Ciphers: Making and Breaking

Ralph Morelli

Trinity College, Hartford

(ralph.morelli@trincoll.edu)

Smithsonian Institute
October 31, 2009

Trinity College
HARTFORD CONNECTICUT

This presentation was created using Open Office 3.0,
free and open source software.
http://www.openoffice.org/

Today's Themes

- Crypto Cat vs. Mouse: Cryptographer vs. Cryptanalyst
- Theoretical security vs. practical security
- Implementation, implementation, implementation
- The key is security.

Part I: Classical Cryptology

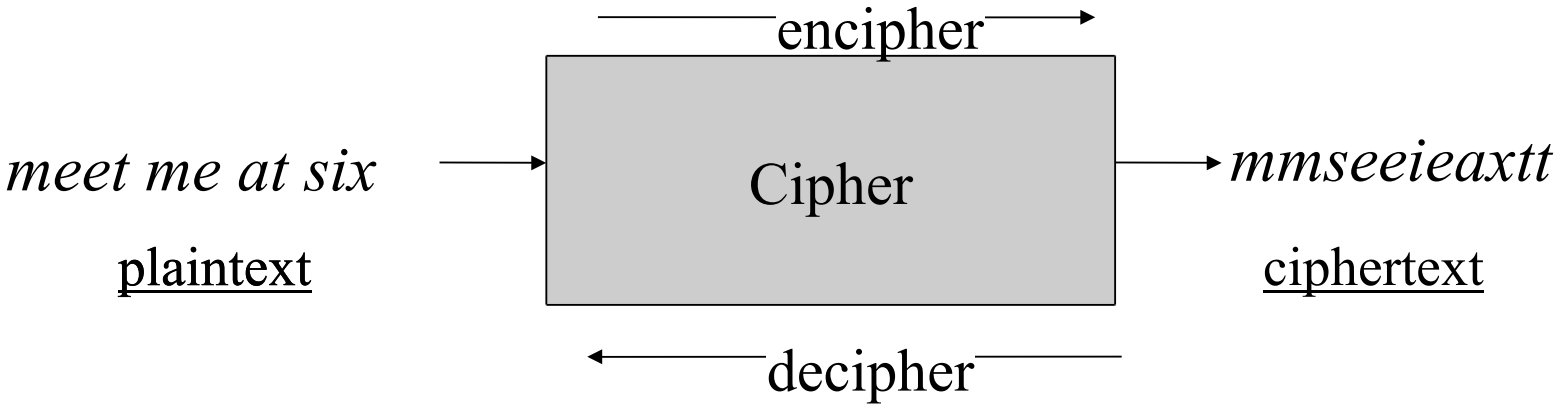
Basic Terminology

Secret Writing

- Cryptography – making secret messages.
- Cryptanalysis – breaking secret messages.
- Cryptology – cryptography and cryptanalysis.
- Steganography – concealing messages.

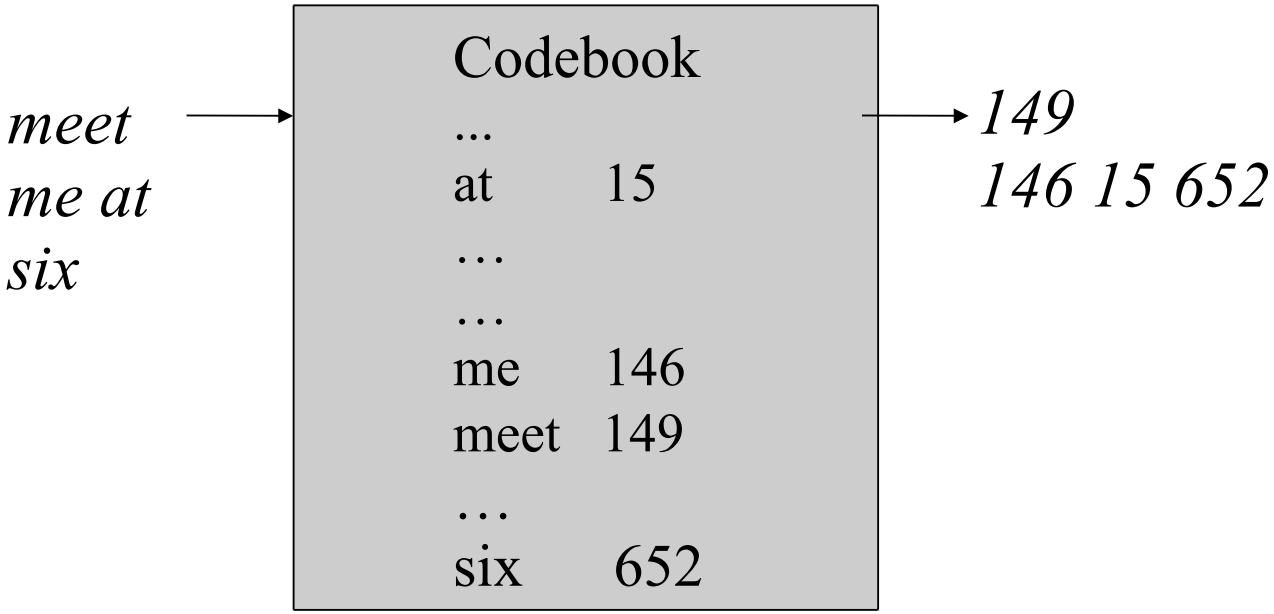
Cipher

- Cipher – a method for transforming a message.

