Reed–Solomon Codes

e.g. PDF417 bar code
$q=929,\ \alpha=3,\ T=4$

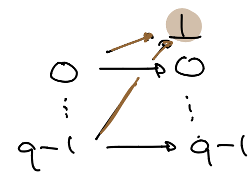
Alphabet: $\mathcal{A} = \mathbb{F}_q$ for q prime ← prime power ok, too

↳ $\{0,\dots,q-1\}$ with $+$ and $\cdot$ modulo q
(finite field with q elements)

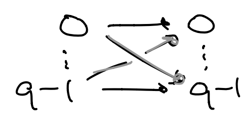
* Strange? No! e.g. with $q=257$ can encode 1 byte per symbol
* large q protects naturally against "burst errors"

Parameters: $K<N<q$ and $\alpha \in \mathbb{F}_q$

* overhead: $T := N-K$



* Can correct up to T erasures (= known error locations)
  or up to $\frac{T}{2}$ errors at unknown locations



* α should be a "generator": $\mathbb{F}_q = \{0, \alpha, \alpha^2, \dots, \alpha^{q-1}=1\}$

  any nonzero element is power of α

  always exists! e.g. $\mathbb{F}_3 = \{0, 2, 2^2=1\}$, $\mathbb{F}_5 = \{0, 2, 2^2=4, 2^3=3, 2^4=1\}$

  all equalities today are "modulo q"

↳ generator polynomial: $G = (Z-\alpha)\cdots(Z-\alpha^T)$

variable of the polynomial

Encoder: Input: $s^K \in \mathcal{A}^K$

* $P \leftarrow s_1 + s_2 Z + \dots + s_K Z^{K-1}$
* $R \leftarrow P \cdot Z^T \bmod G$ — remainder of poly division (see ex. class); degree $<T$ (= degree of G)
* $M \leftarrow P\cdot Z^T - R$ — degree $N-1$ & leading coeffs $s_K,\dots,s_1$
* $x^N \leftarrow$ coefficients of M — i.e. $M = x_1 + x_2 Z + \dots + x_N Z^{N-1}$

By construction:

* $x^N = [x_1,\dots,x_T, \overbrace{s_1,\dots,s_K}^{\text{source message}}]$

M and $P\cdot Z^T$ differ in degree $<T$ only!

* M is multiple of G — we subtracted the remainder!

$\Rightarrow$ "parity checks" $\boxed{M(\alpha) = \ldots = M(\alpha^T) = 0}$ ⊛

$\boxed{\text{ex:}}$ $K=1$, $N=3$, $q=5$ and $\alpha=2$

↳ $T=2$ & $G = (Z-2)(Z-4) = Z^2 - Z - 2$ (mod 5!)

To encode $s \in \mathbb{F}_5$:
* $P \leftarrow s$
* $R \leftarrow s \cdot Z^2 \bmod G = s \cdot Z^2 - s \cdot G = s \cdot Z + 2s$
* $M \leftarrow s \cdot Z^2 - R = s \cdot Z^2 - s \cdot Z - 2s$
* $x^N \leftarrow [-2s, -s, s]$ ⤳ linear code!

as claimed above

How to decode?

Imagine we receive $y^N \in \mathcal{A}^N$.

Interpret as coeffs of polynomial:

$$R = M + E$$

with error polynomial

$$E = \sum_{k=1}^{c} e_k Z^{i_k}$$

#errors

locations $\in \{0, \ldots, N-1\}$ grr...

mismatch

Two settings:

* Erasures: e_k unknown, c and i_k known
* General errors: everything unknown

What do we know? ⊛ implies:

$$(\mathrm{I}) \left\{ \begin{aligned} E(\alpha) &= \sum_{k=1}^{c} e_k \alpha^{i_k} = R(\alpha) \\ &\;\;\vdots \\ E(\alpha^T) &= \sum_{k=1}^{c} e_k \alpha^{T \cdot i_k} = R(\alpha^T) \end{aligned} \right\}$$

T linear equations in unknowns $e_1, \ldots, e_c$

... if locations $i_1, \ldots, i_c$ known

This solves the problem for <u>erasure errors</u>: Can correct $\boxed{G \le T \text{ erasures}}$

ex: $x^N = [\cancel{-2s}, -s, \cancel{s}]$

imagine $T=2$ erasure errors, e.g. $y^N = [0, -s, 0]$.

