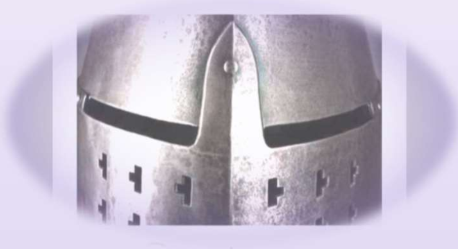


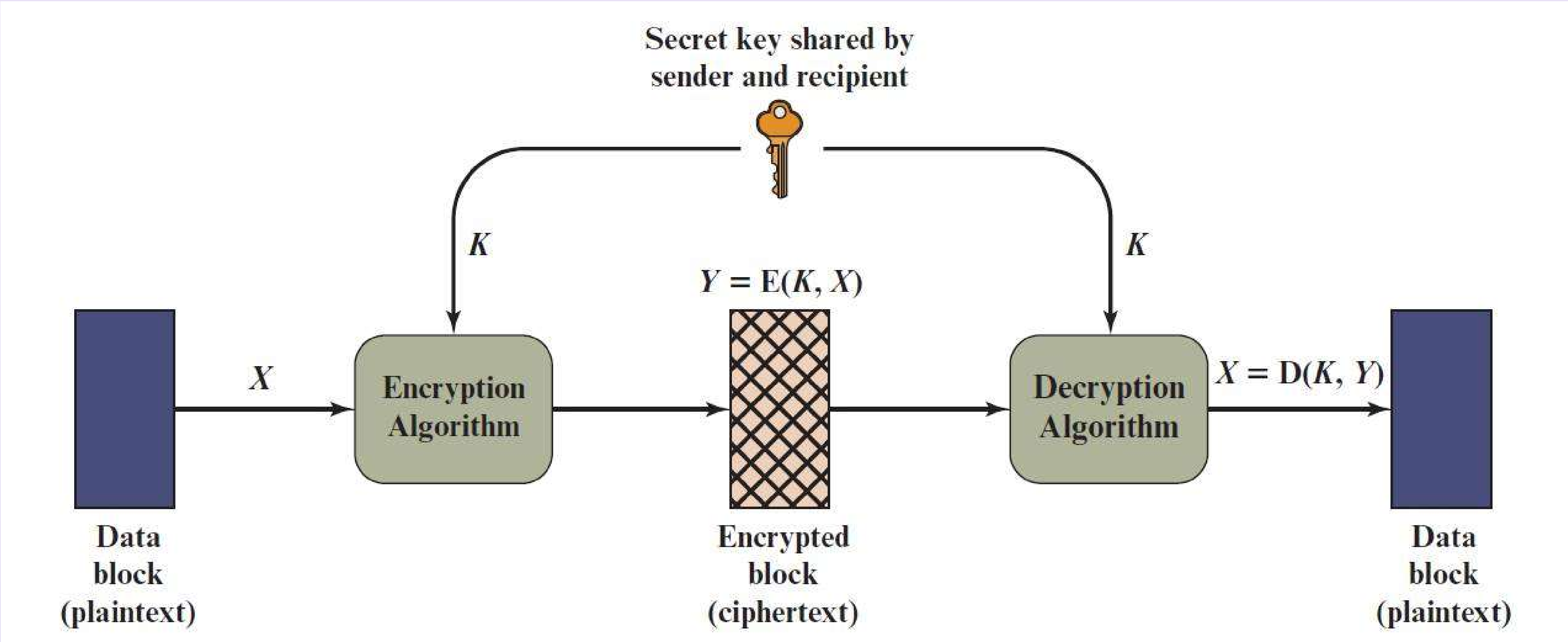
Chapter 3

Classical Encryption Techniques

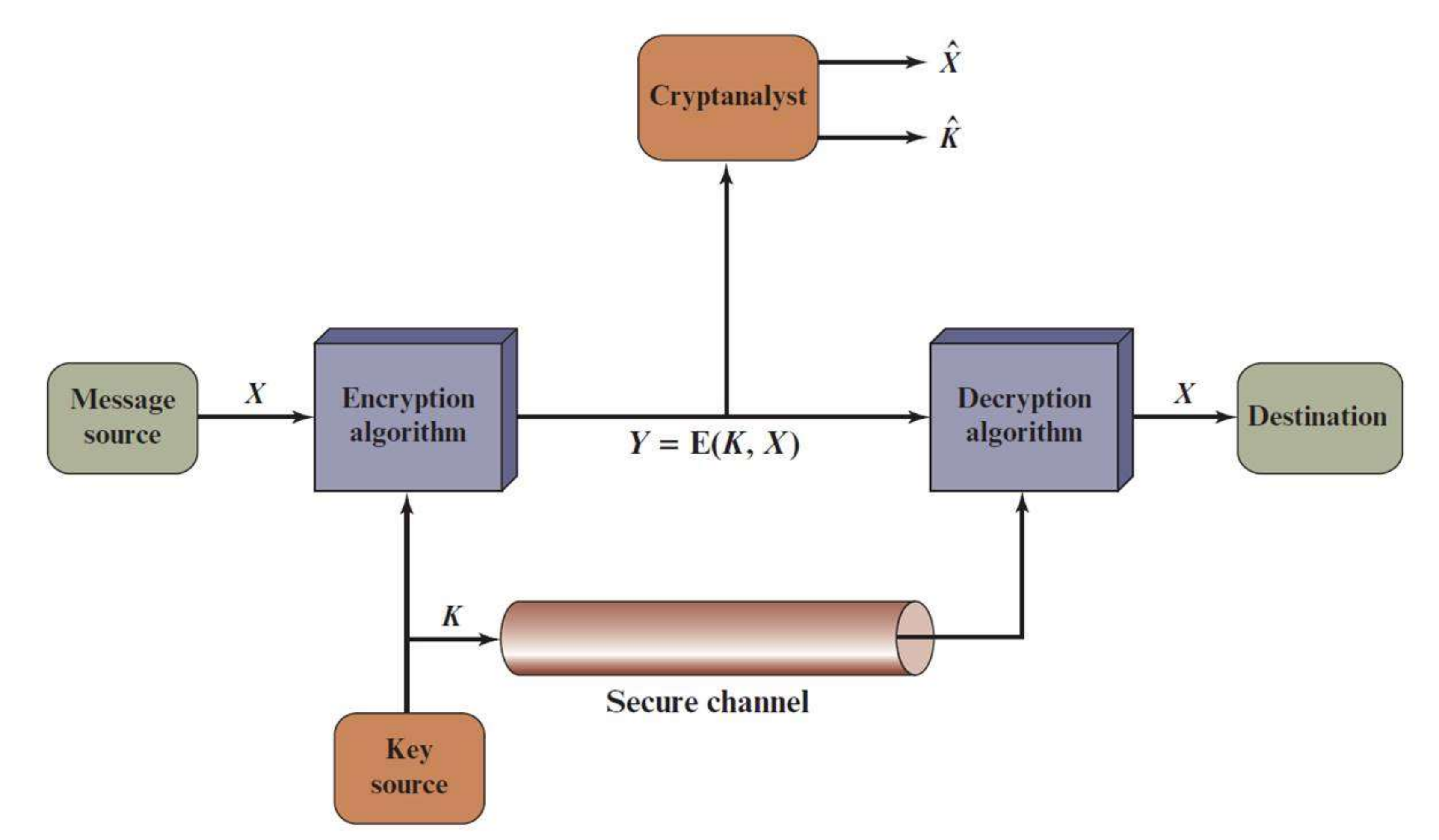
Definitions

- Plaintext: the original message
- Ciphertext: the coded message
- Secret key: a secret used for encryption/decryption
- Encryption/enciphering: converting plaintext to ciphertext using secret key
- Decryption/deciphering: restoring plaintext from ciphertext using secret key
- Cipher/cryptosystem: a system of encryption/decryption
- Cryptography: the study of designing cryptographic schemes
- Cryptanalysis: the study of analyzing ciphertexts without keys
- Cryptology: cryptography and cryptanalysis

Symmetric cryptosystem: usage model



Symmetric cryptosystem: attack model



Symmetric cryptosystem: attack types

- Brute-force attack
  - Use computing power to try every possible key on a ciphertext until an intelligible (meaningful) plaintext is decrypted
  - On average, half of all possible keys must be tried to achieve success
- Cryptanalysis
  - Analyze and exploit the cipher algorithm with some knowledge about plaintext and ciphertext
  - Attempt to deduce plaintext or key from some known information
    - Ciphertext-only attack
    - Known plaintext attack
    - …

Type of attack	Information known to cryptanalyst
Ciphertext Only	• Encryption algorithm • Ciphertext
Known Plaintext	• Encryption algorithm • Ciphertext • One or more plaintext–ciphertext pairs formed with the secret key
Chosen Plaintext	• Encryption algorithm • Ciphertext • Plaintext message chosen by cryptanalyst, together with its corresponding ciphertext generated with the secret key
Chosen Ciphertext	• Encryption algorithm • Ciphertext • Ciphertext chosen by cryptanalyst, together with its corresponding decrypted plaintext generated with the secret key
Chosen Text	• Encryption algorithm • Ciphertext • Plaintext message chosen by cryptanalyst, together with its corresponding ciphertext generated with the secret key • Ciphertext chosen by cryptanalyst, together with its corresponding decrypted plaintext generated with the secret key

Security levels

- Unconditionally/perfectly secure
  - It is impossible to decrypt ciphertext no matter how much time a cryptanalyst spends
  - The information of ciphertext is simply not there without the used secret key
- Computationally secure
  - The time of breaking ciphertext exceeds the useful lifetime of encrypted information
  - Technically, the problem of breaking ciphertext is not poly-time computable

