



## Introduction

Since the beginning of written language, humans have wanted to share information secretly. The information could be orders from a general in times of war, a message between secret admirers, or information regarding some of the world's most villainous crimes.

Suppose that someone wants to send a message to a receiver, and wants to be sure that no-one else can read the message. However, there is the possibility that someone else opens the letter or hears the electronic communication.

## Definitions and Illustrations

First we will settle upon the meaning of *cryptography*,

*Cryptography* is the science of information security or it is the art and science of secret writing.

*Cryptography* has, as its etymology, *kryptos* from the Greek, meaning *hidden*, and *graphein*, meaning *to write*.

In the basic communication scenario, we assume that it as a game between three parties:

a *sender* (e.g., an embassy); we denote her by **Alice**,

a *receiver* (e.g., the government office) we denote him by **Bob**, and

an *opponent* (or a hacker e.g., a spy) we denote him by **Charlie**.

The original message is called the *plaintext* or *cleartext* (always we use small letters).

The disguised message is called the *ciphertext* (always we use capital letters).

The final message, encapsulated and sent, is called a *cryptogram*.

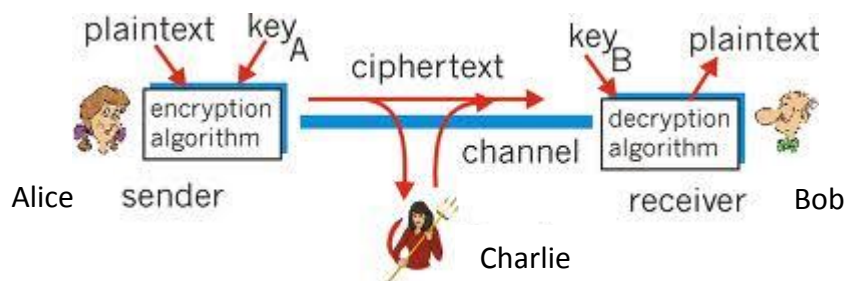
The process of transforming plaintext into ciphertext is called *encryption* or *enciphering*.

The reverse process of turning ciphertext into plaintext is called *decryption* or *deciphering*.

The person who enciphers the message is known as the *encipherer*.

The study of mathematical techniques for attempting to defeat cryptographic methods is called *cryptanalysis*.

The term *cipher* is a method for enciphering and deciphering.



## 1. Greek Ciphers

### 1.1 Skytale Cipher

The need to make a message confidential goes back to millions of years before the invention of computer science. The first technique of cryptography of which we have news, goes back to the ninth century B.C with the **Skytale** system which was used by the Spartans.

It consisted in a cylinder with a strip of parchment wound around it on which is written a message. The ancient Greeks and the Spartans in particular, are said to have used this cipher to communicate during military campaigns.

Once the ribbon was removed from the cylinder it was “impossible” to read the message, unless it was rewound on a cylinder with the same diameter, as shown in the following figures



## 1.2. Polybius Cipher

Another example of Greek cryptography is the Polybius cipher. This cipher has the disadvantage of doubling the length of the message.

	1	2	3	4	5
1	A	B	C	D	E
2	F	G	H	I&J	K
3	L	M	N	O	P
4	Q	R	S	T	U
5	V	W	X	Y	Z

### Polybius square

In the Polybius cipher, each letter is replaced by the position in which it appears in the Polybius square, using first the row number and then the column number. For example, **D** would be replaced with **14**. To decipher a message you find the letter that intersects the specified row and column.

**Example.** Encipher the message **this is easy to break**

T	H	I	S	I	S	E	A	S	Y	T	O	B	R	E	A	K
44	23	24	43	24	43	15	11	43	54	44	34	12	42	15	11	25

It's a special case of a general class of ciphers known as substitution ciphers, where a given letter is always substituted for with the same symbol wherever it appears.

Of course, we could use a six-by-six square, which would allow for 26 letters and 10 digits. Alternatively, numbers can be spelled out. A six-by-six square would be used for languages written in the Cyrillic alphabet.

The Polybius square is sometimes called a Polybius checkerboard. The letters may be placed in the square in any order; for example, the keyword **DERANGEMENT** could be used to rearrange the letters like so:

	1	2	3	4	5
1	D	E	R	A	N
2	G	M	T	B	C
3	F	H	I&J	K	L
4	O	P	Q	S	U
5	V	W	X	Y	Z

**Exercise:** Encipher **this is easy to break**. Decrypt your answer.

## 2. Substitution and Transposition Ciphers

With a *transposition* cipher, we permute the places where the plaintext letters sit. What this means is that we do not change the letters but rather move them around, transpose them, without introducing any new letters.

A *substitution* uses the same substitution across the entire message. For example, if you know that the letter A is enciphered as the letter K, this will hold true for the entire message. These types of messages can be cracked by using frequency analysis, educated guesses or trial and error.

### 2.1. Message Reversal:

In this cipher the ciphertext will be the reverse of the original message. It is a transposition cipher.

For example, to encrypt the message “**computer**”.

The plaintext is **computer**.

The ciphertext will be **RETUPMOC**.

to encrypt the message “**agent one meet me in the zoo**”.

The ciphertext will be: **TNEGA ENOTE EMEMN IEHTO OZ**.

### 2.2. Permutation ciphers

The Permutation Cipher is a transposition cipher.

The Permutation Cipher has been used for hundreds of years. The main idea of this cipher is permutation of the positions of letters.

In the Permutation Cipher, Alice and Bob need to choose the length *m* of the permutation and then a random permutation as the key.

**Example**

Suppose Alice and Bob decide that  $m = 6$  and use the permutation,

$$\pi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 4 & 3 & 1 & 6 & 2 & 5 \end{pmatrix}.$$

**Encryption**

Alice wants to send the plaintext:

he walked up and down the passage two or three times.

Alice first divides the plaintext into groups of size 6 (we call these groups blocks):

hewalk edupan ddownt hepass agetwo orthre etimes

then performs the permutation on each of the groups and obtains the ciphertext:

**WLEHKAUADENPOND DDTWPSEHSAEWGAOTTRROEH IETESM.**

**Decryption**

When Bob received that ciphertext, he divides the text into blocks of size 6 and for each block he makes the permutation

$$\pi^{-1} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 3 & 5 & 2 & 1 & 6 & 4 \end{pmatrix}.$$

Then he obtains the plaintext.