Known locations

$R = -sZ \qquad E = e_1 Z^{0} + e_2 Z^{2} = e_1 + e_2 Z^2$

$$\begin{aligned} E(2) &= e_1 + e_2 4 \doteq R(2) = -2s \\ E(4) &= e_1 + e_2 \doteq R(4) = s \end{aligned} \quad \Rightarrow \quad e_1 = 2s, \quad e_2 = -s, \quad E = 2s - sZ^2$$

$\Rightarrow M = R - E = -2s - sZ + sZ^2 \,\hat{=}\, [-2s, -s, s]$ ☺

Decoder for erasures: <u>Input</u>: $y^N \in \mathcal{A}^N$, error locations $i_1, \dots, i_G$

* $R \leftarrow y_1 + y_2 Z + \dots + y_N Z^{N-1}$
* Solve (1) for $e_1, \dots, e_G$
* $E \leftarrow e_1 Z^{i_1} + \dots + e_G Z^{i_G}$
* $M \leftarrow R - E$
* $\hat{S}^K \leftarrow$ leading K coeffs of M (i.e. $\hat{S}_1 = m_{N-K+1}, \dots, \hat{S}_K = m_N$)

What if locations unknown? Consider <u>locator polynomial</u>:

Should all be distinct: need $N \le q-1$

$$L := \prod_{k=1}^{G} (1 - Z\alpha^{i_k}) = 1 + L_1 Z + \dots + L_G Z^G$$

Roots are α^{-i_k} for $k = 1, \dots, G$. How to determine L?

$$0 = \sum_k e_k \alpha^{i_k(j+G)} \underbrace{L(\alpha^{-i_k})}_{=0}$$

$$= E(\alpha^{j+G}) + L_1 E(\alpha^{j+G-1}) + \dots + L_G E(\alpha^{j}) \qquad \text{for } j = 1, 2, \dots$$

But: $E(\alpha)=R(\alpha), \dots, E(\alpha^T)=R(\alpha^T)$:

$$(2)\quad \begin{bmatrix} R(\alpha^{C}) & \cdots & R(\alpha) \\ \vdots & & \vdots \\ R(\alpha^{2C-1}) & \cdots & R(\alpha^{C}) \end{bmatrix} \begin{bmatrix} L_1 \\ \vdots \\ L_C \end{bmatrix} = \begin{bmatrix} -R(\alpha^{C+1}) \\ \\ -R(\alpha^{2C_1}) \end{bmatrix}$$

← linear system for $L_1, \dots, L_C$

... as long as $2C_1 \leq T$, i.e., $\boxed{C_1 \leq \frac{T}{2} \text{ errors}}$.

Still don't know C_1 — so just try from $C_1 = \lfloor \frac{T}{2} \rfloor, \dots, 1$ until (2) unique solution.

Once we know L: search roots α^{-i_k} ⇝ i_k ⇝ e_k ⇝ E.

$\boxed{\text{ex:}}$ $S=1$ is encoded in $x^N=[-2,-1,1]$

Assume we receive $y^N=[-2,-1,0]$ ⇝ $R=-2-z$

$$\left.\begin{aligned} R(\alpha) &= 1 \neq 0 \\ R(\alpha^2) &= -1 \neq 0 \end{aligned}\right\} \Longrightarrow \text{error(s) happened.}$$

Try $\boxed{C_1=1}$:

* Determine L: $\quad (2): \overset{1}{R(\alpha)} \cdot L_1 = -\overset{-1}{R(\alpha^2)}$

$\Longrightarrow L_1 = 1$, i.e. $L = 1+z$

* Determine error locations: $\quad L$ has root $g_1 = -1 = 4 = \alpha^2 = \alpha^{-2}$