Substitution and transposition

- Substitution: letters of plaintext are replaced by other letters
  - Caesar cipher
  - Monoalphabetic cipher
  - Playfair cipher
  - Hill cipher
  - Vigenere cipher
  - Vernam cipher
  - One-time pad
  - Enigma

- Transposition: letters of plaintext are re-shuffled into different positions
  - Rail fence cipher
  - Row transposition cipher

Caesar cipher

- Simplest and earliest substitution cipher
- $A \Rightarrow 0, B \Rightarrow 1, \ldots, Z \Rightarrow 25$
- Key: $k, 1 \leq k \leq 25$
- Encryption: $C = E(k, P) = (P + k) \bmod 26$
- Decryption: $P = D(k, C) = (C - k) \bmod 26$
- Example,

plaintext: `meet me after the toga party`
ciphertext: `PHHW PH DIWHU WKH WRJD SDUWB`

Caesar cipher: brute-force attack

KEY	PHHW	PH	DIWHU	WKH	WRJD	SDUWB
1	oggv	og	chvgt	vjg	vqic	rctva
2	nffu	nf	bgufs	uif	uphb	qbsuz
3	meet	me	after	the	toga	party
4	ldds	ld	zesdq	sgd	snfz	ozqsx
5	kccr	kc	ydrcp	rfc	rmey	nyprw
6	jbbq	jb	xcqbo	qeb	qldx	mxoqv
7	iaap	ia	wbpan	pda	pkcw	lwnpu
8	hzzo	hz	vaozm	ocz	ojbv	kvmot
9	gyyn	gy	uznyl	nby	niau	julns
10	fxxm	fx	tymxk	max	mhzt	itkmr
11	ewwl	ew	sxlwj	lzw	lgys	hsjlq
12	dvvk	dv	rwkvi	kyv	kfxr	grikp
13	cuuj	cu	qvjuh	jxu	jewq	fqhjo
14	btti	bt	puitg	iwt	idvp	epgin
15	assh	as	othsf	hvs	hcuo	dofhm
16	zrrg	zr	nsgre	gur	gbtn	cnegl
17	yqqf	yq	mrfqd	ftq	fasm	bmdfk
18	xppe	xp	lqepc	esp	ezrl	alcej
19	wood	wo	kpdob	dro	dyqk	zkbdi
20	vnnc	vn	jocna	cqn	cxpj	yjach
21	ummb	um	inbmz	bpm	bwoi	xizbg
22	tlla	tl	hmaly	aol	avnh	whyaf
23	skkz	sk	glzkx	znk	zumg	vgxze
24	rjjy	rj	fkyjw	ymj	ytlf	ufwyd
25	qiix	qi	ejxiv	xli	xske	tevxc

intelligible message

- Caesar cipher can be used on compressed texts
  - Compressed texts have no intelligible words
- Example: a compressed text

```
˜+Wµ"— Ω–O)≤4{∞‡¸ ë~Ω%ràu·¯Í ◊¯Z-
Ú≠2Ò#Åæ∂ œ«q7¸Ωn·®3N◊Ú ŒzʼY–ƒ∞Í[±Û_ èΩ,<NO¬±«ˇxã  Åä£èü3Å
x}ö§kºÂ
_yÍ ^ΔÉ] ˛¤ J/˚iTê&ı 'c<uΩ-
 ÄD(G WÄC~y_ïõÄW PÔı«ÎÜ†ç],¤¡ˇÌ^üÑπ˘≈ˇLˇ9Ogfl0˘&Œ≤ ¬≤ ØÔ§˝:
 ˇŒ!SGqèvo^ ú\¸S>h<-*6ø‡%x´˝|fiÓ#≈˜my‰ˇ≥ñP‹,ƒi Áj  Å◊¿˝Zù-
Ω¨Õ¯6Œÿ{% „ΩÊó ¸ï π÷Áî˚úO2çSÿ´O-
2Äfllßi /@^”∏Kºª PŒπ¸úé^´3∑˘öˇÔZÌ"Y¬ŸΩœY› Ω+eô/˙<K£¿*÷˜"≤û˜
B ZøK˜Qßÿüƒ˛!ÒflÎzsS/]›ÈQ ü
```

