Cryptography Notes

Alessandro Vulcu

13 ottobre 2025

Indice

1 History and Why Mathematics Came In

1.1 Definition of Cryptography

The word **cryptography** originates from the Ancient Greek words *κρυπτός* (*kryptós*, "hidden, secret") and *γράφειν* (*graphein*, "to write"). Therefore, cryptography literally means *"the art of writing in secret"*.

Common misspellings. Crypthography, Criptography, Kryptography.

Cryptography has always been deeply connected with the need for secure communication. For thousands of years, rulers and generals relied on efficient communication to govern and lead armies, while fearing interception. This led to the invention of **codes** and **ciphers**: techniques for disguising a message so that only the intended recipient can understand it. Over the centuries, an arms race emerged between **code makers** and **code breakers**.

(Source: Singh, The Code Book*)*

1.2 Terminology

- **Code:** replaces entire words with other words/symbols (e.g., "attack at dawn" → "JUPITER").
- **Cipher:** replaces individual letters by a rule/pattern (e.g., a→b, b→c gives "attack at dawn" → "buubdl bu ebxo").
- **Plaintext:** original message (lowercase).
- **Ciphertext:** encrypted message (uppercase).

In modern usage, *decoding* and *deciphering* are often used interchangeably. Today, virtually all practical systems are ciphers (codes are mostly obsolete).

1.3 Classical Cryptography

The earliest forms of cryptography are **classical ciphers**, e.g., the *Caesar cipher*, where each letter is shifted by a fixed amount.

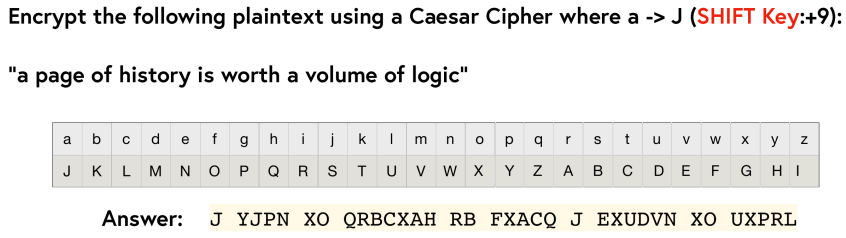

Figura 1: Caesar cipher: each letter is shifted by a fixed number of positions.

1.4 Example of Decryption

Decrypt the ciphertext `DVVKZECFSSPRKKVE` assuming a Caesar cipher.

Solution: `meet in lobby at ten` (Key: $a \to R$, Shift $+17$).

This shows that a Caesar cipher has only **26** possible keys (one per shift). A full monoalphabetic substitution, instead, admits $26!$ keys:

$$26! = 26 \times 25 \times \cdots \times 2 \times 1 \approx 4.03 \times 10^{26}.$$

1.5 Monoalphabetic Substitution Ciphers

The **Caesar cipher** is a special case of **monoalphabetic substitution**. A general substitution can be any permutation of the 26 letters, hence:

$$26! = 403{,}291{,}461{,}126{,}605{,}635{,}584{,}000{,}000.$$

Even with 10^{18} ops/s, a brute-force search would take on the order of years, showing that *pure brute force* is impractical—yet these ciphers are still weak due to *statistics*.

1.6 The First Known Cryptanalyst: Al-Kindi

In the 9th century, **Al-Kindi** wrote *"A Manuscript on Deciphering Cryptographic Messages"*, introducing **frequency analysis**. Between 800–1200 AD, Arab scholars made major advances while Europe was in the Dark Ages; Al-Kindi laid foundations for cryptanalysis centuries before Europe.

1.7 Frequency Analysis in English Texts

Letters occur with different probabilities. In English, very common letters include:

e, t, a, o, i, n, s, h, r, d, l, u, m, w, c.

Digrams (pairs) like `th`, `he`, `in` are frequent:

`th`: 3.87%, `he`: 3.69%, `in`: 2.36%.

1.8 Decrypting Using Frequency Analysis

A monoalphabetic ciphertext can be attacked by matching ciphertext frequencies to language statistics (letters, digrams, trigrams), iteratively hypothesizing mappings to reconstruct plaintext. Computers automate these comparisons efficiently.

1.9 Example: Decrypting a Long Ciphertext

`LIVITCSWPIYVEWHEVSRIQMXLEYVEOIEWHRXEXIPFEMVEWHKVSTYLXZIXLIKIIXPIJVSZEYPE...`

Frequency analysis reveals the plaintext progressively; for illustration:

> *"This course aims to introduce you to the principles and techniques of securing computers and networks, focusing on Internet security. You will learn both theoretical and practical aspects of cryptography, algorithms, and real-world security protocols."*

1.10 Unit 2 — Modern Cryptography: Polyalphabetic Ciphers

1.10.1 Introduction to Polyalphabetic Ciphers

The **Vigenère cipher** (Bellaso, 1550; later attributed to Vigenère) applies multiple shifting alphabets driven by a repeating keyword (e.g., key $2, 5, 4, 7$). The same plaintext letter may encrypt to different ciphertext letters depending on position.

