1 Code and Distance

1.1 Definition: Code

A code C of block length n over a finite alphabet Σ is any subset of Σ^n.

$$C \subseteq \Sigma^n$$

"the set of all possible codewords"

e.g. $\Sigma = \{0, 1\}$, $n = 5$, $C = \{00000, 11111, 00001\}$

1.2 Definition: Dimension of a Code

Given a code $C \subseteq \Sigma^n$, C has dimension k defined by:

$$k = log_{|\Sigma|}|C|$$

"n is the size of any codeword"

"k is the size of any decoded codeword"

Note: $k \leq n$

e.g. $\Sigma = \{0, 1\}$, $n = 5$, $C = \{00000, 00001, 00010, 00011, 00100, 00101, 00110, 00111\}$ then, $k = log_2(8) = 3$

1.3 Definition: Rate of a code

Given a code $C \subseteq \Sigma^n$ with dimension k, C has rate R defined by:

$$R = \frac{k}{n}$$

"R is the ratio of non-redundent bits, higher is better, lower means more rendency"

1.4 Definition: Hamming Distance

The Hamming Distance between two equal length strings is the number of elementwise differences.

$$d_H = |\{i \mid x_i \neq y_i\}|$$

e.g. $d_H(bbb, aaa) = 3$

e.g. $d_H(xyz, abc) = 3$

1.5 Definition: Minimum distance of a code

Given a code $C \subseteq \Sigma^n$, C's minimum distance d is the smallest distance between any two codewords in C.

$$d = min\{d_H(i, j) \mid i, j \in C, i \neq j\}$$

1.6 Note: [n,k,d]

Given a code C, with dimension k, block length n and minimum distance d, we call C an $[n, k, d]$ code. These statements hold:

1. The maximum number of errors that an [n,k,d] code can correct is $\lfloor \frac{d-1}{2} \rfloor$.
2. The maximum number of errors that an [n,k,d] code can detect is $d - 1$

2 Linear Codes

2.1 Field

Field Axioms

The field axioms are generally written in additive and multiplicative pairs.

name	addition	multiplication
associativity	$(a+b)+c=a+(b+c)$	$(a\,b)\,c=a\,(b\,c)$
commutativity	$a+b=b+a$	$a\,b=b\,a$
distributivity	$a\,(b+c)=a\,b+a\,c$	$(a+b)\,c=a\,c+b\,c$
identity	$a+0=a=0+a$	$a\cdot 1=a=1\cdot a$
inverses	$a+(-a)=0=(-a)+a$	$a\,a^{-1}=1=a^{-1}\,a$ if $a\neq 0$

2.2 Finite(Galois) Field

Is a field that contains a finite number of elements.

e.g. $GF(2)$ is a finite field with elements $\{0,1\}$ addition defined as XOR and standard multiplication.

2.3 Definition: Linear Code

We say that $C \subseteq \Sigma^n$ is a linear code if C is a linear subspace of Σ^n where Σ is a finite field. i.e.:

1. $0 \in C$
2. $\forall a, b \in C, a + b \in C$

3 Error Correction

Any codeword with enough noise can be any other codeword... we want to be resilient to small noise.

3.1 Definition: Encoding Function

Given a code $C \subseteq \Sigma^n$, a mapping $E : \Sigma^k \to C$ is called an encoding function.

3.2 Definition: Decoding Function

Given a code $C \subseteq \Sigma^n$, a mapping $D : \Sigma^n \to \Sigma^k$ is called an decoding function.
Note: D is not injective ... since many codewords may get error corrected to the same decoded codeword

3.3 Definition: Error Correction

Given a code $C \subseteq \Sigma^n$, and an integer $t \in \mathbb{Z}$. C is said to be a t error-correcting code if there exists a decoding function D, such that for any $m \in \Sigma^k$ and any noise $\epsilon \in \Sigma^n$ with at most t errors, $D(E(m) + \epsilon) = m$.

3.4 Example: Repetition Code

Recall : $GF(2)$ is a finite field with elements $\{0, 1\}$ addition defined as XOR and standard multiplication.

$C_{3,rep}$ is a 1-error correcting code. Suppose:

$$C_{3,rep} = \{(0, 0, 0), (1, 1, 1)\} \subseteq GF(2)^3$$

$$E : GF(2) \to C_{3,rep}$$

$$D : GF(2)^3 \to GF(2)$$

With functions:

$$E(x_1) = (x_1, x_1, x_1)$$

$$D(x_1, x_2, x_3) = (majority(x_1, x_2, x_3))$$

e.g. $m = 0, \epsilon = (0, 1, 0)$, then $D(E(m) + \epsilon) = D(0, 1, 0) = 0$

e.g. $m = 1, \epsilon = (0, 1, 0)$, then $D(E(m) + \epsilon) = D(1, 0, 1) = 1$