Algebraic curves, error correcting codes and post-quantum cryptography

Gábor P. Nagy

Department of Algebra, Budapest University of Technology and Economics (Hungary)

Mathematical Modelling Seminar
Oct 29, 2019

Outline

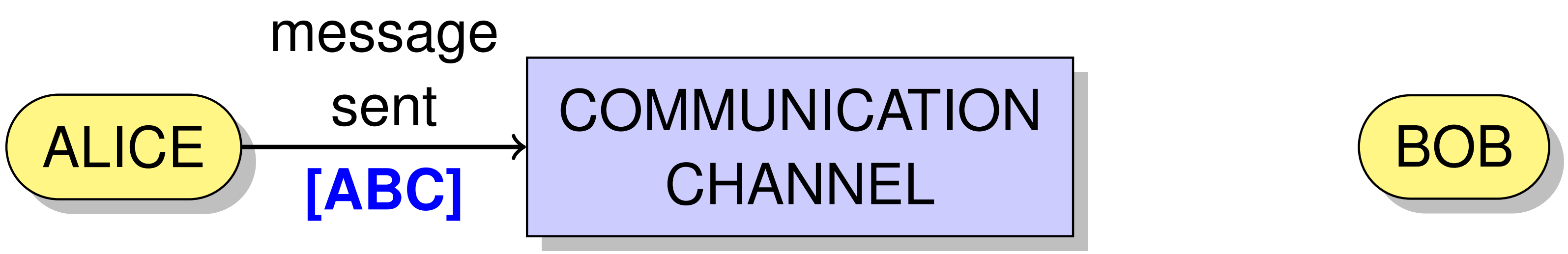
1. Communication on noisy channels
2. Error correction codes
3. Algebraic-geometric codes
4. Post-quantum cryptography

Outline

The scheme of communication

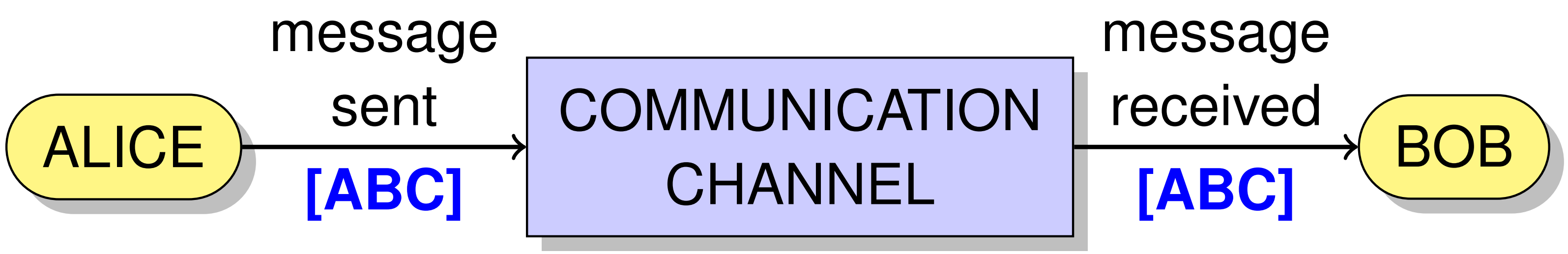
- The message can be: *text, picture, sound, measurement data, etc.*
- The communication channel can be: *one way, two way, data transmission, data storage, etc.*

The scheme of communication



- The message can be: *text, picture, sound, measurement data, etc.*
- The communication channel can be: *one way, two way, data transmission, data storage, etc.*

The scheme of communication



- The message can be: *text, picture, sound, measurement data, etc.*
- The communication channel can be: *one way, two way, data transmission, data storage, etc.*

The scheme of communication

- The message can be: *text, picture, sound, measurement data, etc.*
- The communication channel can be: *one way, two way, data transmission, data storage, etc.*

The scheme of communication

- The message can be: *text, picture, sound, measurement data, etc.*
- The communication channel can be: *one way, two way, data transmission, data storage, etc.*

The scheme of communication

- The message can be: *text, picture, sound, measurement data, etc.*
- The communication channel can be: *one way, two way, data transmission, data storage, etc.*

Digitization, digital reformatting

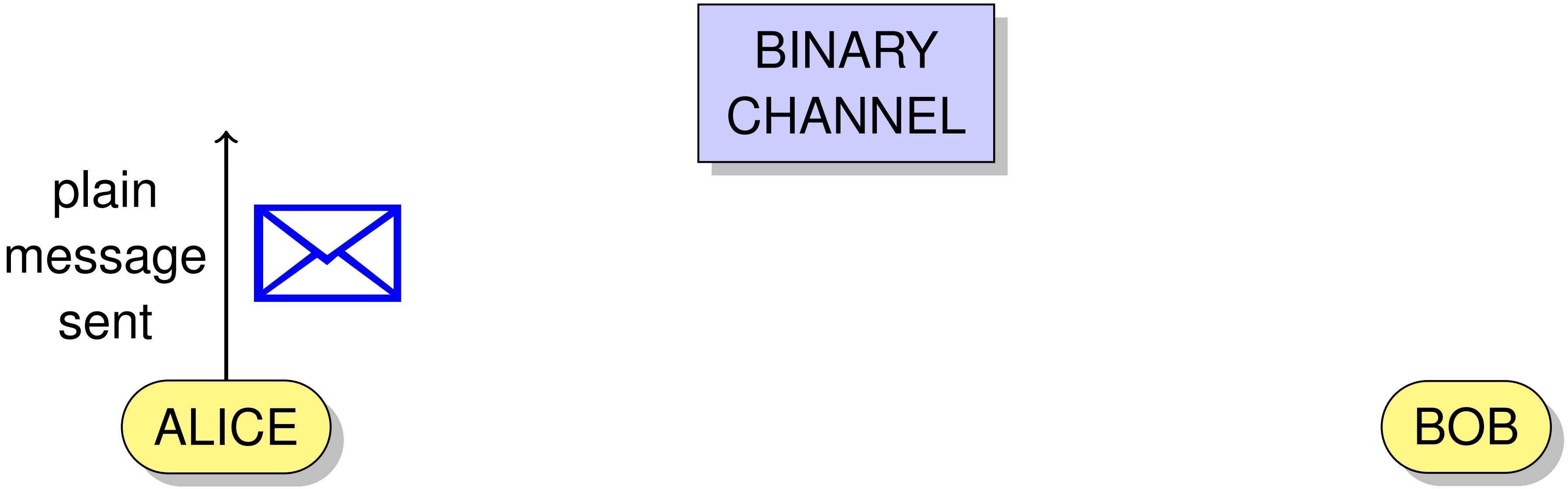
INFORMATION
text, picture, voice, . . .

DIGITALIZATION ⟷

0-1 SEQUENCES
bits, bytes, . . .

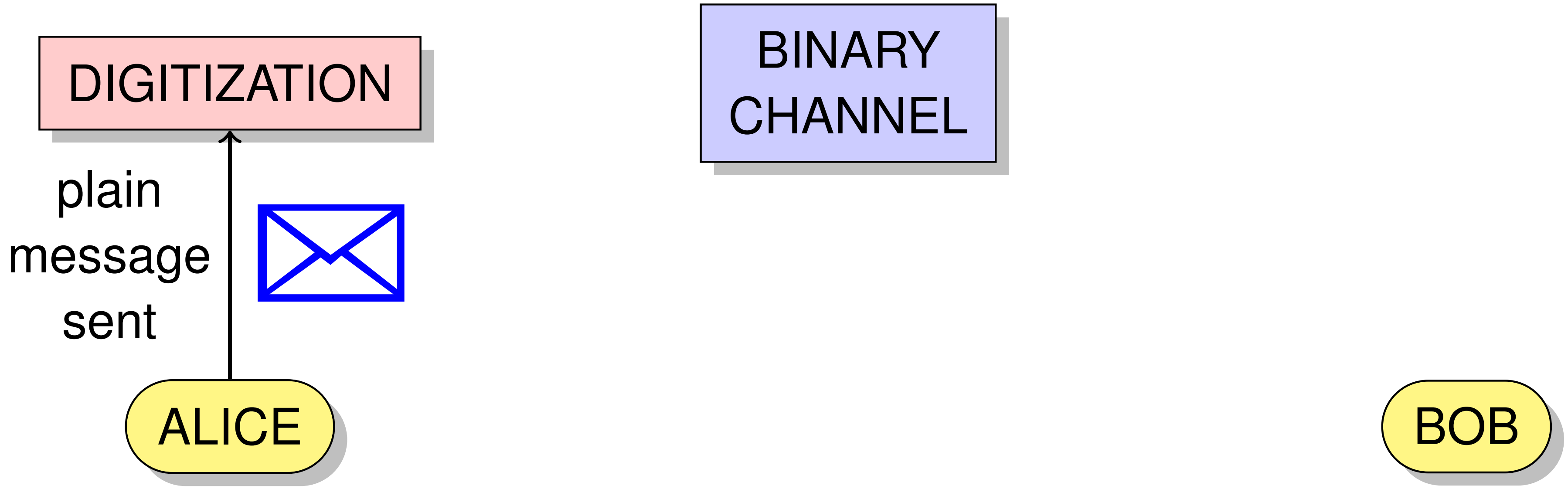
John von Neumann
(1903-1957)
Hungarian
mathematician

Digitization, digital reformatting (cont.)



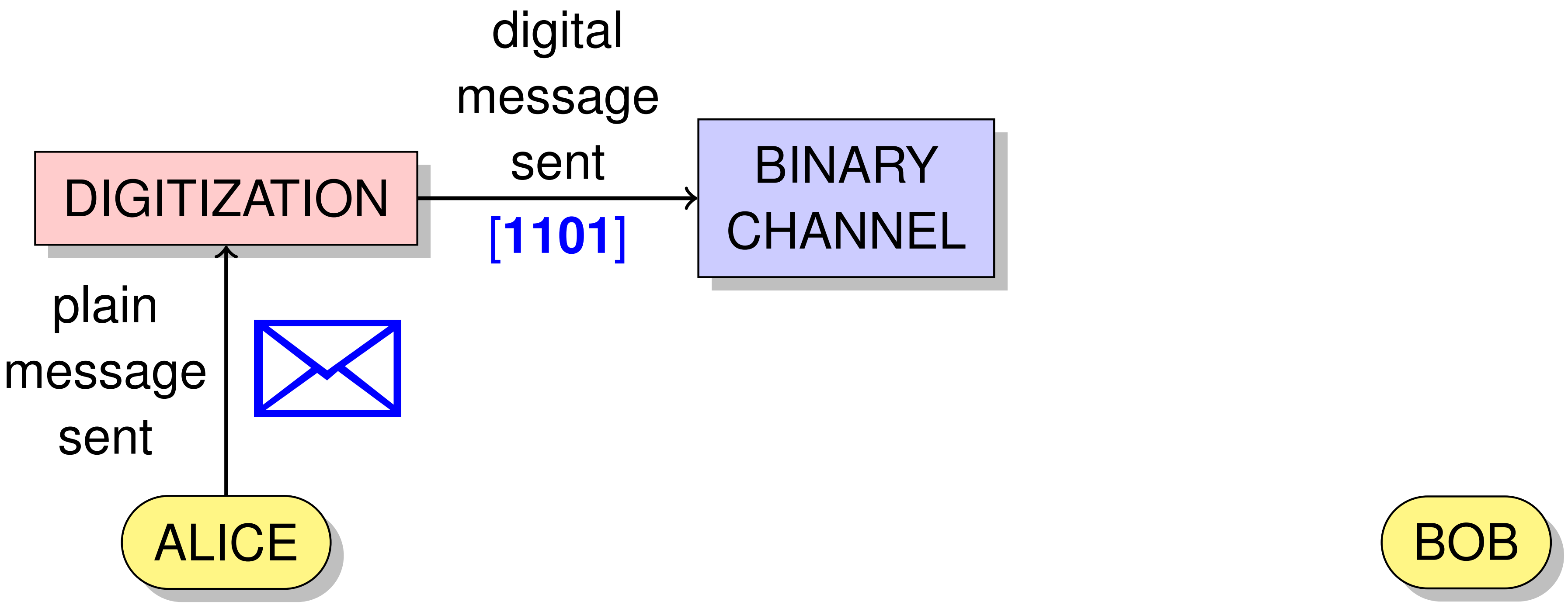
- We are not interested in different digitization techniques.
- We will assume that our messages are 0/1 sequences of fixed length.

Digitization, digital reformatting (cont.)



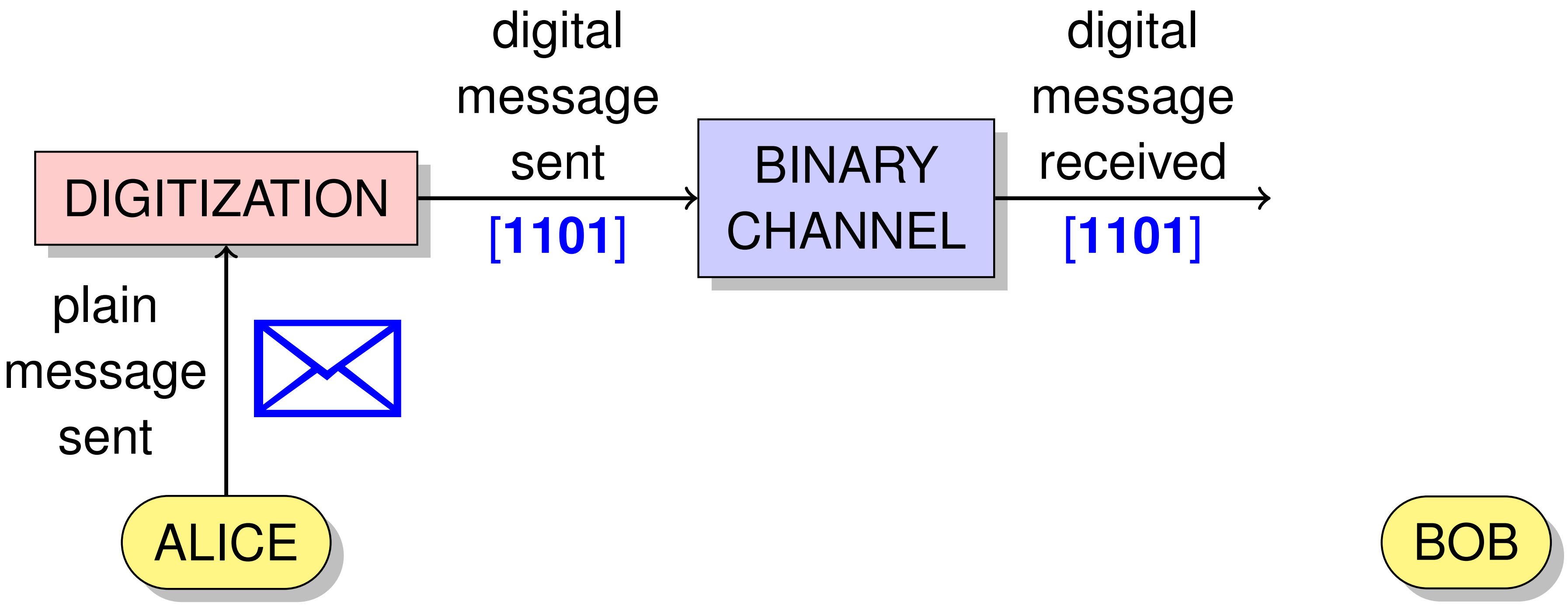
- We are not interested in different digitization techniques.
- We will assume that our messages are 0/1 sequences of fixed length.

Digitization, digital reformatting (cont.)