1.10.2 Breaking the Vigenère Cipher

Kasiski (19th c.) introduced a method to infer the *key length* by detecting repeated sequences and factoring distances between repeats; common factors suggest candidate lengths (e.g., 3 and 6).

1.10.3 Frequency Analysis per Key Position

Once the period is known, split the ciphertext into groups by position modulo the period; each group is a Caesar cipher and can be cracked by frequency analysis. Example outcome (period $= 3$): shifts $(+2,\ 0,\ +19)$.

1.10.4 Revealing the Plaintext

Decrypting with the recovered key yields readable plaintext. The method's weakness: long reuse of a key exposes periodic patterns; one key per message is impractical.

1.11 Historical Context: Codebreakers and the Evolution of Cryptography

States have long employed **codebreakers**. Examples: **Mary, Queen of Scots** (1587) and the **Zimmermann Telegram** (1917) show strategic impact of cryptanalysis.

2 Cipher Machines

2.1 The Rise of Cryptographic Machines (1920–1970)

Radio/telegraph required secrecy over open channels; manual ciphers were replaced by **mechanical devices** enabling faster/stronger encryption.

2.2 The Enigma Machine

Invented by **Scherbius** (1918), later adopted by the German military. Enigma scrambles letters via electromechanical **rotors**; after each keypress, rotor stepping changes the mapping (like an odometer).

2.3 The Rotors Mechanism

Three rotors each implement a 26-letter permutation. Pressing a key advances rotor(s); thus the same plaintext letter encrypts differently at different times.

2.4 The Reflector and the Plugboard

A **reflector** returns the signal through the rotors (same setup for enc/dec). The **plugboard** swaps up to six letter pairs, adding confusion.

Path: Key $\rightarrow$ Plugboard $\rightarrow$ Rotors $\rightarrow$ Reflector $\rightarrow$ Rotors $\rightarrow$ Plugboard $\rightarrow$ Lampboard.

2.5 Configuration Space of Enigma

- Rotor positions: $26^3 = 17{,}576$.
- Rotor choices/order from 5: $5 \cdot 4 \cdot 3 = 60$.
- Plugboard (6 pairs): $\dfrac{26!}{14!\,2^6} \approx 1.0039 \times 10^{11}$.

Total $\approx 2 \times 10^{17}$ configurations.

2.6 Cracking Enigma — The Polish Breakthrough

Rejewski, Różycki, Zygalski (1932) recovered rotor wirings using mathematics and intelligence. The doubled *message key* created patterns exploitable via cycle structure; a large catalog enabled daily-key recovery in minutes (from 1934).

2.7 Turing and the British Bombe

German upgrades (1938) increased complexity. **Turing** designed the **Bombe**, exploiting stereotyped headers to test configurations quickly; crossword-style recruitment highlighted logic/language skills.

2.8 The Lorenz Cipher and the Birth of the Computer

High-level traffic used **Lorenz SZ40**. **Max Newman** led development of **Colossus** (1943), the first programmable digital computer, to attack Lorenz traffic. Post-war secrecy delayed public recognition.

3 Modern Cryptography (1970–)

3.1 From Mechanical to Digital Cryptography

Electronic computers enabled fast encryption/decryption but also brute-force attacks. Text is encoded in **ASCII**:

$$\mathrm{A} = 01000001,\ \mathrm{B} = 01000010,\ \mathrm{C} = 01000011,\ \ldots$$

3.2 Lucifer and the Birth of DES

IBM's **Lucifer** (1971) evolved into **DES**, both **symmetric-key** ciphers combining substitution and transposition.

3.3 The Problem of Private Keys

Symmetric systems suffer from **key distribution**. Historical practice (e.g., Enigma daily keys) shows the logistical and security burden of sharing secrets safely.

3.4 The Idea of Public Keys

Goal:

> *How can Alice and Bob communicate securely over an eavesdropped channel without a pre-shared key?*

The "double-lock" analogy motivates, but cryptographic ops need not commute—so a different idea is needed.

3.5 Two Keys: One Private, One Public

Diffie & Hellman (1976) proposed separate **public** and **private** keys: encrypt with the recipient's public key; decrypt with their private key.

3.6 The Diffie–Hellman Key Exchange

Public parameters: large prime p, base g. Alice picks a, sends $A = g^a \bmod p$; Bob picks b, sends $B = g^b \bmod p$. Shared secret:

$$K = g^{ab} \bmod p = (g^a)^b \bmod p = (g^b)^a \bmod p.$$

Security relies on the hardness of the discrete logarithm problem.