In the above example we can see that the first **e** is encrypted as **L**, the second **e** is encrypted as **U** and the third **e** is encrypted as **S**.

**2.3. Bifid Cipher:**

The Bifid Cipher uses a Polybius Square to encipher a message in a way that makes it fairly difficult to decipher without knowing the secret. This is because each letter in the ciphertext message is dependent upon two letters from the plaintext message. As a result, frequency analysis of letters becomes much more difficult.

The first step is to use the Polybius Square to convert the letters into numbers. We will be writing the numbers vertically below the message.

	1	2	3	4	5
1	A	B	C	D	E
2	F	G	H	I&J	K
3	L	M	N	O	P
4	Q	R	S	T	U
5	V	W	X	Y	Z

### Example

Suppose Alice and Bob decide to share the message “**secret message**” in a secret way using Bifid Cipher,

**411414 3144121**

**353254 2533125**

The numbers are now read off horizontally and grouped into pairs as follow

**41 14 14 31 44 12 13 53 25 42 53 31 25**

The Polybius Square is used again to convert the numbers back into letters which gives us our ciphertext.

**QDDL TB CXKR XLK**

Since the first letter in the plaintext is encoded into the first and middle letters of the ciphertext, the recipient of the message must have the entire message before they can decode it. This means that if part of the ciphertext is discovered by a third party, it is unlikely that they will be able to crack it. To decipher a bifid encrypted message, you first convert each letter into its corresponding number via the Polybius Square. Now, divide the long string of numbers into two equal rows. The digit in the top row and the digit in the bottom row will together reference the decoded letter in the Polybius Square.

## 2.4. Rail-Fence Transposition

**Example** To encrypt the following message

**Anyone who looks at us the wrong way twice will surely die.**

We simply write the text moving back and forth in a zigzag fashion from the top line to the bottom line:

A Y N W O O K A U T E R N W Y W C W L S R L D E  
N O E H L O S T S H W O G A T I E I L U E Y I

and then read across the top line first to get the ciphertext:

**AYNWO OKAUT ERNWy WCWLS RLDEN OEHL O STSHW OGATI EILUE YI**

The “fence” needn’t be limited to two tiers. We could encipher the same message as follows:

A                V                K                T                N                W                L                L  
N            E H            O S            S H            O G            T I            I L            E Y  
Y N            O O            A U            E R            V Y            C V            S R            D E  
O            L            T            V            A            E            U            I

to get the ciphertext:

**AWKTN WLLNE HOSSH OGTII LEYYN OOAUE RWYCW SRDEO LTWAE UI**

For the decryption, first we write the expected zigzag form and complete it by lines, before reading as zig-zag, as following

X                X                X                X                X                X                X  
X            X X            X X            X X            X X            X X            X X            X X  
X X            X X            X X            X X            X X            X X            X X  
X                X                X                X                X                X                X

The **key** of this cipher is just the number of **rails**.



## 2.5. Pigpen cipher

This system is known as the pigpen cipher (or Freemason's cipher), because the letters are separated like pigs in a pen. It is also called the Masonic cipher, as the Society of Freemasons has made use of it. It was used in the civil wars in England in the 17th century and even as recently as the U.S. Civil War by prisoners sending messages to friends.

The pigpen cipher uses graphical symbols assigned according to a key similar to the following diagram. It is a substitution cipher.

A	B	C	J	K	L
D	E	F	M	N	O
G	H	I	P	Q	R
T		U	X		Y

**Example.** Using the pigpen cipher the message "x marks the spot" is encrypted as follows

## 2.6. Francis Bacon's cipher

To encode a message, each letter of the plaintext is replaced by a group of five of the letters 'A' or 'B'. This replacement is done according to the alphabet of the Baconian cipher, shown below.

a	AAAAA	g	AABBA	n	ABBAA	t	BAABA
b	AAAAB	h	AABBB	o	ABBAB	u-v	BAABB
c	AAABA	i-j	ABAAA	p	ABBBA	w	BABAA
d	AAABB	k	ABAAB	q	ABBBB	x	BABAB
e	AABAA	l	ABABA	r	BAAAA	y	BABBA
f	AABAB	m	ABABB	s	BAAAB	z	BABBB

**Example** Encrypt the message “math”

Plaintext: **math**

Cipher text: **ABABBAAAAABAABAAABBB**

## 2.7. Caesar Cipher

The Caesar cipher is one of the earliest examples of a substitution cipher used by **Julius Caesar** in the Gallic wars (58-50)BC. Each letter of a given plaintext, the information to be encrypted, is substituted with another letter some given number of positions from it in the alphabet. For example, if we had an alphabet comprised of the standard 26 letters in the English alphabet and swapped each letter with the letter three places after it in the alphabet; we would have the following Caesar cipher

Plaintext	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
Ciphertext	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C

Alphabet shifted by 3 spaces.

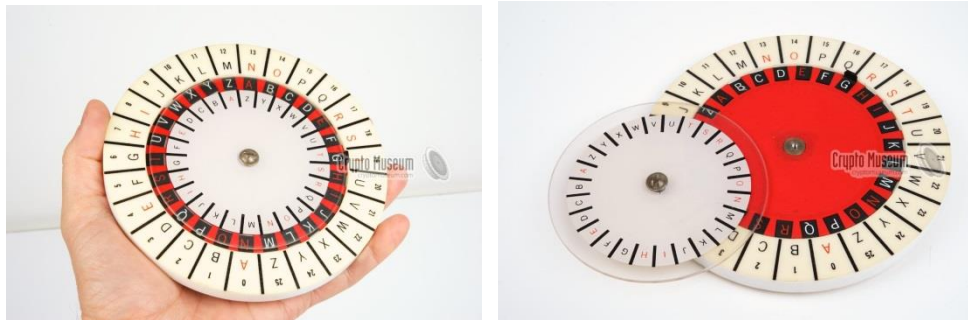
**Example.** Encrypt the following message

**Caesar was a great soldier**

The ciphertext is

**FDHVDU ZDV D JUHDW VROGLHU**

The following Figure represents two concentric rings of which the outer one is free to rotate. If the outer ring is moved so that the (D) coincides with (a), then enciphering with the Caesar cipher is achieved by regarding the inner circle as representing the plain text letters and replacing each letter by the one outside it.