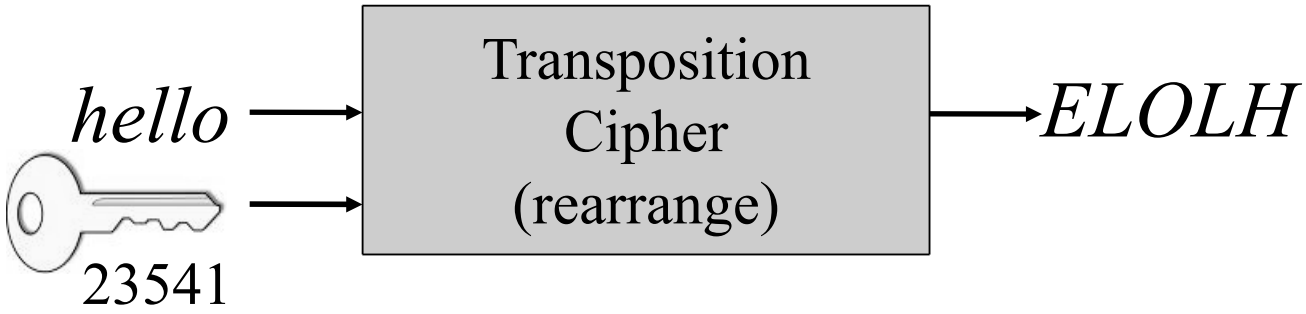


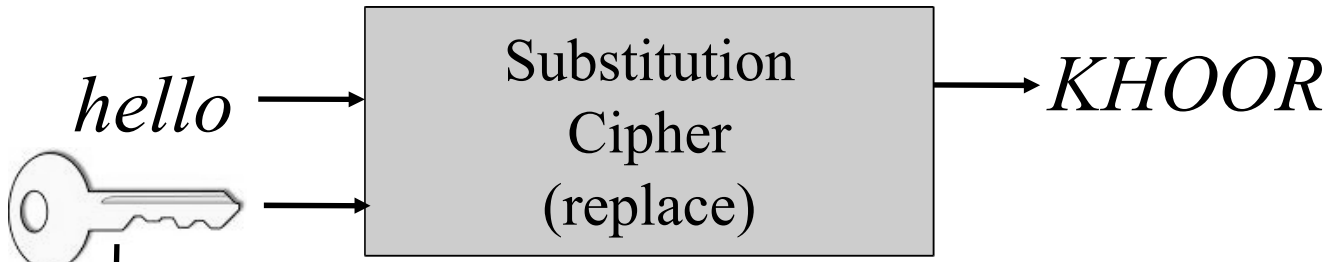
Code

- Code – a system in which codewords replace plaintext words or symbols using a codebook.



Types of Cipher

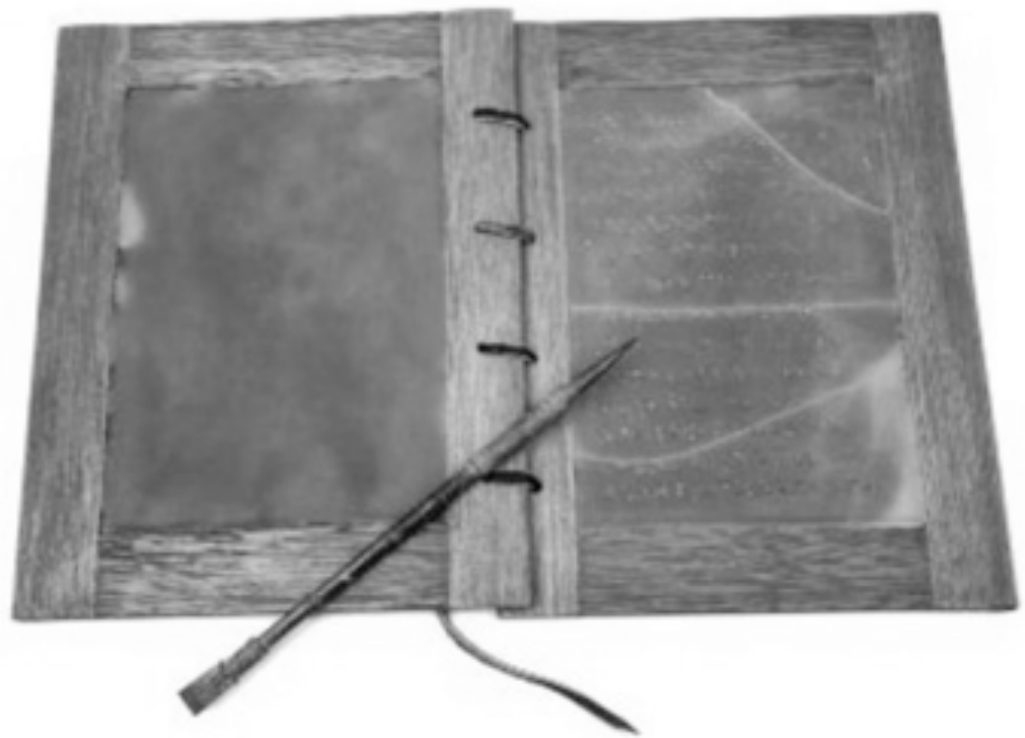





```
Plaintext alphabet:  a b c d e f g h i j k l m n o p q r s t u v w x y z
Cipher alphabet:     D E F G H I J K L M N O P Q R S T U V W X Y Z A B C
```

Earliest Ciphers

Herodotus (484-425 B.C.)