3.7 The Digital Signature Problem

DH provides secrecy but not authentication (vulnerable to MITM). Digital signatures address authenticity and integrity (concept by Diffie–Hellman; efficient schemes came later).

3.8 The RSA Cryptosystem

Rivest–Shamir–Adleman (1978) introduced **RSA** for public-key encryption and signatures, relying on the hardness of factoring $N = pq$. (C. Cocks, GCHQ, discovered a similar idea in 1973—declassified in 1997.)

3.9 Conclusion

From Caesar to RSA, cryptography evolved with mathematics and computation. As computing advances, cryptography leans more on number theory and complex problems to stay ahead.

"History and why Mathematics came in Cryptography." — End of Lesson 1

4 Divisions

4.1 Divisibility

Naturals and Integers.

$$\mathbb{N} = \{0, 1, 2, 3, \ldots\}, \quad \mathbb{Z} = \{\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots\}$$

Definition. Let $a, b \in \mathbb{Z}$ with $b \neq 0$. We say that **b divides a** (and write $b \mid a$) if there exists $c \in \mathbb{Z}$ such that:

$$a = bc.$$

Example.

$$2 \mid 4 \quad \text{but} \quad 2 \nmid 5.$$

—

4.1.1 Even and Odd Numbers

Definition. An integer is **even** if it is divisible by 2, otherwise it is **odd**.

Examples.

$$-4 \text{ is even}, \quad 13 \text{ is odd}.$$

By convention, the divisors of a natural number are its positive divisors.

$$\text{Example: divisors of } 6 \text{ are } 1, 2, 3, 6.$$

—

4.2 Euclidean Division

Proposition 4.1 (Euclidean Division)**.** *Let $a, b \in \mathbb{N}$ with $b > 0$. Then there exist unique integers q (quotient) and r (remainder) such that:*

$$a = bq + r, \qquad 0 \leq r < b.$$

Example.

$$a = 27,\ b = 7 \Rightarrow q = 3,\ r = 6, \quad \text{since } 27 = 7 \cdot 3 + 6.$$

Observation. b divides a if and only if $r = 0$.

Another Example.

$$a = 7 \times 4 - 5 = 7 \times 3 + 2,$$

so $q = 3$ and $r = 2$.

—

4.3 Euclidean Division in Practice

$$a = bq + r, \quad 0 \leq r < b \quad \Longleftrightarrow \quad \frac{a}{b} = q + \frac{r}{b}, \quad 0 \leq \frac{r}{b} < 1.$$

Hence:

$$q = \left\lfloor \frac{a}{b} \right\rfloor, \quad r = a - b \left\lfloor \frac{a}{b} \right\rfloor.$$

Example.

$$a = 15476, \quad b = 137 \Rightarrow q = \left\lfloor \frac{15476}{137} \right\rfloor, \quad r = 15476 - 137q.$$

—

4.4 Binary Representation

Example. Every natural number m can be expressed in binary form:

$$112 = (1110000)_2 = 1 \times 2^6 + 1 \times 2^5 + 1 \times 2^4 + 0 \times 2^3 + 0 \times 2^2 + 0 \times 2^1 + 0 \times 2^0.$$

Definition. Let $m \in \mathbb{N}$. We write:

$$m = (m_{B-1} \dots m_1 m_0)_2$$

if

$$m = m_0 + m_1 2 + \cdots + m_{B-1} 2^{B-1}, \quad m_i \in \{0, 1\}.$$

Each m_i is called a **bit**.

—

4.5 Binary Representation and Euclidean Division

Binary representation arises naturally from repeated division by 2:

$$a = 2q + r, \quad r \in \{0, 1\}, \quad q = \left\lfloor \frac{a}{2} \right\rfloor.$$

Applying division recursively gives:

$$a = (r_{B-1} \dots r_1 r_0)_2,$$

where the r_i are remainders.

Example.

$$\begin{aligned} 112 &= 2 \times 56 + 0 \\ 56 &= 2 \times 28 + 0 \\ 28 &= 2 \times 14 + 0 \\ 14 &= 2 \times 7 + 0 \\ 7 &= 2 \times 3 + 1 \\ 3 &= 2 \times 1 + 1 \\ 1 &= 2 \times 0 + 1 \end{aligned}$$

Hence $112 = (1110000)_2$.

