Lecture 13: Reed-Solomon Codes with an Example

The Field $\mathbb{GF}[2]$

- Let $(\mathbb{F}, +, \cdot)$ be a field such that $|\mathbb{F}| = 2$
- Let $\mathbb{F} = \{0, 1\}$
- We define $a + b := (a + b) \mod 2$
- We define $a \cdot b := (a \cdot b) \mod 2$
- Note that $-a = a$, for $a \in \mathbb{F}$

The Field $\mathbb{GF}[8]$

- Let $(\mathbb{F}, +, \cdot)$ be a field such that $|\mathbb{F}| = 8$
- Let $\mathbb{F}$ be the set of all polynomials in X that have coefficients in $\mathbb{GF}[2]$ with degree < 3
- Concretely,
  $\mathbb{F} = \{0, 1, X, X + 1, X^2, X^2 + 1, X^2 + X, X^2 + X + 1\}$
- We can represent these elements as numbers with 3-bit binary representation, i.e. $\{0, 1, 2, \ldots, 7\}$
- For $f(X), g(X) \in \mathbb{F}$, we define
  $f(X) + g(X) := (f_0 + g_0) + (f_1 + g_1)X + (f_2 + g_2)X^2$
- For $f(X), G(X) \in \mathbb{F}$, we define
  $f(X) + g(X) := (\ f(X) \cdot g(X)\ ) \mod (X^3 + X + 1)$

- For example, $(X^2 + 1) \cdot (X + 1) = X^3 + X^2 + X + 1 = X^2 \mod X^3 + X + 1$
- And $(X + 1)^{-1} = (X^2 + X)$
- Henceforth, we will write the elements as $\{0, 1, 2, \ldots, 7\}$
- So, in this representation, the above two statements correspond to $5 \cdot 3 = 4$ and $3^{-1} = 6$

- Let $\mathcal{F}_{4,8}$ be the set of all polynomials with degree < 4 and each coefficient of the polynomial is in $\mathbb{GF}[8]$
- That is, $\{F_0 + F_1 Z + F_2 Z^2 + F_3 Z^3 \colon F_0, F_1, F_2, F_3 \in \mathbb{GF}[8]\}$
- The set of all messages $\mathcal{M}$ corresponds to

$$\left\{(F_0, F_1, F_2, F_3) \;:\; F_0, F_1, F_2, F_3 \in \mathbb{GF}[8]\right\}$$

- So, the size of the message space is $|\mathcal{M}| = \left|\mathbb{GF}[8]\right|^4 = 8^4$
- The encoding of the message (F_0, F_1, F_2, F_3) is the evaluation of the function $F(Z) = \sum_{k=0}^{k=3} F_k Z^k$ at every $Z \in \mathbb{GF}[8]$
- That is, we output

$$\left(F(0), F(1), \dots, F(7)\right)$$

- Note that the code is 8 elements in $\mathbb{GF}[8]$ and each element in $\mathbb{GF}[8]$ is represented by 3-bits. So, the codeword is represented by $8 \cdot 3 = 24$ bits

- So, the encoding function
  $\mathrm{Enc}\colon (F_0, F_1, F_2, F_3) \mapsto (F(0), F(1), F(2), \ldots, F(7))$
- In other words, it takes 12-bit input and provides 24-bit output

Claim

The following set is a vector space

$$\{\mathrm{Enc}(F) \colon F \in \mathcal{M}\}$$

- Let F, G be two polynomials in $\mathcal{M}$. Interpret $(F(0), \ldots, F(7))$ and $(G(0), \ldots, G(7))$ as vectors. Their sum is identical to $(H(0), \ldots, H(7))$, where $H = F + G$.
- Let $\alpha \in \mathbb{GF}[8]$. Note that $\alpha \cdot (F(0), \ldots, F(7))$ is the vector $(H(0), \ldots, H(7))$, where $H = \alpha F$.

- Now, we can claim that every $\mathrm{Enc}(F)$ can be written as a linear combination of 4 basis vectors. For example, if $F = F_0 \cdot (1) + F_1 \cdot (Z) + F_2 \cdot (Z^2) + F_3 \cdot (Z^3)$, then we have $\mathrm{Enc}(F) = F_0 \cdot \mathrm{Enc}(1) + F_1 \cdot \mathrm{Enc}(Z) + F_2 \cdot \mathrm{Enc}(Z^2) + F_3 \cdot \mathrm{Enc}(Z^3)$
- Note that $\mathrm{Enc}(Z^i) = (0^i, 1^i, 2^i, \ldots, 7^i)$
- So, we can conclude that $\mathrm{Enc}(F)$ can be computed by the following matrix multiplication

$$\begin{pmatrix} F_0 & F_1 & F_2 & F_3 \end{pmatrix} \cdot \begin{pmatrix} 0^0 & 1^0 & 2^0 & \ldots & 7^0 \\ 0^1 & 1^1 & 2^1 & \ldots & 7^1 \\ 0^2 & 1^2 & 2^2 & \ldots & 7^2 \\ 0^3 & 1^3 & 2^3 & \ldots & 7^3 \end{pmatrix}$$

- The matrix $G = \begin{pmatrix} 0^0 & 1^0 & 2^0 & \ldots & 7^0 \\ 0^1 & 1^1 & 2^1 & \ldots & 7^1 \\ 0^2 & 1^2 & 2^2 & \ldots & 7^2 \\ 0^3 & 1^3 & 2^3 & \ldots & 7^3 \end{pmatrix}$ is the *generator matrix* of the code

Claim

If F is not the 0 message, then $\mathrm{Enc}(F)$ can have at most 3 zeros.

Because F is a non-zero polynomial of degree (at most) 3, it can have at most 3 zeros.

Claim

For two distinct polynomials F and G, the $\mathrm{Enc}(F)$ *and* $\mathrm{Enc}(G)$ *can match at at most* 3 *places*

Note that $\mathrm{Enc}(F-G) = \mathrm{Enc}(F) - \mathrm{Enc}(G)$, and $\mathrm{Enc}(F-G)$ can have at most 3 zeros

Claim

Given 4 evaluations of the polynomial F at distinct points, we can uniquely recover the polynomial F

Using Lagrange Interpolation

Think: Generalize this discussion to polynomials of degree $< d$ with coefficients in a field $\mathbb{F}$. The encoding evaluates the polynomial at all elements of $\mathbb{F}$.

- How long are the messages?
- How long are the codewords?
- What is the generator matrix?
- How many positions can two different codewords have identical entries?