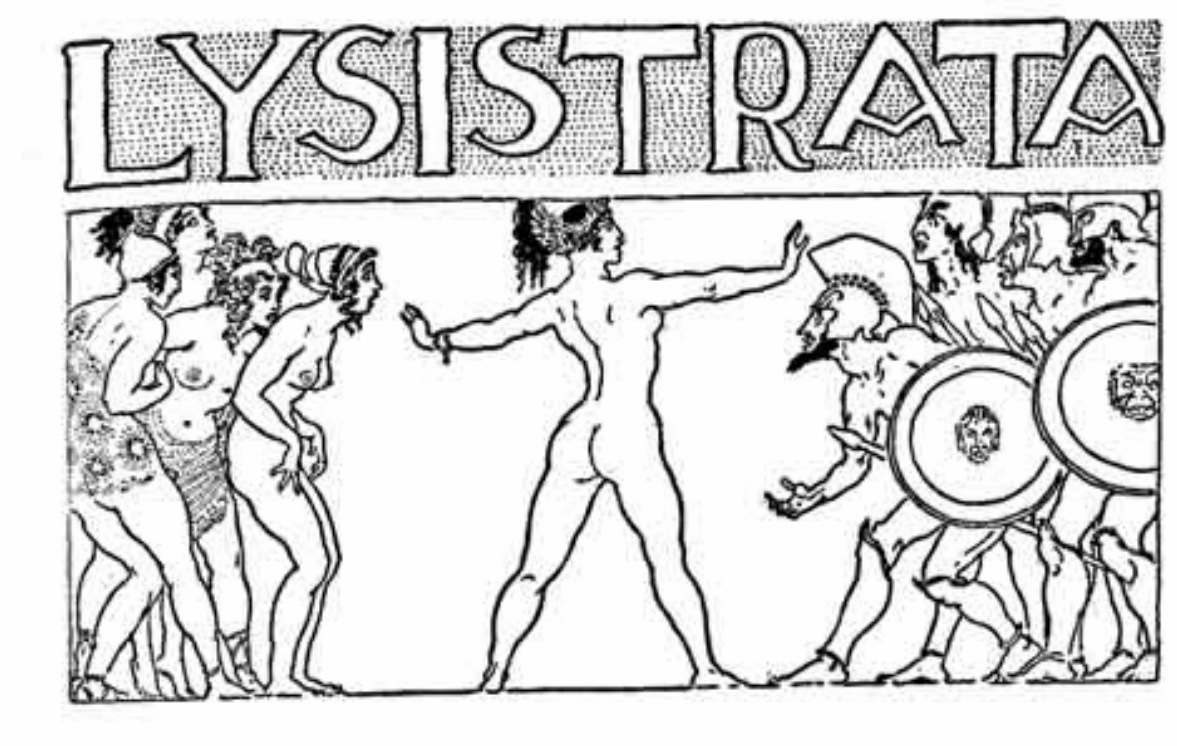


- Persian War (~480 BC), Demaratus' secret message to the Spartans was hidden under the wax on a wooden tablet (steganography).

The Scytale

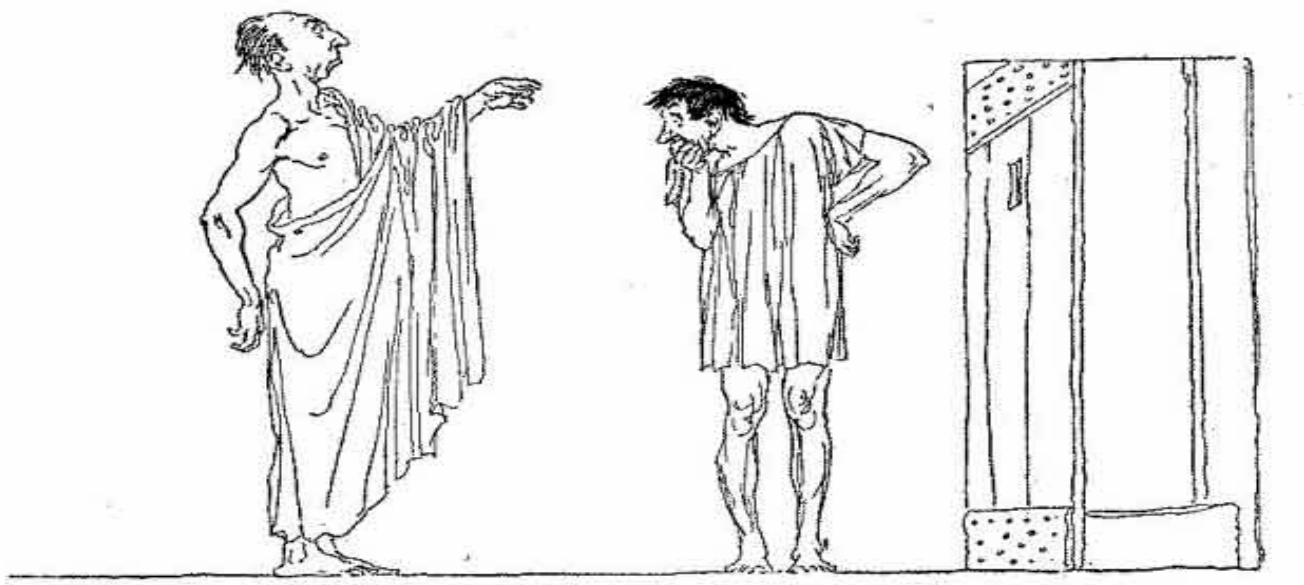
- According to Plutarch (120 AD), the Scytale was a Greek transposition cipher.
- Some scholars dispute this.

