Code-based cryptography

Alice $\xrightarrow{m}$ encode(m) $\overset{\text{noisy channel}}{\rightsquigarrow}$ m^* $\longrightarrow$ Bob

Q: How to ensure that Bob can recover corrupted message?

Ex (Repetition Code). Alphabets $= \{0,1\}$

Fix $k = 4$ (repetition number)

Alice wants send a 2-bit message "10"

Alice $\rightsquigarrow$ "10101010" $\overset{\text{noise}}{\rightsquigarrow}$ "11101000" $\longrightarrow$ Bob

3 "1" more likely $\leftarrow$ odd: there are 3 "1"'s, 1 "0"'s

even: there are 3 "0"'s, 1 "1"

if in noisy channel, $p < 1/2$ probability of bit flipped.

1-bit "0" 00000100 $\rightsquigarrow$ know "0" original message

prob $= (1-p)^7 p$ $\quad p < 1/2$

"close" to 1

Ex ISBN for books.

0-674-57629-2

$a_1 - a_2 a_3 a_4 - a_5 a_6 a_7 a_8 a_9 - a_{10}$

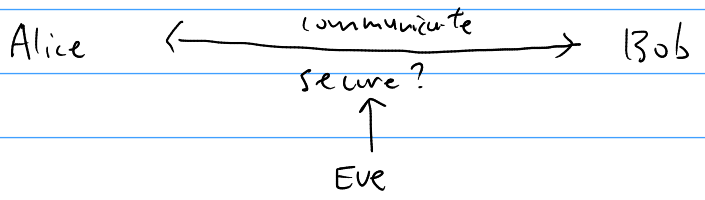
first 9 digit contains information about the book

$a_{10} \equiv a_1 + 2a_2 + 3a_3 + \cdots + 9a_9 \pmod{11}$

- check single digit errors
- check transposition error (2 positions swapped)

Applications (Error-correcting codes): ISBN's, deep space image, cryptography, etc.

Cryptography



CBC:
- algebraic approach (Reed-Solomon codes, algebraic geometry codes, Goppa codes, etc.)
- probabilistic approach (convolution, MPPC, LDPC, polar)

Finite field:

finite number of elements — commutative ring with unity, every nonzero element has multiplicative inverse.

(e.g. $\mathbb{R}$ real numbers, $\mathbb{C}$ complex numbers)

$\mathbb{F}_p = \{0, 1, \dots, p-1\}$, $\mathbb{F}_q$ where $q = p^r$

Polynomial rings: $R[x] = \{a_0 + a_1 x + a_2 x^2 + \dots + a_n x^n\}$

$\mathbb{F}_q[x]$, $\mathbb{F}_q[x, y]$, ...

Def

- A **code** of length n is $C \subseteq F^n$ (e: codeword)
  where $F = \{\text{alphabets}\}$
- **Hamming distance**, $d_H(x, y)$ $\quad x, y \in C \subseteq F^n$

  = # entries of x, y that are different.

($x = (0,0,0)$, $y = (0,1,0)$ $\in \mathbb{F}_2^3$, $d_H(x,y) = 1$)

- Hamming weight of codeword, $wt(x)$ $\quad x \in C$

  ii

  $d_H(x,0)$

- Hamming distance of a code C,
  $d_H(C) = \min_{x \neq y} \{d_H(x,y)\}$

Exercise d_H is a metric:

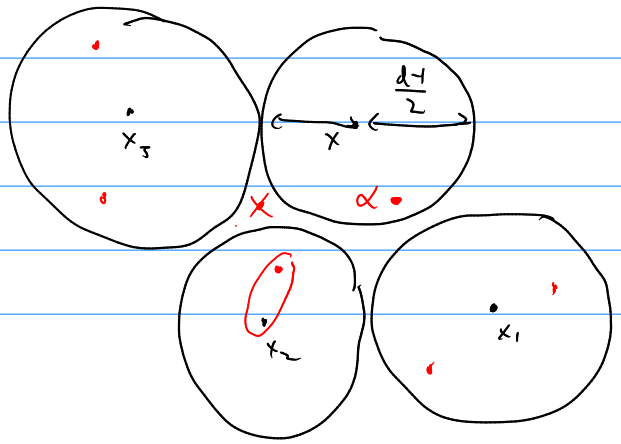
- $d_H(x,x) \geq 0$
- $d_H(x,y) = d_H(y,x)$
- $d_H(x,z) + d_H(z,y) \leq d(x,y)$

Rmk A code is a subset of F^n.
A linear code is a subspace of $\mathbb{F}_q^n$.
$\rightsquigarrow$ dimension of a code

n = length of code
k = dim of code
d = Hamming distance of code

To design a good code, want large k, d (wrt n)

k to be large: transmit more information
d large: $B_{\frac{d-1}{2}}(x) = \{ c \in C : d_H(x,c) \leq \frac{d-1}{2} \}$



$\alpha \in \mathbb{F}_q^n$, $\alpha \in B_{\frac{d-1}{2}}(x)$

$\Downarrow$

$d_H(\alpha, x) \leq \frac{d-1}{2}$