Monoalphabetic cipher

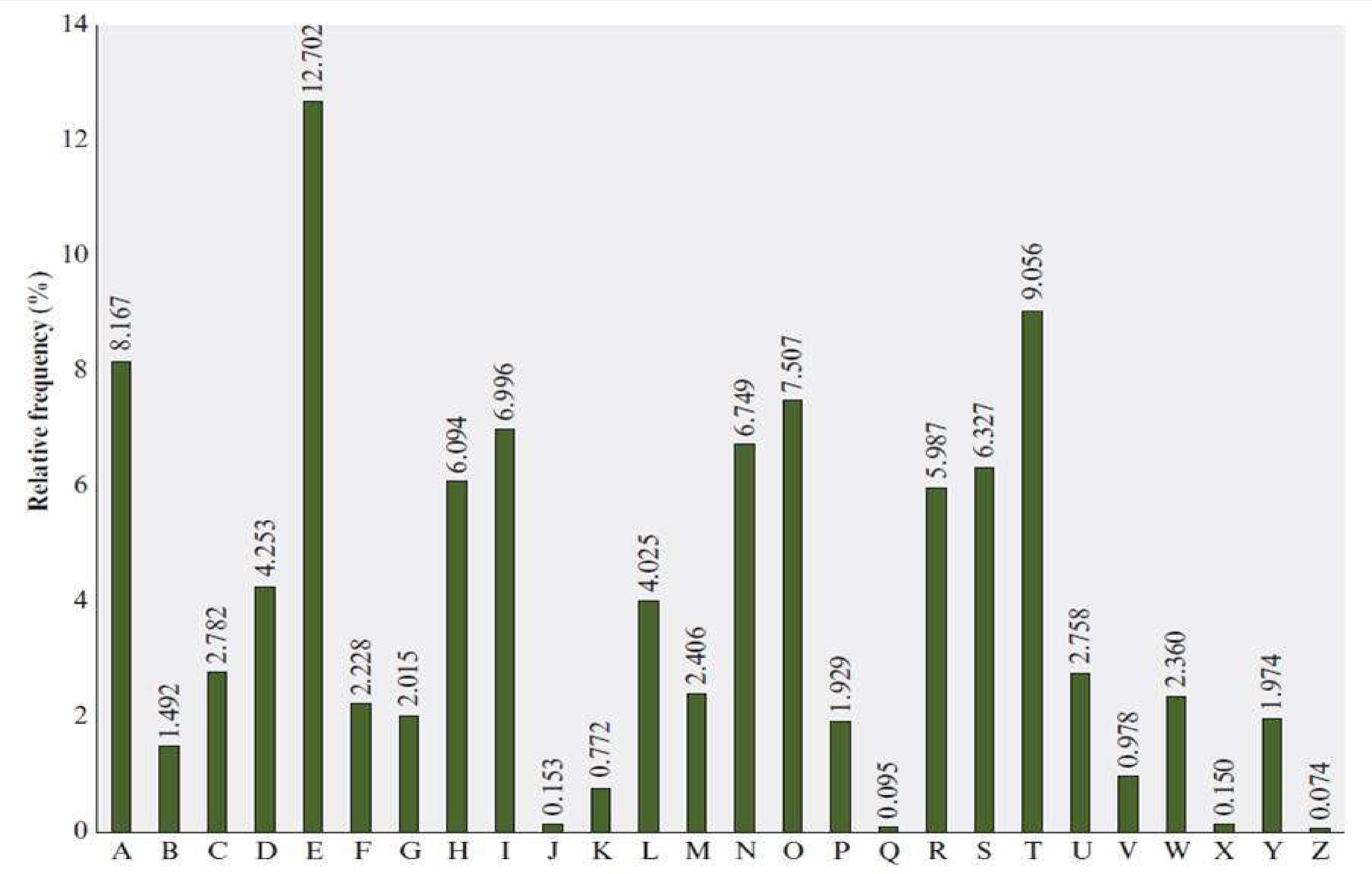
- A substitution cipher
- Key: a permutation $\boldsymbol{p}$ of $0, 1, 2, \dots, 25$
  - E.g., $\boldsymbol{p}(0, 1, 2, 3, \dots, 25) = (23, 9, 29, 18, \dots, 3)$
- Encryption: $C = E(\boldsymbol{p}, P)$=$\boldsymbol{p}$(P), $0 \leq P \leq 25$
- Decryption: $P = D(\boldsymbol{p}, C) = \boldsymbol{p}^{-1}(C)$
- There 26! possible keys: $26! \approx 4 \times 10^{26}$