Lysistrata



- Aristophanes (446 – 386 BC).
- Comic anti war (Peloponnesian) play.

The Herald and the Scytale



- MAGISTRATE: Then why do you turn aside and hold your cloak so far out from your body? Is your groin swollen with stress of travelling?
- HERALD: It's my despatch cane.
- MAGISTRATE: Of course--a Spartan cane!

Skytale Encryption/Decryption

1) Attach paper strip to one end of the dowel.

2) Wind it tightly and fasten the other end.

3) Write each line of the message along dowel length.

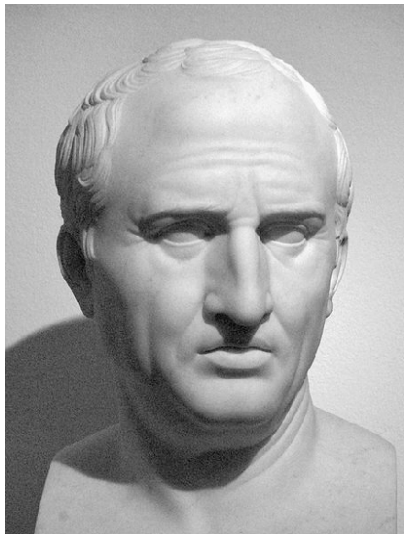
4) Remove and transmit the message.

5) To decrypt, repeat steps 1-4 on the same sized dowel.

Skytale Cryptanalysis

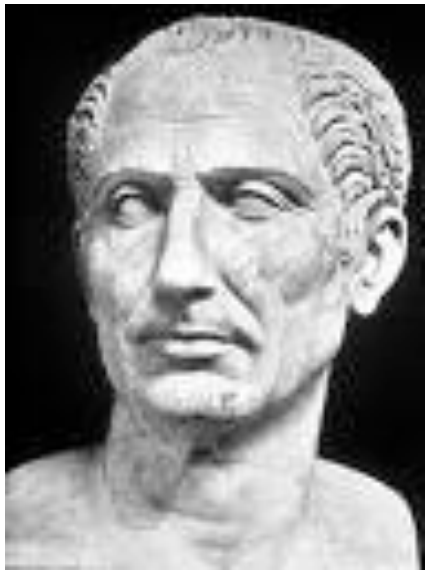
- What is the key for this cipher?
- Brute force attack: try all possible keys.
- How many possible keys?
  - The more keys the more secure.
- A cipher or a communication device?

The Caesar Cipher



```
a b c d e f g h i j k l m
D E F G H I J K L M N O P

n o p q r s t u v w x y z
Q R S T U V W X Y Z A B C
```



- Julius Caesar (~ 50 B.C., Gallic Wars, chapter 48) reported using secret writing to communicate with Cicero.
- Suetonius reported that Caesar used an alphabet with a shift of 3.

ROT13

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

Q: How can you tell an extrovert from an introvert at NSA?

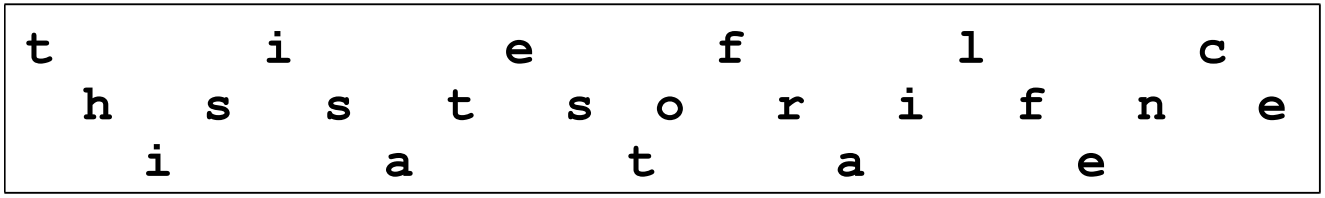
A: Va gur ryringbef, gur rkgebireg ybbxf ng gur BGURE thl'f fubrf.

In the elevators, the extrovert looks at the OTHER guy's shoes.

