Welcome! Everything is fine.

When you're connected by zoom:

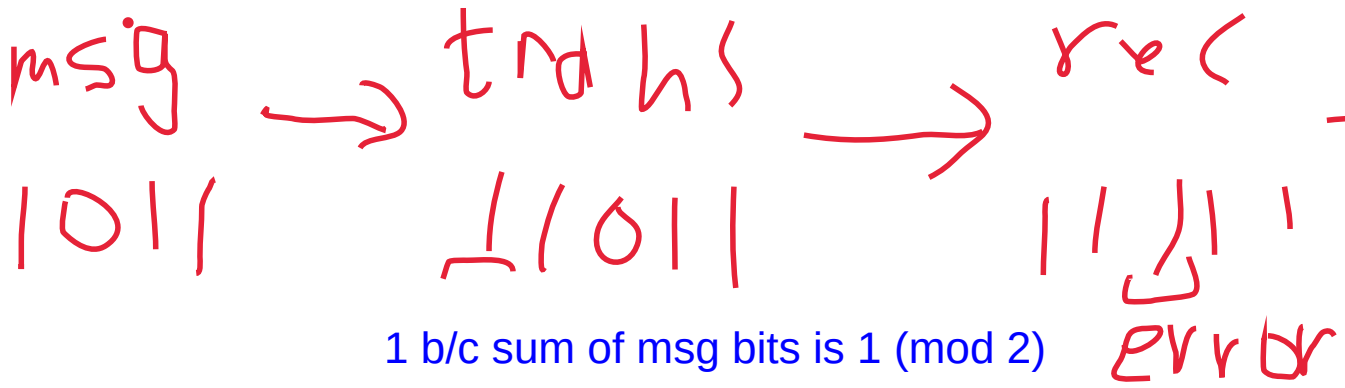
- Use a laptop or desktop with a large screen so you can read these words clearly.
- To conserve bandwidth, please turn off your camera.
- Please mute your microphone unless I call on you.
- Please have the chat window open to ask questions.

$H = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ x_0 & x_1 & x_2 & x_3 & x_4 \end{pmatrix}$

$Hx = 0$ is $x_0 + x_1 + \cdots + x_4 = 0 \pmod 2$

Last time

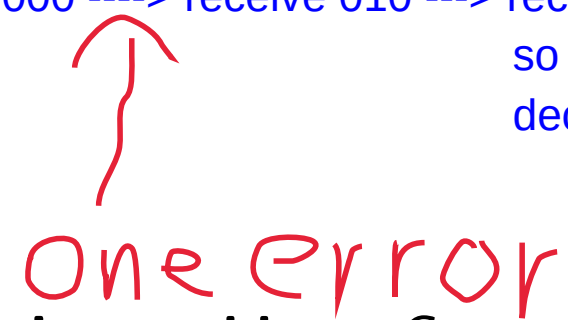
- Parity check code of length $n+1$ detects one error



sum is 1 (mod 2) so receiver knows something went wrong.
So PC detects 1 error.

- Repetition code of length 3 corrects one error

msg 0 ---> transmit 000 ----> receive 010 ---> receiver reasons that 1 error more likely than 2, so most likely intended transmission is 000, so decode msg as 0.



- **Motivating problem:** Can we transmit bitstrings in blocks of some length n, with one error-correction per block, at a cost of less than 3 transmitted bits per 1 message bit?

(Binary linear) codes

Recall: "bit" is element of $\mathbf{F}_2$, "bitstring of length n" is element of $\mathbf{F}_2^n$. (written horizontally, e.g., 01110010 instead of an 8 x 1 column vector)

Definition

(a binary code of length n)

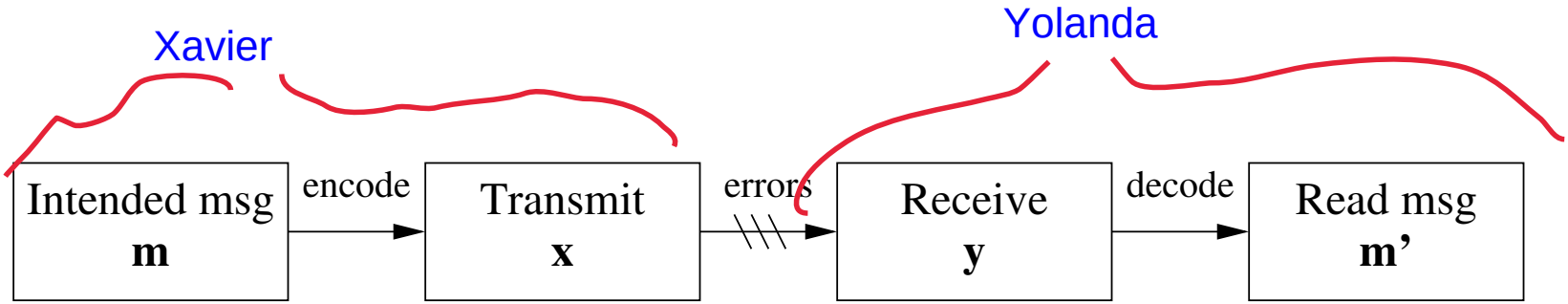
A **code** is a subset $\mathcal{C}$ of $\mathbf{F}_2^n$. Elements (vectors) of a code are called **codewords**, and again, we think of codewords as the words that are possible correctly transmitted messages.