Now that we have our machine, it is clear that we can use it to obtain other ciphers. We can rotate the outside ring until any chosen letter is outside (a) to get a different substitution alphabet and then encipher our plaintext message using the same algorithm. In this way we can use our machine to get 26 different ciphers of which the Caesar cipher is merely one example. These ciphers are called the *additive ciphers* or the *translation*.

## 2.8. Keyword Cipher (or the Key-Phrase Cipher)

The Keyword Cipher is identical to the Caesar Cipher with the exception that the substitution alphabet used can be represented with a keyword. To create a substitution alphabet from a keyword, you first write down the alphabet. Below this you write down the keyword (omitting duplicate letters) followed by the remaining unused letters of the alphabet.

<b>a</b>	<b>b</b>	<b>c</b>	<b>d</b>	<b>e</b>	<b>f</b>	<b>g</b>	<b>h</b>	<b>i</b>	<b>j</b>	<b>k</b>	<b>l</b>	<b>m</b>	<b>n</b>	<b>o</b>	<b>p</b>	<b>q</b>	<b>r</b>	<b>s</b>	<b>t</b>	<b>u</b>	<b>v</b>	<b>w</b>	<b>x</b>	<b>y</b>	<b>z</b>
<b>K</b>	<b>E</b>	<b>Y</b>	<b>W</b>	<b>O</b>	<b>R</b>	<b>D</b>	<b>A</b>	<b>B</b>	<b>C</b>	<b>F</b>	<b>G</b>	<b>H</b>	<b>I</b>	<b>J</b>	<b>L</b>	<b>M</b>	<b>N</b>	<b>P</b>	<b>Q</b>	<b>S</b>	<b>T</b>	<b>U</b>	<b>V</b>	<b>X</b>	<b>Z</b>

To encipher a plaintext message, you convert all letters from the top row to their corresponding letter on the bottom row (A to K, B to E, etc). These types of simple substitution ciphers can be easily cracked by using frequency analysis and some educated guessing.

### Example.

Plaintext: this is a secret message

Cipher text: QABP BP k POYNOQ HOPPKDO

## 3. Polyalphabetic Ciphers

In a polyalphabetic cipher, the substitution may change throughout the message. In other words, the letter **A** may be encoded as the letter **K** for part of the message, but latter on it might be encoded as the letter **W**.

### 3.1.Vigenère Cipher

In a Caesar Cipher, each letter of the alphabet is shifted along some number of places; for example, in a Caesar cipher of shift 3, A would become D, B would become E and so on. The **Vigenère** cipher consists of using several Caesar ciphers in sequence with different shift values.

To encipher, a table of alphabets can be used, called a **Vigenère square**. It consists of the alphabet written out **26** times in different rows, each alphabet shifted cyclically to the left compared to the previous alphabet, corresponding to the **26** possible Caesar ciphers. At different points in the encryption process, the cipher uses a different alphabet from one of the rows. The alphabet used at each point depends on a repeating keyword.

**Example.** Encrypt the message “**Attack at dawn**” using the key word “**lemon**”.

The person sending the message repeats the keyword “lemon” until it matches the length of the plaintext, as follows

**LEMONLEMONLE**

**Plaintext : a t t a c k a t dawn**

**Keyword: LEMONLEMONLE**

Each letter is encoded by finding the intersection in the grid between the plaintext letter and keyword letter. For example, the first letter of the plaintext, **A**, is enciphered using the alphabet in row **L**, which is the first letter of the key. This is done by looking at the letter in row **L** and column **A** of the **Vigenere** square, namely **L**. Similarly, for the second letter of the plaintext, the second letter of the key is used; the letter at row **E** and column **T** is **X**. The rest of the plaintext is enciphered in a similar fashion.

**Ciphertext: LXFOPVEFRNHR**

**Plaintext**

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
A	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
B	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A
C	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B
D	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C
E	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D
F	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E
G	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F
H	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G
I	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H
J	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I
K	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J
L	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K
M	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L
N	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M
O	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N
P	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
Q	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
R	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
S	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
T	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
U	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
V	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
W	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V
X	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W
Y	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X
Z	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y

**Vigenère Square**  
© WORDS & UNWORDS



Decryption is performed by finding the position of the ciphertext letter in a row of the table, and then taking the label of the column in which it appears as the plaintext. For example, in row **L**, the ciphertext **L** appears in column **A**, which taken as the first plaintext letter. The second letter is decrypted by looking up **X** in row **E** of the table; it appears in column **T**, which is taken as the plaintext letter.

## The Playfair Cipher

The scheme was invented in **1854** by **Charles Wheatstone**, but bears the name of **Lord Playfair** who promoted the use of the cipher.

The Playfair cipher uses a **5** by **5** table containing a key word or phrase. For example, if we use the key word “**key word**”, we get the following Digraph cipher, where the letters **I** and **J** are considered as a single entity.

<b>K</b>	<b>E</b>	<b>Y</b>	<b>W</b>	<b>O</b>
<b>R</b>	<b>D</b>	<b>A</b>	<b>B</b>	<b>C</b>
<b>F</b>	<b>G</b>	<b>H</b>	<b>I/J</b>	<b>L</b>
<b>M</b>	<b>N</b>	<b>P</b>	<b>Q</b>	<b>S</b>
<b>T</b>	<b>U</b>	<b>V</b>	<b>X</b>	<b>Z</b>

To encipher a message, first we divide it into pairs of letters. Pairs of letters are enciphered according to the following rules.

- (A) If two letters are in the same row, then their ciphertext equivalents are immediately to their right. For instance, **VZ** in plaintext is **XT** in ciphertext. (This means that if one is at the right or bottom edge of the table, then one “wraps around” as indicated in the example.).
- (B) If two letters are in the same column, then their cipher equivalents are the letters immediately below them. For example, **AP** in plaintext is **HV** in ciphertext, and **KM** in plaintext is **RT** in ciphertext.
- (C) If two letters are on the corners of a diagonal of a rectangle formed by them, then their cipher equivalents are the letters in the opposite corners and same row as the plaintext letter. For instance, **UL** in plaintext becomes **ZG** in ciphertext and **YZ** in plaintext is **OV** in ciphertext.
- (D) If the same letter occurs as a pair in plaintext, then we agree by convention to put a **Z** between them and encipher.