Monoalphabetic cipher: cryptanalysis

- A ciphertext

KVFNO OGMDD QBKSE BKRSN CKMKG QYQKC SSFBA NTEJB ABYIB QKGQE
TRKAQ BGKYF SMTJJ SFBEG DDBRS OKRRB MDBQK GQOJK MBSFB VNQJA
FKRBU BQIMN VMGRK MGMYQ BAGEJ BRGDF SVFBS FBQNM JKMAN QGMSF
BKGQR TYFDQ KUGSX ABCXG MDOQN ONQSG NMRYN LEGMB AVGSF SFBDB
MTGMB JXBMH NXKEJ BBWOB QGBMY BNCCJ XGMDG MNMBF KUBVN MGSJB
DGNMR NCKAL GQBQR RGMYB GSRCG QRSYN LLBQY GKJUN XKDBG MZVBG
MTJJM TJJRG BEBMR NGSVK RMNRT QOQGR BSFKS SFBQB VKRVG ABROQ
BKAYN MRSBQ MKSGN MKSSF BCBEQ TKQXB GMRUG BQKMM NTMYB LBMSS
FKSKG QETRG MSBMA RSNYB KRBOQ NATYS GNMNC SFBKA QBGKY FSMTJ
JGMZV BGMTJ JBGMR MBTMB CCBYS GUBJX OJKYG MDKMB WOGQX AKSBN
MKMKG QYQKC SSFKS VKRNM YBRBB MKRSF BCTST QBNCK UGKSG NMETS
FNVPT GYIJX KQBKA QBGKY FSMTJ JRDNG MDSNU KMGRF CQNLN TQRIG
BRGRV GABRO QBKAK CCBYS GNMCN QSFBF TDBKG QYQKC SBMNT DFSNI
BBOGS CJXGM DVBJJ GMSNG SRANS KDBGM SFBVK XLKMX YJKRR GYOJK
MBRYN MSGMT BCJXG MDVBJ JEBXN MASFB GQRBQ UGYBJ GCB

- Observation
  - Letter frequencies are un-balanced in normal texts
  - Frequencies do not change in ciphertext



- 1-Letter frequencies in ciphertext

A:20	B:93	C:22	D:20	E:12	F:30
G:70	H:1	I:5	J:34	K:59	L:7
M:68	N:45	O:15	P:1	Q:47	R:42
S:56	T:24	U:9	V:17	W:2	X:15
Y:27	Z:2				

  - Inference: $\{B, G, M, K, S, Q/N\} \rightarrow \{E, T, A, O, I, N/S\}$
- Further observations
  - “th” has highest frequency of two-letter diagram
    - Since sf: 15 is largest, sf $\rightarrow$ th $\Longrightarrow s \rightarrow t, f \rightarrow h$
  - “the”, “that” occur often
    - sfb: 8 $\rightarrow$ the $\Longrightarrow b \rightarrow e$
    - sfks $\rightarrow$ that $\Longrightarrow k \rightarrow a$
  - …

- Decrypted ciphertext

```
A whopping great beast of an aircraft, the double-decker
Airbus A380 -- the biggest passenger airplane the world has
ever known -- is an incredible sight whether on land or in
the air.
   Such gravity-defying proportions combined with the
genuinely enjoyable experience of flying in one have won it
legions of admirers since its first commercial voyage in
2007.
   So it was no surprise that there was widespread
consternation at the February 14 announcement that Airbus
intends to cease production of the A380 in 2019, effectively
placing an expiry date on an aircraft that was once seen as
the future of aviation.
   But how quickly are A380s going to vanish from our skies?
Is widespread affection for the huge aircraft enough to keep
it flying well into its dotage, in the way many classic
planes continue flying well beyond their service life?
```

- **Note: spaces and special characters are omitted in ciphertext**

Playfair cipher

- A two-letter substitution cipher
- Was used as the standard field system by British Army in World War I and U.S. Army and other Allied forces during World War II
- Key: a 5 x 5 matrix of letters
- Encryption: fill in letters of keyword (minus duplicates) from left to right and from top to bottom, then fill in the remainder of the matrix with the remaining letters in alphabetic order