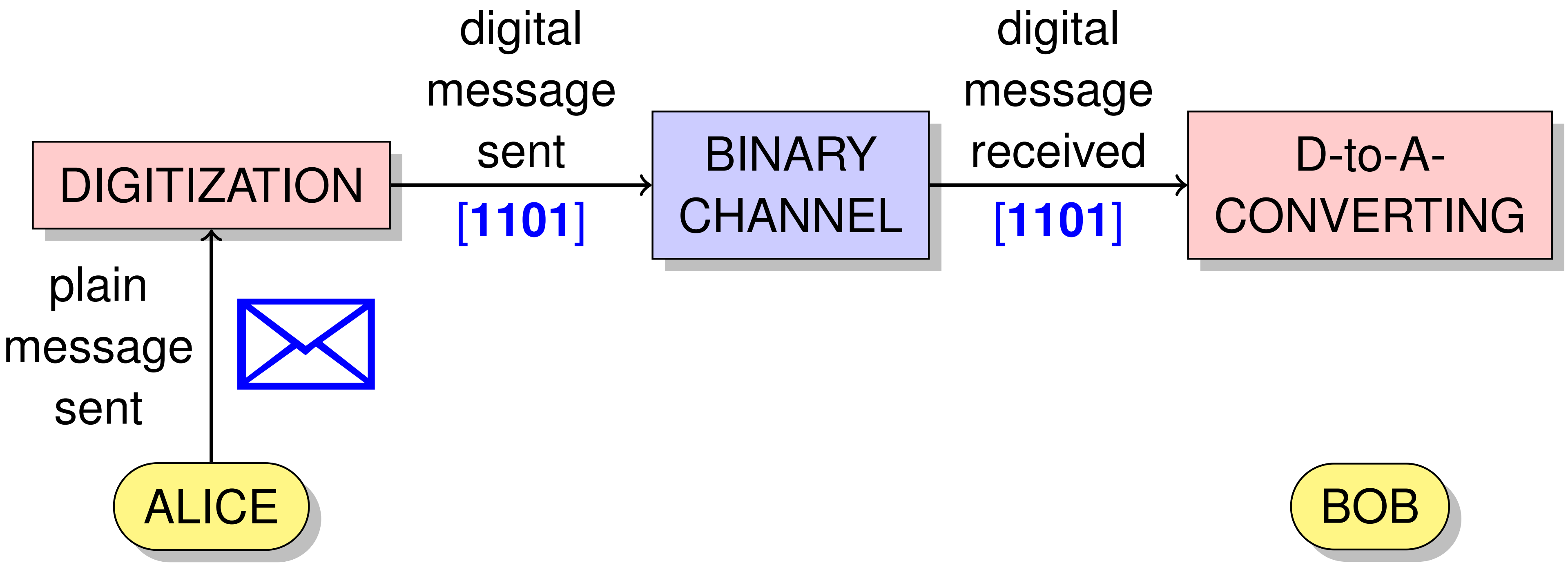
- We are not interested in different digitization techniques.
- We will assume that our messages are 0/1 sequences of fixed length.

Digitization, digital reformatting (cont.)



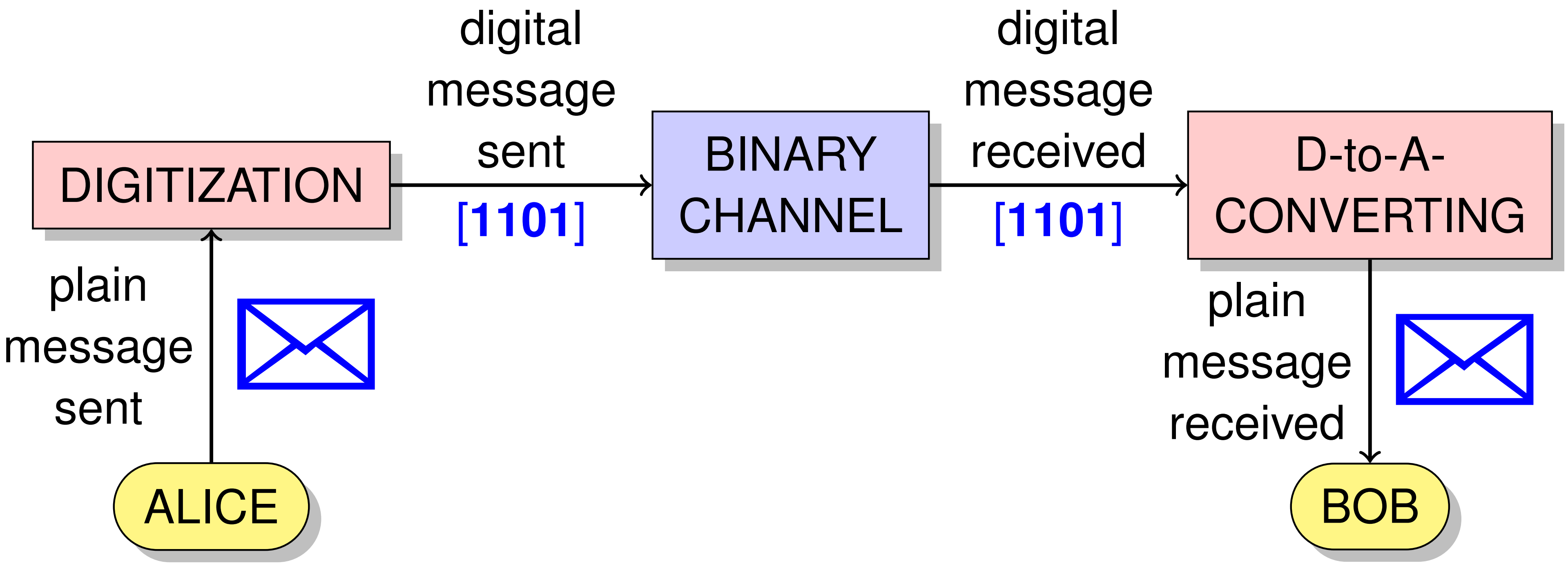
- We are not interested in different digitization techniques.
- We will assume that our messages are 0/1 sequences of fixed length.

Digitization, digital reformatting (cont.)



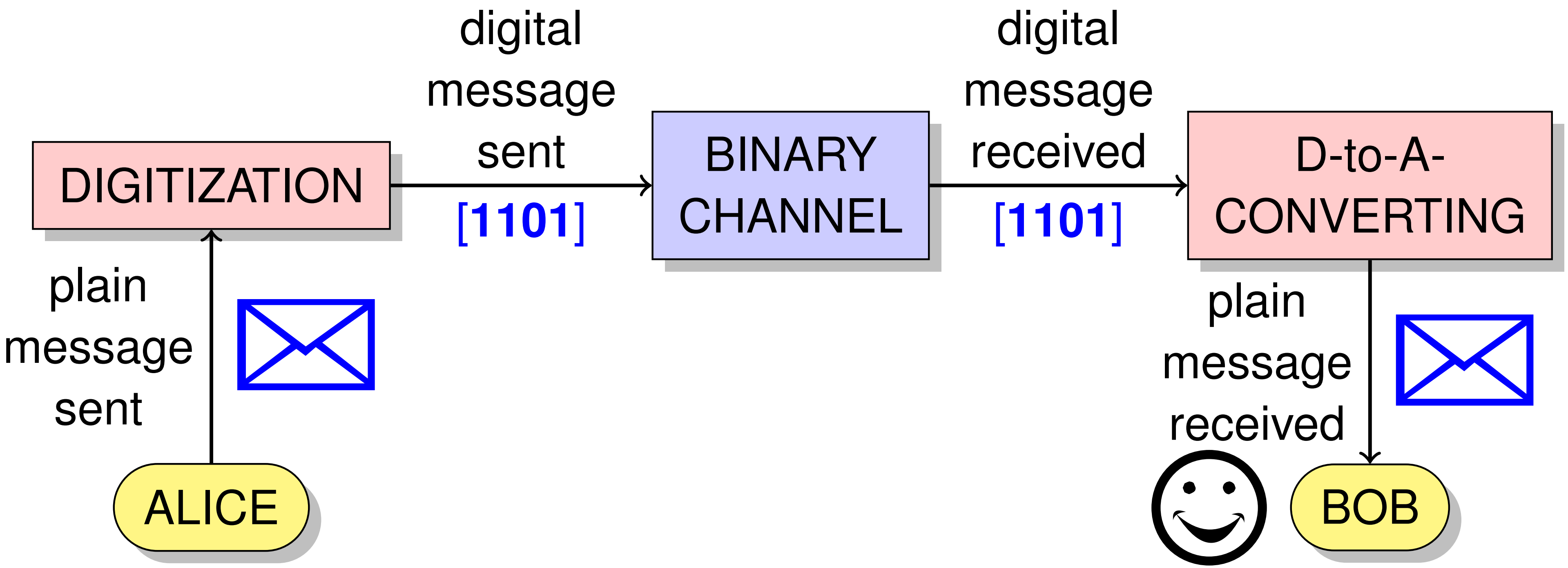
- We are not interested in different digitization techniques.
- We will assume that our messages are 0/1 sequences of fixed length.

Digitization, digital reformatting (cont.)



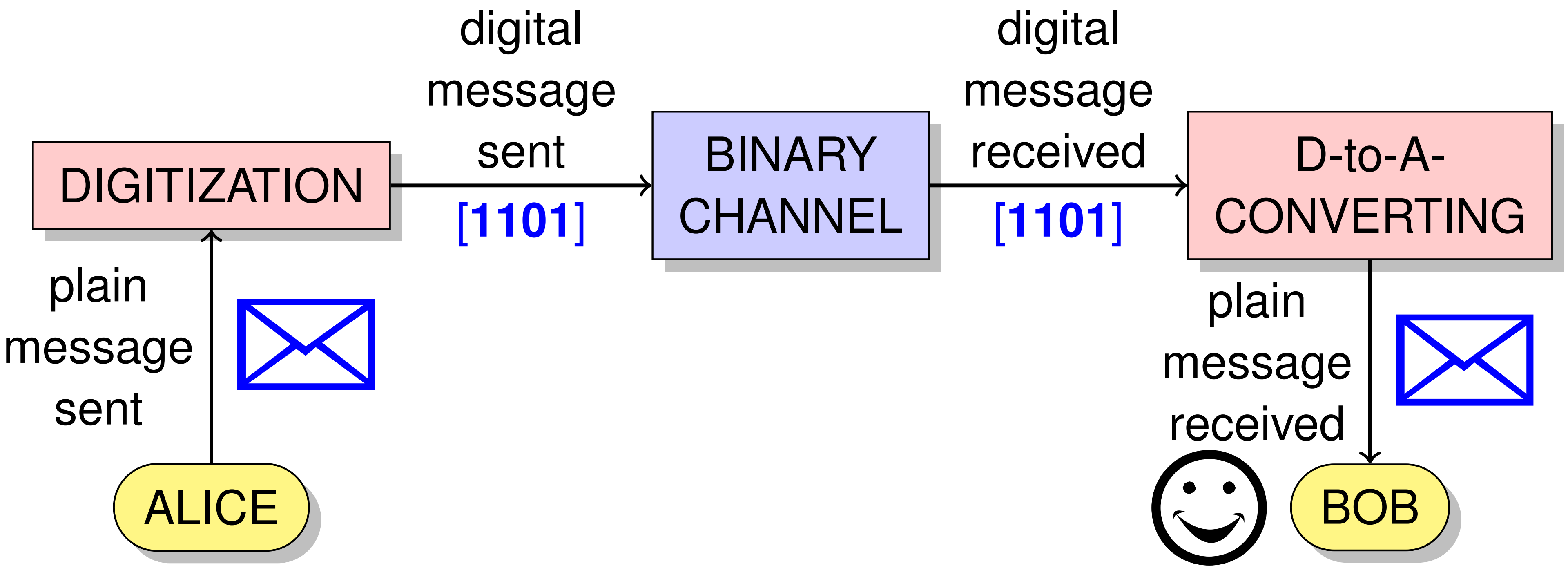
- We are not interested in different digitization techniques.
- We will assume that our messages are 0/1 sequences of fixed length.

Digitization, digital reformatting (cont.)



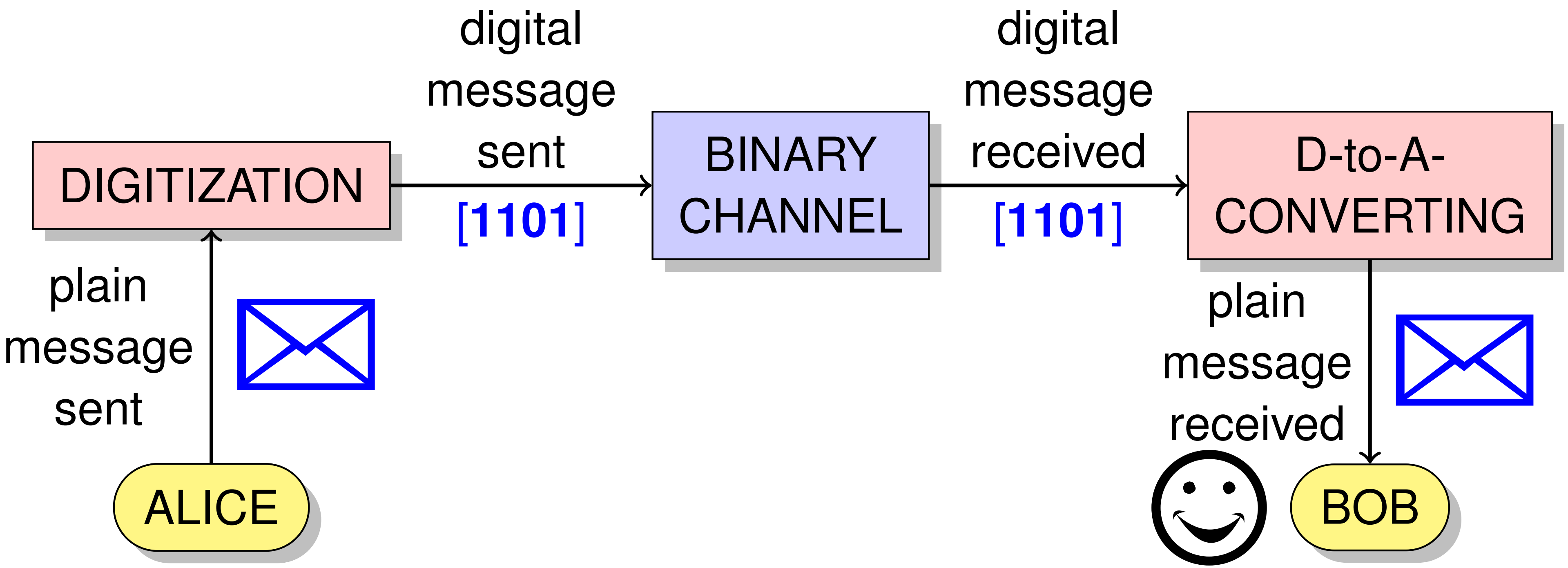
- We are not interested in different digitization techniques.
- We will assume that our messages are 0/1 sequences of fixed length.

Digitization, digital reformatting (cont.)



- We are not interested in different digitization techniques.
- We will assume that our messages are 0/1 sequences of fixed length.

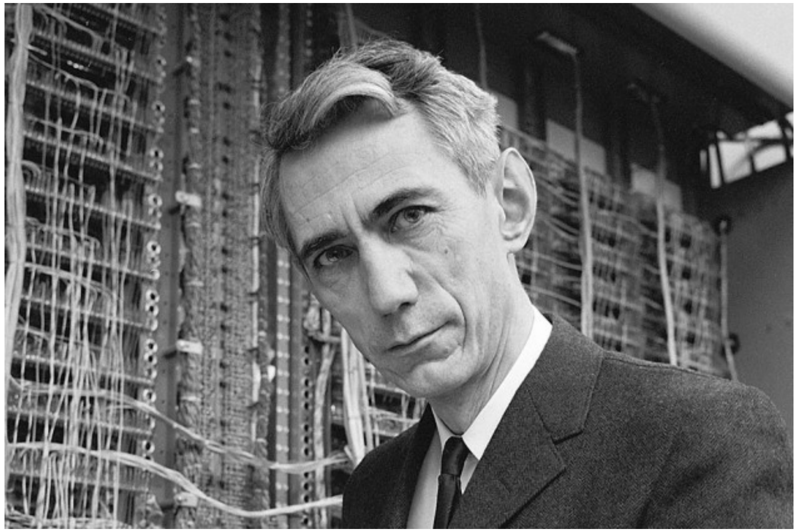
Digitization, digital reformatting (cont.)

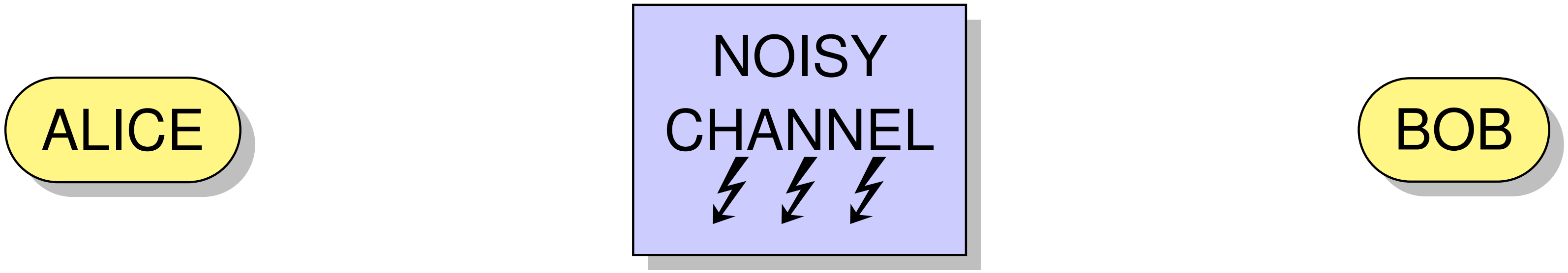


- We are not interested in different digitization techniques.
- We will assume that our messages are 0/1 sequences of fixed length.

