1.1 Hamming Distance and Hamming Weight

Consider any set $F$. The Hamming distance $d_H(\vec{x}, \vec{y})$ between two vectors $\vec{x}, \vec{y} \in F^n$ is the number of positions where $\vec{x}, \vec{y}$ differ. That is, denoting $\vec{x} = (x_1, x_2, \ldots, x_n)$, $\vec{y} = (y_1, y_2, \ldots, y_n)$, we can write

$$d_H(\vec{x}, \vec{y}) = |\{i : i \in \{1, 2, \ldots, n\}, x_i \neq y_i\}|.$$

The Hamming distance is a metric on $F^n \times F^n$ since it satisfies the following properties for any $\vec{x}, \vec{y}, \vec{z} \in F^n$:

1. Non-negativity, $d_H(\vec{x}, \vec{y}) \geq 0$,
2. Symmetry $d_H(\vec{x}, \vec{y}) \geq 0$,
3. $d_H(\vec{x}, \vec{y}) = 0 \Leftrightarrow \vec{x} = \vec{y}$,
4. Triangle Inequality $d_H(\vec{x}, \vec{y}) + d_H(\vec{y}, \vec{z}) \geq d_H(\vec{x}, \vec{z})$.

For a set $F$ with a reference element 0, the Hamming weight of a vector $\vec{x}$ is the number of non-zero elements in $\vec{x}$, i.e., it is $d_H(\vec{x}, (0, 0, \ldots, 0))$.

1.2 The $t$-error and $t$-erasure channels

A channel is a mathematical model of data communication and data storage. A channel consists of an input alphabet $F$, and output alphabet $\Phi$, and a positive integer parameter $n$ known as the length of the channel. The channel consists of an input $\vec{x} \in F^n$ and an output $\vec{y} \in \Phi^n$. An adversarial channel can be specified by subsets $\Phi_{\vec{x}} \subseteq \Phi^n$, $\vec{x} \in F^n$; the set $\Phi_{\vec{x}}$ specifies the set of possible output vectors, when the input to the channel is $\vec{x}$. The channel can be viewed as the action of an adversary who observes the input $\vec{x}$ chooses an output $\vec{y}$ arbitrarily from the set $\Phi_{\vec{x}}$.

1.2.1 $t$-error channel

For any vector $\vec{s} \in F^n$, let $B(\vec{s}, r)$, be a ball of radius $r$ around $\vec{s}$. That is, $B(\vec{s}, r) = \{\vec{z} \in F^n : d_H(\vec{z}, \vec{s}) \leq r\}$. For instance, note that if $F = \{0, 1\}$ then the size of the ball is $|B(\vec{s}, r)| = \sum_{i=0}^{t} \binom{n}{i}$. A $t$-error channel over a finite alphabet $F$ is specified as, $\Phi = F$ and $\Phi_{\vec{x}} = B(\vec{x}, t)$. 

1-1
1.2.2 \( t \)-erasure channel

A \( t \)-erasure channel has \( F = \{0, 1\} \), and \( \Phi = \{0, 1, \epsilon\} \), where \( \epsilon \) is the erasure symbol. Let \( S \) be the set of sequences in \( \Phi^n \) with at most \( t \) erasures, i.e., \( S = \{(y_1, y_2, \ldots, y_n): |\{i: y_i = \epsilon\}| \leq t\} \).

\[
\Phi(x_1, x_2, \ldots, x_n) = \{(y_1, y_2, \ldots, y_n): y_i \in \{x_i, \epsilon\}, i = 1, 2, \ldots n \} \cap S.
\]

1.3 An \( (n, k) \) error correcting code.

For a given channel, an \( (n, k) \) and a decoder \( D \). 1.3 of generality, assume that \( x \) rate are called codewords. The \( n, k, d \) Consider an \( (n, k, d) \) code with \( \Phi \nabla \) of the code is defined to be \( \frac{k}{n} \) and the minimum distance \( d_{\min} \) of the code is defined to be the Hamming distance between the two closest codewords. An error correcting code with minimum distance \( d \) is sometimes referred to as an \( (n, k, d) \) code.

We say that an error correcting code \( C \) corrects errors over an adversarial channel if there exists a decoder \( D \) such that \( D(\hat{y}) = \hat{x} \) for all \( \hat{y} \in \Phi_\hat{x}, \hat{x} \in C \). Equivalently, an error correcting code \( C \) corrects errors if \( \hat{x}, \hat{x}' \in C \Rightarrow \Phi_{\hat{x}} \cap \Phi_{\hat{x}'} = \{\} \). For the \( t \)-error channel, we say that an error correcting code \( C \) detects errors if \( \hat{x}' \notin \Phi_{\hat{x}} \) for every \( \hat{x}, \hat{x}' \in C \).

**Theorem 1.1** An \( (n, k, d) \) code \( C \) corrects errors over a \( t \)-error channel if and only if \( d \geq 2t - 1 \).

**Proof:** Consider an \( (n, k, d) \) code with \( d \geq 2t + 1 \). Let \( \hat{y} \in \Phi_{\hat{x}}, \hat{x} \in C \). For \( \hat{x}' \neq \hat{x}, \hat{x}' \in C \), we have from the triangle inequality,

\[
\begin{align*}
   d_H(\hat{x}, \hat{y}) + d_H(\hat{x}', \hat{y}) &\geq d_H(\hat{x}, \hat{x}') \\
   \Rightarrow d_H(\hat{x}, \hat{y}) &\geq d_H(\hat{x}, \hat{x}') - d_H(\hat{x}', \hat{y}) \\
   \Rightarrow d_H(\hat{x}', \hat{y}) &\geq 2t + 1 - t = t + 1 \\
   \Rightarrow \hat{y} &\notin \Phi_{\hat{x}'}
\end{align*}
\]

Therefore, the code corrects \( t \) errors.