(E) If a single letter remains at the end of the plaintext, then a **Z** is added to it to complete the digraph.

**Example.** Using the Playfair cipher presented above, we encipher the message “**do not trust them**” as follow:

Alice first divides the plaintext into pairs of letters as following

**do no tt ru st th em**

**do** is enciphered as **CE**

**no** is enciphered as **SE**

We put **z** between **tt**’ s to get “ **tz tr us tz th em**”.

**tz** is enciphered as: **UT**

**tr** is enciphered as: **KF**

**us** is enciphered as: **ZN**

**tz** is enciphered as: **UT**

**th** is enciphered as: **VF**

**em** is enciphered as: **KN**

The final ciphertext will be

**CESEUTKFZNUTVFKN**

<b>K</b>	<b>E</b>	<b>Y</b>	<b>W</b>	<b>O</b>
<b>R</b>	<b>D</b>	<b>A</b>	<b>B</b>	<b>C</b>
<b>F</b>	<b>G</b>	<b>H</b>	<b>I/J</b>	<b>L</b>
<b>M</b>	<b>N</b>	<b>P</b>	<b>Q</b>	<b>S</b>
<b>T</b>	<b>U</b>	<b>V</b>	<b>X</b>	<b>Z</b>

**Decryption** We decipher the ciphertext: **CESEUTKFZNUTVFKN**

Since **CE** sits on a diagonal with corresponding letters **do**, then that digraph is the plaintext. Similarly employing all the rules, we get **donotztrustzthem** and once we remove the extraneous letters **z**, occurring twice, since **tt** occurs twice, we get the plaintext **do not trust them**.



**Definition.** Let  $m$  and  $k$  be two integers, then we say that  $m \equiv k \pmod{n}$ , if  $n \mid (m - k)$ , that is,  $n$  divides the quantity  $m - k$  evenly.

**Definition.** Let  $k$  be an integer, then  $k^{-1} \pmod{n}$ , is the number  $m$  that is in the range  $0 \leq m < n$ , such that  $m.k \equiv 1 \pmod{n}$  if  $m$  exists. If no such number  $m$  exists then we say that  $k$  does not have an inverse modulo  $n$  or that  $k$  is not invertible modulo  $n$ .

### Examples

- (1)  $2^{-1} \pmod{5} \equiv 3$ , since  $2.3 \equiv 1 \pmod{5}$ .
- (2)  $12^{-1} \pmod{17} \equiv 10$ , since  $12.10 \equiv 1 \pmod{17}$ .
- (3)  $11^{-1} \pmod{26} \equiv 19$ , since  $11.19 \equiv 1 \pmod{26}$ .
- (4)  $4^{-1} \pmod{26}$ , does not exist since there is no number  $b$  between 0 and 25 with  $4.b \equiv 1 \pmod{26}$ .

**Theorem** Let  $k$  be an integer, then  $k^{-1} \pmod{n}$  exists if and only if the greatest common divisor of  $k$  and  $n$  is 1, that is,  $\gcd(k, n) = 1$ .

### Examples

- (1)  $2^{-1} \pmod{5}$ , exists since  $(2, 5) = 1$ .
- (2)  $12^{-1} \pmod{17}$ , exists since  $(12, 17) = 1$ .
- (3)  $11^{-1} \pmod{26}$ , exists since  $(11, 26) = 1$ .
- (4)  $4^{-1} \pmod{26}$ , does not exist since  $(4, 26) \neq 1$ .

Our next goal is to generalize the above modular calculations to matrices. Taking a matrix that has only integer entries **mod**  $n$  is simple, we just take each entry **mod**  $n$ .

### Example

$$\begin{pmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{pmatrix} (\text{mod } 5) = \begin{pmatrix} 1 & 4 & 2 \\ 2 & 0 & 3 \\ 3 & 1 & 4 \end{pmatrix}.$$

### Notes.

- (1) Addition, subtraction, and multiplication **mod**  $n$  is just as easy. First do the addition, subtraction, or multiplication as you would normally and then take the result **mod**  $n$ , that is, take each entry **mod**  $n$ .
- (2) Matrix inverses are a little more different.

### Example.

$$(1) \begin{pmatrix} 2 & 1 & -1 \\ 3 & 5 & 2 \end{pmatrix} \begin{pmatrix} 2 & -4 \\ -1 & -7 \\ 5 & 9 \end{pmatrix} (\text{mod } 5) = \begin{pmatrix} -2 & -24 \\ 11 & -29 \end{pmatrix}.$$

$$(2) \text{ If } A = \begin{pmatrix} 2 & 1 & 4 \\ 3 & 0 & 2 \\ 1 & 3 & 4 \end{pmatrix}, \text{ find } A^{-1} (\text{mod } 5).$$

**Solution** We know that,  $A^{-1} = \frac{1}{|A|} \text{adj}(A)$ .

$$\text{Hence, } A^{-1} = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 1 & 0 & 3 \end{pmatrix} (\text{mod } 5). \quad (\text{Check})$$

## Affine Cipher

The *affine cipher* encrypts by multiplying the plaintext by one part of the key followed by addition of another part of the key.

The *affine cipher* is a special case of substitution ciphers and it is a generalization of the **additive cipher (Caesar Cipher)**.

## The Algorithm

When **encrypting**, we first convert all the letters to numbers ('a'=0, 'b'=1,..., 'z'=25), as in the following table:

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25

In this case, the **Caesar cipher** will be equivalent to the following formula:

$$c \equiv p + \beta \pmod{26}, 1 \leq \beta \leq 26.$$

In which,  $1 \leq p \leq 26$  represents the letter and  $c$  is the ciphertext.

The 'key' for the **Affine cipher** consists of 2 numbers, we'll call them  $\alpha$  and  $\beta$ . The following discussion assumes the use of a 26 character alphabet ( $m = 26$ ) and  $\alpha$  should be chosen to be relatively prime to  $m$ . For example, 15 and 26 have no factors in common, so 15 is an acceptable value for  $\alpha$ , however 12 and 26 have factors in common so 12 cannot be used for a value of  $\alpha$ . The ciphertext letter  $c$ , for any given letter  $p$  is given as follows (remember  $p$  is the number representing a letter):

$$c \equiv \alpha p + \beta \pmod{m}, 1 \leq \alpha \leq m, 1 \leq \beta \leq m.$$

The decryption function is:

$$p \equiv \alpha^{-1}(c - \beta)(\text{mod } m),$$

where  $\alpha^{-1}$  is the multiplicative inverse of  $\alpha$  in the group of integers modulo  $m$ .