Communication on noisy channel

Claude Shannon
(1916-2001)
US mathematician

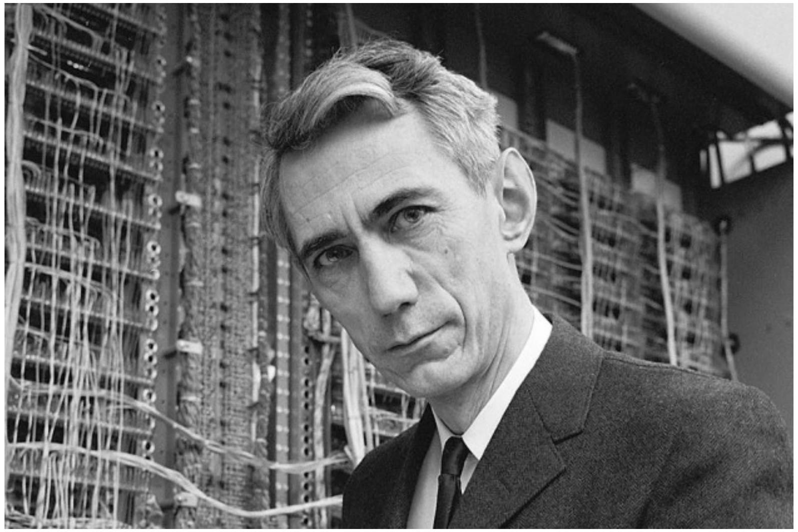


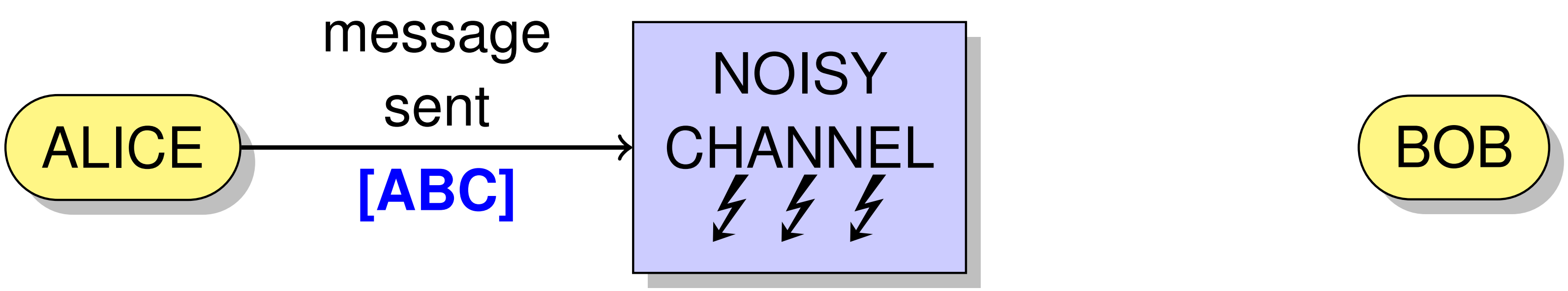


- Simple noise modell: **Binary Symmetric Channel** with fixed bit error ratio.

Communication on noisy channel

Claude Shannon
(1916-2001)
US mathematician

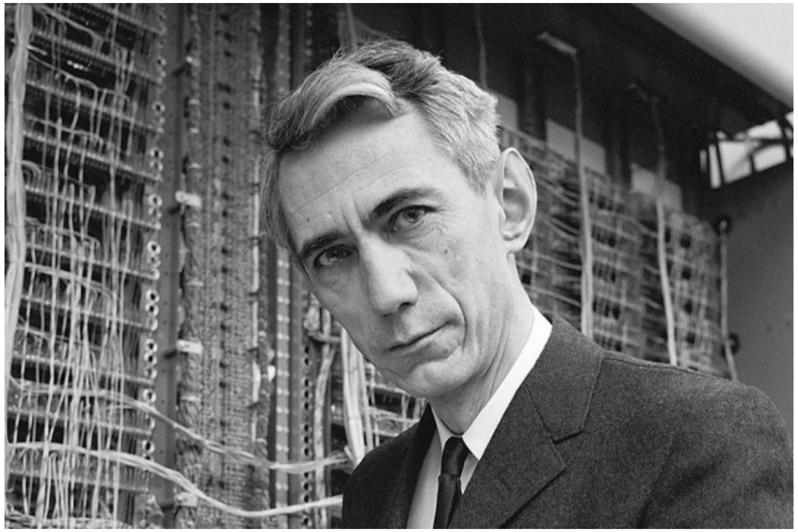


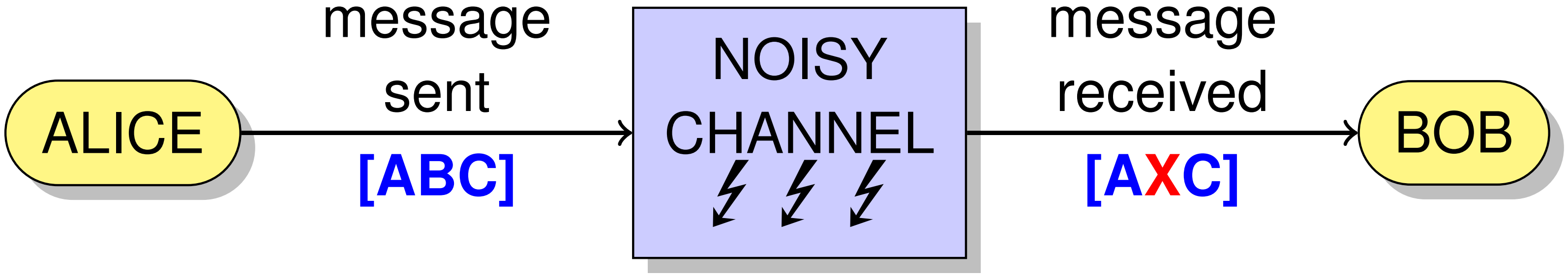


- Simple noise modell: **Binary Symmetric Channel** with fixed bit error ratio.

Communication on noisy channel

Claude Shannon
(1916-2001)
US mathematician

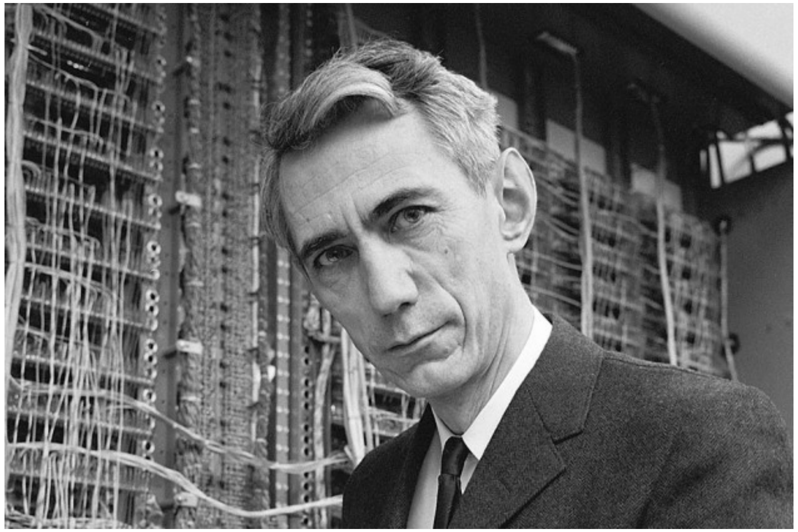


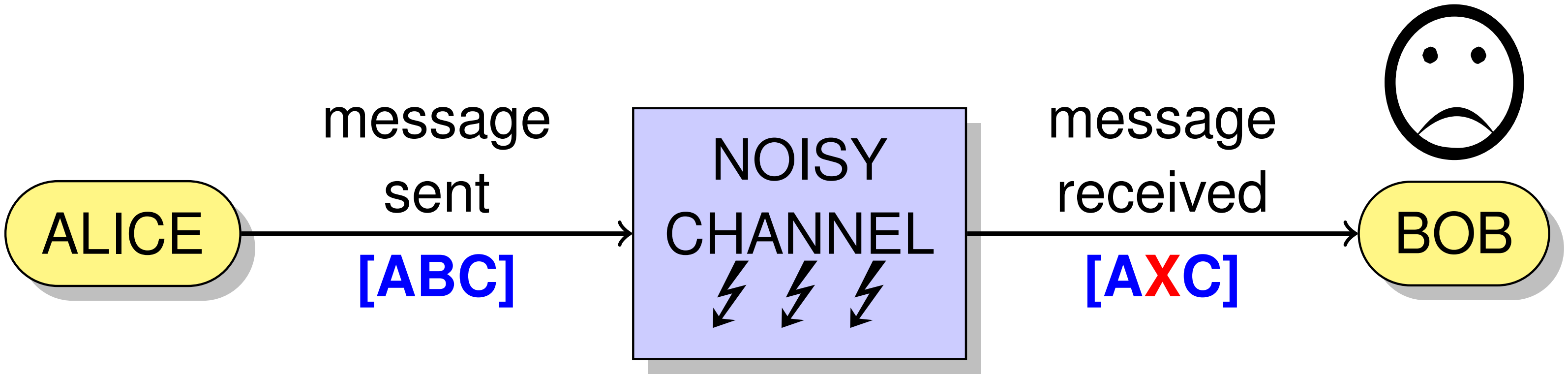


- Simple noise modell: **Binary Symmetric Channel** with fixed bit error ratio.

Communication on noisy channel

Claude Shannon
(1916-2001)
US mathematician

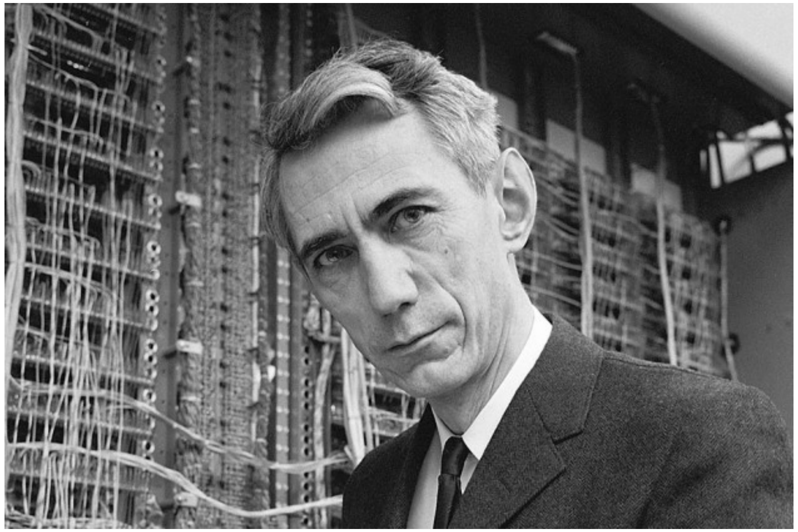


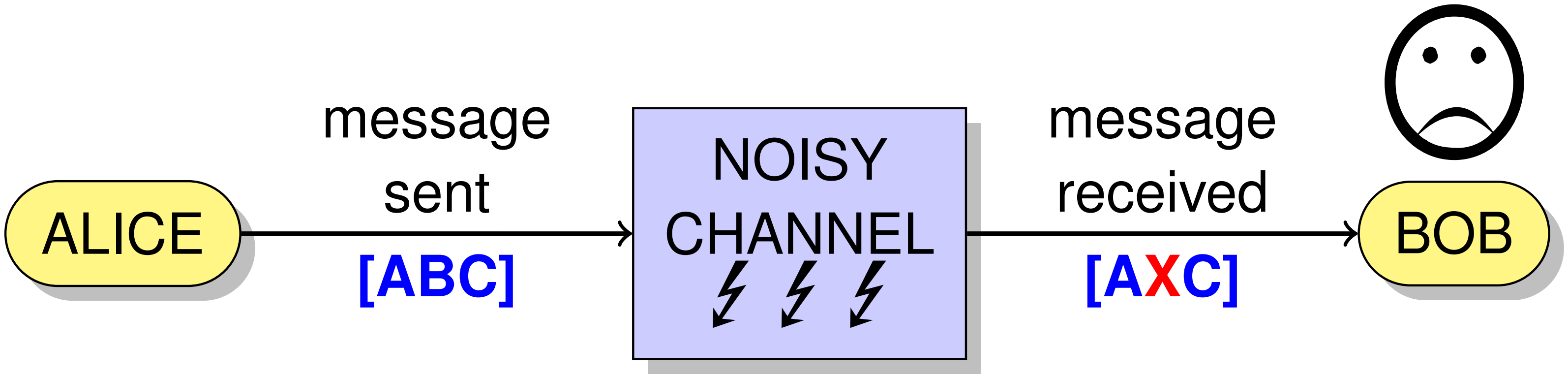


- Simple noise modell: **Binary Symmetric Channel** with fixed bit error ratio.

Communication on noisy channel

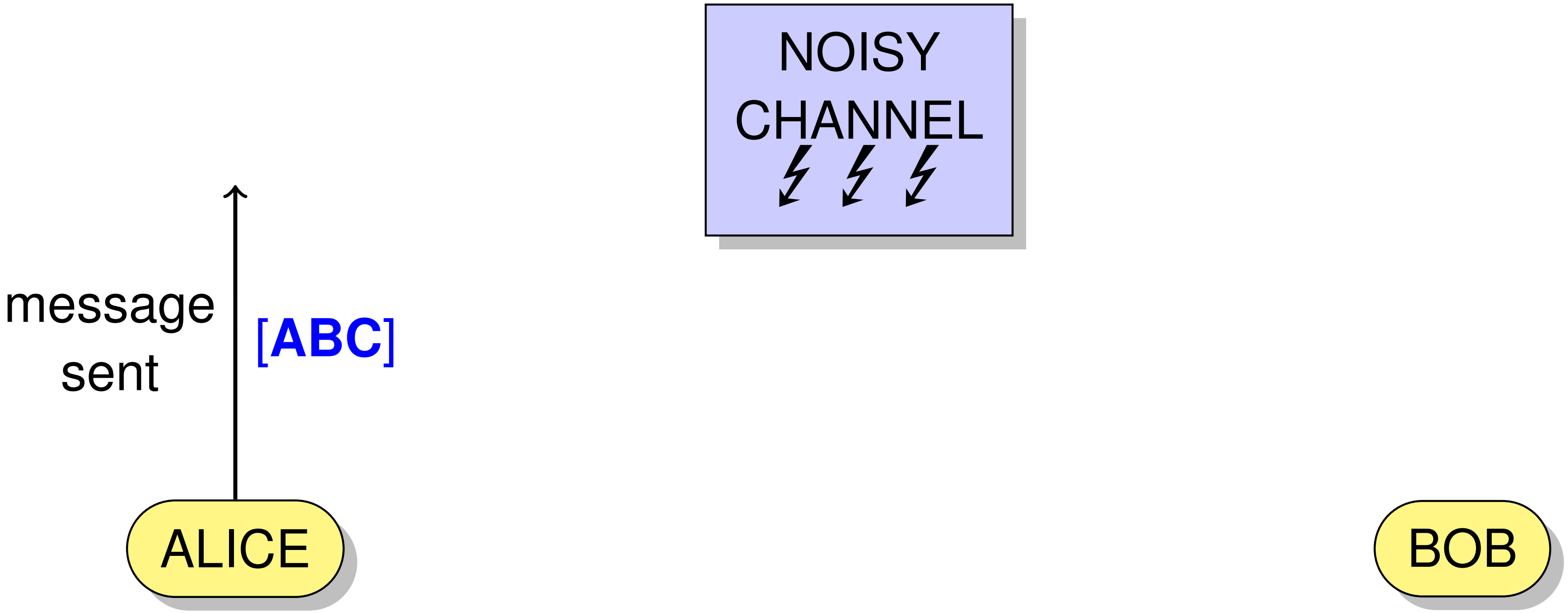
Claude Shannon
(1916-2001)
US mathematician