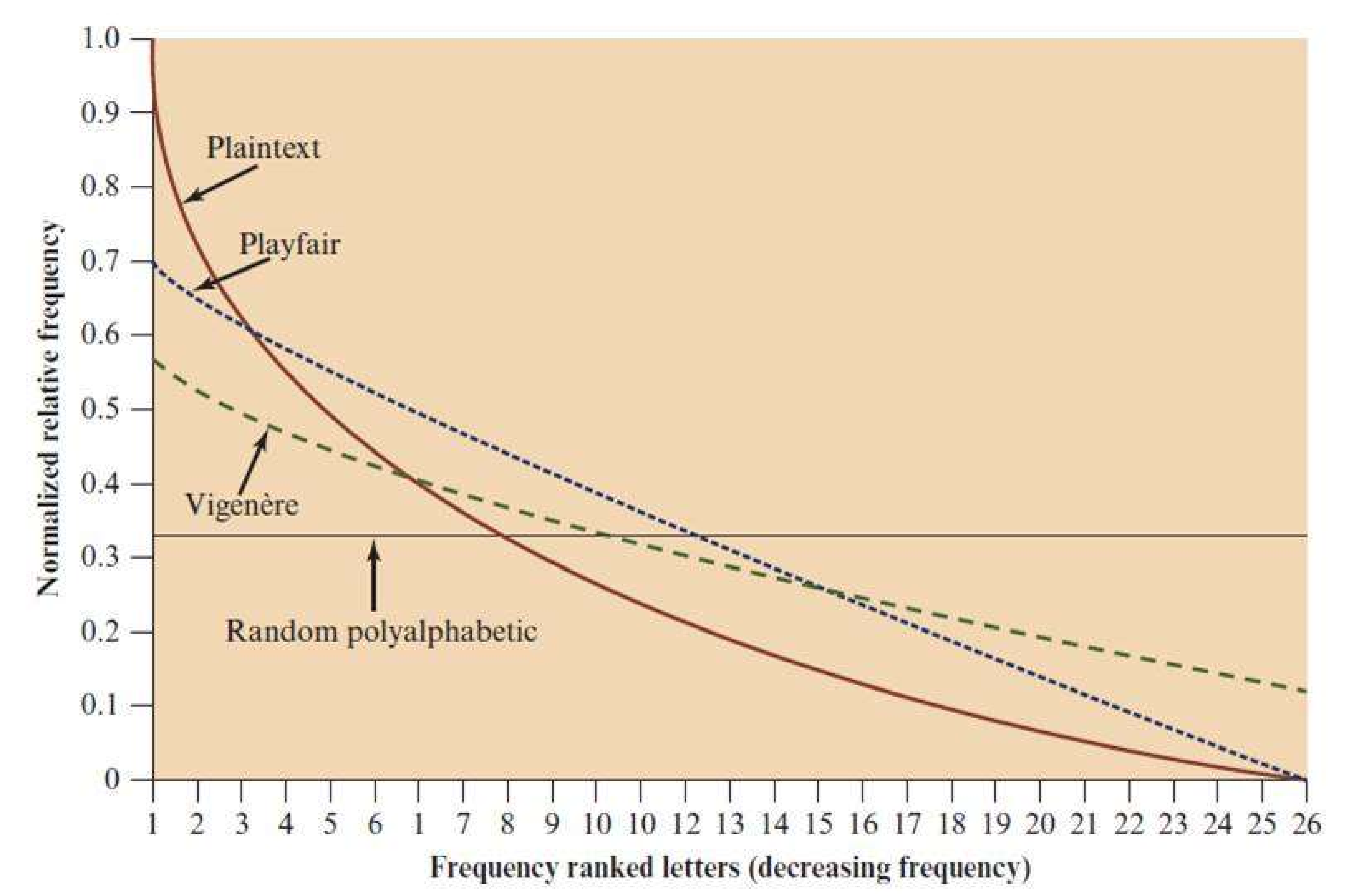
- Example: keyword = Monarchy

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

- mn → OA, ny→YG, eq→GL , pi→SF

```
    at ta ck at fo ur pm
→   RS SR DE RS PH ZM LO
```

- Cryptanalysis: letter frequencies are still un-balanced


1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
Normalized relative frequency
Plaintext
Playfair
Vigenère
Random polyalphabetic
1 2 3 4 5 6 1 7 8 9 10 10 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26
Frequency ranked letters (decreasing frequency)

Hill cipher

- A substitution cipher using linear algebra
- Hide multiple-letter frequencies
- Key: an invertible $n \times n$ matrix in mod 26
- Example : $n = 3$
  - K= $\begin{bmatrix} k_{11} & k_{12} & k_{13} \\ k_{21} & k_{22} & k_{23} \\ k_{31} & k_{32} & k_{33} \end{bmatrix}$
  - $C = PK \bmod 26$, where $P = [p_1 \quad p_2 \quad p_3]$
  - $P = CK^{-1} \bmod 26$
- Strong against ciphertext-only attack
- Broken under known plaintext attack, $n = 3$
  - Given 3 pairs of (P, C), solve the linear equations of $C = PK$

Polyalphabetic ciphers

- Use a set of monoalphabetic substitutions
- A key is used to pick up a substitution for a letter in different positions
- Examples
  - Vigenere cipher
  - Vernam cipher
  - One-time pad

Vigenère cipher

- A polyalphabetic substitution cipher
- The set of monoalphabetic substitutions:
  26 Caesar ciphers
- Substitution X: $a \to X, b \to X + 1, \ldots$
- Key: a repeated keyword, as long as plaintext

