac{nq^{k-1}(q-1)}{q^k-1}\).

Let \(1 \leq i \leq n\) s.t. there is a codeword with \(i\)'th component nonzero. Since \(C\) has dimension \(k\) (there are \(q^k\) codewords), the subspace of \(C\) consisting of codewords with \(i\)'th component zero has dimension \(k-1\) (there are \(q^{k-1}\) such codewords). By counting over \(i\), each of those codewords have weight \(n(q-1)\), so the total weight is \(n(q-1)q^{k-1}\). But the minimum distance is the minimum nonzero weight, therefore the total weight is at least \(d(q^k-1)\).
A linear code $C \in \mathbb{F}_q^n$ is called **cyclic** if for any $c = (c_0, \ldots, c_{n-1}) \in C$, the shifted codeword $c' = (c_{n-1}, c_0, \ldots, c_{n-2}) \in C$, too.

Clearly, Repetition code, Parity Check code and Hamming $(7, 4, 3)$ are cyclic.
For $c = (c_0, ..., c_{n-1}) \in C$, the codeword polynomial of $c$ is defined as $c(x) = c_0 + c_1 x + ... + c_{n-1} x^{n-1} \in \mathbb{F}_q[x]/(x^n - 1)$. The set of all codeword polynomials is $C(x) = \{c(x) \in \mathbb{F}_q[x]/(x^n - 1) : c \in C\}$.

Using codeword polynomials, we can characterize cyclic codes as follows:
A linear code $C \in \mathbb{F}_q^n$ is cyclic if $c(x) \in C(x) \Rightarrow xc(x) \in C(x)$
Theorem

A linear code $C \in \mathbb{F}_q^n$ is cyclic if and only if $C(x)$ is an ideal of $\mathbb{F}_q[x]/(x^n - 1)$.

Proof.

Suppose $C(x)$ is an ideal of $\mathbb{F}_q[x]/(x^n - 1)$, then clearly for any $c(x) \in C(x)$ we have $x \cdot c(x) \in C(x)$.

Conversely, assume that $C$ is cyclic. Take $f(x) \in \mathbb{F}_q[x]/(x^n - 1)$ and consider

$$f(x) \cdot c(x) = f_0 \cdot c(x) + f_1 x \cdot c(x) + ... + f_{n-1} x^{n-1} \cdot c(x) \in C(x)$$

since $C$ is cyclic, we have $x^i \cdot c(x) \in C(x)$ for $0 \leq i \leq n - 1$. Also, $c(x) + c'(x) \in C(x)$ for $c(x), c'(x) \in C(x)$, since $C$ is linear.

Hence, $C(x)$ is an ideal of $\mathbb{F}_q[x]/(x^n - 1)$. \qed
Theorem

Let $C \in \mathbb{F}_q^n$ be a linear cyclic. There exists an uniquely determined monic polynomial $g(x) \in C(x)$ of minimum degree such that

1. $(g(x)) = C(x)$
2. $g(x) | x^n - 1$

Proof.

1. Obvious, since $\mathbb{F}_q[x]/(x^n - 1)$ is a PID.
2. $x^n - 1 = g(x)h(x) + r(x) \implies r(x) = -g(x)h(x) \mod x^n - 1 \implies r(x) \in (g(x))$. But then $r(x) = 0$ since $\deg r(x) < \deg g(x)$. Therefore $g(x) | x^n - 1$. 
This theorem gives us the opportunity to define the cyclic codes of length \( n \) using monic polynomials of degree \( \leq n - 1 \), which divide \( x^n - 1 \).

Such a polynomial \( g(x) \) of minimum degree, described as in the theorem is called the \textbf{generator polynomial} of the cyclic code \( C \), where \( (g(x)) = C(x) \) in \( \mathbb{F}_q[x]/(x^n - 1) \).

That is, \( c \in C \Rightarrow c(x) = g(x)f(x) \) for some \( f(x) \in \mathbb{F}_q[x]/(x^n - 1) \)
i.e. every codeword \( c \in C \) is generated by the coefficients of the polynomial \( g(x) \), it is a linear combination of the coefficient vector of \( g(x) \) and all its cyclic shifts up to the n’th order.
Corollary Let \( (g(x)) = C(x) \) be the generator polynomial of the cyclic code \( C \) and \( g(x)h(x) = x^n - 1 \). Then \( c \in C \) if and only if \( c(x)h(x) = 0 \mod x^n - 1 \).