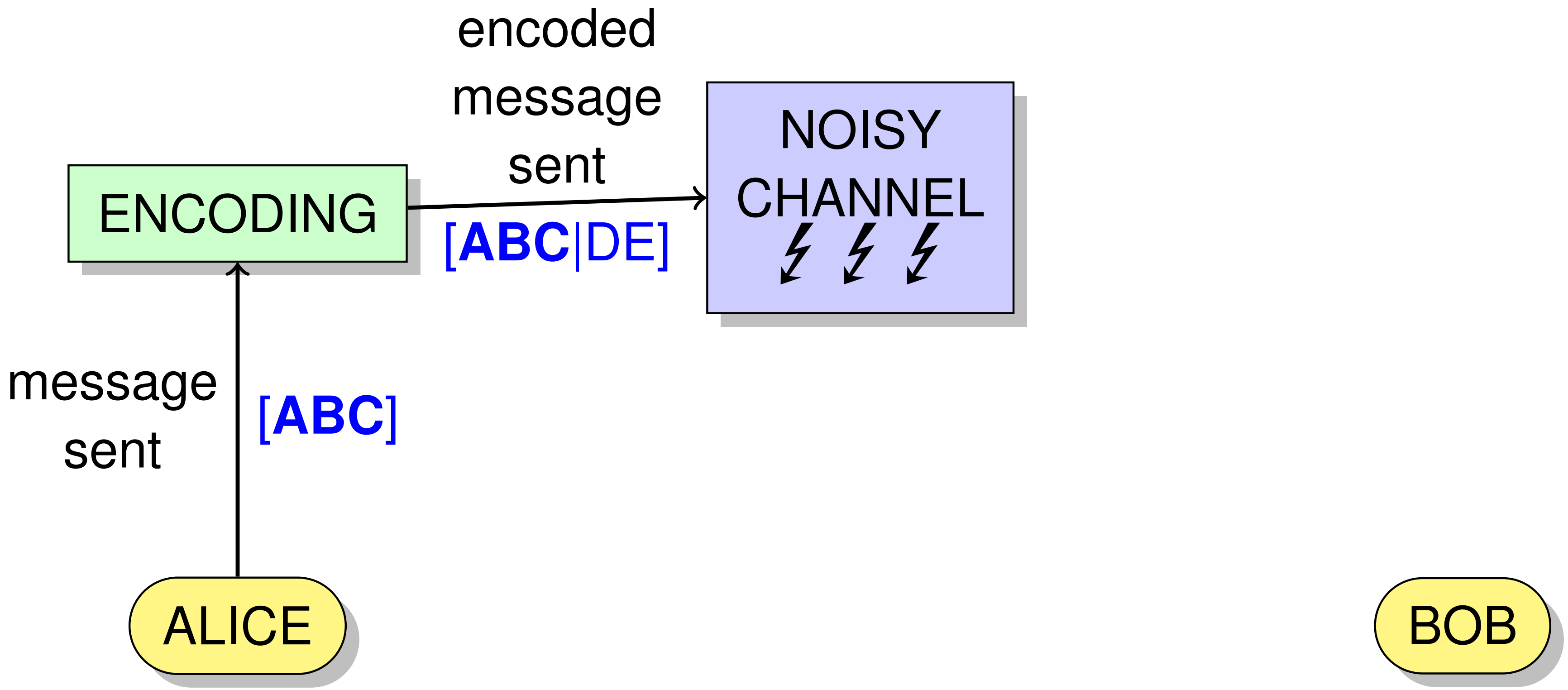
- Simple noise modell: **Binary Symmetric Channel** with fixed bit error ratio.

Error correction on noisy communication channel



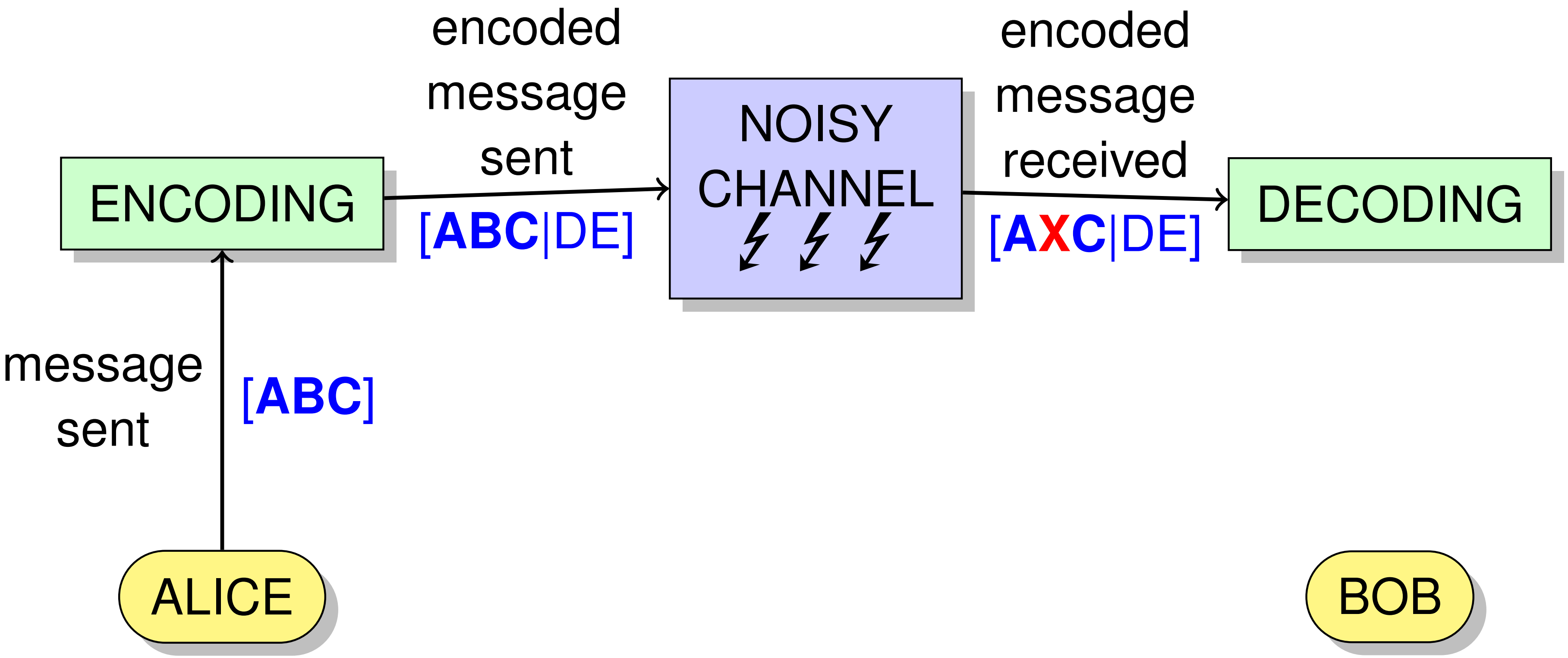
- Example: **3-fold repetition code:** $0 \mapsto 0|00$, $1 \mapsto 1|11$.
- **Majority/Nearest neighbor/Maximum likelihood** Encoding:
  $0|00, 1|00, 0|10, 0|01 \mapsto 0|00 \mapsto \mathbf{0}$
  $1|10, 1|01, 0|11, 1|11 \mapsto 1|11 \mapsto \mathbf{1}$.

Error correction on noisy communication channel



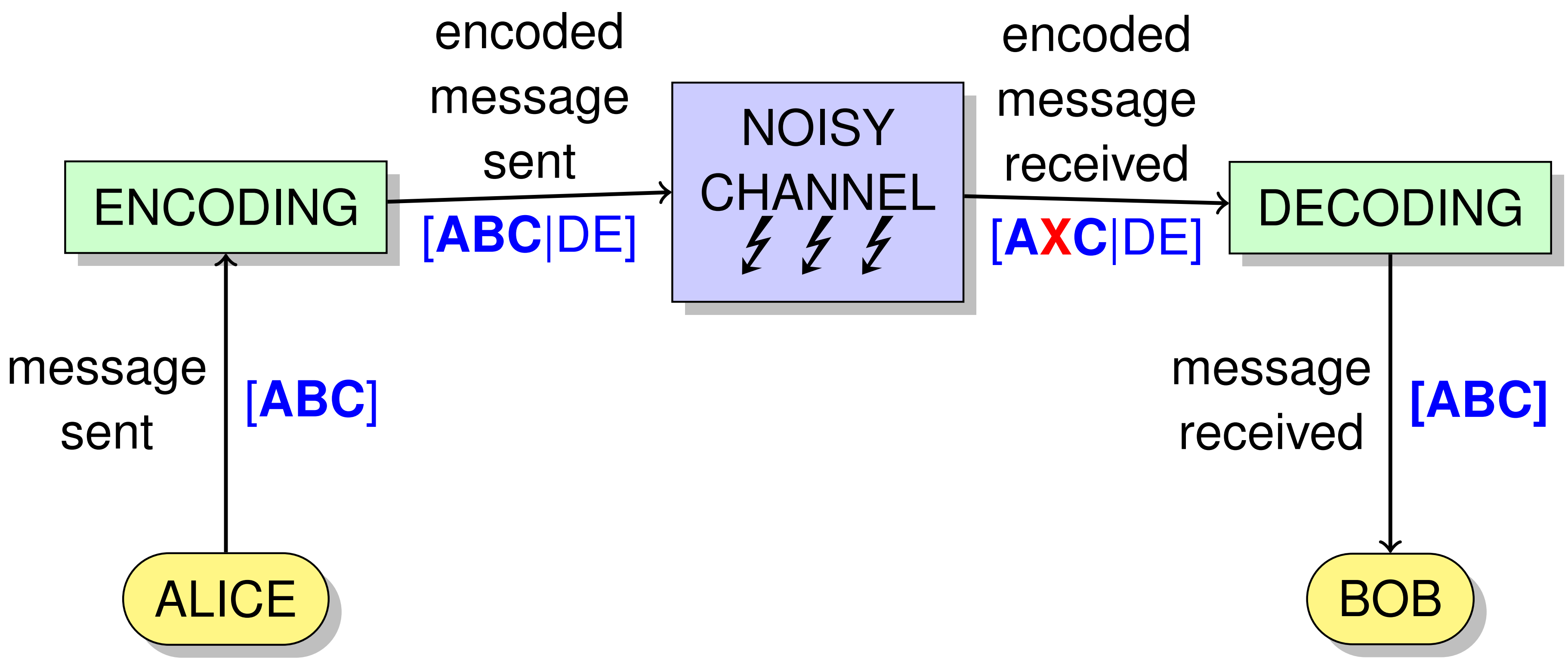
- Example: **3-fold repetition code:** $0 \mapsto 0|00,\ 1 \mapsto 1|11$.
- **Majority/Nearest neighbor/Maximum likelihood** Encoding:
  $0|00, 1|00, 0|10, 0|01 \mapsto 0|00 \mapsto \mathbf{0}$
  $1|10, 1|01, 0|11, 1|11 \mapsto 1|11 \mapsto \mathbf{1}$.

Error correction on noisy communication channel



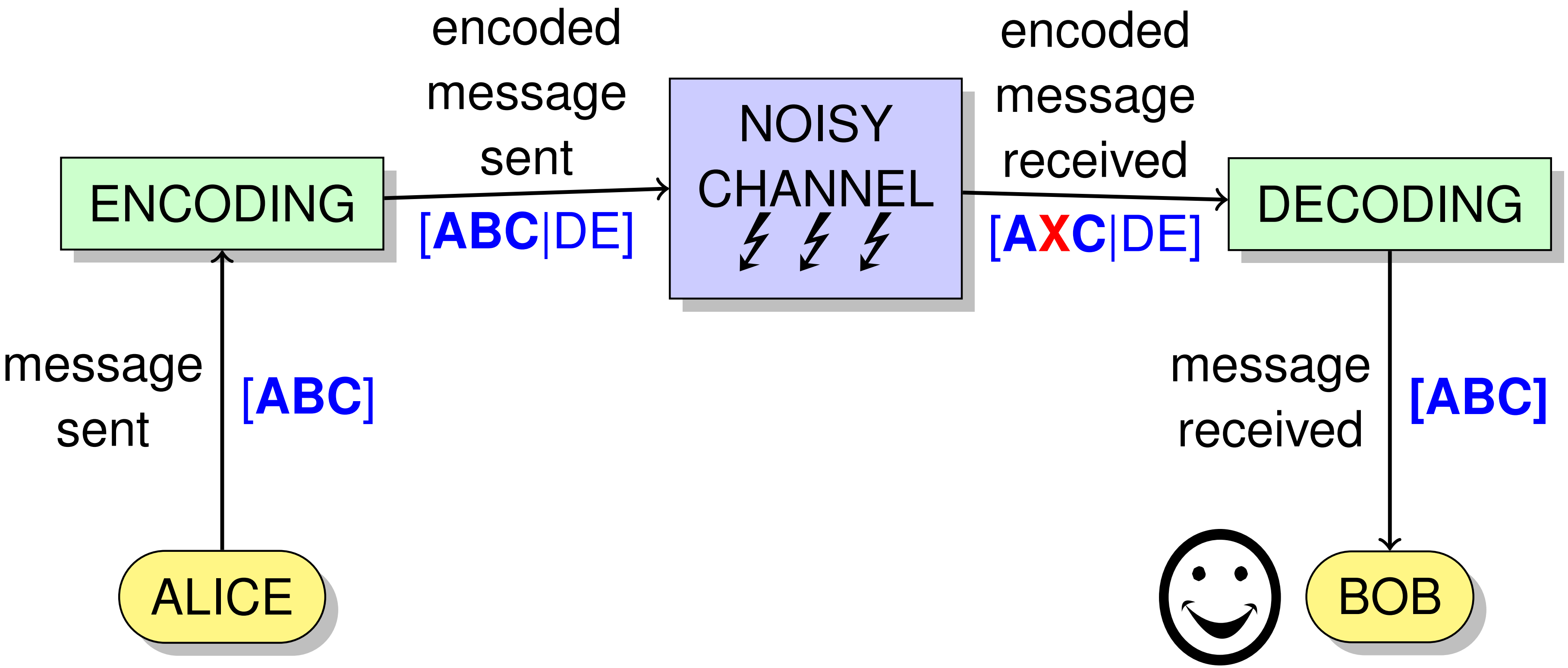
- Example: **3-fold repetition code:** $0 \mapsto 0|00$, $1 \mapsto 1|11$.
- **Majority/Nearest neighbor/Maximum likelihood** Encoding:
  $0|00, 1|00, 0|10, 0|01 \mapsto 0|00 \mapsto \mathbf{0}$
  $1|10, 1|01, 0|11, 1|11 \mapsto 1|11 \mapsto \mathbf{1}$.

Error correction on noisy communication channel



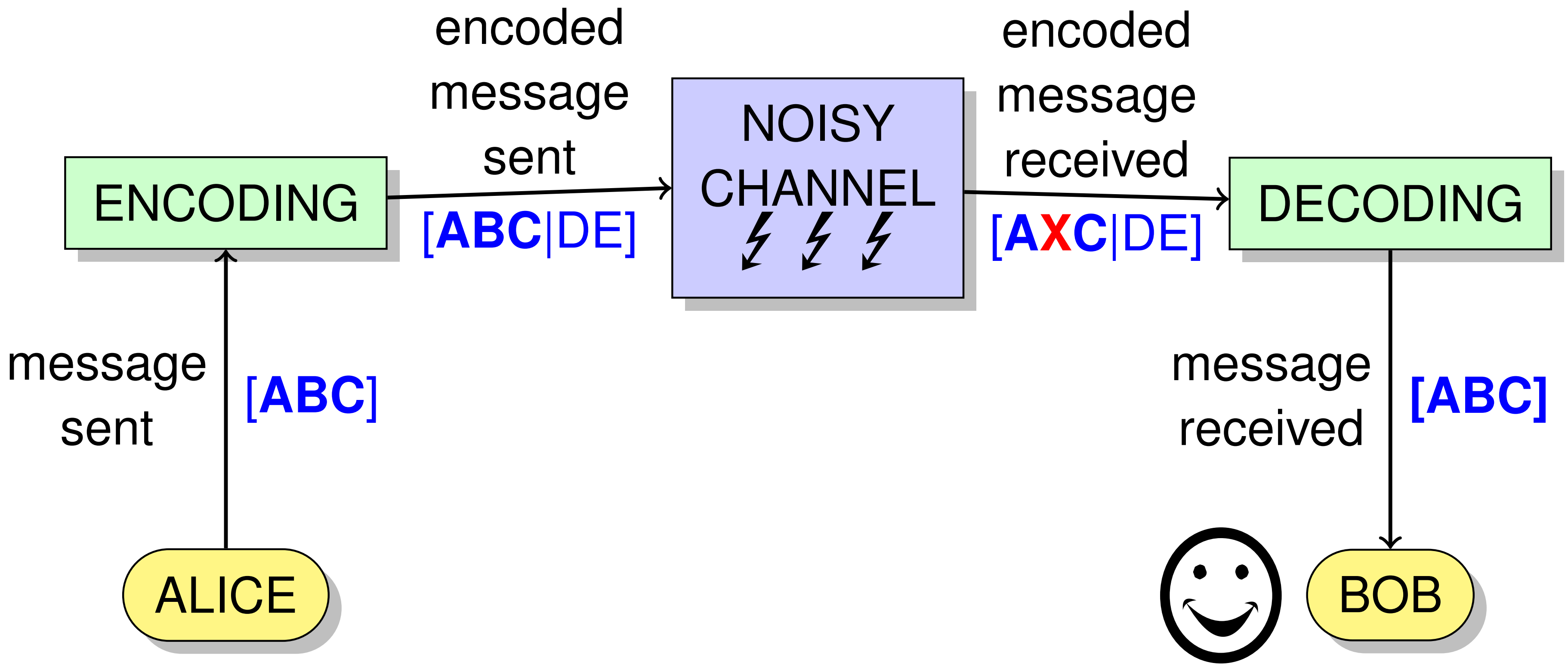
- Example: **3-fold repetition code:** $0 \mapsto 0|00$, $1 \mapsto 1|11$.
- **Majority/Nearest neighbor/Maximum likelihood** Encoding:
  $0|00, 1|00, 0|10, 0|01 \mapsto 0|00 \mapsto \mathbf{0}$
  $1|10, 1|01, 0|11, 1|11 \mapsto 1|11 \mapsto \mathbf{1}$.