Some Simple Transposition Ciphers

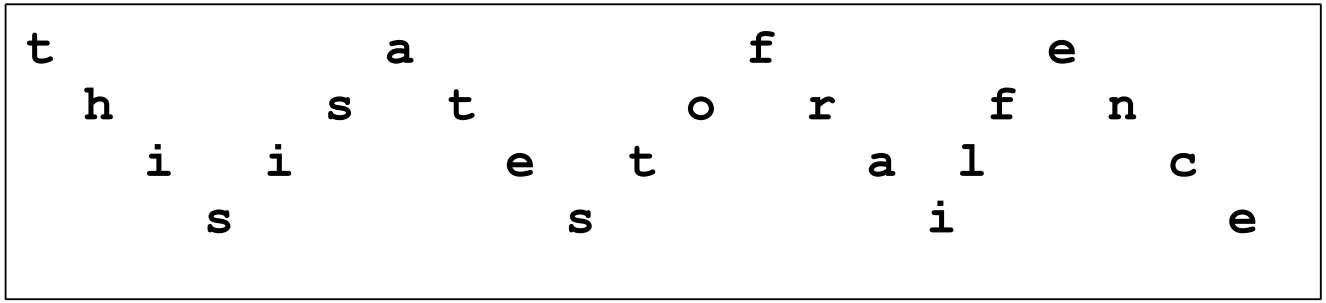
Rail Fence Cipher

Plain: thisisatestofrailfence



Depth 3

Crypto: TIEFLCHSSTSORIFNEIATAE



Depth 4

Crypto: TAFEHSTORFNIIETALCSSIE

Route Cipher

Plain: thisatestoftheroutecipher

```
t h i s a t
e s t o f t
h e r o u t
e c i p h e
r z y x w v
```

Reverse horizontal: TASIHT TFOTSE TUOREH EHPICE VWXYR

Down vertical : TEHER HSECZ ITRIY SOOPX AFUHW TTTEV

Diagonal: T AT SFT IOUE HTOHV TSRPW EEIX HCY EZ R

Spiral counterclockwise: VETTT ASIHT EHER ZYXW HUF OTS ECIPOR

Columnar Transposition

Plain: thisisatestofcolumntransposition

```
T H I S I S
A T E S T O
F C O L U M
N T R A N S
P O S I T I
O N Z Y X W

6 3 2 4 1 5
```

Keyword = ZEBRAS = 632415

Ciphertext: ITUNTX IEORSZ HTCTON SSLAIY SOMSIW TAFNPO

Cryptanalysis

- Brute force attack – test every possible key.
- Factoring – break the message into rectangles.
- Frequency analysis – count letters, vowels.
- Probable word attack – look for probable words.
- Digram, trigram analysis – look for frequent 2- and 3-letter sequences.

Abū Yūsuf Ya‘qūb ibn Isḥāq al-Kindī

- 801-873 CE
- “The Arab Philosopher”
- Arab Iraqi polymath.
- *A Manuscript on Deciphering Cryptographic Messages* (Ottoman Archives, Instanbul)
- First explanation of frequency analysis.
- Dark Ages in the West

Attacking Transpositions

- Factor – to identify possible rectangles.
  - 64 letters could be 8 x 8 or 16 x 4 or four 4 x 4.
- Arrange – into rectangle.
- Count the vowels – 40% vowels, 12% *e*'s
  - Rows and columns should have ~ 40% vowels.
- Rearrange – rows and/or columns.
  - Digrams, trigrams, probable sequences guide the rearrangements.

Columnar Transposition Case

Message: ETTUH OMEAW EXETE STDEH TIGYT HLKHS MAODT O

Letters = 36 Possible squares 6x6, 9x4, 12x3

```
E M E E T M   3
T E T H H A   2
T A E T L O   3
U W S I K D   2
H E T G H T   1
O X D Y S O   2

3 3 2 2 0 3
```

```
E W E K   2
T E H H   1
T X T S   0
U E I M   3
H T G A   1
O E Y O   4
M S T D   0
E T H D   1
A D L O   2

5 3 3 3
```

- Analysis:
  - In the 6x6 square, the 40% rule gives 2-3 vowels per row or column.
  - The rows seem better, so rearrange the columns.

Let's Solve It !

IELTB OSTAA RIUZN IDFEM XKSMO FPYIH WRMVT TEOKA EEHRR EYCR

- Hint 1: Count the letters
  - 49
- Hint 2: Factor into a rectangle
  - 7 x 7
- Hint 3: Place 7 letters, L to R, on 7 strips.
- Hint 4: Rearrange the columns.
  - X, Y, Z ??

I think there is a world market for maybe five computers.
- Thomas J. Watson, IBM, 1943

Hands On Example

E M E E T M	E M E E T M	M E E T M E	M E E T M E	M E E T M E
T E T H H A	T E T H H A	E T H H A T	E T T H A H	A T T H E H
T A E T L O	E A T T L O	A T T L O E	A T E L O T	O T E L A T
U W S I K D	S W U I K D	W U I K D S	W U S K D I	D U S K W I
H E T G H T	T E H G H T	E H G H T T	E H T H T G	T H T H E G
O X D Y S O	D X O Y S O	X O Y S O D	X O D S O Y	O O D S X Y

- Analysis
  - AE (row 3) don't go, swap cols 1 & 3.
  - MEET ME in row 1 by moving col 1 to end
  - HH (row 2) don't go, swap cols 3 & 6
  - THA becomes THE (row 2) by swapping cols 1 and 5.