Such a polynomial \( h(x) \) is said to be the \textbf{parity check polynomial} of the cyclic code \( C \).
Let \((g(x)) = C(x)\) be the generator polynomial of the cyclic code \(C\) where \(g(x) = g_0 + g_1x + ... + g_{r-1}x^{r-1} + x^r\). Then the corresponding generator matrix for the code \(C\) is:

\[
G = \begin{pmatrix}
g_0 & g_1 & g_2 & \cdots & 1 & 0 & 0 & 0 & 0 & \cdots & 0 \\
0 & g_0 & g_1 & \cdots & g_{r-1} & 1 & 0 & 0 & \cdots & 0 \\
0 & 0 & g_0 & \cdots & g_{r-2} & g_{r-1} & 1 & 0 & \cdots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & 0 & \cdots & g_0 & g_1 & g_2 & \cdots & 1 & 0 \\
0 & 0 & 0 & \cdots & 0 & g_0 & g_1 & g_2 & \cdots & 1
\end{pmatrix}
\]
Parity Check Matrix

Let $h(x)$ be the parity check polynomial of the cyclic code $C$ where $h(x) = h_0 + h_1x + \ldots + h_{k-1}x^{k-1} + x^k$ and $k = n - r$.

\[
\begin{pmatrix}
1 & h_{k-1} & h_{k-2} & \ldots & h_0 & 0 & 0 & 0 & \ldots & 0 \\
0 & 1 & h_{k-1} & \ldots & h_1 & h_0 & 0 & 0 & \ldots & 0 \\
0 & 0 & 1 & \ldots & h_2 & h_1 & h_0 & 0 & \ldots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & 0 & \ldots & h_{k-1} & h_{k-2} & h_{k-3} & \ldots & h_0 & 0 \\
0 & 0 & 0 & \ldots & 0 & h_{k-1} & h_{k-2} & h_{k-3} & \ldots & h_0
\end{pmatrix}
\]

Observe that the matrix construction is the same as above, but with respect to the reciprocal polynomial of $h(x)$, namely $h^{-1}(x) = x^k \cdot h(x^{-1}) = 1 + h_{k-1}x + \ldots + h_1x^{k-1} + h_0x^k$. 
Binary Hamming code

Let \( n = 2^k - 1 \), where \( k \geq 2 \). The binary Hamming Code \( H_2(n, k) \) has the parity check matrix \( H \) such that the columns of \( H \) are the binary representations of integers \( 1, 2, \ldots, 2^k - 1 \).

Hamming Code \( H_2(7, 3) \)

\[
H = \begin{pmatrix}
0 & 0 & 0 & 1 & 1 & 1 & 1 \\
0 & 1 & 1 & 0 & 0 & 1 & 1 \\
1 & 0 & 1 & 0 & 1 & 0 & 1
\end{pmatrix}
\]

Rearrangement into systematic form:

\[
H = \begin{pmatrix}
1 & 1 & 1 & 0 & 1 & 0 & 0 \\
0 & 1 & 1 & 1 & 0 & 1 & 0 \\
1 & 0 & 1 & 1 & 0 & 0 & 1
\end{pmatrix}
\]
Binary Hamming code

Any two columns of $H$ are linearly independent, but the sum of any two columns gives another column, hence the minimum distance is $d = 3$. Therefore $H_2(n, k)$ is an $[n, n - k, 3]$-Code where $n = 2^k - 1$.

**Theorem**

A binary cyclic code of length $n = 2^k - 1$ whose generator polynomial is a primitive polynomial over $\mathbb{F}_{2^k}$ is equivalent to the binary Hamming code $H_2(n, k)$. 
Binary Hamming code