Error correction on noisy communication channel



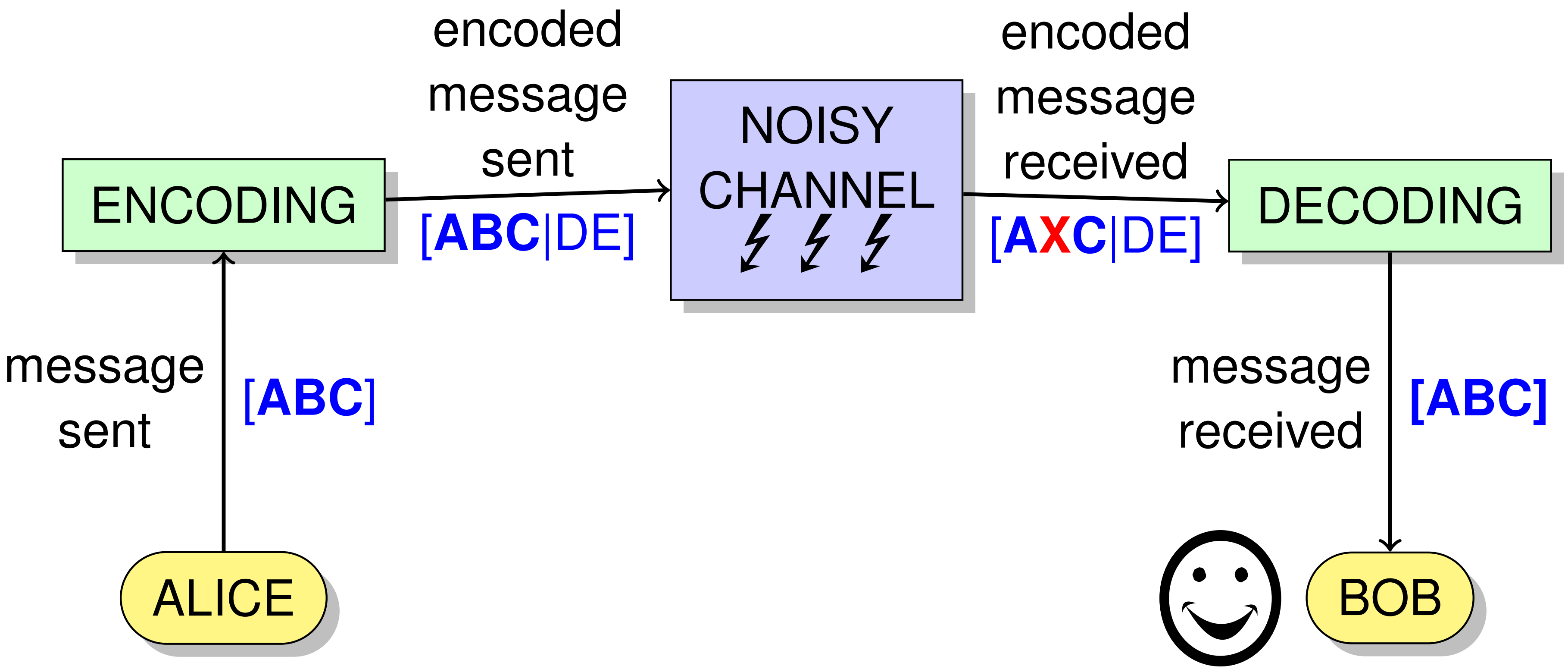
- Example: **3-fold repetition code:** 0 ↦ 0|00, 1 ↦ 1|11.
- **Majority/Nearest neighbor/Maximum likelihood** Encoding:
  0|00, 1|00, 0|10, 0|01 ↦ 0|00 ↦ **0**
  1|10, 1|01, 0|11, 1|11 ↦ 1|11 ↦ **1**.

Error correction on noisy communication channel



- Example: **3-fold repetition code:** $0 \mapsto 0|00$, $1 \mapsto 1|11$.
- **Majority/Nearest neighbor/Maximum likelihood** Encoding:
  $0|00, 1|00, 0|10, 0|01 \mapsto 0|00 \mapsto \mathbf{0}$
  $1|10, 1|01, 0|11, 1|11 \mapsto 1|11 \mapsto \mathbf{1}$.

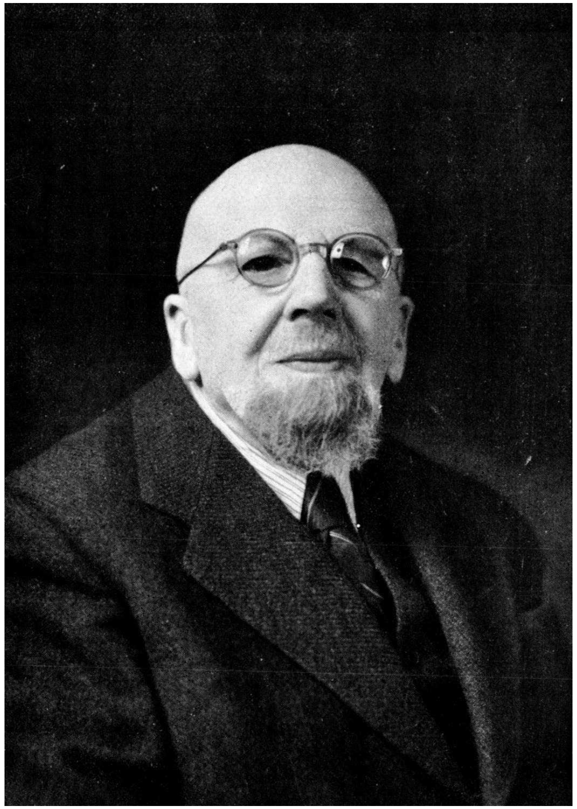
Error correction on noisy communication channel



- Example: **3-fold repetition code:** $0 \mapsto 0|00$, $1 \mapsto 1|11$.
- **Majority/Nearest neighbor/Maximum likelihood** Encoding:
  $0|00, 1|00, 0|10, 0|01 \mapsto 0|00 \mapsto \mathbf{0}$
  $1|10, 1|01, 0|11, 1|11 \mapsto 1|11 \mapsto \mathbf{1}$.

Example: QR codes

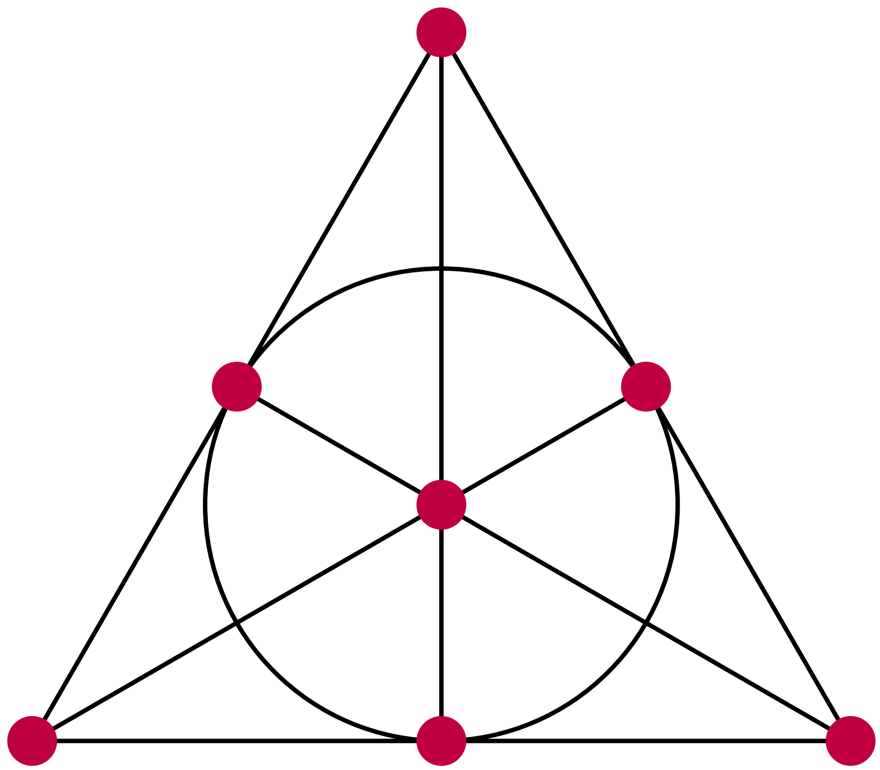
Example: QR codes



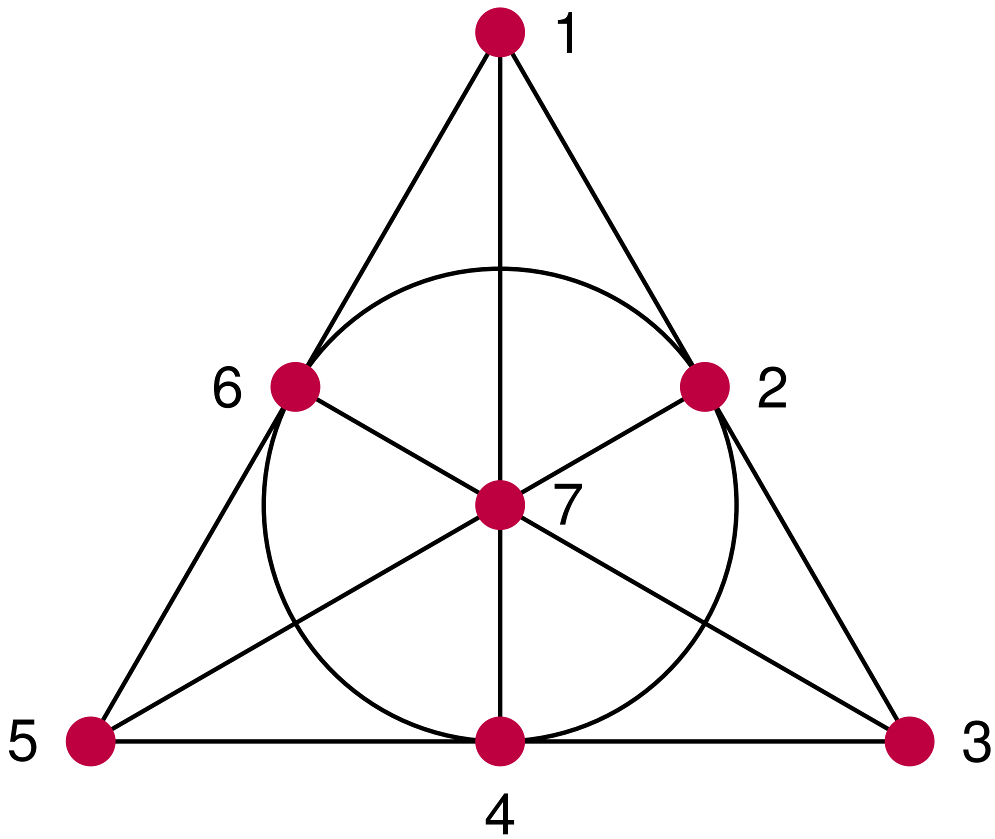
Gino Fano
(1871-1952)
Italian mathematician

Richard Hamming
(1915-1998)
US mathematician

The Fano plane: 7 points, 7 „lines”



The Fano plane: 7 points, 7 „lines”



$\{1, 2, 3\}$

$\{3, 4, 5\}$

$\{1, 5, 6\}$

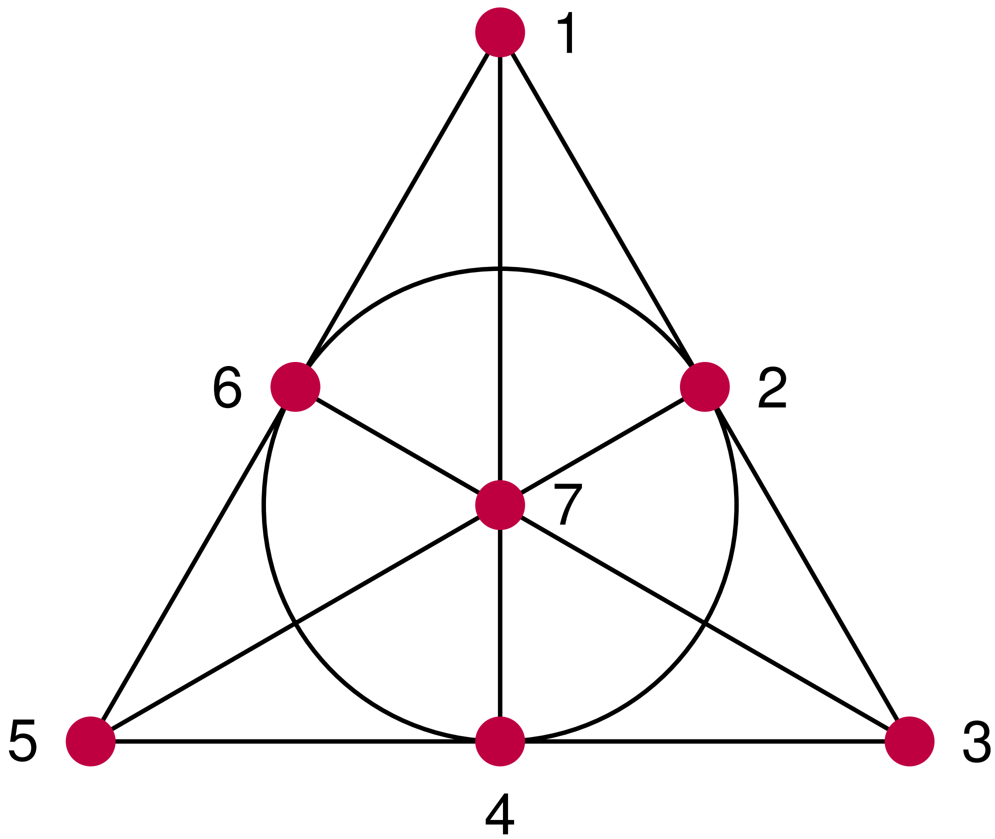
$\{1, 4, 7\}$

$\{2, 5, 7\}$

$\{3, 6, 7\}$

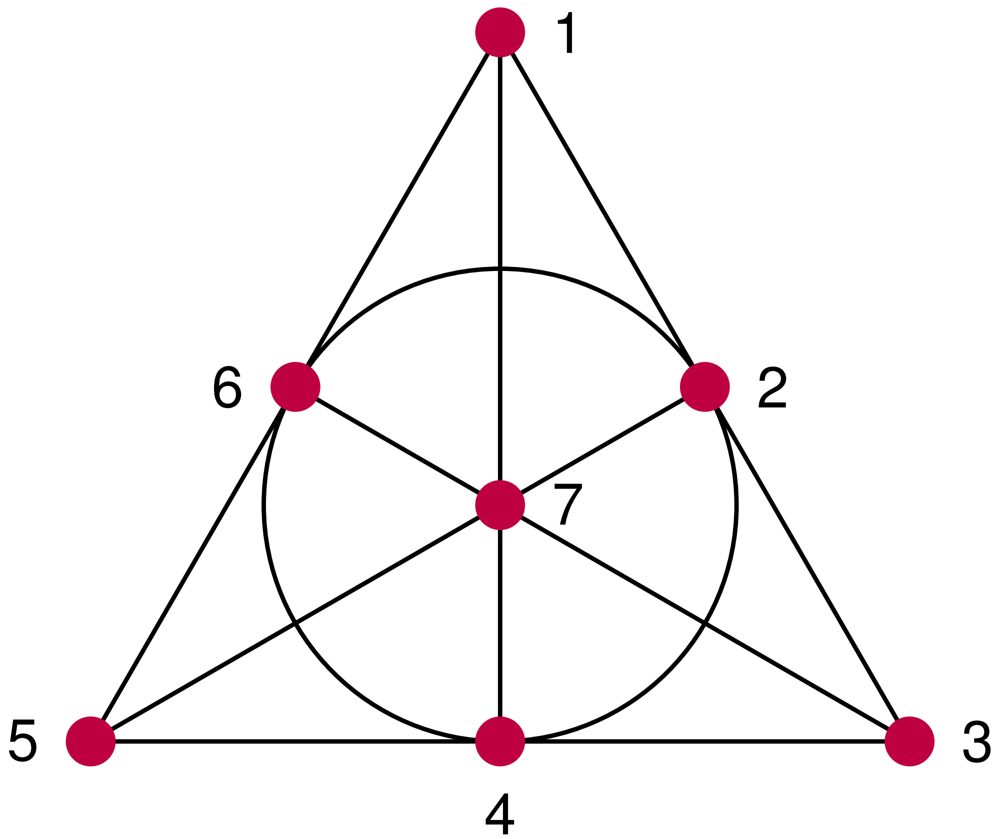
$\{2, 4, 6\}$

The Fano plane: 7 points, 7 „lines”