↳ location $i_1 = 2$ → $E = e z^2$

* Determine E and correct: $\quad (1): E(\alpha) = 1 \Longrightarrow e=-1$, $E=-z^2$

$E(\alpha^2) = -1$

$\Longrightarrow M = R - E = -2 - z + z^2 \triangleq [-2,-1,1]$

In general, have the following algorithm:

Decoder for general errors: Input: $y^N \in \mathcal{A}^N$

* $R \leftarrow y_1 + y_2 Z + \ldots + y_N Z^{N-1}$
* If $R(\alpha) = \ldots = R(\alpha^T) = 0$:

  $M \leftarrow R$

  else:

  For $C = \lfloor \frac{T}{2} \rfloor, \ldots, 1$:

  If Det $= 0$ in ②: Continue

  Solve ② for $L_1, \ldots, L_C$

  $L \leftarrow 1 + L_1 Z + \ldots + L_C Z^C$

  $S_1, \ldots, S_C \leftarrow$ roots of L ← — search

  For $k = 1, \ldots, C$:

  $i_k \leftarrow$ number in $\{0, \ldots, N-1\}$ s.th. $S_k = \alpha^{-i_k}$ $= \alpha^{q-1-i_k}$ ← — search/ look up

  Solve ① for $e_1, \ldots, e_C$

  $E \leftarrow \sum_{k=1}^{C} e_k Z^{i_k}$

  $M \leftarrow R - E$

  Break

* $\hat{S}^K \leftarrow$ leading K coeffs of M (i.e. $\hat{S}_1 = m_{N-K+1}, \ldots, \hat{S}_K = m_N$)

Appendix: Why does ① have a unique solution if $\boxed{G \leq T}$?

We use linear algebra, which works the same over $\mathbb{F}_q$ as over $\mathbb{R}$ or $\mathbb{C}$.

Consider the following $T \times T$-matrix, where $\beta_1, \dots, \beta_T$ are arbitrary:

$$B = \begin{bmatrix} \beta_1 & \cdots & \beta_T \\ \vdots & & \\ \beta_1^T & & \beta_T^T \end{bmatrix}$$

* $\det(B)$ is polynomial of degree $1+2+\dots+T = \frac{T(T+1)}{2}$ in $\beta_1, \dots, \beta_T$
* $\det(B) = 0$ if $\beta_i = 0$ $\Longrightarrow$ $\beta_i \overset{\text{divides}}{\mid} \det B$
* $\det(B) = 0$ if $\beta_i = \beta_j$ $\Longrightarrow$ $\beta_i - \beta_j \overset{\text{divides}}{\mid} \det B$

$\Longrightarrow$ $\det(B)$ proportional $\beta_1 \cdots \beta_T \prod_{i<j} (\beta_i - \beta_j)$ (same degree!)

RESULT: $\boxed{\text{If } \beta_1, \dots, \beta_T \text{ distinct and nonzero then } \underbrace{B \text{ is invertible}}}$

In particular: all columns linearly independent!

Now note that our linear system ① is of the following form:

$$\underbrace{\begin{bmatrix} \alpha^{i_1} & \cdots & \alpha^{i_C} \\ \vdots & & \vdots \\ (\alpha^{i_1})^T & \cdots & (\alpha^{i_C})^T \end{bmatrix}} \begin{bmatrix} e_1 \\ \vdots \\ e_C \end{bmatrix} \overset{!}{=} \begin{bmatrix} R(\alpha) \\ \\ \\ R(\alpha^T) \end{bmatrix} \qquad (C \leq T)$$

linearly independent columns, since $\beta_k = \alpha^{i_k}$ distinct and nonzero!

indeed: since α is "generator",

$$\mathbb{F}_q = \{0, \underbrace{\alpha, \dots, \alpha^{q-1} = 1}_{\text{all distinct}}\}$$

and $0 \leq i_1 \neq \dots \neq i_C \leq N-1 < q-1$

THUS: $\boxed{\text{linear system has unique solution}}$

(solution exists by assumption)