**Proof.**

Let $\xi$ be a primitive element of $\mathbb{F}_{2^k}$ and 
$p(x) = (x - \xi)(x - \xi^2) \cdots (x - \xi^{2^{m-1}})$ be its minimal polynomial over $\mathbb{F}_2$. Consider the cyclic code $C$ generated by $p(x)$: the parity check matrix $H$ has $j$’th column $(c_0, c_1, \ldots, c_{m-1})^T \in \mathbb{F}_2^m$ if $\xi^{j-1} = c_0 + c_1\xi + \cdots + c_{m-1}\xi^{m-1}$ for $j = 1, 2, \ldots, 2^m - 1$. If $a = (a_0, a_1, \ldots, a_{n-1})$ and $a(x) = a_0 + a_1x + \cdots + a_{n-1}x^{n-1}$ then $H \cdot a^T = a(\xi)$ where $1, \xi, \ldots, \xi^{m-1}$ is a basis for $\mathbb{F}_2^m$. So, $H \cdot a^T = 0$ if and only if $a(\xi) = 0$ which means $p(x)|a(x)$. Since the columns of $H$ are permutations of binary representations of the elements $1, 2, \ldots, 2^k - 1$, $C$ is equivalent to $H_2(n, k)$.
Ternary Hamming code

Hamming Code $H_3(4, 2)$

$$H = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 2 \end{pmatrix}$$

Hamming Code $H_3(13, 3)$

$$H = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 2 & 2 & 2 \\ 1 & 0 & 1 & 2 & 0 & 1 & 2 & 0 & 1 & 2 & 0 & 1 & 2 \end{pmatrix}$$
q-ary Hamming code

The Hamming code $H_q(n, k)$ over $\mathbb{F}_q$ is an $(\frac{q^k-1}{q-1}, \frac{q^k-1}{q-1} - k, 3)$-code with parity check matrix $H$ whose columns are pairwise linearly independent.
q-ary Hamming code

Remarks

- Hamming Codes are single error correcting codes, since $d = 3$.
- They are perfect, i.e. they attain the Sphere-Packing bound (Hamming bound).

Check if $q^{n-k} \sum_{i=0}^{1} \binom{n}{i} (q-1)^i = q^n$ where $t = 1 = \left\lfloor \frac{d-1}{2} \right\rfloor$:

$$q^{n-k} \sum_{i=0}^{1} \binom{n}{i} (q-1)^i = q^k(1 + n(q - 1))$$

$$= q^{n-k}(1 + \frac{q^k - 1}{q - 1}(q - 1))$$

$$= q^{n-k}q^k = q^n$$
BCH codes

Let $b$ be a nonnegative integer, $\xi \in \mathbb{F}_{q^m}$ a primitive $n$’th root of unity, where $m = \text{ord}_n q$.

The BCH code over $\mathbb{F}_q$ of length $n$ and designed distance $d$, where $2 \leq d \leq n$, is the cyclic code defined by the roots $\xi^b, \xi^{b+1}, ..., \xi^{b+d-2}$ of the generator polynomial.

$$g(x) = \text{lcm}\{m^b(x), m^{b+1}(x), ..., m^{b+d-2}(x)\}$$

where $m^i(x)$ is the minimum polynomial of $\xi^i$ over $\mathbb{F}_q$.

If $b = 1$, then the code is called narrow-sense BCH code. If $n = q^m - 1$, then we have a primitive BCH code.
BCH codes

Theorem

The minimum distance of a BCH code of designed distance $d$ is at least $d$.

The BCH code is in the null space of the matrix:

$$
H = \begin{pmatrix}
1 & \xi^b & \xi^{2b} & \ldots & \xi^{(n-1)b} \\
1 & \xi^{b+1} & \xi^{2(b+1)} & \ldots & \xi^{(n-1)(b+1)} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
1 & \xi^{b+d-2} & \xi^{2(b+d-2)} & \ldots & \xi^{(n-1)(b+d-2)}
\end{pmatrix}
$$

To show: Any $d - 1$ columns of $H$ are linearly independent. (Use determinant)
Example

Consider the factorization of $x^{15} - 1$ over $\mathbb{F}_2$:

$x^{15} - 1 = (x+1)(x^2+x+1)(x^4+x+1)(x^4+x^3+1)(x^4+x^3+x^2+x+1)$

- $(x^4 + x + 1) \rightarrow$ Binary Hamming $H_2(15, 11)$
- $(x^4 + x + 1)(x^4 + x^3 + x^2 + x + 1) \rightarrow$ 2-error-correcting BCH
- $(x^2 + x + 1)(x^4 + x + 1)(x^4 + x^3 + x^2 + x + 1) \rightarrow$ 3-error-correcting BCH
- $(x+1)(x^2+x+1)(x^4+x+1)(x^4+x^3+1)(x^4+x^3+x^2+x+1) \rightarrow$ maximum length
Example

