Today we will see the main underlying problem that motivates error-correcting codes.

Then, we will introduce the basics of Fields, and

Finally, we will introduce Reed-Solomon Codes.
The sender wants to send an $n$-bit message $m$ to the receiver.

But the communication channel they are using is not reliable.

The channel flips every bit transmitted over it independently with probability $\varepsilon$.

If the sender transmits the message $m$ as is over the channel, note that if any bit gets flipped, the receiver will not receive the correct message.

So, the probability that all bits are correctly transmitted is $(1 - \varepsilon)^n$, which is exponentially low.

How can the sender reliably communicate to the receiver?
The sender has a message \( m \)

The sender uses an encoding algorithm \( \text{Enc}(\cdot) \) to compute the encoding of the message \( m \), i.e., \( c = \text{Enc}(m) \)

The message \( c \) is transmitted over the channel and the receiver receives the (possibly) altered message \( \tilde{c} \)

The receiver applies a decoding algorithm \( \text{Dec}(\cdot) \) on \( \tilde{c} \) to recover the message, i.e., \( \tilde{m} = \text{Dec}(\tilde{c}) \)

We want to ensure that the probability of correctly recovering the message is at least, say, 0.99
First Encoding Scheme: Repetition Code

- Suppose \( m \in \{0, 1\} \)
- Suppose \( \text{Enc}(m) = mmm \)
- Suppose \( \text{Dec}(\tilde{m}_1, \tilde{m}_2, \tilde{m}_3) = \text{maj}(\tilde{m}_1, \tilde{m}_2, \tilde{m}_3) \)

Note that the probability that the message is correctly recovered is:

\[
\binom{3}{0}(1 - \varepsilon)^3 + \binom{3}{1}\varepsilon(1 - \varepsilon)^2
\]

In this case, the encoding function repeated the input message 3 times.

Think: Given \( \varepsilon \) and the probability of successful transmission 0.99, how many times should the encoding function repeat the message?
Suppose the receiver receives the erroneous string $\tilde{c}$ from the channel.

What message should it decode to?

The best decoding algorithm (ignoring efficiency of the decoding algorithm) is the Maximum Likelihood Decoding.

Let $\mathcal{M}$ be the set of all messages.

Suppose the message $m$ is encoded as $\text{Enc}(m)$ by the encoding function.

We can compute the probability $p(\tilde{c}|\text{Enc}(m))$, i.e., the probability that the channel input $c = \text{Enc}(m)$ was altered into the received string $\tilde{c}$.

Output $m \in \mathcal{M}$ such that $p(\tilde{c}|\text{Enc}(m))$ is maximum.

For specific codes, there are algorithms that are more efficient.
Note that when $\varepsilon = 1/2$, there is no way to reliably transmit a message, because all messages are equally likely conditioned on the received string $\tilde{c}$.

Note that when $\varepsilon = 0$, it is trivial to transmit messages reliably.

As $\varepsilon$ increases from 0 to 1/2, we expect the task of transmitting message to get “increasingly difficult.” Alternately, their reliability continues to decrease.

When $\varepsilon > 1/2$ the starts to get more “reliable!” Note that $\varepsilon = \delta$ and $\varepsilon = 1 - \delta$ are (roughly) “identical channels” and, intuitively, their qualities are identical.

When $\varepsilon = 1$, it is again trivial to transmit messages over the channel.
Abstract Algebra: Fields

A field \((\mathbb{F}, +, \cdot)\) is a set of elements \(\mathbb{F}\) endowed with two operations + (addition) and \(\cdot\) (multiplication) that satisfies the following conditions:

- **Closure:** For all \(a, b \in \mathbb{F}\), we have \(a + b \in \mathbb{F}\) and \(a \cdot b \in \mathbb{F}\).
- **Commutativity:** For all \(a, b \in \mathbb{F}\), we have \(a + b = b + a\) and \(a \cdot b = b \cdot a\).
- **Associativity:** For all \(a, b, c \in \mathbb{F}\), we have \((a + b) + c = a + (b + c)\) and \((a \cdot b) \cdot c = a \cdot (b \cdot c)\).
- **Identities:** There exists unique elements 0, 1 \(\in \mathbb{F}\) such that, for all \(a \in \mathbb{F}\), we have \(a + 0 = a\) and \(a \cdot 1 = a\).
- **Inverses:** For every \(a \in \mathbb{F}\), there exists a unique element \((-a) \in \mathbb{F}\) such that \(a + (-a) = 0\), and for every \(a \in \mathbb{F}\), if \(a \neq 0\), there exists a unique element \(a^{-1} \in \mathbb{F}\) such that \(a \cdot a^{-1} = 1\).
- **Distributivity:** For every \(a, b, c \in \mathbb{F}\), we have \(a \cdot (b + c) = a \cdot b + a \cdot c\).
Example: Infinite Fields

Let \( \mathbb{Q} \) be the set of all rationals. Then \((\mathbb{Q}, +, \cdot)\) is a field, where the operations are defined as follows:

- \( \frac{a}{b} + \frac{c}{d} = \frac{ad + bc}{bd} \), and
- \( \frac{a}{b} \cdot \frac{c}{d} = \frac{ac}{bd} \).