To find a multiplicative inverse, we need to find a number  $x$  such that:

$$\alpha x \equiv 1 (\text{mod } m).$$

The easiest way to solve this equation is to search each of the numbers **1** to **25**, and see which one satisfies the equation. Thus,  $\alpha$  must be in the set:

$$\alpha \in \{1, 3, 5, 7, 9, 11, 15, 17, 19, 21, 23, 25\}$$

For instance, if  $\alpha = 3$ , then  $\alpha^{-1} = 9$ , since  $3 \cdot 9 = 27 \equiv 1 (\text{mod } 26)$ .

**Example** Encipher the message **attack** with  $\alpha = 9$  and  $\beta = 13$ .

First, we will convert the plaintext to numbers as follows

$$\text{attack} = p_1, p_2, p_3, p_4, p_5, p_6 = 0, 19, 19, 0, 2, 10.$$

Then, the ciphertext will be given by

$$c \equiv 9p + 13 (\text{mod } 26),$$

Hence,  $c_1, c_2, c_3, c_4, c_5, c_6 = 13, 2, 2, 13, 5, 25 = \text{NCCNFZ}$ .

**Decryption** First, we will convert the ciphertext **NCCNFZ** to numbers as follows:

$$\text{NCCNFZ} = c_1, c_2, c_3, c_4, c_5, c_6 = 13, 2, 2, 13, 5, 25.$$

We have  $\alpha^{-1} \equiv 3 (\text{mod } 26)$ .

$$p \equiv \alpha^{-1}(c - \beta)(\text{mod } 26).$$

Therefore,

$$p_1 = 3(13 - 13)(\text{mod } 26) \equiv 0, \text{ and}$$

$$p_2 = 3(2 - 13)(\text{mod } 26) \equiv 19.$$

Similarly,  $p_3, p_4, p_5, p_6 = 19, 0, 2, 10$ .

**Definition** The set of all possible keys is called the *key space*.

The key space in the **additive cipher** equals **26**.

The key space in the **affine cipher** is given by

$$(\text{Number of values for } \beta) \times (\text{Number of values for } \alpha) = (26 \times 12) = 312.$$

A key space with **312** elements can, of course, still be searched exhaustively. In addition, the affine cipher has the same weakness as the shift and substitution cipher: The mapping between plaintext letters and ciphertext letters is fixed. Hence, it can easily be broken with **letter frequency analysis**.

### **Letter Frequency Analysis**

In cryptanalysis, frequency analysis is the study of the frequency of letters or groups of letters in a cipher text. The method is used as an aid to breaking classical ciphers.

**Letter Frequency analysis** is based on the fact that, in any given stretch of written language, certain letters and combinations of letters occur with varying frequencies. Moreover, there is a characteristic distribution of letters that is roughly the same for almost all samples of that language. For instance, given a section of English language **E, T, A** and **O** are the most common, while **Z, Q** and **X** are rare. Likewise, **TH, ER, ON,** and **AN** are the most common pairs of letters (termed bigrams or digraphs), and **SS, EE, TT,** and **FF** are the most common repeats and the **THE** most common trigram in the English language. The nonsense phrase "**ETAOIN SHRDLU**" represents the 12 most frequent letters in typical English language text.

In some ciphers, such properties of the natural language plaintext are preserved in the cipher text, and these patterns have the potential to be exploited in a cipher text-only attack.

Letter Frequency Analysis is a cryptanalysis technique of studying the frequency that letters occur in the encrypted cipher text. In English, certain letters are more **28** commonly used than others.

This fact can be used to take educated guesses at deciphering a Monoalphabetic Substitution Cipher.

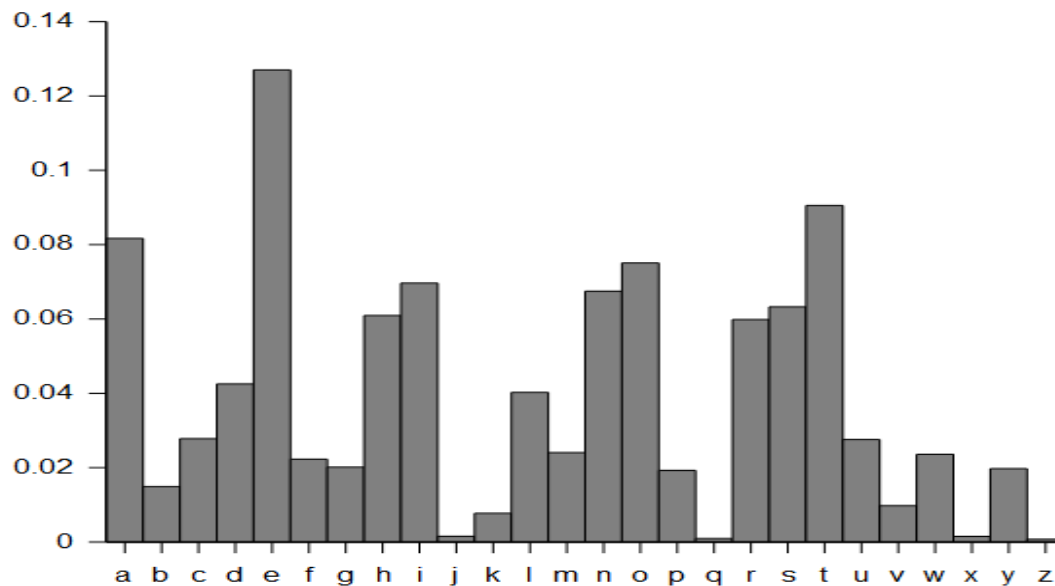
Here is the alphabet in order of the frequency that each letter is used.