Take $m^1(x) = x^4 + x + 1$. We know that this polynomial is primitive, so $\xi$ being a root of $m^1(x)$, we have $<\xi> = \mathbb{F}_{2^4}^*$. Thus we have the parity check matrix:

$$
\begin{pmatrix}
1 & \xi & \xi^2 & \xi^3 & \xi^4 & \xi^5 & \xi^6 & \xi^7 & \xi^8 & \xi^9 & \xi^{10} & \xi^{11} & \xi^{12} & \xi^{13} & \xi^{14} \\
\end{pmatrix}
$$

Since $\mathbb{F}_{2^4}$ has dim 4 over $\mathbb{F}_2$, we can represent $\xi^i$'s as linear combinations of $\{1, \xi, \xi^2, \xi^3\}$. If we write the coefficients with respect to that basis to the columns of $H$, we get the following:

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

This example represents Binary Hamming $H_2(15, 11)$ Code as a narrow sense BCH Code: we have $\xi$ as root, so $b = 1$. Here we have set $m^1(x) = x^4 + x + 1 = g(x)$ as the generator polynomial, which is primitive, therefore $n = 15 = 2^4 - 1$ and $H_2(15, 11)$ is a primitive BCH code. Observe that $\xi^2$ is also a root, but $\xi^3$ not. So, $m^1(x)$ is characterized by $\xi, \xi^2 \Rightarrow d = 3. (\xi^b, \xi^{b+1}, ..., \xi^{b+d-2})$
Example

Now, again on the same example let $b = 1$ and take the designed distance to be $d = 4$. We’ve seen that $\xi, \xi^2$ are roots of $x^4 + x + 1$, but $\xi^3$ is a root of $x^4 + x^3 + x^2 + x + 1$.

So, the narrow sense BCH code with $d = 4$ is generated by $g(x) = (x^4 + x + 1)(x^4 + x^3 + x^2 + x + 1)$, where $m^1(x) = x^4 + x + 1$ and $m^3(x) = x^4 + x^3 + x^2 + x + 1$, they are relatively prime so their lcm is their product.

This is also the generator polynomial for $d = 5$, since $\xi^4$ is also a root of $x^4 + x + 1$. 
Example

The dimension is $15 - \deg(g(x)) = 7$. We replace the parity check matrix by:

$$H = \begin{pmatrix} 1 & \xi^2 & \xi^3 & \ldots & \xi^{14} \\ 1 & \xi^3 & \xi^6 & \ldots & \xi^{42} \end{pmatrix}$$

Again by representing the powers of $\xi$ and $\xi^3$ as coefficient vectors:

$$H = \begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 1 & 0 & 1 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 & 0 & 1 & 0 & 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 & 1 & 0 & 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 & 1 & 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 & 1 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 & 1 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & 1 & 1 \\ \end{pmatrix}$$
Reed-Solomon Codes

A BCH code over $\mathbb{F}_q$ of length $n = q - 1$ is called **Reed-Solomon Code**:

A Reed-Solomon Code (RS Code) is a primitive BCH code of length $n = q - 1$ over $\mathbb{F}_q$. The generator polynomial of such a code is in the form $g(x) = \prod_{i=0}^{d-1} (x - \xi^i)$ where $\xi$ is primitive in $\mathbb{F}_q$. ($<\xi> = \mathbb{F}^*_q$)

RS Codes have dimension $k = n - \deg(g(x)) = n - d + 1$. Recall that Singleton Bound asserts $d \leq n - k + 1$, hence RS Codes attain this bound. Such codes are called **Maximum Distance Separable (MDS)** Codes.
Remark

Since the code length of a Reed-Solomon Code is $n = q - 1$, there is no binary RS Code...
Let $g(x)$ be the generator polynomial of the Reed Solomon code $C$

where $g(x) = \prod_{i=0}^{d-1} (x - \xi^i) = g_0 + g_1 x + \ldots + g_{d-2} x^{d-2} + x^{d-1}$.