Message: MEET ME AT THE HOTEL AT DUSK WITH THE GOODS

Shortcomings of Transposition

- A minor transmission error could garble the message.

Substitution Ciphers

Simple Substitution

- Definition – letters are replaced by other letters or symbols.
- Simple substitution – 26! possible keys.
  - Permutations of the alphabet.
- 26! = 26 * 25 * 24 * … * 2 * 1 ≅ 10^{26} keys.

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

The Argenti's of Rome

- Giovanni Battista Argenti – cipher secretary for Pope Sixtus V in 1589.
- Matteo Argenti (nephew) – cipher secretary for 6 popes through 1602.
- Matteo authored 135-page cipher manual.
- Invented *mnemonic keyword* technique:

p	i	e	t	r	o	a	b	c	d	f	g	h	l	m	n	q	s	u	z
10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29

Analyzing Simple Substitution

- Recognizing simple substitution.
  - Non-standard letter frequencies.
- Frequency Analysis.
  - Given enough text (> 50 letters).
  - Most frequent cipher letters correspond to most frequent plaintext letters.
- Other Attacks.
  - Probable word.
  - *N*-gram – bigram, trigram, etc.

Characteristics of English Text

- High Frequency Group:
  - E T A O N I R S H (70% of text)
  - 12 10 8 8 7 7 6 7 5 (individual %)
- T O S – frequent initial and final letters
- A I H – frequent initial but not final
- E N R – frequently final but less so as initials.

Vowel Solution Method

- Count letter frequencies.
- Separate vowels and consonants.
  - Vowels are high frequency.
- Identify individual vowels.
  - *E* is most frequent;
  - *E* almost never touches *O;*
  - *EE* and *OO* are relatively frequent;
  - *A* follows but never precedes *E, etc.*
- Identify recognizable consonants.
  - *H* precedes and never follows vowels;
  - *TH* and *HE* and *HA* are common bigrams, etc.

```
 5 10 11  3  3  2  1  9  4  4  2 11 10  1 10  6  4  9  6  2  1  3 11  4
 F  D  R  J  N  U  H  V  X  X  U  R  D  M  D  S  K  V  S  O  P  J  R  K
                                                 ----                 ----
 5 10  3  5  5  3  4  1  6 11 11  9  3  1  3 11  3 10  2 11  3 10  5  9
 Z  D  Y  F  Z  J  X  G  S  R  R  V  T  Q  Y  R  W  D  A  R  W  D  F  V
 ----------                    ----       -------     -------

11  4  9 10 11  4  9  3 10  5  6  5  5 10  3  5 11 10  3  3  9  2  9  3
 R  K  V  D  R  K  V  T  D  F  S  Z  Z  D  Y  F  R  D  N  N  V  O  V  T
 -------     -------                    ----------                 ----
                    -----
 6  4  6  2  3  9  5 11
 S  X  S  A  W  V  Z  R
```

Helen Fouche Gaines, p. 73ff.

Frequency Counts

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
2			10		5	1	1		3	4		1	3	2	1	1	11	6	3	2	9	3	4	3	5

High Frequency Contacts

R	D	V	S	F	Z	K	X
D.J	F.R	H.X	D.K	..D	K.D	S.V	V.X
U.D	R.M	K.S	V.O	Y.Z	F.J	R.Z	X.U
J.K	M.S	R.T	G.R	D.V	S.Z	R.V	J.G
S.R	Z.Y	F.R	F.Z	D.S	Z.D	R.V	S.S
R.V	W.A	K.D	T.X	Y.R	V.R		
Y.W	W.F	K.T	X.A				
A.W	V.R	N.O					
V.K	T.F	O.T					
D.K	Z.Y	W.Z					
F.D	R.N						
Z..							