Note that \((\mathbb{Z}, +, \cdot)\) is not a field, where \( \mathbb{Z} \) is the set of all integers.

Note that \((\mathbb{C}, +, \cdot)\) is a field, where \( \mathbb{C} \) is the set of all complex numbers.
Example: Finite Fields

Let \( \mathbb{Z}_p \), represent the set of all integers \( \{0, \ldots, p - 1\} \). For prime \( p \), \((\mathbb{Z}_p, +, \cdot)\) is a finite field where we define:

- \( a + b = (a + b) \mod p \) (i.e., integer addition mod \( p \)), and
- \( a \cdot b = (ab) \mod p \) (i.e., integer multiplication mod \( p \)).

The only non-triviality is to argue that every \( a \in \mathbb{Z}_p \) such that \( a \neq 0 \) has a unique inverse. The proof is left as an exercise. Hint: Show that \( a^{p-2} \) is the inverse of \( a \).
Let $n = p^\alpha$, where $p$ is a prime and $\alpha$ is a positive integer.

Let $\mathbb{F}$ be the set of all polynomials in $X$ of degree $< \alpha$ such that the coefficients of each term in the polynomial is in $\mathbb{Z}_p$.

So, the tuple $(a_0, \ldots, a_{\alpha-1}) \in \mathbb{Z}_p^\alpha$ can be equivalently interpreted as the polynomial $\sum_{i=0}^{\alpha-1} a_i X^i$.

So, elements of $\mathbb{F}$ can be interpreted either as the tuple $(a_0, \ldots, a_{\alpha-1})$ or the polynomial $\sum_{i=0}^{\alpha-1} a_i X^i$.

The sum of two polynomials is defined as follows:

$$(a_0, \ldots, a_{\alpha-1}) + (b_0, \ldots, b_{\alpha-1}) := (a_0 + b_0, \ldots, a_{\alpha-1} + b_{\alpha-1})$$
Let \( \Pi(X) \) be a monic polynomial with degree \( \alpha \) and coefficients in \( \mathbb{Z}_p \). Suppose \( \Pi(X) \) does not have any roots in \( \mathbb{Z}_p \).

The product of two polynomials \( (a_0, \ldots, a_{\alpha-1}) \) and \( (b_0, \ldots, b_{\alpha-1}) \) is given by the polynomial

\[
\left( \sum_{i=0}^{\alpha-1} a_i X^i \right) \cdot \left( \sum_{i=0}^{\alpha-1} b_i X^i \right) \mod \Pi(X)
\]

Think: What is the unique inverse of the polynomial represented by \( (a_0, \ldots, a_{\alpha-1}) \)?
Suppose we want to define a field of size $8 = 2^3$

We have $p = 2$ and $\alpha = 3$

So, $\mathbb{F}$ is the following set

$$\left\{ 0, 1, X, X + 1, X^2, X^2 + 1, X^2 + X, X^2 + X + 1 \right\}$$

We use the irreducible polynomial $\Pi(X) = X^3 + X + 1$

Sum of two polynomials is defined naturally

Product of two polynomials is defined by multiplying them and then taking $\mod \Pi(X)$

What are the inverses of each element in $\mathbb{F}$
Suppose the message is \((m_0, \ldots, m_{k-1}) \in \mathbb{F}^k\)

Consider the polynomial \(M(Z) = \sum_{i=0}^{k-1} Z^i\)

Let \(\mathbb{F} = \{e_0, e_1, \ldots, e_{|\mathbb{F}|-1}\}\)

The encoding of \((m_0, \ldots, m_{k-1})\) is defined to be

\[
\left( M(e_0), M(e_1), \ldots, M(e_{|\mathbb{F}|-1}) \right)
\]

Think: “Sum of two different codewords” is the codeword corresponding to the “sum of the two corresponding messages”
Think: Two different codewords have Hamming distance at least $|\mathbb{F}| - (k - 1)$. If the Hamming distance is $\leq |\mathbb{F}| - k$, then the difference of the codewords has $\geq k$ zeros. But a degree $(k - 1)$ polynomial can have at most $(k - 1)$ zeros, unless it is the zero-polynomial. So, the difference of the two codewords is the evaluation of the zero-polynomial at the field elements. This implies that the corresponding messages were identical. Hence contradiction.