Then the corresponding generator matrix is:

$$G = \begin{pmatrix}
g_0 & g_1 & \cdots & g_{d-2} & 1 & 0 & 0 & 0 & \cdots & 0 \\
0 & g_0 & g_1 & \cdots & g_{d-2} & 1 & 0 & 0 & \cdots & 0 \\
0 & 0 & g_0 & g_1 & \cdots & g_{d-2} & 1 & 0 & \cdots & 0 \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & \cdots & g_0 & g_1 & \cdots & g_{d-2} & 1 & 0 \\
0 & 0 & 0 & \cdots & 0 & g_0 & g_1 & \cdots & g_{d-2} & 1
\end{pmatrix}$$
Classical Definition

Let $x_1, \ldots, x_n \in \mathbb{F}_q$ distinct elements and $f(x) \in \mathbb{F}_q[x]$ of degree less than $k$, where $1 \leq k \leq n \leq q - 1$. The Reed-Solomon code consists of the codewords as $n$-tuples of values obtained by evaluating every such $f(x)$ at each $x_i$, i.e. $C = \{(f(x_1), \ldots, f(x_n)) : f(x) \in \mathbb{F}_q[x], \deg(f(x)) < k\}$

$f(x) = a_0 + a_1x + \ldots + a_{k-1}x^{k-1}$ and $(x_1, \ldots, x_n) = (1, \xi, \xi^2, \ldots, \xi_{n-1})$ since RS Code is a primitive BCH code of length $n = q - 1$ over $\mathbb{F}_q$, where $<\xi> = \mathbb{F}_q^*$. We can represent the encoding as:

$$
\begin{pmatrix}
1 & 1 & 1 & 1 & 1 \\
1 & \xi & \xi_2 & \ldots & \xi_{n-1} \\
1 & \xi_2 & \xi_4 & \ldots & \xi_2(n-1) \\
\vdots & \vdots & \vdots & \ldots & \vdots \\
1 & \xi_{k-1} & \xi_2(k-1) & \ldots & \xi(n-1)(k-1)
\end{pmatrix}
$$
Equivalence of two Characterizations

\[ f = (f(x_1), \ldots, f(x_n)) = (f(1), f(\xi), \ldots, f(\xi^{n-1})) \]

where \( f(x) = a_0 + a_1 x + \ldots + a_{k-1} x^{k-1} \).

Put \( a = (a_0, a_1, \ldots, a_{k-1}) \)

\[ f_i = a(\xi^i) \]

\[ = a_0 + a_1 \xi^i + \ldots + a_{k-1} (\xi^i)^{k-1} \]

\[ = \sum_{j=0}^{k-1} a_j \xi^{ij} \]

Substitute for \( 1 \leq r \leq n - k \)

\[ f(\xi^r) = \sum_{i=0}^{n-1} f_i(\xi^r)^i \]

\[ = \sum_{i=0}^{n-1} \left( \sum_{j=0}^{k-1} a_j \xi^{ij} \right) \xi^{ri} \]

\[ = \sum_{j=0}^{k-1} a_j \sum_{i=0}^{n-1} (\xi^j \xi^r)^i \]

\[ = \sum_{j=0}^{k-1} a_j \sum_{i=0}^{n-1} \gamma_{jr}^i \]

where \( \gamma_{jr} = \xi^{j+r} \) and \( 1 \leq j + r \leq n - 1 \)

\[ \sum_{j=0}^{n-1} \gamma_{jr}^i = 1 + \gamma_{jr} + \gamma_{jr}^2 + \ldots + \gamma_{jr}^{n-1} = \frac{1 - \gamma_{jr}^n}{1 - \gamma_{jr}} = 0 \text{ since } n = q - 1. \]
Another Representation of Cyclic Codes

Let $m = \text{ord}_n p$, $p$: prime and $< \xi > = \mathbb{F}_q^*$ where $q = p^k$. Take another primitive element $\beta \in \mathbb{F}_q^*$. Then the set:

$$V = \{ c(\xi) := (\text{Tr}(\xi), \text{Tr}(\xi \beta), ..., \text{Tr}(\xi \beta^{n-1})) , \xi \in \mathbb{F}_q^* \}$$

is an $[n, k]$ cyclic code.
Reed-Muller Codes

\[ RM_q(k, m) = \{ \text{eval}(p(\alpha_1, \ldots, \alpha_m)) : p(x_1, \ldots, x_m) \in \mathbb{F}_q[x_1, \ldots, x_m], \deg(f(x)) < k \} \]

where \( \mathbb{F}_q[x_1, \ldots, x_m] \) is the polynomial ring of \( m \)-variate polynomials over \( \mathbb{F}_q \) of degree less than \( k \) and \( \text{eval}(p(\alpha_1, \ldots, \alpha_m)) \) is the values obtained by evaluating every \( p \) at each \((\alpha_1, \ldots, \alpha_m) \in \mathbb{F}_q^m\).

