Def^n Let $u, v \in A^n$.

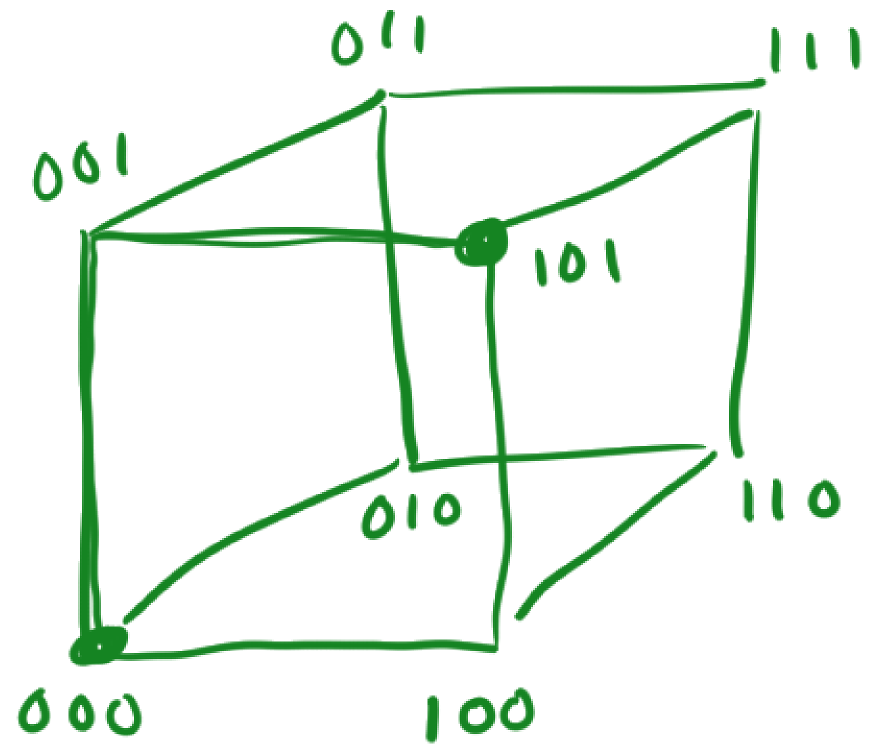
The Hamming distance $d(u,v)$ = # of positions where u & v differ.

eg. $u = (1,0,1,1,0)$ and $v = (1,1,1,0,0)$

then $d(u,v) = 2$.

Typically, decoding means finding the codeword at closest Hamming distance from the received message.

binary length 3 code



Prop^n $d(u,v)$ is a metric on A^n, meaning:

1. $d(u,v) \geq 0$ and $d(u,v) = 0 \iff u = v$
2. $d(u,v) = d(v,u) \quad \forall\, u,v$
3. $d(u,v) \leq d(u,w) + d(w,v) \quad \forall\, u,v,w$ (Δ inequality)

Proof: Exercise.

