Polynomials (over $\mathbb{F}_p$) in Coding Theory

- **Basic:** A(lice) ⟶ B(ob)

  Send $m \in \{0,1\}^N$ over a channel having t many bit-errors.

  • How to correctly communicate in minimum bits?

Trivial: • $N \cdot (2t+1)$ many bits suffice; with each bit repeated $(2t+1)$ times.

• Bob takes a majority vote / block to find the correct bit!

Clever algebraic solution: Reed & Solomon (1960) gave a code requiring $O(N \log N)$ bits, to correct $\approx N/2$ many errors!

- RS codes are widely used in:
  1) mass storage systems, eg. HD, CD, DVD, distributed online storage etc.
  2) bar codes;
  3) deep space & satellite communications.

Reed-Solomon Code

- View the message as a function : $\mathbb{F}_q \to \mathbb{F}_q$.
  ↳ (univariate polynomial)

  A sends its evaluations to B.

Encoding: 1) Break the N-bit msg m into k blocks each of size b-bits.
View these blocks as $d_0, \dots, d_{k-1} \in \mathbb{F}_{2^b}$.

2) Define $P(x) := d_0 + d_1 x + \dots + d_{k-1} x^{k-1} \in \mathbb{F}_{2^b}[x]$.

3) Pick n distinct points $e_0, \dots, e_{n-1} \in \mathbb{F}_{2^b}$.

Send code $(c_0, \dots, c_{n-1}) := (P(e_0), \dots, P(e_{n-1})) \in (\mathbb{F}_{2^b})^n$.

▷ Encoding is a <u>linear map</u> : $\{0,1\}^N = (\mathbb{F}_{2^b})^k \longrightarrow \{0,1\}^{bn} = (\mathbb{F}_{2^b})^n$.

(∵ $P(e)+P'(e) = (P+P')(e)$)

▷ The map can be computed in time $\leq \tilde{O}(nb)$.

[∵ For $n = 2^b - 1$, we can use FFT ideas!]

▷ If $\bar{c} := (c_0, \ldots, c_{n-1})$ <u>doesn't</u> get corrupted then Bob can find P, by interpolation, if $2^b \geq n \geq k$.

- How does Bob decode m from a <u>corrupted</u> version <u>$\bar{c}'$</u> of $\bar{c}$? Let $P(e_{i_1}), \ldots P(e_{i_t})$ be wrong.

Decoding RS [Peterson 1960]

- Idea: Consider the error-locator $Q := \prod_{j=1}^{t} (x - e_{i_j})$.

$\Rightarrow$ $(c'_j - c_j) \cdot Q(e_j) = 0$, $\forall 0 \le j \le n-1$.

$\Rightarrow$ $P(e_j) \cdot Q(e_j) = c'_j \cdot Q(e_j)$, "

$\Rightarrow$ [L.Sys.] $R(e_j) = c'_j \cdot Q(e_j)$, "

where $R(x) := P(x) \cdot Q(x)$

→ Coefficients of R & Q are the unknowns.

R: deg = k-1+t

Q: deg = t & manic

$\Rightarrow$ ▷ # unknowns $= (k-1+t)+1+t = (k+2t)$.

▷ # linear eqns $= n$

Claim: · $\forall$ solutions R', Q' of the system: $Q' \mid R'$
(as long as $n \geq k+2t$).

· Thus, $R'(x) / Q'(x) = P(x)$.

Pf: · Let $2^b \geq n \geq k+2t$.

· The system has ≥ 1 solution (eg. $R = P \cdot Q$ & Q).

· Let Q', R' be some other solution.

· Consider $\Delta(x) := R' - P \cdot Q'$. $\deg \leq k-1+t < n-t$

· Δ vanishes on $(n-t)$ points in $\{e_0, \dots, e_{n-1}\}$.

$$\Rightarrow \quad \Delta = R' - P \cdot Q' = 0$$

$$\Rightarrow \quad P = R'(x) / Q'(x) \; . \qquad \square$$

▷ Error-tolerance $:= t \leq \frac{n-k}{2} = \frac{n}{2} \cdot \left(1 - \frac{k}{n}\right)$

– $2^b \geq n \geq k + 2t$ & $\underline{N} = b \cdot k$.

R given parameter

→ k, b, n could be fixed for any desired $t < n/2$.

▷ Time to solve the special lr. sys. $\leq \tilde{O}(nb)$.

(∵ with $n = 2^b - 1$ we can use FFT ideas, to invert the matrix!)

– Example fixing of parameters:

$$b = \lg N, \quad k = N/\lg N, \quad n = N$$

$$\Rightarrow \quad t \leq \frac{n-k}{2} \leq \frac{N}{2} \cdot \left(1 - \frac{1}{\lg N}\right).$$

$\approx$ 50% error-correction in $\mathbb{F}_{2^b}$ - alphabet.

▷ RS-code is of length $= N \lg N$ & corrects up to $\frac{N}{2}\left(1 - \frac{1}{\lg N}\right)$ errors. Time $\leq \tilde{O}(N)$.