$\{1, 2, 3\}$	**[1 1 1 0 0 0 0]**
$\{3, 4, 5\}$	**[0 0 1 1 1 0 0]**
$\{1, 5, 6\}$	**[1 0 0 0 1 1 0]**
$\{1, 4, 7\}$	**[1 0 0 1 0 0 1]**
$\{2, 5, 7\}$	**[0 1 0 0 1 0 1]**
$\{3, 6, 7\}$	**[0 0 1 0 0 1 1]**
$\{2, 4, 6\}$	**[0 1 0 1 0 1 0]**

The Fano plane: 7 points, 7 „lines”



matrix M

$\{1,2,3\}$	[1 1 1 0 0 0 0]
$\{3,4,5\}$	[0 0 1 1 1 0 0]
$\{1,5,6\}$	[1 0 0 0 1 1 0]
$\{1,4,7\}$	[1 0 0 1 0 0 1]
$\{2,5,7\}$	[0 1 0 0 1 0 1]
$\{3,6,7\}$	[0 0 1 0 0 1 1]
$\{2,4,6\}$	[0 1 0 1 0 1 0]

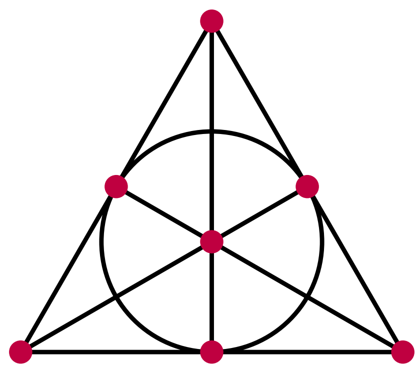
The codewords of the Hamming code

0	0	0	0	0	0	0
1	1	1	0	0	0	0
0	0	1	1	1	0	0
1	0	0	0	1	1	0
1	0	0	1	0	0	1
0	1	0	0	1	0	1
0	0	1	0	0	1	1
0	1	0	1	0	1	0
0	0	0	1	1	1	1
1	1	0	0	0	1	1
0	1	1	1	0	0	1
0	1	1	0	1	1	0
1	0	1	1	0	1	0
1	1	0	1	1	0	0
1	0	1	0	1	0	1
1	1	1	1	1	1	1

- $1 + 7 + 7 + 1 = 16$ bit sequences of length 7.
- YELLOW: All 0's and all 1's.
- RED: The matrix M of the Fano plane
- BLUE: The complementer matrix of M.

Claim

Any two codewords of the Hamming code differ in at least 3 positions.



The Hamming code: computer memory error correction

0	0	0	0	0	0	0	0
14	1	1	1	0	0	0	0
3	0	0	1	1	1	0	0
8	1	0	0	0	1	1	0
9	1	0	0	1	0	0	1
4	0	1	0	0	1	0	1
2	0	0	1	0	0	1	1
5	0	1	0	1	0	1	0
1	0	0	0	1	1	1	1
12	1	1	0	0	0	1	1
7	0	1	1	1	0	0	1
6	0	1	1	0	1	1	0
11	1	0	1	1	0	1	0
13	1	1	0	1	1	0	0
10	1	0	1	0	1	0	1
15	1	1	1	1	1	1	1

Claim 1

The first four bits of the codewords contain all 0/1 vectors of length 4 precisely once.

- GREEN: Information bits
- SÁRGA: Parity check bits

Claim 2

The Hamming code can detect 2 errors and correct 1 error.

Claim 3

The Hamming code is linear over $\mathbb{F}_2$

Outline

Basic concepts

Definition: Error correction codes over a finite alphabet

Let Q be a finite set and n a positive integer. Any subset C of the Cartesian product Q^n is called a code of length n over the alphabet Q.

- The elements of C are called codewords.
- The encoding map is a $1-1$ correspondence between the set of messages $\mathcal{M}$ and C.
- The channel noise is a random map from C to Q^n, uniform on each component.
- The decoding map is a 2-step procedure.
- Step 1 (hard): a function from Q^n to $C \cup \{?\}$.
- Step 2 (easy): the inverse of the encoding function, mapping $C \cup \{?\}$ to $\mathcal{M} \cup \{?\}$.
- Output „?” means uncorrectable transmission error (erasure).

Basic concepts

Definition: Error correction codes over a finite alphabet

Let Q be a finite set and n a positive integer. Any subset C of the Cartesian product Q^n is called a code of length n over the alphabet Q.

- The elements of C are called codewords.
- The encoding map is a $1-1$ correspondence between the set of messages $\mathcal{M}$ and C.
- The channel noise is a random map from C to Q^n, uniform on each component.
- The decoding map is a 2-step procedure.
- Step 1 (hard): a function from Q^n to $C \cup \{?\}$.
- Step 2 (easy): the inverse of the encoding function, mapping $C \cup \{?\}$ to $\mathcal{M} \cup \{?\}$.
- Output „?” means uncorrectable transmission error (erasure).

Basic concepts

Definition: Error correction codes over a finite alphabet

Let Q be a finite set and n a positive integer. Any subset C of the Cartesian product Q^n is called a code of length n over the alphabet Q.

- The elements of C are called codewords.
- The encoding map is a $1-1$ correspondence between the set of messages $\mathcal{M}$ and C.
- The channel noise is a random map from C to Q^n, uniform on each component.
- The decoding map is a 2-step procedure.
- Step 1 (hard): a function from Q^n to $C \cup \{?\}$.
- Step 2 (easy): the inverse of the encoding function, mapping $C \cup \{?\}$ to $\mathcal{M} \cup \{?\}$.
- Output „?” means uncorrectable transmission error (erasure).

Basic concepts

Definition: Error correction codes over a finite alphabet

Let Q be a finite set and n a positive integer. Any subset C of the Cartesian product Q^n is called a code of length n over the alphabet Q.

- The elements of C are called codewords.
- The encoding map is a $1-1$ correspondence between the set of messages $\mathcal{M}$ and C.
- The channel noise is a random map from C to Q^n, uniform on each component.
- The decoding map is a 2-step procedure.
- Step 1 (hard): a function from Q^n to $C \cup \{?\}$.
- Step 2 (easy): the inverse of the encoding function, mapping $C \cup \{?\}$ to $\mathcal{M} \cup \{?\}$.
- Output „?" means uncorrectable transmission error (erasure).

Basic concepts

Definition: Error correction codes over a finite alphabet

Let Q be a finite set and n a positive integer. Any subset C of the Cartesian product Q^n is called a code of length n over the alphabet Q.

- The elements of C are called codewords.
- The encoding map is a $1-1$ correspondence between the set of messages $\mathcal{M}$ and C.
- The channel noise is a random map from C to Q^n, uniform on each component.
- The decoding map is a 2-step procedure.
- Step 1 (hard): a function from Q^n to $C \cup \{?\}$.
- Step 2 (easy): the inverse of the encoding function, mapping $C \cup \{?\}$ to $\mathcal{M} \cup \{?\}$.
- Output „?” means uncorrectable transmission error (erasure).

Basic concepts

Definition: Error correction codes over a finite alphabet

Let Q be a finite set and n a positive integer. Any subset C of the Cartesian product Q^n is called a code of length n over the alphabet Q.

- The elements of C are called codewords.
- The encoding map is a $1-1$ correspondence between the set of messages $\mathcal{M}$ and C.
- The channel noise is a random map from C to Q^n, uniform on each component.
- The decoding map is a 2-step procedure.
- Step 1 (hard): a function from Q^n to $C \cup \{?\}$.
- Step 2 (easy): the inverse of the encoding function, mapping $C \cup \{?\}$ to $\mathcal{M} \cup \{?\}$.
- Output „?” means uncorrectable transmission error (erasure).

Basic concepts

Definition: Error correction codes over a finite alphabet

Let Q be a finite set and n a positive integer. Any subset C of the Cartesian product Q^n is called a code of length n over the alphabet Q.

- The elements of C are called codewords.
- The encoding map is a $1-1$ correspondence between the set of messages $\mathcal{M}$ and C.
- The channel noise is a random map from C to Q^n, uniform on each component.
- The decoding map is a 2-step procedure.
- Step 1 (hard): a function from Q^n to $C \cup \{?\}$.
- Step 2 (easy): the inverse of the encoding function, mapping $C \cup \{?\}$ to $\mathcal{M} \cup \{?\}$.
- Output „?” means uncorrectable transmission error (erasure).

Hamming distance and nearest neighbor decoding

Definition

- For two tuples $\boldsymbol{x} = (x_1, \ldots, x_n)$ and $\boldsymbol{y} = (y_1, \ldots, y_n)$ the Hamming distance
$$d_H(\boldsymbol{x}, \boldsymbol{y}) = |\{i \mid x_i \neq y_i\}|$$
is the number of position where $\boldsymbol{x}, \boldsymbol{y}$ differ.
- The minimum distance of the code $C \subseteq Q^n$ is
$$d(C) = \min\{d_H(\boldsymbol{x}, \boldsymbol{y}) \mid \boldsymbol{x}, \boldsymbol{y} \in C, \boldsymbol{x} \neq \boldsymbol{y}\}.$$
- The map $D : Q^n \to C \cup \{?\}$ is a nearest neighbor decoding, if $D(\boldsymbol{x})$ is one of the nearest codewords to $\boldsymbol{x}$ w.r.t. the Hamming distance.

Theorem

The Hamming distance defines a metric in the geometric sense. Any code can detect $d(C) - 1$ and correct $\lfloor \frac{d(C)-1}{2} \rfloor$ errors per codewords.

Codes with good parameters

Definition: Information rate, error correction rate

- The number of information symbols per codeword is approx. $\log |C|$.
- The information rate of C is $R = \frac{\log |C|}{n}$.
- The error correction rate of C is $\delta = \frac{d(C)}{n}$.

Remarks.

- Mathematicians look for codes with high information and error correcting rates.
- The Singleton bound restricts $R + \delta \leq 1 + \frac{1}{n}$.
- Engineers compare codes using their BER curves.
- In fact, the package error ratio p^* of the code is a function of bit error ratio p of the channel.
- For the Hamming code of length 7, we have

$$p^* = 1 - (1-p)^7 - 7p(1-p)^6 \approx 21p^2 + o(3).$$

Codes with good parameters

Definition: Information rate, error correction rate

- The number of information symbols per codeword is approx. $\log|C|$.
- The information rate of C is $R = \frac{\log|C|}{n}$.
- The error correction rate of C is $\delta = \frac{d(C)}{n}$.

Remarks.

- Mathematicians look for codes with high information and error correcting rates.
- The Singleton bound restricts $R + \delta \leq 1 + \frac{1}{n}$.
- Engineers compare codes using their BER curves.
- In fact, the package error ratio p^* of the code is a function of bit error ratio p of the channel.
- For the Hamming code of length 7, we have

$$p^* = 1 - (1-p)^7 - 7p(1-p)^6 \approx 21p^2 + o(3).$$

Codes with good parameters

Definition: Information rate, error correction rate

- The number of information symbols per codeword is approx. $\log|C|$.
- The information rate of C is $R = \frac{\log|C|}{n}$.
- The error correction rate of C is $\delta = \frac{d(C)}{n}$.

Remarks.

- Mathematicians look for codes with high information and error correcting rates.
- The Singleton bound restricts $R + \delta \leq 1 + \frac{1}{n}$.
- Engineers compare codes using their BER curves.
- In fact, the package error ratio p^* of the code is a function of bit error ratio p of the channel.
- For the Hamming code of length 7, we have

$$p^* = 1 - (1-p)^7 - 7p(1-p)^6 \approx 21p^2 + o(3).$$

Codes with good parameters

Definition: Information rate, error correction rate

- The number of information symbols per codeword is approx. $\log|C|$.
- The information rate of C is $R = \frac{\log|C|}{n}$.
- The error correction rate of C is $\delta = \frac{d(C)}{n}$.

Remarks.

- Mathematicians look for codes with high information and error correcting rates.
- The Singleton bound restricts $R + \delta \leq 1 + \frac{1}{n}$.
- Engineers compare codes using their BER curves.
- In fact, the package error ratio p^* of the code is a function of bit error ratio p of the channel.
- For the Hamming code of length 7, we have
$$p^* = 1 - (1-p)^7 - 7p(1-p)^6 \approx 21p^2 + o(3).$$

Codes with good parameters

Definition: Information rate, error correction rate

- The number of information symbols per codeword is approx. $\log |C|$.
- The information rate of C is $R = \frac{\log |C|}{n}$.
- The error correction rate of C is $\delta = \frac{d(C)}{n}$.

Remarks.

- Mathematicians look for codes with high information and error correcting rates.
- The Singleton bound restricts $R + \delta \leq 1 + \frac{1}{n}$.
- Engineers compare codes using their BER curves.
- In fact, the package error ratio p^* of the code is a function of bit error ratio p of the channel.
- For the Hamming code of length 7, we have

$$p^* = 1 - (1 - p)^7 - 7p(1 - p)^6 \approx 21p^2 + o(3).$$

Good news on binary linear codes

Nothing is *easier* than to produce good binary linear codes:

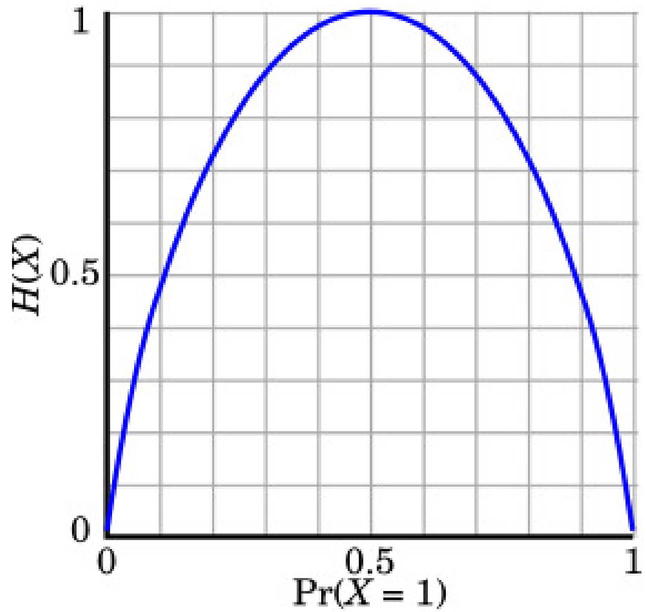
Theorem (Shannon's Noisy-Channel Coding Theorem 1948)

Define the binary entropy function

$$H(p) = -p\log_2 p - (1-p)\log_2(1-p).$$

Fix constants $0 < p < 1/2$, $0 < R < 1 - H(p)$ *and* $\varepsilon > 0$. *Then:*

- *for n sufficiently big,*
- *the „**random**" binary linear code*
- *of length n and rate R satisfies*
- $p^* \leq \varepsilon$.