Conversely, consider an \( (n, k, d) \) code with \( d \leq 2t + 1 \). Then, there exist two codewords \( \bar{x} = (x_1, x_2, \ldots, x_n) \) an \( \bar{x}' = (x'_1, x'_2, \ldots, x'_n) \) such that \( d_H(\bar{x}, \bar{x}') \leq 2t \). Without loss of generality, assume that \( \bar{x}, \bar{x}' \) agree on the first \( n - 2t \) co-ordinates. Consider the vector \( \bar{y} = (y_1, y_2, \ldots, y_n) \) constructed as follows:

\[
y_i = \begin{cases} 
  x_i & 1 \leq i \leq n - t \\
  x'_i & n - t + 1 \leq i \leq n 
\end{cases}, i = 1, 2, \ldots, n
\]

Then \( d_H(\bar{y}, \bar{x}), d_H(\bar{y}, \bar{x}') \leq t \). Therefore, \( \bar{y} \in \Phi_{\bar{x}} \) and \( \bar{y} \in \Phi_{\bar{x}'} \). Since \( \Phi_{\bar{x}} \cap \Phi_{\bar{x}'} \neq \{\} \), the channel cannot correct \( t \) errors.

**Theorem 1.2** An \( (n, k, d) \) code \( C \) corrects errors over a \( t \)-erasure channel if and only if \( d \geq t - 1 \).

**Proof:**

Consider an \( (n, k, d) \) code with \( d \geq t + 1 \). Let \( \bar{y} \in \Phi_{\bar{x}}, \bar{x} = (x_1, x_2, \ldots, x_n) \in C \). We argue that \( \bar{y} \notin \Phi_{\bar{x}} \) for any codeword \( \bar{x}' = (x'_1, x'_2, \ldots, x'_n) \neq \bar{x} \). Note that since \( \bar{x}' \in C \), we know \( d_H(\bar{x}, \bar{x}') \geq t + 1 \). Without loss of generality, assume that \( x_i \neq x'_i, i = 1, 2, \ldots, t + 1 \). Denoting \( \bar{y} = (y_1, y_2, \ldots, y_n) \), since \( \bar{y} \) has at most \( t \)
erasures, we know that there exists a co-ordinate \( c \in \{1, 2, \ldots, t + 1\} \) among the first \( t + 1 \) co-ordinates such that \( y_c \neq \epsilon \). Since \( \vec{y} \in \Phi_{\vec{x}}, y_c = x_c \). However, \( x_c \neq x'_c \). Therefore \( y_c \neq x'_c \) and therefore, \( \vec{y} \notin \Phi_{\vec{x'}} \).

Conversely, consider an \((n, k, d)\) code with \( d \leq t \). Then there exist two codewords \( \vec{x}, \vec{x}' \) that agrees upon at least \( n - t \) co-ordinates. Without loss of generality, assume that \( \vec{x}, \vec{x}' \) agree upon the first \( n - t \) co-ordinates. Then a sequence \( \vec{y} \in \Phi^n \) which agree with \( \vec{x}, \vec{x}' \) in the first \( n - t \) co-ordinates and has erasures in the last \( t \) co-ordinates lies in both \( \Phi_{\vec{x}} \) and \( \Phi_{\vec{x}'} \). Therefore the code cannot correct \( t \) erasures. \( \blacksquare \)

**Theorem 1.3** An \((n, k, d)\) code \( C \) detects errors over a \( t \)-error channel if and only if \( d \geq t - 1 \).

Proof is an exercise.

### 1.4 Probabilistic Channel

A probabilistic channel with finite input alphabet \( F \) and finite output alphabet \( \Phi \) consists, for every \( \vec{x} \in F^n \), a probability measure \( \mathbb{P}_{Y|X}(\cdot|\vec{x}) \) on \( \Phi^n \). We consider memoryless channels, which have the form

\[
\mathbb{P}_{Y|X}(y_1, y_2, \ldots, y_n|x_1, x_2, \ldots, x_n) = \prod_{i=1}^n P_{Y|X}(y_i|x_i),
\]

where, for every \( x \in F \), \( P_{Y|X}(\cdot|x) \) is a probability measure on \( \Phi \).

#### 1.4.1 Binary Symmetric Channel

\( F = \Phi = \{0, 1\} \), with cross-over probability \( p \).

\[
P_{Y|X}(y|x) = \begin{cases} 
1 - p & \text{if } y = x \\
p & \text{if } y \neq x 
\end{cases}
\]

Note equivalently that \( \mathbb{P}_{Y|X}(\vec{y}|\vec{x}) = p^d_{H}(\vec{y}, \vec{x}) (1 - p)^{n-d_{H}(\vec{y}, \vec{x})} \).

#### 1.4.2 Binary Symmetric Channel

\( F = \{0, 1\}, \Phi = \{0, 1, \epsilon\} \) with erasure probability \( p \).

\[
P_{Y|X}(y|x) = \begin{cases} 
1 - p & \text{if } y = x \\
p & \text{if } y = \epsilon 
\end{cases}
\]

Note equivalently that

\[
\mathbb{P}_{Y|X}(\vec{y}|\vec{x}) = p^{\text{number of erasures in } \vec{y}} (1 - p)^{n-\text{number of erasures in } \vec{y}} \prod_{i=1}^N \mathbb{I}_{y_i \in \{x_i, \epsilon\}}
\]

where \( \vec{y} = (y_1, y_2, \ldots, y_n), \vec{x} = (x_1, x_2, \ldots, x_n) \) and \( \mathbb{I} \) is an indicator function.