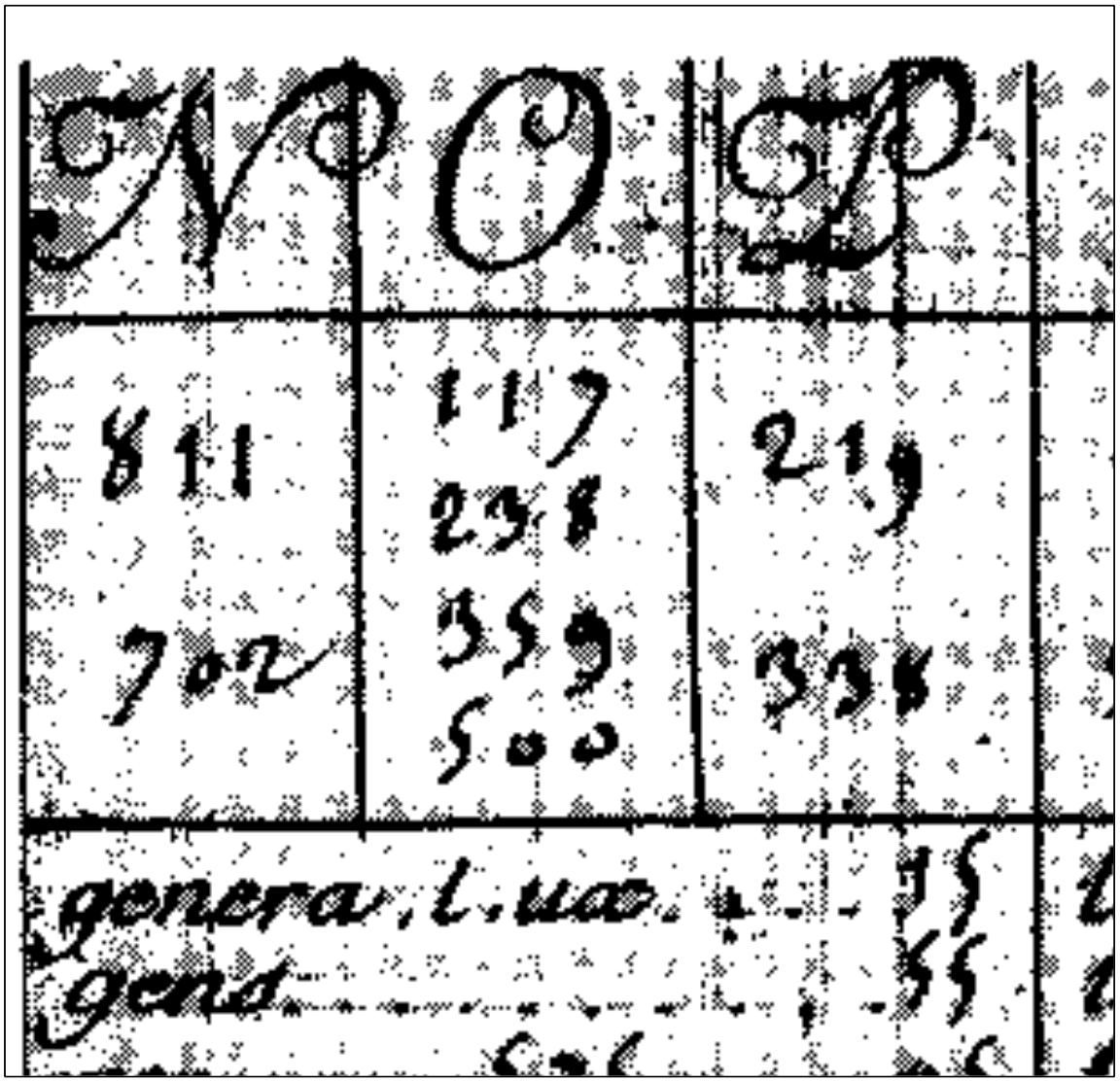
Low Frequency Contacts

G	H	M	P	Q
X.S	U.V	D.D	O.J	T.Y

Moderate Frequency Contacts

J	N	T	W	Y	A	O	U
R.N	J.U	V.Q	R.D	D.F	D.R	S.P	N.H
P.R	D.N	V.D	R.D	Q.R	S.W	V.V	X.R
Z.X	N.V	V.S	A.V	D.F			

Resisting Frequency Analysis



Nomenclator

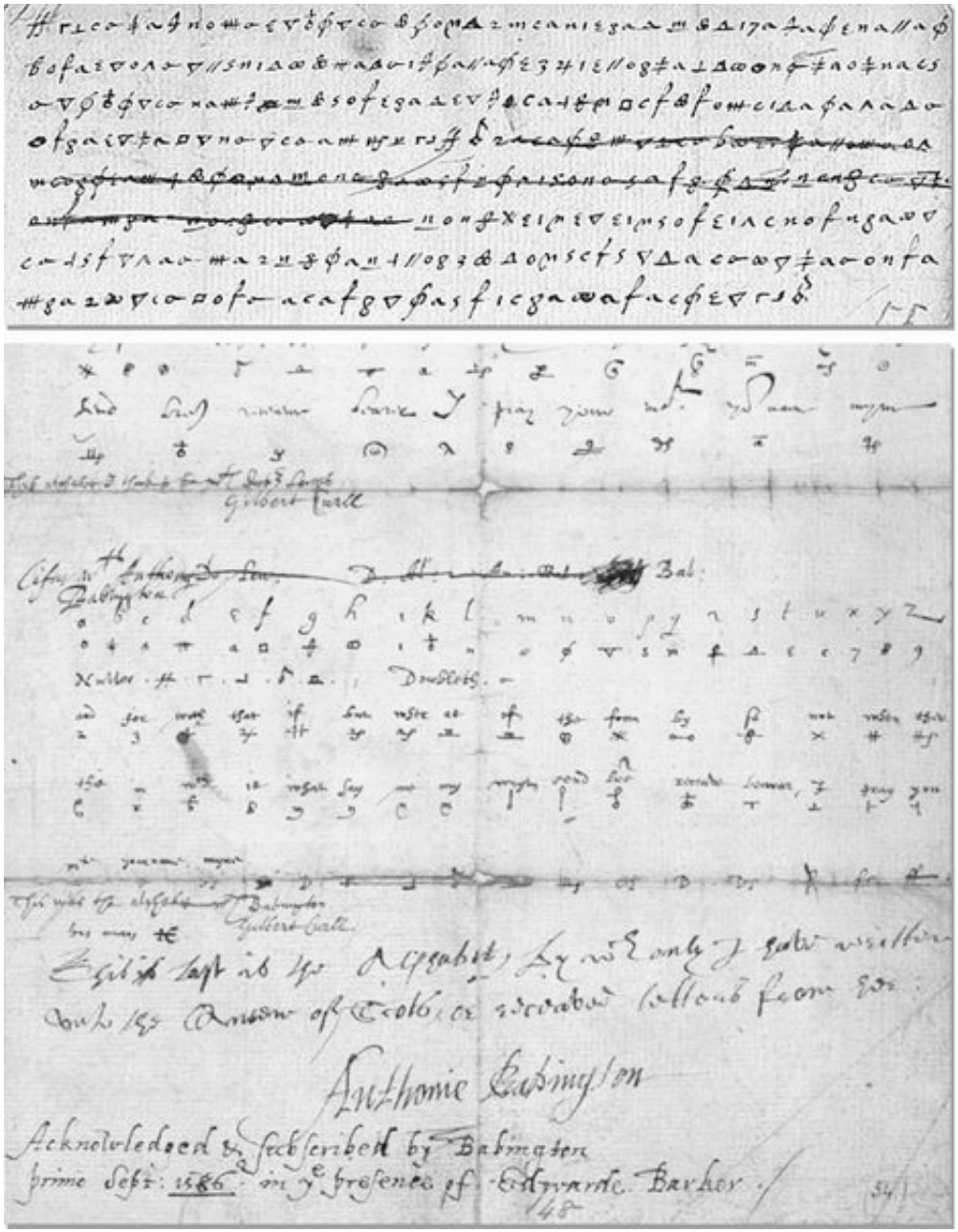
from

Rossignol's

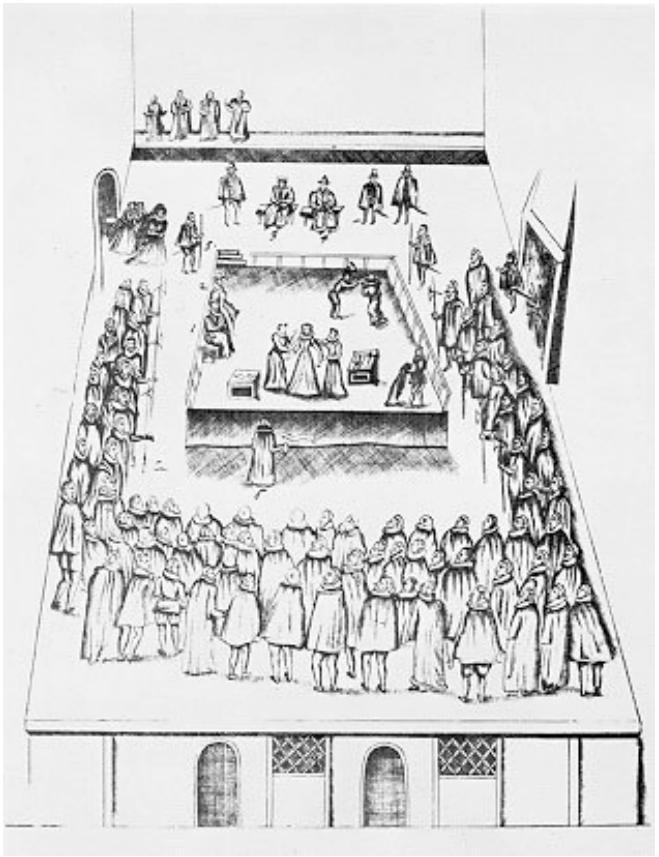
Great Cipher

The Nomenclator

- Early 15th – late 18th centuries.
- Homophones – plaintext letters map to more than one ciphertext letter.