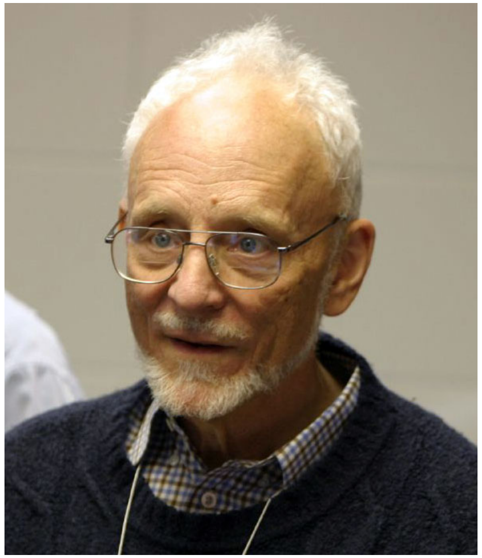
Bad news on binary linear codes

It is almost *hopeless* to make use of random binary linear codes:

Theorem (Berlekamp, McEliece, van Tilborg, 1978)

The following problem is ***NP-complete:*** *Given a $k \times n$ binary matrix A, a binary vector $\mathbf{y}$ and an integer $w > 0$. Let C be the subspace spanned by the rows of A. Is there an element $\mathbf{c} \in C$ such that $d_H(\mathbf{c}, \mathbf{y}) \leq w$?*

Robert McEliece
(1942-2019)



Elwyn Berlekamp
(1940-2019)

Henk van Tilborg
(1947-)

Aspects of decoding of linear codes

- Let $C \leq \mathbb{F}_2^n$ be a binary linear code of length n and dimension k.
- Let $\boldsymbol{x} \in C$ be the sent codeword and $\boldsymbol{y} = \boldsymbol{x} + \boldsymbol{e}$ the received word with error $\boldsymbol{e}$.
- With **hard-decision decoding** we have $\boldsymbol{y}, \boldsymbol{e} \in \mathbb{F}_2^n$.
- Efficient decoding algorithms when C has some algebraic and/or combinatorial structure: *Golay code, Reed-Solomon code, LDPC codes.*
- With **soft-decision decoding** we have $\boldsymbol{y} \in [0,1]^n$.
- Easiest example for the *repetition code:*

$$\text{decode to} \begin{cases} \mathbf{1} & \text{if} \quad \sum y_i \geq 0.5 \\ \mathbf{0} & \text{if} \quad \sum y_i < 0.5 \end{cases}$$

- Further examples: *Viterbi, turbo code.*

Outline

Linear codes over finite fields

Definition: Finite field $\mathbb{F}_q$ of order q

- Let p be a prime, n a positive integer and $q = p^n$ a prime power.
- There is a (unique) algebraic structure $\mathbb{F}_q$ of order q, endowed with four operations

$$x + y, \quad x - y, \quad x \cdot y, \quad x/y.$$

- The operation satisfy the usual arithmetic axioms.

Definition: Linear code

Let C be a linear subspace of $\mathbb{F}_q^n$. Then C is a linear code of length n over the alphabet $\mathbb{F}_q$.

- If $k = \dim C$ then $|C| = q^k$ and $R = k/n$.
- C is may be given by generators (generator matrix) or by a system of linear equations (parity check matrix).
- Encoding function is **matrix calculus:** fast and easy $\mathbb{F}_q^k \to \mathbb{F}_q^n$.

Generalized Reed-Solomon codes

- Let q be a prime power, n, k nonnegative integers such that $1 \le k \le n \le q$.
- Let $\boldsymbol{\alpha} = \{\alpha_1, \ldots, \alpha_n\}$ be n distinct elements of $\mathbb{F}_q$, $\boldsymbol{v} = (v_1, \ldots, v_n)$ a nonzero vector of $\mathbb{F}_q^n$ with $v_i \neq 0$ for all i.

Definition

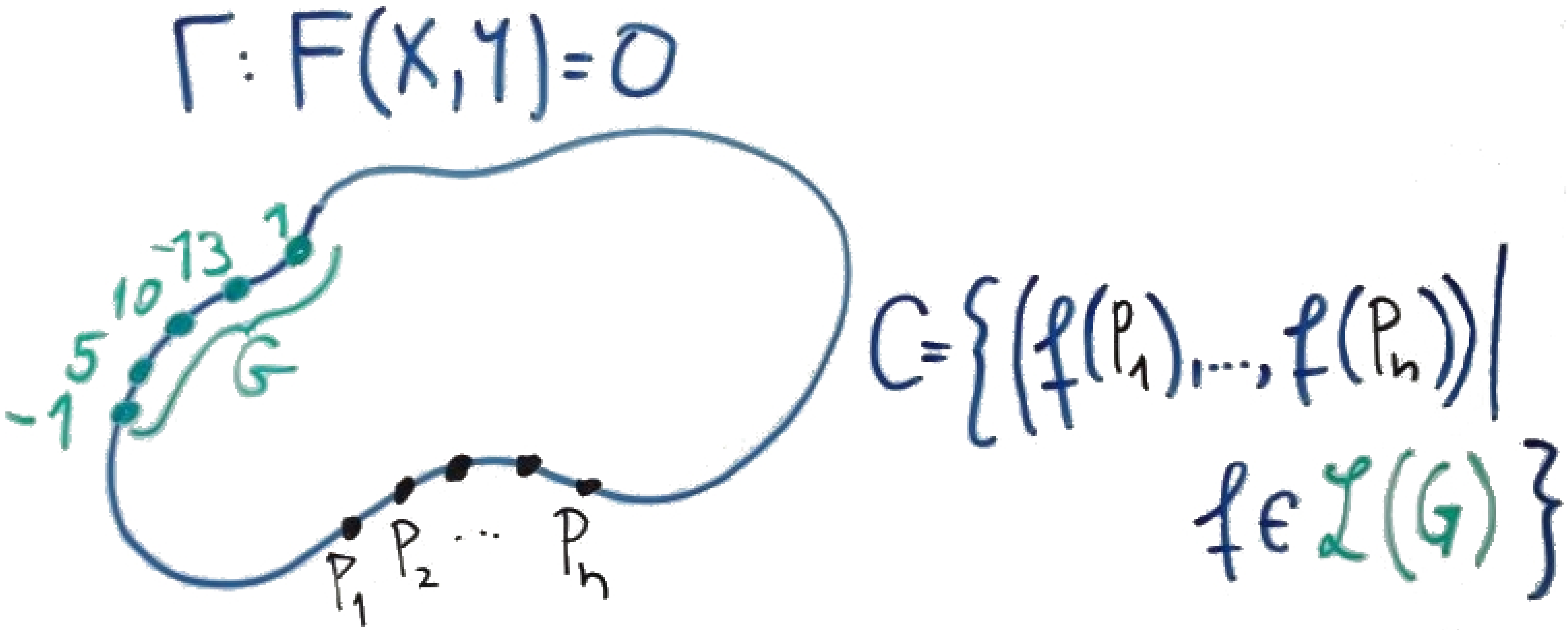
The Generalized Reed-Solomon code, denoted by $\mathbf{GRS}_k(\alpha, \boldsymbol{v})$ consists of all vectors

$$(v_1 f(\alpha_1), v_2 f(\alpha_2), \ldots, v_n f(\alpha_n)),$$

where $f(z)$ is a polynomial over $\mathbb{F}_q$ of degree less than k.

- A rich class of codes with an **efficient decoding** up to $(n-k)/2$ errors.
- Used in QR codes with $q = 256$.

Algebraic-geometric codes and curves over finite fields



- An algebraic plane curve Γ is given by a polynomial $F(X, Y) = 0$ over the finite field $\mathbb{F}_q$.
- **Hard:** points, divisors G, functions, evaluation, Riemann-Roch space $\mathscr{L}(G)$.
- **Advantage to RS:** More than q points, longer codes.

Outline

Motivation

- In this section, we present a public key cryptosystem that was proposed by McEliece in 1978.
- Its **security** is based on the hardness of binary decoding.
- In the last decades, this system was **not used** because **(1)** the keys are large, **(2)** the encrypted messages are long, and **(3)** there are not many safe binary codes beside binary BCH and Goppa codes.
- However, this system is one of the few which resists the quantum attack by Peter Shor (1994).
- The **recent progress** in the construction of **quantum computers** indicates that in 30 years, the recently used cryptosystems (RSA, ECC, etc.) will have to be replaced.

Motivation

- In this section, we present a public key cryptosystem that was proposed by McEliece in 1978.
- Its **security** is based on the hardness of binary decoding.
- In the last decades, this system was **not used** because **(1)** the keys are large, **(2)** the encrypted messages are long, and **(3)** there are not many safe binary codes beside binary BCH and Goppa codes.
- However, this system is one of the few which resists the quantum attack by Peter Shor (1994).
- The **recent progress** in the construction of **quantum computers** indicates that in 30 years, the recently used cryptosystems (RSA, ECC, etc.) will have to be replaced.

Motivation

- In this section, we present a public key cryptosystem that was proposed by McEliece in 1978.
- Its **security** is based on the hardness of binary decoding.
- In the last decades, this system was **not used** because **(1)** the keys are large, **(2)** the encrypted messages are long, and **(3)** there are not many safe binary codes beside binary BCH and Goppa codes.
- However, this system is one of the few which resists the quantum attack by Peter Shor (1994).
- The **recent progress** in the construction of **quantum computers** indicates that in 30 years, the recently used cryptosystems (RSA, ECC, etc.) will have to be replaced.

Motivation

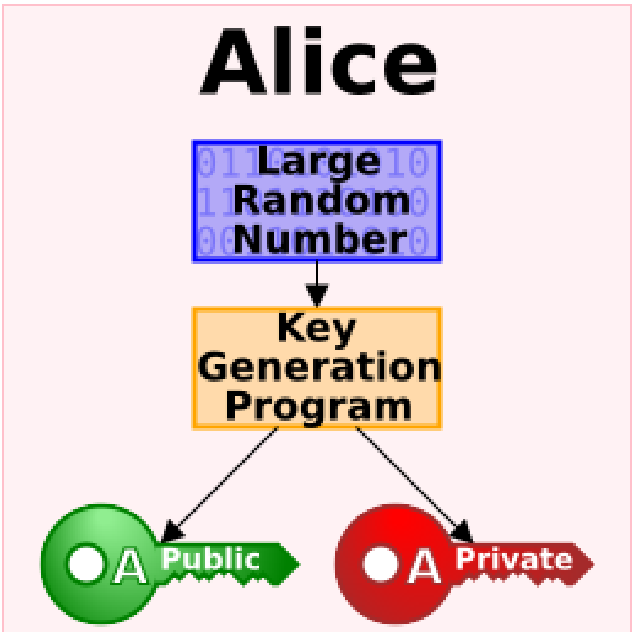
- In this section, we present a public key cryptosystem that was proposed by McEliece in 1978.
- Its **security** is based on the hardness of binary decoding.
- In the last decades, this system was **not used** because **(1)** the keys are large, **(2)** the encrypted messages are long, and **(3)** there are not many safe binary codes beside binary BCH and Goppa codes.
- However, this system is one of the few which resists the quantum attack by Peter Shor (1994).
- The **recent progress** in the construction of **quantum computers** indicates that in 30 years, the recently used cryptosystems (RSA, ECC, etc.) will have to be replaced.

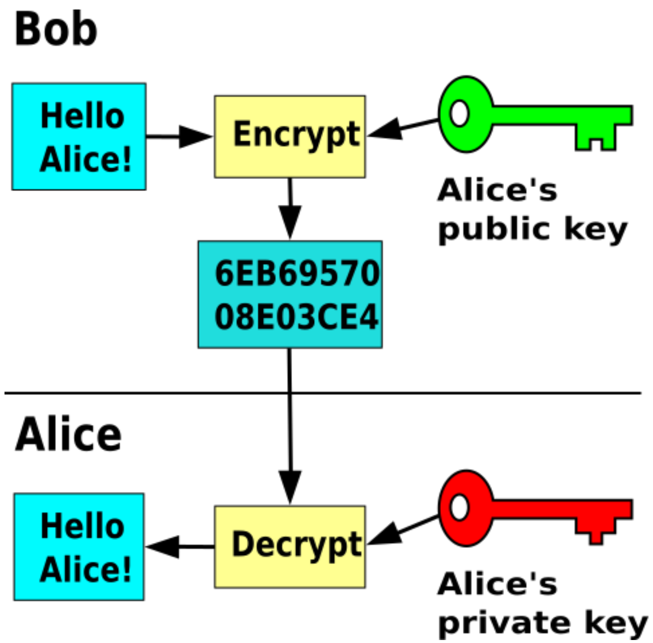
Motivation

- In this section, we present a public key cryptosystem that was proposed by McEliece in 1978.
- Its **security** is based on the hardness of binary decoding.
- In the last decades, this system was **not used** because **(1)** the keys are large, **(2)** the encrypted messages are long, and **(3)** there are not many safe binary codes beside binary BCH and Goppa codes.
- However, this system is one of the few which resists the quantum attack by Peter Shor (1994).
- The **recent progress** in the construction of **quantum computers** indicates that in 30 years, the recently used cryptosystems (RSA, ECC, etc.) will have to be replaced.

Public key (asymmetric) cryptography

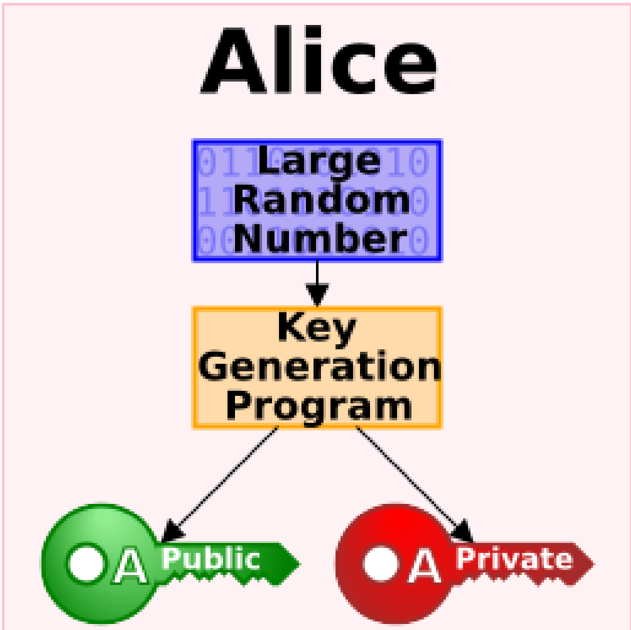
- In a **public key (or asymmetric) cryptosystem,** each user X has two keys,
- a private key $K_D(X)$ and a public key $K_E(X)$.
- If Bob wants to send a message m to Alice, he encrypts it to m' using Alice's public key $K_E(\text{Alice})$.
- For the decryption, Alice uses her private key $K_D(\text{Alice})$.

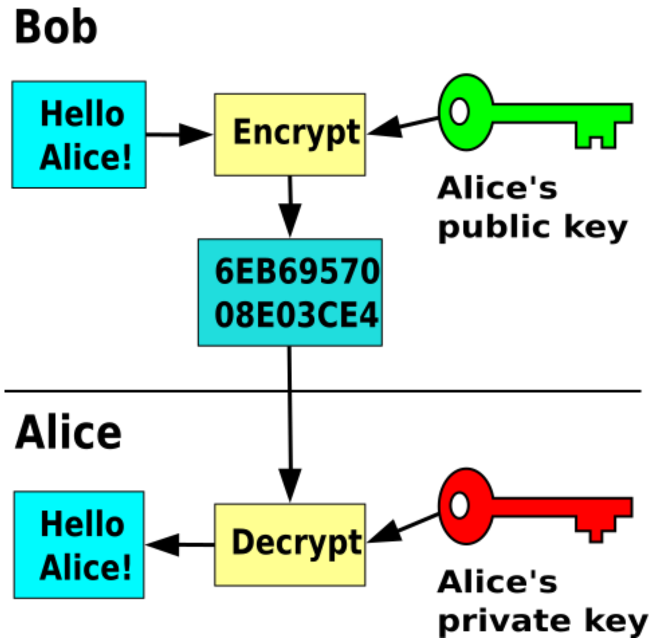




Public key (asymmetric) cryptography

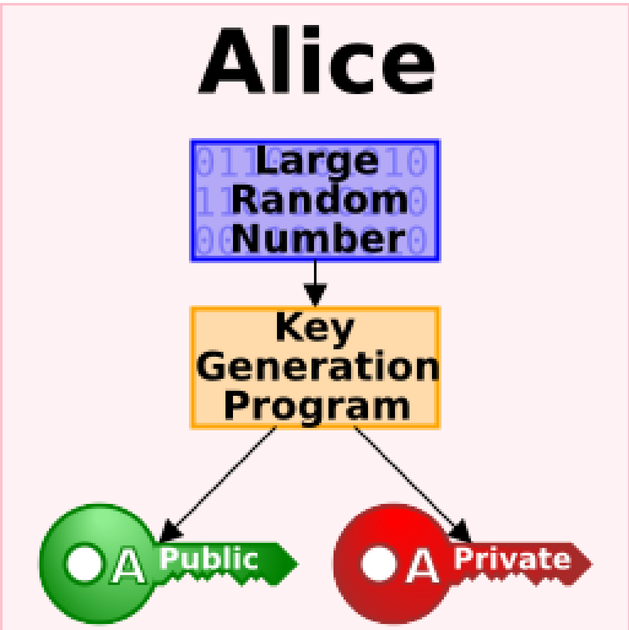
- In a **public key (or asymmetric) cryptosystem,** each user X has two keys,
- a private key $K_D(X)$ and a public key $K_E(X)$.
- If Bob wants to send a message m to Alice, he encrypts it to m' using Alice's public key $K_E(\text{Alice})$.
- For the decryption, Alice uses her private key $K_D(\text{Alice})$.