(Ex: repetition code length 3, C = {000,111} , n=3)

Definition

A **binary linear code $\mathcal{C}$ of length** n is a subspace $\mathcal{C}$ of $\mathbf{F}_2^n$.

Standard framework

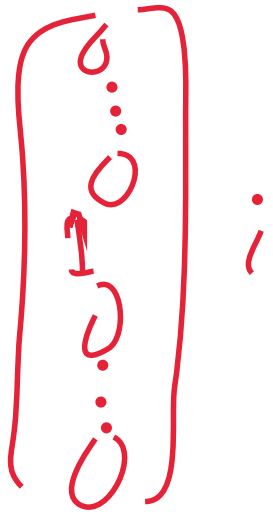


Rep code:

Sender Xavier, receiver Yolanda.

1. Xavier wants to send msg **m**. (a bitstring of some length t)
2. Xavier **encodes m** to codeword $\mathbf{x} \in \mathcal{C}$. (a bitstring of length n)
3. Xavier transmits **x**, Yolanda receives **y**. No errors: $\mathbf{y} = \mathbf{x}$; errors: $\mathbf{y} \neq \mathbf{x}$.
4. Yoland **decodes y** as $\mathbf{m}'$; success means $\mathbf{m}' = \mathbf{m}$. Often:
   - 4.1 Yolanda **corrects y** to $\mathbf{y}' \in \mathcal{C}$.
   - 4.2 Yolanda **reads** $\mathbf{y}'$ as a message $\mathbf{m}'$

Linear algebra of errors



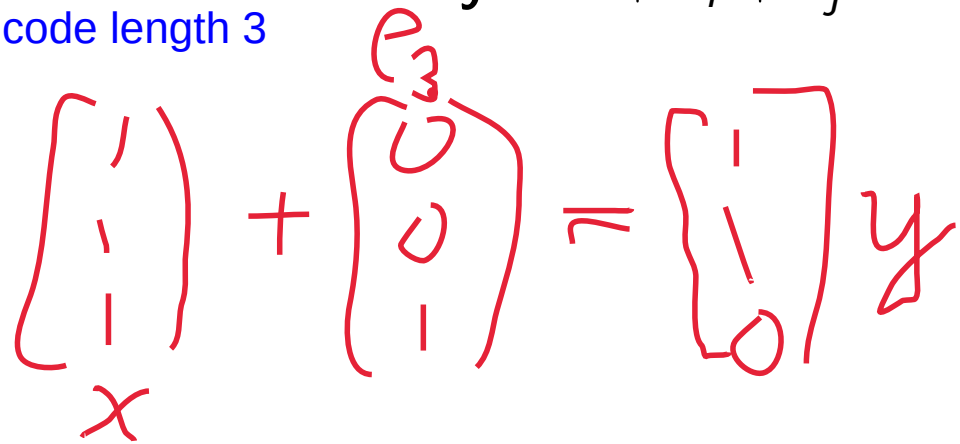
$\mathbf{e}_i$ has ith coordinate 1 and other coordinates 0.
One error flipping bit i modelled by:

$$\mathbf{y} = \mathbf{x} + \mathbf{e}_i. \tag{1}$$

(Adding 1 changes 0 to 1 and 1 to 0.)
Two errors flipping bits i and j:

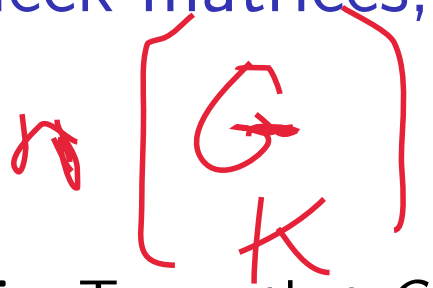
$$\mathbf{y} = \mathbf{x} + \mathbf{e}_i + \mathbf{e}_j. \tag{2}$$

rep code length 3



(algebraic model of flipping bit 3)

Generator matrices, parity check matrices, syndrome



columns have height n

Definition

Let G be an $n \times k$ matrix over $\mathbf{F}_2$. To say that G is the **generator matrix** of a binary linear code $\mathcal{C}$ of length n means that $\mathcal{C} = \mathrm{Col}(G)$, so codewords are lin combs of columns

Definition

Let H be a $k \times n$ matrix over $\mathbf{F}_2$. To say H is the **parity check matrix** of a binary linear code $\mathcal{C}$ of length n means that $\mathcal{C} = \mathrm{Null}(H)$, so codewords are $\mathbf{x}$ such that $H\mathbf{x} = \mathbf{0}$.

maybe tells you what went wrong?