- Code book – names of people and places are replaced by a symbol.
- Breakable by mid 16th century.
- Rossignol's Great Cipher (1620s) wasn't broken until 1893 by Etienne Bazeries who guessed that symbols represented syllables and not letters.

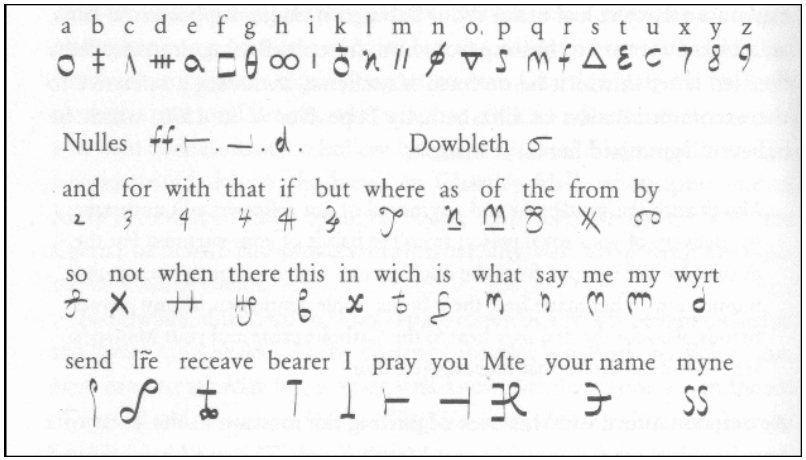


Mary Queen of Scots Cipher



Babington Plot

Mary, Queen of Scots
1542-1587



Mary Queen of Scots Nomenclator
(Singh,*The Code Book,* p. 38

Queen Elizabeth
1533-1603

Thomas Babington
1561-1586

Sir Francis Walsingham
1532-1590