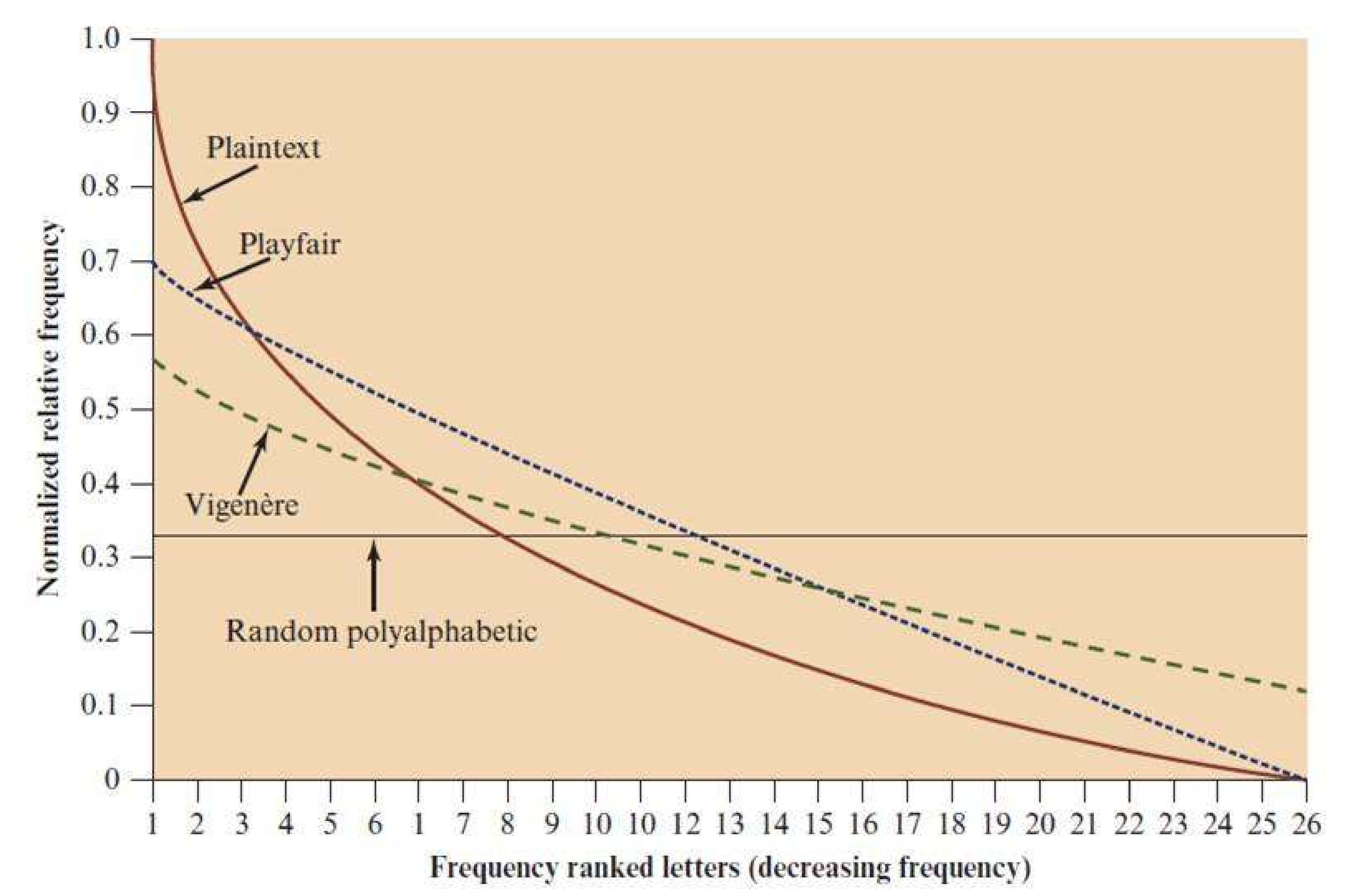
- Example
  - keyword: deceptive
  - Message: we are discovered save yourself
  - Keyword: deceptive
  - key: deceptivedeceptivedeceptive
  - plaintext and ciphertext

```
   wearediscoveredsaveyourself
→  ZICVTWQNGRZGVTWAVZHCQYGLMGJ
```

- Autokey system
  - Keyword is concatenated with plaintext to provide a running key
  - Keyword: deceptive
  - key: deceptivewearediscoveredsav
  - plaintext and ciphertext

```
   wearediscoveredsaveyourself
→  ZICVTWQNGKZEIIGASXSTSLVVWLA
```

- Still vulnerable to cryptanalysis
  - key and plaintext share the same frequency distribution of letters
  - a statistical analysis can be applied