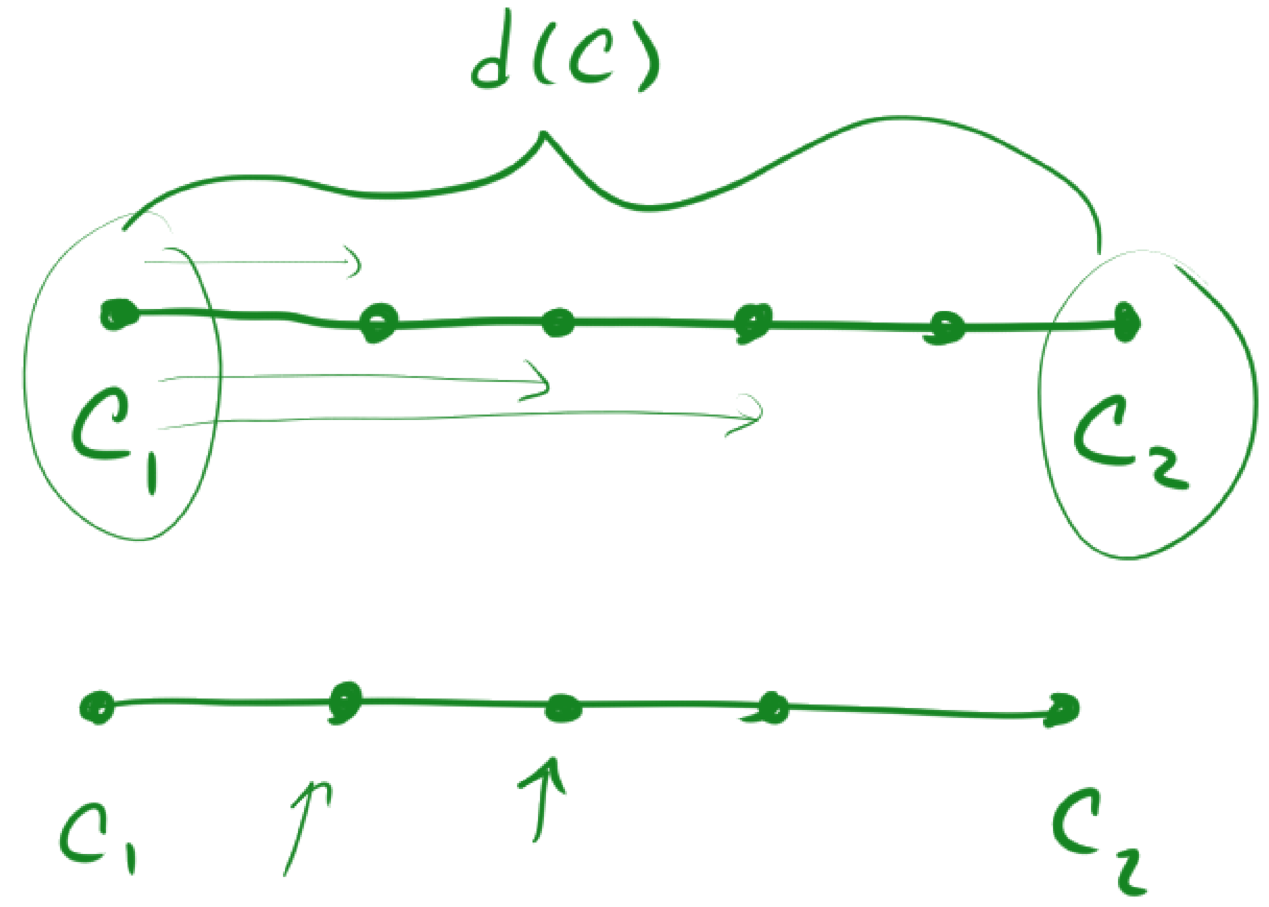
Def^n. The minimum distance of a code $\mathcal{C}$ is $d(\mathcal{C}) = \min\{d(u,v) : u, v \in \mathcal{C},\ u \neq v\}$.

"Nearest neighbour decoding": For a received message m, find the <u>closest</u> codeword, call this the <u>decoded</u> codeword.

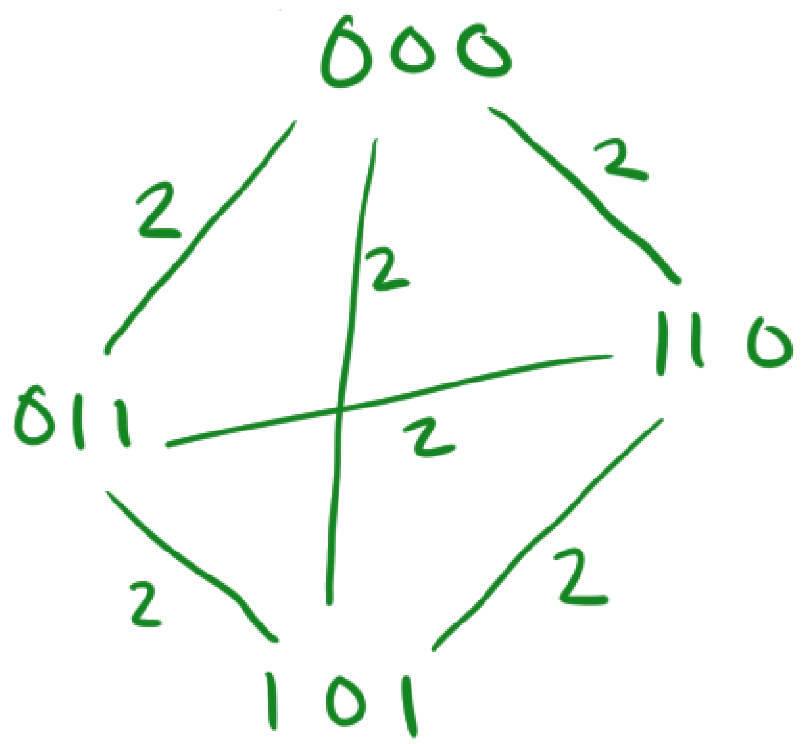
Def^n. A code $\mathcal{C}$ can

① detect s errors if $d(\mathcal{C}) \geq s+1$.

② correct t errors if $d(\mathcal{C}) \geq 2t+1$.



Ex.