- \( RM_q(k, m) \) has length \( n = q^m \)
- \( p(x_1, \ldots, x_m) = \sum_{i_1+i_2+\ldots+i_k < k} m_{i_1,i_2,\ldots,i_k} \prod_{j=1}^k x_j^{i_j} \) with \( 0 \leq i_j \leq q - 1 \), and \( k \leq m(q - 1) \).
- \( RM_q(k, m) \) is a \([q^m, \binom{m+k}{m}, (1 - \frac{k}{q})q^m]\)-code.
- For \( k = 1 \) \( RM_q(k, 1) \) is equivalent to \( RS_q[n, k + 1] \) where \( n = q - 1 \).
\( \mathbb{F}_2[x_1, \ldots, x_m] \) consists of \textit{Boolean Polynomials} in \( m \) variables, which are linear combinations of \textit{Boolean Monomials} \( p = x_1^{r_1}x_2^{r_2}\cdots x_m^{r_m} \) where \( r_i \in \mathbb{N}, 1 \leq i \leq n \).

e.g. \( q = x_1 + x_3 + x_1x_2 + x_1x_2x_3 \) is a Boolean polynomial of degree 3, where \( x_1, x_3, x_1x_2, x_1x_2x_3 \) are monomials.

The set of all monomials in \( m \) variables over \( \mathbb{F}_2 \) is:
\[
M = \{1, x_1, x_2, x_m, x_1x_2, x_1x_3, \ldots, x_{m-1}x_m, x_1x_2x_3, \ldots, x_1x_2\cdots x_m\}.
\]
Consider $RM_2(2, 2)$. We will evaluate every polynomial in 2 variables $f(x_1, x_2) \in \mathbb{F}_2[x_1, x_2]$ at $(1, 1), (1, 0), (0, 1), (0, 0)$. We have $n = 2^2 = 4$, the codewords are of the form $f(1, 1), f(1, 0), f(0, 1), f(0, 0)$.

$M = \{1, x_1, x_2, x_1x_2\}$ is our basis for polynomials of degree $\leq 2$.

<table>
<thead>
<tr>
<th>Monomials</th>
<th>$(1, 1)$</th>
<th>$(1, 0)$</th>
<th>$(0, 1)$</th>
<th>$(0, 0)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p = 1$</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>$p = x_1$</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$p = x_2$</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>$p = x_1x_2$</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
Example

\(RM_2(2, 3) : n = 2^3\)
\(M = \{1, x_1, x_2, x_1x_2, x_1x_3, x_2x_3\}\) is the basis for polynomials of degree \(\leq 2\).
Evaluate at
\((1, 1, 1), (1, 1, 0), (1, 0, 1), (1, 0, 0), (0, 1, 1), (0, 1, 0), (0, 0, 1), (0, 0, 0)\)

\(G = \begin{pmatrix}
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\
1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\
1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\
1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{pmatrix}\)
Properties

- \( RM_2(k, m) \) has length \( n = 2^m \), dimension \( \sum_{i=1}^{k} \binom{m}{i} \),

- and minimum distance \( 2^{m-k} \),

- \( RM_2(k + 1, m + 1) = \{(f, f + g) : f \in RM_2(k + 1, m), g \in RM_2(k, m)\} \),

- the dual of \( RM_2(k, m) \) is \( RM_2(m - k - 1, m) \).

- \( RM_2(0, m) \) is repetition code \([2^m, 1, 2^m]\), \( RM_2(m - 1, m) \) is parity check code \([2^m, 2^m - 1, 2]\).

- \( RM_2(1, m) \) is called Hadamard code \([2^m, m + 1, 2^{m-1}]\).