**E, T, A, O, I, N, S, R, H, L, D, C, U, M, F, P, G, W, Y, B, V, K, X, J, Q, Z**

a	b	c	d	e	f	g	h	i	j	k	l	m
8.2	1.5	2.8	4.3	12.7	2.2	2.0	6.1	7.0	0.2	0.8	4.0	2.4
n	o	p	q	r	s	t	u	v	w	x	y	z
6.7	7.5	1.9	0.1	6.0	6.3	9.1	2.8	1.0	2.4	0.2	2.0	0.1

and ranked in order:

e	t	a	o	i	n	s	h	r	d	l	u	c
12.7	9.1	8.2	7.5	7.0	6.7	6.3	6.1	6.0	4.3	4.0	2.8	2.8
m	w	f	y	g	p	b	v	k	x	j	q	z
2.4	2.4	2.2	2.0	2.0	1.9	1.5	1.0	0.8	0.2	0.2	0.1	0.1



Now, we may classify the frequencies into **5** groups, as follow

(I) **e**

(II) **t,a,o,i,n,s,h,r**

(III) **d,l**

(IV) **c,u,m,w,f,g,y,p,b**

(V) **v,k,j,x,q,z**

(1)The word “**the**” has a **tremendous effect** on these statistics. It is the main reason for the high frequency of **t**, **h**, **th**, **he** and **the**. If the word “**the**” were deleted from the plain text then **t** would fall below some of the other letters in group (II), while **h** would fall into group (III). Furthermore **th** and **he** would no longer be the most popular bigrams.

(2) Over half the words in the English language end in **e**, **s**, **d** or **t**.

(3) About half the words in English language begin with **t**, **a**, **s** or **w**.

**Example**

In the key-phrase cipher, the key may be consists of a key word letter together with one extra **special letter**. In this case, the plaintext alphabet is written out and the key phrase then written underneath beginning at the special letter, but with no letter of the phrase repeated.

If the key phrase is **MY LITTLE FINGER** and the special letter is **h**, then we get

Plaintext																									
a	b	c	d	e	f	g	<u>h</u>	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z
Q	S	U	V	W	X	Z	<u>M</u>	<u>Y</u>	<u>L</u>	<u>I</u>	<u>T</u>	<u>E</u>	<u>F</u>	<u>N</u>	<u>G</u>	<u>R</u>	A	B	C	D	H	J	K	O	P
Ciphertext																									

Now, if the ciphertext is given by

YKHLBA JCZ SVIJ JZB TZVHI JCZ VHJ DR IZXKHLBA VSS RDHEI  
 DR YVJV LBXSKYLBA YLALJVS IFZZXC CVI LEFHDNZY EVBLRDSY  
 JCZ FHLEVHT HZVIDB RDH JCLI CVI WZZB JCZ VYNZBJ DR  
 ELXHDZSZXJHDBLXI JCZ XDEFSZQLJT DR JCZ RKBXJLDBI JCVJ  
 XVB BDP WZ FZHRDHEZY WT JCZ EVXCLBZ CVI HLIZB  
 YHVEVJLXVSST VI V HZIKSJ DR JCLI HZXZBJ YZNZSDFEZBJ LB  
 JZXCBDSDAT EVBT DR JCZ XLFCZH ITIJZEI JCVJ PZHZ DBXZ  
 XDBILYZHZY IZXKHZ VHZ BDP WHZVMVWSZ.

We see that, the encrypted message contain **338** letters with the following frequencies:

Letter	A	B	C	D	E	F	G	H	I	J	K	L
Frequency	5	24	19	23	12	7	0	24	21	29	6	21

Letter	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
Frequency	1	3	0	3	1	11	14	8	0	27	5	17	12	45



This means that the order of popularity is

**Z, J, V, B, H, D, I, L, C, X, S, Y, E, R, T, F, K, A, W, N, P, M, Q, U, G, O.**

- (1) Since the frequency of **Z** is significantly greater than all others we deduce that **Z** must correspond to **e**.
- (2) The next most popular is **J** but, since **29** is not much larger than **27**, we hesitate before assuming that **J** corresponds to **t**. The three word **JCZ** occur eight times and **J** is **t** and **Z** is **e** then **JCZ** is **the**.
- (3) Since the single letter word **V** occurs we know that **V** must be **a** or **i**. But **a** and **i** are both in group (II) which means we need some extra information before deciding between them. That extra information is provided by the existence of the cipher word **JCVJ**.
- (4) Since there is no English word **thit**, we now know **V** must be **a**. Consider the fourth **JZB**. We know **J** is **t**, **Z** is **e** and **B** is one of **o, n, r, i, s**. Trying each possibility for **B** gives plain text words of **teo, ten, ter, tei, tes**. Clearly the plaintext word should be **ten** which indicates **B** is **n**.
- (5) We now note the two letter word **VI**. Since **V** is **a**, **I** must be **s** or **n**. But we already know that **B** is **n** and this means that **I** must be **s**. The cipher word **VHJ** forces **H** to be **r** and **JCLI** forces **L** to be **i**. Thus **D** must be **o**.

### Plaintext

a	b	c	d	e	f	g	h	i	j	k	L	m	n	o	p	q	r	s	t	u	v	w	x	y	z
V				Z			C	L					B	D			H	I	J						

### Ciphertext

It is clear that the special letter was **f**. Since there are three gaps between **V** and **Z** and there are three letters between them in the alphabet we know that **W** is **b**, **X** is **c** and **Y** is **d**. Also there are two gaps between **D** and **H** and to deduce that two of **E, F** and **G** must be **p** and **q**. **DR** is **of**, **WT** is **by** and **BDP** is **now** .....

### Plaintext

a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z
V	W	X	Y	Z	R	A	C	L	O	M	S	E	B	D	F	G	H	I	J	K	N	P	Q	T	U

### Ciphertext

Thus we have a key phrase cipher with phrase **RACAL COMSEC** and special letter **f**. Then the message is:

### Plaintext

during the last ten years the art of securing all forms of data including digital speech has improved manifold the primary reason for this has been the advent of microelectronics the complexity of the functions that can now be performed by the machine has risen dramatically as a result of this recent development in technology many of the cipher systems that were once considered secure are now breakable.