Can detect 1 error

Can correct no errors

Basic Parameters

code C: $\left\{ \begin{array}{l} \text{length } n \text{ (characters in a codeword)} \\ M \text{ codewords} \\ d = d(C) \text{ (minimum distance)} \end{array} \right\}$ "(n, M, d) - code"

The code rate / information rate of a q-ary code is

$$R = \frac{\log_q M}{n} = \frac{\text{symbols needed to specify a codeword}}{\text{transmitted symbols for a codeword}}$$

$$= \text{info}/\text{space}$$

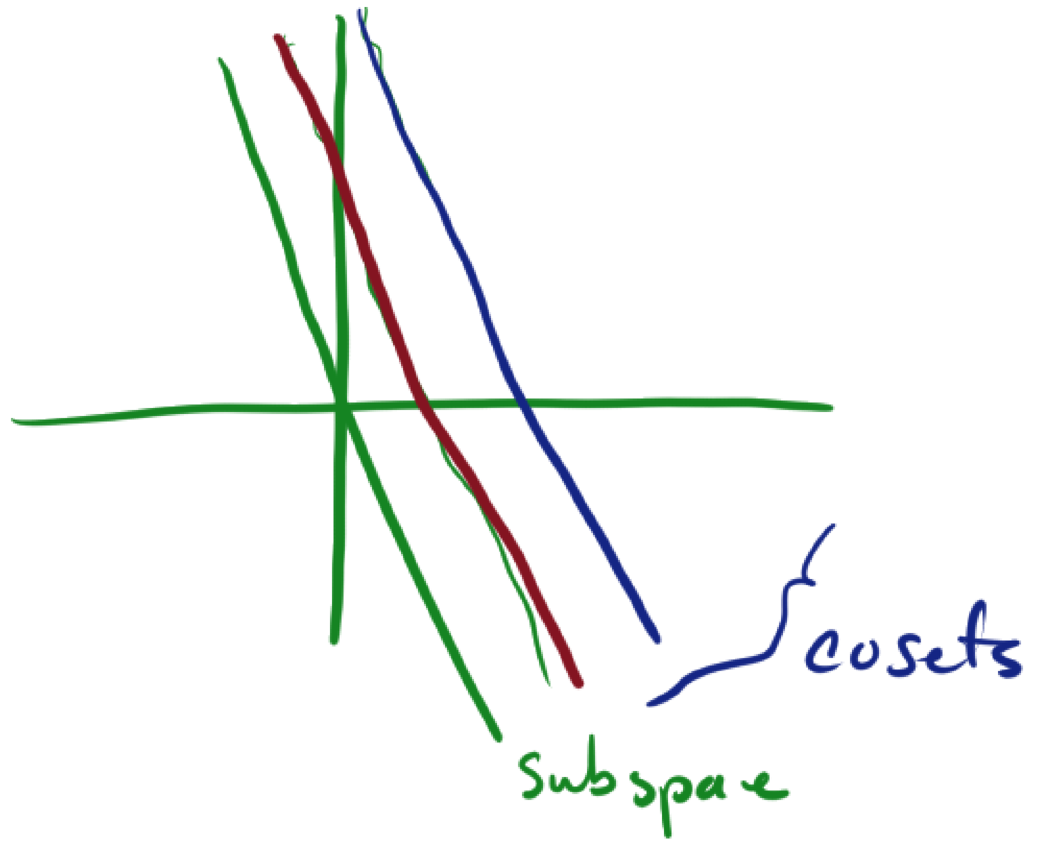
Ex. $\{(0,0,0), (1,1,1)\}$ is a binary $(3, 2, 3)$-code.

with $R = \dfrac{\log_2 2}{3} = \dfrac{1}{3}$.

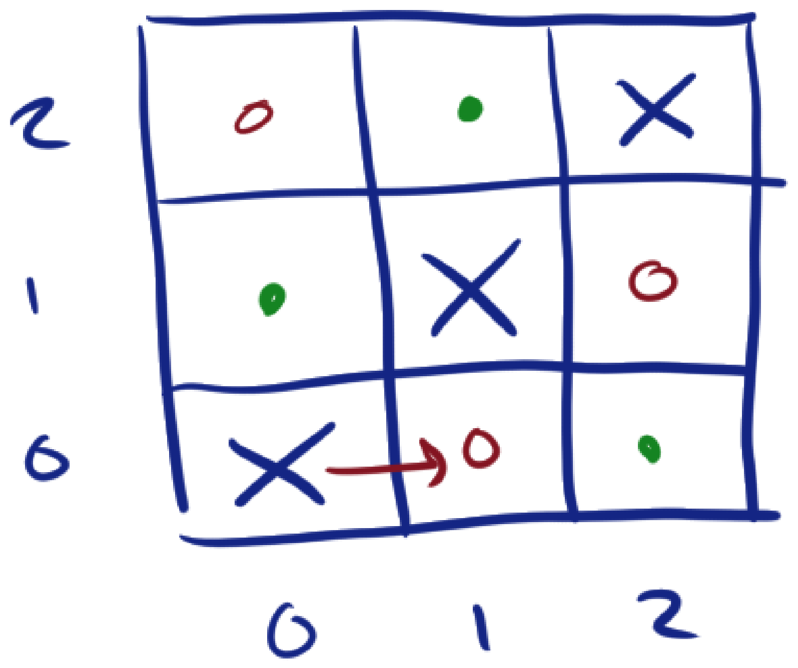
Linear Codes: these live in a vector space over finite field.