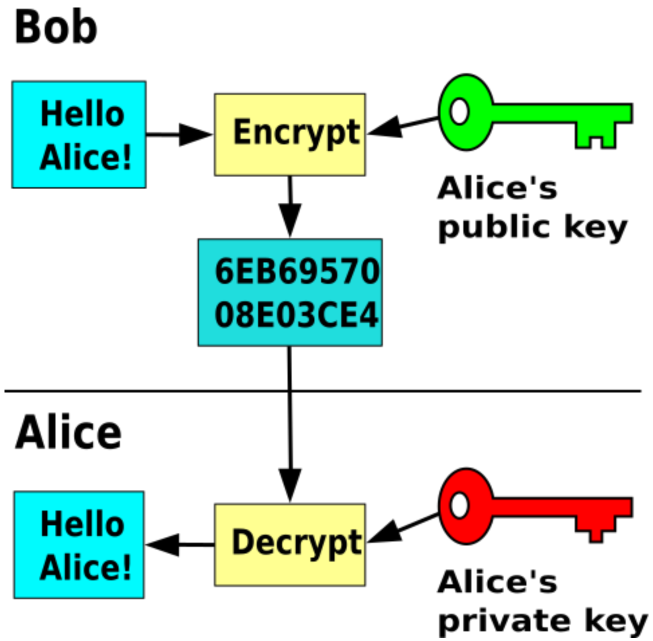




Public key (asymmetric) cryptography

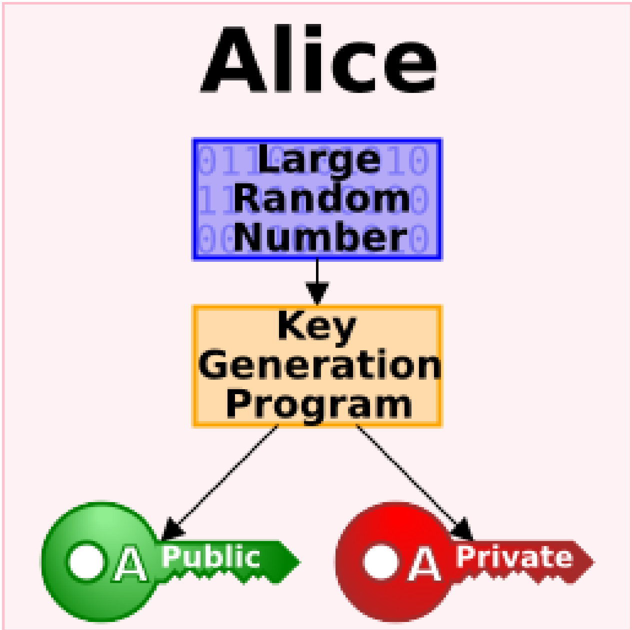
- In a **public key (or asymmetric) cryptosystem,** each user X has two keys,
- a private key $K_D(X)$ and a public key $K_E(X)$.
- If Bob wants to send a message m to Alice, he encrypts it to m' using Alice's public key $K_E(\text{Alice})$.
- For the decryption, Alice uses her private key $K_D(\text{Alice})$.

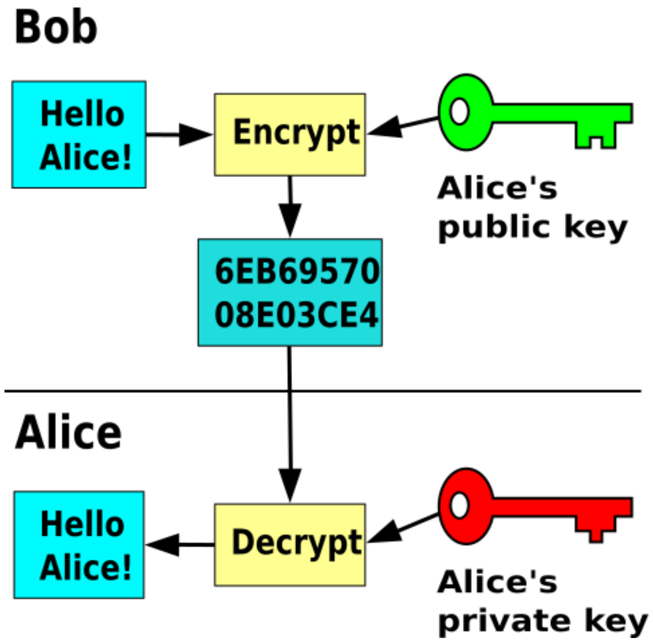




Public key (asymmetric) cryptography

- In a **public key (or asymmetric) cryptosystem,** each user X has two keys,
- a private key $K_D(X)$ and a public key $K_E(X)$.
- If Bob wants to send a message m to Alice, he encrypts it to m' using Alice's public key $K_E(\text{Alice})$.
- For the decryption, Alice uses her private key $K_D(\text{Alice})$.





McEliece Cryptosystem

- The **McEliece Cryptosystem** is based on a **binary linear code** C of length n and dimension k, which has a **fast algorithm** correcting up to t errors per code word. Let G denote the $n \times k$ generator matrix of C.
- **Creation of Alice's keys** She picks a random $k \times k$ invertible matrix S and a random $n \times n$ permutation matrix P. Her *private key* is the pair (S, P) and her *public key* is the $n \times k$ matrix $G' = SGP$.
- **Encryption** Assume that Bob's message is $\boldsymbol{m} \in \mathbb{F}_2^k$. Bob picks a random binary vector $\boldsymbol{e} \in \mathbb{F}_2^n$ of weight t and computes the encrypted message $\boldsymbol{m}' = \boldsymbol{m}G' + \boldsymbol{e}$.
- **Decryption** First Alice computes

$$\boldsymbol{m}'P^{-1} = (\boldsymbol{m}G' + \boldsymbol{e})P^{-1} = \boldsymbol{m}SG + \boldsymbol{e}',$$

  where $\boldsymbol{m}SG \in C$ and $\boldsymbol{e}' = \boldsymbol{e}P^{-1}$ is an error vector of weight t.
- Now, using the *fast decoding method,* Alice determines $\boldsymbol{m}S$ and $\boldsymbol{e}'$. Finally, Alice computes the message $\boldsymbol{m} = (\boldsymbol{m}S)S^{-1}$.

McEliece Cryptosystem

- The **McEliece Cryptosystem** is based on a **binary linear code** C of length n and dimension k, which has a **fast algorithm** correcting up to t errors per code word. Let G denote the $n \times k$ generator matrix of C.
- **Creation of Alice's keys** She picks a random $k \times k$ invertible matrix S and a random $n \times n$ permutation matrix P. Her *private key* is the pair (S, P) and her *public key* is the $n \times k$ matrix $G' = SGP$.
- **Encryption** Assume that Bob's message is $\boldsymbol{m} \in \mathbb{F}_2^k$. Bob picks a random binary vector $\boldsymbol{e} \in \mathbb{F}_2^n$ of weight t and computes the encrypted message $\boldsymbol{m}' = \boldsymbol{m}G' + \boldsymbol{e}$.
- **Decryption** First Alice computes

$$\boldsymbol{m}'P^{-1} = (\boldsymbol{m}G' + \boldsymbol{e})P^{-1} = \boldsymbol{m}SG + \boldsymbol{e}',$$

  where $\boldsymbol{m}SG \in C$ and $\boldsymbol{e}' = \boldsymbol{e}P^{-1}$ is an error vector of weight t.
- Now, using the *fast decoding method,* Alice determines $\boldsymbol{m}S$ and $\boldsymbol{e}'$. Finally, Alice computes the message $\boldsymbol{m} = (\boldsymbol{m}S)S^{-1}$.

McEliece Cryptosystem

- The **McEliece Cryptosystem** is based on a **binary linear code** C of length n and dimension k, which has a **fast algorithm** correcting up to t errors per code word. Let G denote the $n \times k$ generator matrix of C.
- **Creation of Alice's keys** She picks a random $k \times k$ invertible matrix S and a random $n \times n$ permutation matrix P. Her *private key* is the pair (S, P) and her *public key* is the $n \times k$ matrix $G' = SGP$.
- **Encryption** Assume that Bob's message is $\boldsymbol{m} \in \mathbb{F}_2^k$. Bob picks a random binary vector $\boldsymbol{e} \in \mathbb{F}_2^n$ of weight t and computes the encrypted message $\boldsymbol{m}' = \boldsymbol{m}G' + \boldsymbol{e}$.
- **Decryption** First Alice computes

$$\boldsymbol{m}'P^{-1} = (\boldsymbol{m}G' + \boldsymbol{e})P^{-1} = \boldsymbol{m}SG + \boldsymbol{e}',$$

where $\boldsymbol{m}SG \in C$ and $\boldsymbol{e}' = \boldsymbol{e}P^{-1}$ is an error vector of weight t.
- Now, using the *fast decoding method,* Alice determines $\boldsymbol{m}S$ and $\boldsymbol{e}'$. Finally, Alice computes the message $\boldsymbol{m} = (\boldsymbol{m}S)S^{-1}$.

McEliece Cryptosystem

- The **McEliece Cryptosystem** is based on a **binary linear code** C of length n and dimension k, which has a **fast algorithm** correcting up to t errors per code word. Let G denote the $n \times k$ generator matrix of C.
- **Creation of Alice's keys** She picks a random $k \times k$ invertible matrix S and a random $n \times n$ permutation matrix P. Her *private key* is the pair (S, P) and her *public key* is the $n \times k$ matrix $G' = SGP$.
- **Encryption** Assume that Bob's message is $\boldsymbol{m} \in \mathbb{F}_2^k$. Bob picks a random binary vector $\boldsymbol{e} \in \mathbb{F}_2^n$ of weight t and computes the encrypted message $\boldsymbol{m}' = \boldsymbol{m}G' + \boldsymbol{e}$.
- **Decryption** First Alice computes

$$\boldsymbol{m}'P^{-1} = (\boldsymbol{m}G' + \boldsymbol{e})P^{-1} = \boldsymbol{m}SG + \boldsymbol{e}',$$

where $\boldsymbol{m}SG \in C$ and $\boldsymbol{e}' = \boldsymbol{e}P^{-1}$ is an error vector of weight t.
- Now, using the *fast decoding method,* Alice determines $\boldsymbol{m}S$ and $\boldsymbol{e}'$. Finally, Alice computes the message $\boldsymbol{m} = (\boldsymbol{m}S)S^{-1}$.

McEliece Cryptosystem

- The **McEliece Cryptosystem** is based on a **binary linear code** C of length n and dimension k, which has a **fast algorithm** correcting up to t errors per code word. Let G denote the $n \times k$ generator matrix of C.
- **Creation of Alice's keys** She picks a random $k \times k$ invertible matrix S and a random $n \times n$ permutation matrix P. Her *private key* is the pair (S, P) and her *public key* is the $n \times k$ matrix $G' = SGP$.
- **Encryption** Assume that Bob's message is $\boldsymbol{m} \in \mathbb{F}_2^k$. Bob picks a random binary vector $\boldsymbol{e} \in \mathbb{F}_2^n$ of weight t and computes the encrypted message $\boldsymbol{m}' = \boldsymbol{m}G' + \boldsymbol{e}$.
- **Decryption** First Alice computes

$$\boldsymbol{m}'P^{-1} = (\boldsymbol{m}G' + \boldsymbol{e})P^{-1} = \boldsymbol{m}SG + \boldsymbol{e}',$$

  where $\boldsymbol{m}SG \in C$ and $\boldsymbol{e}' = \boldsymbol{e}P^{-1}$ is an error vector of weight t.
- Now, using the *fast decoding method,* Alice determines $\boldsymbol{m}S$ and $\boldsymbol{e}'$. Finally, Alice computes the message $\boldsymbol{m} = (\boldsymbol{m}S)S^{-1}$.

Challenges

- Find codes with good parameters.
- Find codes with effective decoding algorithms.
- Give bounds for the parameters of certain codes.
- Find the true values of the parameters of certain codes.
- Improve the decoding algorithms.
- Make probabilistic decoding algorithms into deterministic ones.
- Understand the structure of subfield subcodes of AG codes.
- Investigate codes w.r.t. to non Hamming distances.
- Sloane's problem (1978): Find a self-dual binary linear code of length 72, dimension 36 and minimum distance 16.

Challenges

- Find codes with good parameters.
- Find codes with effective decoding algorithms.
- Give bounds for the parameters of certain codes.
- Find the true values of the parameters of certain codes.
- Improve the decoding algorithms.
- Make probabilistic decoding algorithms into deterministic ones.
- Understand the structure of subfield subcodes of AG codes.
- Investigate codes w.r.t. to non Hamming distances.
- Sloane's problem (1978): Find a self-dual binary linear code of length 72, dimension 36 and minimum distance 16.

Challenges

- Find codes with good parameters.
- Find codes with effective decoding algorithms.
- Give bounds for the parameters of certain codes.
- Find the true values of the parameters of certain codes.
- Improve the decoding algorithms.
- Make probabilistic decoding algorithms into deterministic ones.
- Understand the structure of subfield subcodes of AG codes.
- Investigate codes w.r.t. to non Hamming distances.
- Sloane's problem (1978): Find a self-dual binary linear code of length 72, dimension 36 and minimum distance 16.

Challenges

- Find codes with good parameters.
- Find codes with effective decoding algorithms.
- Give bounds for the parameters of certain codes.
- Find the true values of the parameters of certain codes.
- Improve the decoding algorithms.
- Make probabilistic decoding algorithms into deterministic ones.
- Understand the structure of subfield subcodes of AG codes.
- Investigate codes w.r.t. to non Hamming distances.
- Sloane's problem (1978): Find a self-dual binary linear code of length 72, dimension 36 and minimum distance 16.

Challenges

- Find codes with good parameters.
- Find codes with effective decoding algorithms.
- Give bounds for the parameters of certain codes.
- Find the true values of the parameters of certain codes.
- Improve the decoding algorithms.
- Make probabilistic decoding algorithms into deterministic ones.
- Understand the structure of subfield subcodes of AG codes.
- Investigate codes w.r.t. to non Hamming distances.
- Sloane's problem (1978): Find a self-dual binary linear code of length 72, dimension 36 and minimum distance 16.

Challenges

- Find codes with good parameters.
- Find codes with effective decoding algorithms.
- Give bounds for the parameters of certain codes.
- Find the true values of the parameters of certain codes.
- Improve the decoding algorithms.
- Make probabilistic decoding algorithms into deterministic ones.
- Understand the structure of subfield subcodes of AG codes.
- Investigate codes w.r.t. to non Hamming distances.
- Sloane's problem (1978): Find a self-dual binary linear code of length 72, dimension 36 and minimum distance 16.

Challenges

- Find codes with good parameters.
- Find codes with effective decoding algorithms.
- Give bounds for the parameters of certain codes.
- Find the true values of the parameters of certain codes.
- Improve the decoding algorithms.
- Make probabilistic decoding algorithms into deterministic ones.
- Understand the structure of subfield subcodes of AG codes.
- Investigate codes w.r.t. to non Hamming distances.
- Sloane's problem (1978): Find a self-dual binary linear code of length 72, dimension 36 and minimum distance 16.

Challenges

- Find codes with good parameters.
- Find codes with effective decoding algorithms.
- Give bounds for the parameters of certain codes.
- Find the true values of the parameters of certain codes.
- Improve the decoding algorithms.
- Make probabilistic decoding algorithms into deterministic ones.
- Understand the structure of subfield subcodes of AG codes.
- Investigate codes w.r.t. to non Hamming distances.
- Sloane's problem (1978): Find a self-dual binary linear code of length 72, dimension 36 and minimum distance 16.

Challenges

- Find codes with good parameters.
- Find codes with effective decoding algorithms.
- Give bounds for the parameters of certain codes.
- Find the true values of the parameters of certain codes.
- Improve the decoding algorithms.
- Make probabilistic decoding algorithms into deterministic ones.
- Understand the structure of subfield subcodes of AG codes.
- Investigate codes w.r.t. to non Hamming distances.
- Sloane's problem (1978): Find a self-dual binary linear code of length 72, dimension 36 and minimum distance 16.