## Other Ciphers and Codes

### Morse code symbols

Morse code symbols are. -, ..., .. -, - .-, -.. -, respectively, which are sufficiently dissimilar so as to avoid confusion. The complete table of Morse code symbols is given in the following

A . -	J . - - -	S . . .	2 . . - - -
B - . . .	K - . -	T -	3 . . . - -
C - . - .	L . . . .	U . . -	4 . . . . -
D - . .	M - -	V . . . -	5 . . . . .
E .	N - .	W . - -	6 - . . . .
F . . . .	O - - -	X - . . .	7 - . . . .
G - - .	P . - . .	Y - . - -	8 - . . . .
H . . . .	Q - - . -	Z - - . .	9 - - . . .
I . .	R . - .	1 . - - - -	0 - - - - -

### International Morse Code



### Example.

**Sofia** is encoded as ... - - - ..-. .. .-

**Eugenia** is encoded as . ..- - -. . -. .. .-

## ASCII Codes

The (**American Standard Code for Information Interchange**) is a character encoding scheme based on the ordering of the English alphabet.

**ASCII** is a code used by computers to represent characters as numbers. This allows computers to store a letter as one byte of information. One byte of information allows you to represent **256** different values, which is enough to encode all the letters (uppercase and lowercase) as well as the numbers **0-9** and other special characters.

ASCII Alphabet Characters

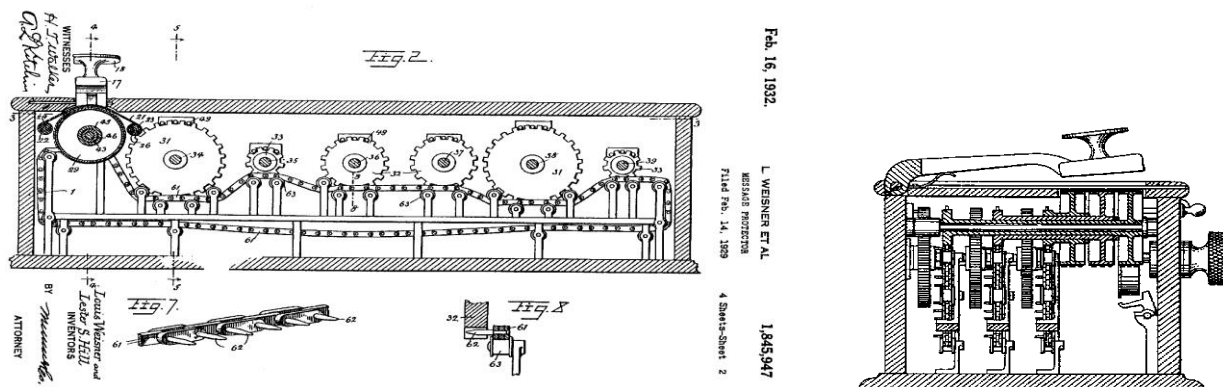
Symbol	Decimal	Binary	Symbol	Decimal	Binary	DEC	BINARY
A	65	1000001	a	97	1100001	0	000 0000
B	66	1000010	b	98	1100010	1	000 0001
C	67	1000011	c	99	1100011	2	000 0010
D	68	1000100	d	100	1100100	3	000 0011
E	69	1000101	e	101	1100101	4	000 0100
F	70	1000110	f	102	1100110	5	000 0101
G	71	1000111	g	103	1100111	6	000 0110
H	72	1001000	h	104	1101000	7	000 0111
I	73	1001001	i	105	1101001	8	000 1000
J	74	1001010	j	106	1101010	9	000 1001
K	75	1001011	k	107	1101011	10	000 1010
L	76	1001100	l	108	1101100	11	000 1011
M	77	1001101	m	109	1101101	12	000 1100
N	78	1001110	n	110	1101110	13	000 1101
O	79	1001111	o	111	1101111	14	000 1110
P	80	1010000	p	112	1110000	15	000 1111
Q	81	1010001	q	113	1110001	16	001 0000
R	82	1010010	r	114	1110010	17	001 0001
S	83	1010011	s	115	1110011	...	...
T	84	1010100	t	116	1110100	30	0011110
U	85	1010101	u	117	1110101	...	...
V	86	1010110	v	118	1110110	127	1111111
W	87	1010111	w	119	1110111	...	...
X	88	1011000	x	120	1111000	...	...
Y	89	1011001	y	121	1111001	...	...
Z	90	1011010	z	122	1111010	...	...

## The Hill Cipher (Linear Algebra in Cryptography)

The *Hill Cipher* was invented in **1929** by **Lester S. Hill**. A Hill cipher is a type of *polygraphic cipher*, where plaintext is divided into groups of letters of a fixed size and then each group is transformed into a different group of letters.

A *block cipher* is a cipher in which groups of letters are enciphered together in equal length blocks. The **Hill cipher** is an example of a block cipher.

**Hill and Louis Weisner** also had an idea to build a machine that would mechanically implement a Hill cipher. They named it the *Message Protector* and patented it. It operated on blocks of six letters and did all of the arithmetic using gears and pulleys. The diagram below is from their patent application.



Hill's cipher machine

### Encryption Algorithm

(1) As in the additive cipher, we first convert all the letters to numbers ('a'=0, 'b'=1,..., 'z'=25).

The key used in the *Hill Cipher* is some kind of  $n \times n$  invertible matrices whose elements are from  $Z_{26}$ . If we want to make the cipher more secure, we are able to work with an alphabet size other than  $m=26$ , but we have to be careful to pick a **key matrix** that is invertible.

**Example** We may work with the following alphabet

a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15

q	r	s	t	u	v	w	x	y	z	,	.	?
16	17	18	19	20	21	22	23	24	25	26	27	28

In his case, we have  $m=29$ .

(2) Select an  $n \times n$  invertible matrix  $E$  whose elements are from  $Z_m$ .

(3) Convert each character to a number between 0 and  $m$ .

(4) Divide this string of numbers up into blocks of size  $n$ . If the message does not break evenly into blocks of size  $n$  we pad the ending of the message with “z”,s and characters, this can be done at random.

(5) Write each block as a column vector of size  $n$ . At this point the message is a sequence of  $n$ -dimensional vectors,  $p_1, p_2, \dots, p_t$ .

(6) Take each of the vectors and multiply them by the encryption matrix  $E$ , so  $c_i = Ep_i \pmod{m}$  (or  $C = EM \pmod{m}$ ). Then we have the ciphertext  $c = c_1c_2 \dots c_t$ .