$\mathbb{F}$ = finite field, e.g. $\mathbb{F}_2 = \mathbb{Z}/2\mathbb{Z}$ often.

Idea: use a subspace as the code.



Eg. $\mathbb{F}_3^2$



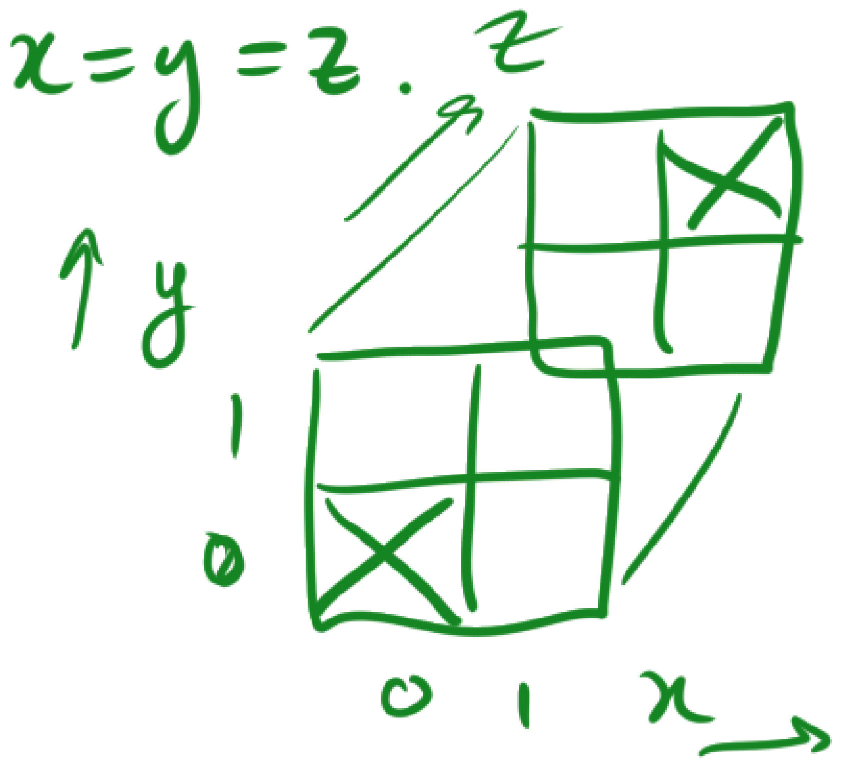
X: line $x=y$ in $\mathbb{F}_3^2$

o: coset (shifted by $(1,0)$)

•: coset

Eg. $\{(0,0,0),(1,1,1)\} \subseteq \mathbb{F}_2^3$ ↙ cardinality 8

Q: linear? Yes. It is the line $x=y=z$.



Def'n A linear code of dimension k and length n over $\mathbb{F}$ = k-dimensional subspace of $\mathbb{F}^n$.

Key Property:

codeword + codeword = codeword

A linear code of dimension k and length n over $\mathbb{F}$ is an "$[n,k]$-code"
or "$[n,k,d]$-code"
when $d(C)=d$.

Note: An $[n,k,d]$-code over $\mathbb{F}_q$
is an $(n, \underbrace{q^k}_{\text{size of the subspace.}}, d)$-code.

Linear Algebra Review: A subspace of dim k is $\left\{ \sum_{i=1}^{k} a_i \vec{v}_i : a_i \in \mathbb{F}_q \right\}$
for some basis $\vec{v}_1, \dots, \vec{v}_k$.

Corollary: cardinality is q^k.

Example. Hamming $[7,4]$ - code.

message: 4 bits $\vec{v} \in \mathbb{F}_2^4$

codewords: 7 bits $\vec{v}G \in \mathbb{F}_2^7$

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

eg. $$(1\,0\,1\,0)\begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{pmatrix} = (\overbrace{\underbrace{1\;0\;1\;0}_{\text{"information symbols"}}}^{\text{orig message}}\;\overbrace{\underbrace{1\;0\;1}_{\text{"check symbols"}}}^{\text{checksum}})$$

Code = rowspace of G = vector subspace of $\mathbb{F}_2^7$ of dim 4.