—

Remark. *If* $m = (m_{B-1} \dots m_1 m_0)_2$*, then* $m < 2^B$*. Indeed:*

$$m = m_0 + m_1 2 + \cdots + m_{B-1} 2^{B-1} \leq 1 + 2 + \cdots + 2^{B-1} = 2^B - 1.$$

—

Proposition 4.2 (Binary Expansion)**.** *Let* $m < 2^B$*. Consider the divisions:*

$$\begin{aligned} m &= 2q_0 + m_0, & 0 \leq m_0 < 2, \\ q_0 &= 2q_1 + m_1, & 0 \leq m_1 < 2, \\ q_1 &= 2q_2 + m_2, & 0 \leq m_2 < 2, \\ &\vdots \\ q_{B-2} &= 2q_{B-1} + m_{B-1}, & 0 \leq m_{B-1} < 2. \end{aligned}$$

Then $q_{B-1} = 0$ *and*

$$m = \sum_{i=0}^{B-1} 2^i m_i.$$

Dimostrazione. By substitution:

$$m = 2q_0 + m_0 = 2(2q_1 + m_1) + m_0 = \cdots = 2^{B-1} m_{B-1} + \cdots + 2m_1 + m_0.$$

□ □

—

Corollary 4.1 (Number of Bits)**.** *A natural number* m *has a binary representation with* $B \geq 1$ *bits*

$$m = (m_{B-1} \dots m_1 m_0)_2 \quad \textit{and} \quad m_{B-1} = 1$$

if and only if

$$2^{B-1} \leq m < 2^B.$$

In particular,

$$B = \lfloor \log_2 m \rfloor + 1.$$

Dimostrazione. If $m < 2^B$, at most B bits are needed. If $B - 1$ bits were enough, then $m < 2^{B-1}$, contradiction. □ □

—

4.6 Example

$$\log_2(368{,}932) = 18.4 \Rightarrow B = \lfloor 18.4 \rfloor + 1 = 19.$$

Hence $368{,}932$ requires 19 bits:

$$368{,}932 = (1011000001000100100)_2.$$

5 Greatest Common Divisor (GCD)

5.1 Definition

Example. The divisors of 12 are $1, 2, 3, 4, 6, 12$. The divisors of 18 are $1, 2, 3, 6, 9$. Their common divisors are $\{1, 2, 3, 6\}$, whose largest is 6:

$$6 := \gcd(18, 12).$$

Remark. *When we write* $\gcd(a, b)$*, we assume* a *or* b *is nonzero.*

Definition. Let $a, b \in \mathbb{Z}$, not both zero. The set of common divisors of a, b has a largest element d, called the **greatest common divisor** of a, b:

$$d = \gcd(a, b).$$

Examples.

$$\gcd(12, 0) = 12, \qquad \gcd(0, 12) = 12.$$

—

5.2 The Euclidean Algorithm

Example. Find $\gcd(119, 259)$:

$$259 = 2 \cdot 119 + 21, \quad 119 = 5 \cdot 21 + 14, \quad 21 = 1 \cdot 14 + 7, \quad 14 = 2 \cdot 7 + 0.$$

The last nonzero remainder is 7:

$$\boxed{\gcd(119, 259) = 7.}$$

Theorem 5.1 (Euclidean Algorithm)**.** *Let* $a, b > 0$ *with* $a \geq b$*. Then* $\gcd(a, b)$ *is found by:*

$$r_0 = a,\ r_1 = b, \quad r_{i-1} = r_i q_i + r_{i+1},\ 0 \leq r_{i+1} < r_i.$$

If $r_{i+1} = 0$*, then* $r_i = \gcd(a, b)$.

Dimostrazione. The sequence $r_0 > r_1 > \cdots \geq 0$ terminates. Let $d = r_{n+1}$ be the last nonzero remainder. Then $d \mid r_i$ for all i, so $d \mid a, b$. If $c \mid a, b$, then $c \mid d$. Hence $d = \gcd(a, b)$. □ □

—

6 The Extended Euclidean Algorithm

Example. Compute $\gcd(a, b)$ as a linear combination:

$$a = 259,\ b = 119$$

$$\begin{aligned} 259 &= 2 \cdot 119 + 21, \\ 119 &= 5 \cdot 21 + 14, \\ 21 &= 1 \cdot 14 + 7, \\ 14 &= 2 \cdot 7 + 0. \end{aligned}$$

Expressing each remainder:

$$\begin{aligned} 21 &= a - 2b, \\ 14 &= b - 5(a - 2b) = -5a + 11b, \\ 7 &= (a - 2b) - (-5a + 11b) = 6a - 13b. \end{aligned}$$

Thus:

$$\boxed{\gcd(a, b) = 7 = 6a - 13b.}$$

Proposition 6.1 (Extended Euclidean Algorithm)**.** *If* $d = \gcd(a, b)$*, then there exist* $u, v \in \mathbb{Z}$ *such that*

$$\boxed{d = au + bv.}$$

Dimostrazione. From the recursive divisions:

$$r_{i-1} = r_i q_i + r_{i+1},$$

each remainder is a linear combination of a, b. By recursion, the final nonzero remainder $r_n = d$ satisfies

$$d = au + bv.$$

□ □