**Decryption Algorithm**

The decryption algorithm is essentially the same as the encryption algorithm, except that we use  $E^{-1}$  in place of  $E$ . Since  $C = EP \pmod{m}$  and  $E$  is invertible  $\pmod{m}$ , we can calculate the message  $P = E^{-1}C \pmod{m}$  (or  $p_i = E^{-1}c_i \pmod{m}$ ).

**Example** Suppose that Alice wants to send Bob the message “**Cryptography is cool**”. Alice chooses the 26 letters alphabet, the block size  $n = 3$  and chooses the encryption matrix  $E$  over  $Z_{26}$  to be,

$$E = \begin{bmatrix} 2 & 3 & 15 \\ 5 & 8 & 12 \\ 1 & 13 & 4 \end{bmatrix}.$$

**Solution**

- Removing the spaces of the plaintext to get **cryptophyiscool**, then convert it to numbers

**2 17 24 15 19 14 6 17 0 15 7 24 8 18 2 14 14 11**

- Dividing this string of numbers up into blocks of length 3 to get

**2 17 24 15 19 14 6 17 0 15 7 24 8 18 2 14 14 11**

$$\text{Hence, } p_1 = \begin{bmatrix} 2 \\ 17 \\ 24 \end{bmatrix}, p_2 = \begin{bmatrix} 15 \\ 19 \\ 14 \end{bmatrix}, p_3 = \begin{bmatrix} 6 \\ 17 \\ 0 \end{bmatrix}, p_4 = \begin{bmatrix} 15 \\ 7 \\ 24 \end{bmatrix}, p_5 = \begin{bmatrix} 8 \\ 18 \\ 2 \end{bmatrix} \text{ and } p_6 = \begin{bmatrix} 14 \\ 14 \\ 11 \end{bmatrix}.$$

$$c_1 = Ep_1 \pmod{26} = \begin{bmatrix} 2 & 3 & 15 \\ 5 & 8 & 12 \\ 1 & 13 & 4 \end{bmatrix} \cdot \begin{bmatrix} 2 \\ 17 \\ 24 \end{bmatrix} \pmod{26} = \begin{bmatrix} 25 \\ 18 \\ 7 \end{bmatrix}.$$

$$c_2 = Ep_2 \pmod{26} = \begin{bmatrix} 2 & 3 & 15 \\ 5 & 8 & 12 \\ 1 & 13 & 4 \end{bmatrix} \cdot \begin{bmatrix} 15 \\ 19 \\ 14 \end{bmatrix} \pmod{26} = \begin{bmatrix} 11 \\ 5 \\ 6 \end{bmatrix}.$$

$$\text{Similarly, } c_3 = \begin{bmatrix} 11 \\ 10 \\ 19 \end{bmatrix}, c_4 = \begin{bmatrix} 21 \\ 3 \\ 20 \end{bmatrix}, c_5 = \begin{bmatrix} 22 \\ 0 \\ 16 \end{bmatrix} \text{ and } c_6 = \begin{bmatrix} 1 \\ 2 \\ 6 \end{bmatrix}.$$

So Alice will send “**ZSHLFGLKTVDUWAQBCG**” to Bob.

**Decryption** We decipher the ciphertext: **ZSHLFGLKTVDUWAQBCG**

First, we will find  $E^{-1}(\text{mod } 26)$ ,

$$E^{-1}(\text{mod } 26) = \begin{bmatrix} 2 & 3 & 15 \\ 5 & 8 & 12 \\ 1 & 13 & 4 \end{bmatrix}^{-1} (\text{mod } 26) = \begin{bmatrix} 10 & 19 & 16 \\ 4 & 23 & 7 \\ 17 & 5 & 19 \end{bmatrix}.$$

Convert the ciphertext into numbers then divide this string of numbers up into blocks of length 3 to get

**25 18 7   11 5 6   11 10 19   21 3 20   22 0 16   1 2 6**

$$c_1 = \begin{bmatrix} 25 \\ 18 \\ 7 \end{bmatrix}, c_2 = \begin{bmatrix} 11 \\ 5 \\ 6 \end{bmatrix}, c_3 = \begin{bmatrix} 11 \\ 10 \\ 19 \end{bmatrix}, c_4 = \begin{bmatrix} 21 \\ 3 \\ 20 \end{bmatrix}, c_5 = \begin{bmatrix} 22 \\ 0 \\ 16 \end{bmatrix} \text{ and } c_6 = \begin{bmatrix} 1 \\ 2 \\ 6 \end{bmatrix}.$$

Finally,  $p_i = E^{-1}c_i (\text{mod } 26)$ .

$$p_1 = \begin{bmatrix} 10 & 19 & 16 \\ 4 & 23 & 7 \\ 17 & 5 & 19 \end{bmatrix} \begin{bmatrix} 25 \\ 18 \\ 7 \end{bmatrix} (\text{mod } 26) = \begin{bmatrix} 2 \\ 17 \\ 24 \end{bmatrix}, p_2 = \begin{bmatrix} 15 \\ 19 \\ 14 \end{bmatrix}, p_3 = \begin{bmatrix} 6 \\ 17 \\ 0 \end{bmatrix}, p_4 = \begin{bmatrix} 15 \\ 7 \\ 24 \end{bmatrix},$$

$$p_5 = \begin{bmatrix} 8 \\ 18 \\ 2 \end{bmatrix} \text{ and } p_6 = \begin{bmatrix} 14 \\ 14 \\ 11 \end{bmatrix}.$$

Now, convert the numbers  $p_1, p_2, p_3, p_4, p_5$  and  $p_6$  to letters to get

**cryptographiscool**

So, Bob adds in a couple spaces to get the plaintext **“cryptography is cool”**.

### Exercise

(1) The ciphertext below was encrypted using a shift cipher. Decrypt the ciphertext without knowledge of the key using a letter frequency analysis

**xultpaajexitltlxaarpjhtiwtgxktghidhipxciwvtgtpilpit ghlxiwiwtxgqadds.**

(2) Decrypt the text below:

**Falszztysyjzyjkywjrztyjztyynaryjkyswarztyegyyj**

Knowing that, an affine cipher was used with  $\alpha = 7$  and  $\beta = 22$ .