$\neq 0$

Hx not 0 means error

Definition

Let H be a parity check matrix for a code $\mathcal{C}$ of length n. For $\mathbf{x} \in \mathbf{F}_2^n$, $H\mathbf{x}$ is the **syndrome** of $\mathbf{x}$. Note codewords in $\mathcal{C}$ are precisely the vectors $\mathbf{x}$ with syndrome equal to $\mathbf{0}$.

Exs: Parity check and repetition

- **Parity check code of length** $n+1$ is nullspace $\mathcal{C}$ of the $1 \times (n+1)$ matrix $H = [1 \ \ldots \ 1]$. In other words, $\mathbf{x} \in \mathbf{F}_2^{n+1}$ is in $\mathcal{C}$ exactly when $H\mathbf{x} = 0$, or:
- **Repetition code of length** n is span $\mathcal{C}$ of the (single) column of the $n \times 1$ generator matrix

$$G = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}. \tag{3}$$

Since the only possible linear combinations of $\{G\}$ are G itself and the zero vector, $\mathcal{C} = \{\mathbf{0}, G\}$.

x is a codeword of PC length 5 exactly when Hx = 0, or:

Hamming 7-code

col 3 upside down is 011, the binary digits of 3.

- $\mathcal{H}_7$ is the nullspace of the parity check matrix

Parity check matrix is in RREF so rank = 3, so nullity = width - rank = 7 - 3 = 4.

$$H_7 = \begin{bmatrix} 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{bmatrix}. \tag{4}$$

x is a codeword of Hamming 7 exactly when: $H_7 x = 0$

- $\mathcal{H}_7$ is the column space of the generator matrix

$\dim(H_7) = 4$

$$G_7 = \begin{bmatrix} 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \tag{5}$$

4 vecs in a basis, so dim = 4

Key point: ith column of H_7 is binary digits of i upside down.

Encoding and decoding Hamming 7

1. Xavier copies four data bits of $\mathbf{m}$ into x_3, x_5, x_6, x_7. (free variables) Other bits x_1, x_2, and x_4 satisfy

$$\begin{aligned} x_1 &= x_3 + x_5 + x_7 \\ x_2 &= x_3 + x_6 + x_7 \\ x_4 &= x_5 + x_6 + x_7. \end{aligned} \tag{6}$$

(leading variables, determined by free variables)

2. Xavier transmits $\mathbf{x}$, Yolanda receives $\mathbf{y}$.
3. Yolanda then decodes $\mathbf{y}$:
   - 3.1 First, Yolanda corrects $\mathbf{y}$ to a codeword $\mathbf{y}'$ as follows. Let $\mathbf{s} = H_7\mathbf{y} \in \mathbf{F}_2^3$ be the syndrome of $\mathbf{y}$.
     If $\mathbf{s} = \mathbf{0}$, then $\mathbf{y}$ is a codeword, so Yolanda chooses $\mathbf{y}' = \mathbf{y}$.
     Otherwise, Yolanda reads $\mathbf{s}$ as the binary digits of a number i and chooses $\mathbf{y}'$ to be $\mathbf{y}$ with its ith bit flipped (i.e., $\mathbf{y}' = \mathbf{y} + \mathbf{e}_i$).
   - 3.2 Yolanda reads the message $\mathbf{m}'$ off of bits 3, 5, 6, and 7 of $\mathbf{y}'$.

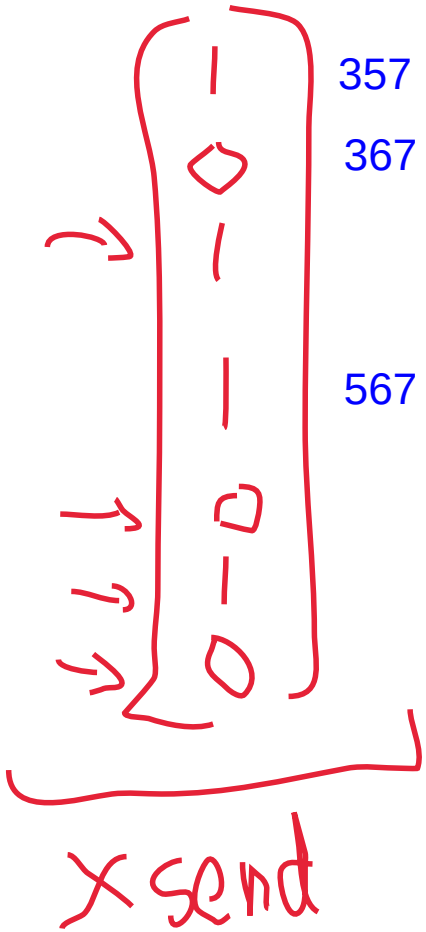
Encoding and decoding Hamming 7: Example

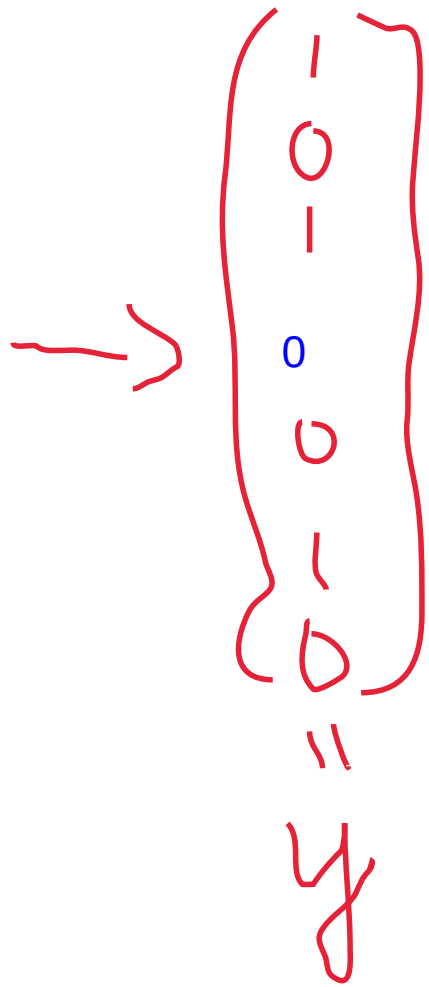
m = 1010
copy to bits 3, 5, 6, 7:

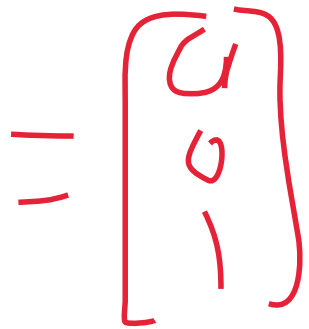
$$H_7 = \begin{bmatrix} 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{bmatrix} \quad (7)$$

s=H_7 y
remember: matrix * vector
is a l.c. of colums w/coeffs
from vector

so s = sum of cols 1, 3, 6







read as 100, the binary digits of the number 4. So Yolanda knows that bit 4 got messed up !!!!!
WHAAAAAAAA

Proof that decoding works

Theorem

If $\mathbf{y} = \mathbf{x} + \mathbf{e}_i$ *(one error in bit* i*), then syndrome* $\mathbf{s} = H_7\mathbf{y}$ *is the binary digits of* i*.*

Proof.

Since $\mathbf{x} \in \mathcal{H}_7 = \text{Null}(H_7)$, $H_7\mathbf{x} = \mathbf{0}$. Therefore,

$$\mathbf{s} = H_7\mathbf{y} = H_7(\mathbf{x} + \mathbf{e}_i) = H_7\mathbf{x} + H_7\mathbf{e}_i = \mathbf{0} + H_7\mathbf{e}_i = H_7\mathbf{e}_i, \quad (8)$$

the ith column of H_7. However, H_7 is the matrix whose ith column is the binary digits of i. □

Hamming 8-code

Definition

Hamming 8-code $\mathcal{H}_8$ is nullspace of

$$H_8 = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{bmatrix}. \tag{9}$$

Delete the first row and first column of H_8, get H_7.
So to be consistent with the Hamming 7-code, we write an arbitrary element of $\mathcal{H}_8$ as $\begin{bmatrix} x_0 \\ x_1 \\ \vdots \\ x_7 \end{bmatrix} \in \mathcal{H}_8$.

Facts of Hamming 8

Theorem

The Hamming 8-code $\mathcal{H}_8$ is the Hamming 7-code $\mathcal{H}_7$, extended by a parity check bit x_0; and $\mathcal{H}_8$ corrects 1 error and detects 2 errors.

Proof is HW.