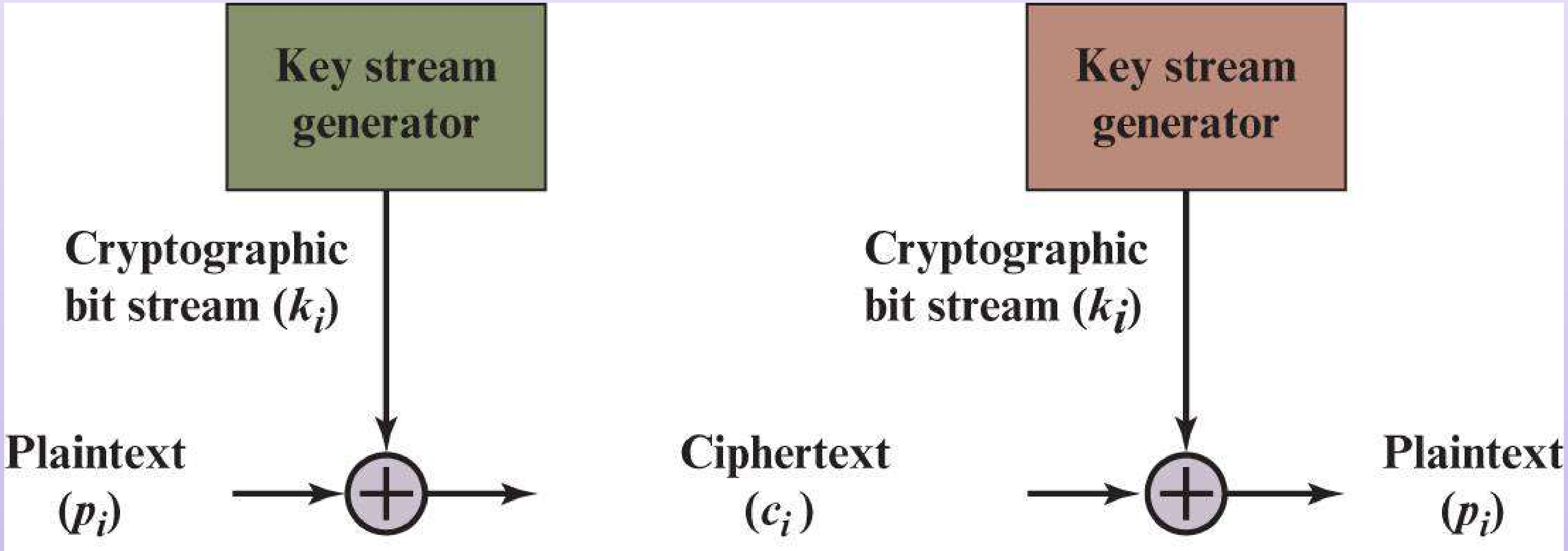
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
Normalized relative frequency
Plaintext
Playfair
Vigenère
Random polyalphabetic
1 2 3 4 5 6 1 7 8 9 10 10 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26
Frequency ranked letters (decreasing frequency)

Defense against letter frequency attack

- Ultimate defense against letter frequency attacks
  - Key length = plaintext length
  - No statistical relation between plaintext and ciphertext
  - Ciphertext is truly random without further information
- Example
  - Vernam cipher
  - One-time pad

Vernam cipher

- Gilbert Vernam, AT&T, 1918



One-time pad

- Improvement to Vernam cipher in two ways
  - Key is truly random and as long as the plaintext
  - Each key is used only once
- Security
  - Unbreakable under ciphertext-only attack
    - perfectly secure, unconditionally secure
    - ciphertext is truly random, no statistical relationship to plaintext
  - Unbreakable under known plaintext attack
    - given a pair (P, C)
    - the used key is $K = P \oplus C$ $(bitwise\ xor)$
    - Since K is not used in any other place, another ciphertext $C' \neq C$ is still unbreakable

- Example

ciphertext	l	p	r	x	a	t	p	q	s	b	f
key stream 1	p	e	c	d	c	r	e	g	w	o	h
plaintext 1	a	t	t	a	c	k	t	w	o	p	m
key stream 2	l	t	c	k	d	y	l	u	r	n	r
plaintext 2	w	i	t	h	d	r	a	w	n	o	w

- There are other intelligible messages for different key streams

One-time pad: analysis

- Let message distribution M be $\Pr[M = m] = p_m,\ m \in \{0,1\}^k$
- Ciphertext C has no statistical relation with plaintext m:

$$\Pr[C = c_1 c_2 \ldots c_k | M = m_1 m_2 \ldots m_k]$$

$$= \Pr[K = (c_1 \oplus m_1)(c_2 \oplus m_2) \cdots (c_k \oplus m_k)]$$

$$= 1/2^k$$

- For any message distribution M, ciphertext is truly random:

$$\Pr[C = c_1 c_2 \ldots c_k] = \sum_{M=m} \Pr[C = c \mid M = m] \Pr[M = m]$$

$$= \sum_{M=m} (1/2^k) p_m = (1/2^k) \sum_{M=m} p_m = 1/2^k$$

One-time pad: difficulties of use

- Hard to produce long truly random keys
- Key distribution problem: sender and receiver are hard to agree on a key, which is used only once
- Useful primarily for low-bandwidth channels requiring very high security
  - E.g., submarine communications

Rail fence cipher

- A transposition cipher
- Plaintext is written down as a sequence of diagonals and then read off as a sequence of rows
- Example: rail fence with depth 2
  - Plaintext: “meet me after the toga party”
  - Encryption:

M	E	M	A	T	R	H	T	G	P	R	Y
E	T	E	F	E	T	E	O	A	A	T	

  - Ciphertext: MEMATRHTGPRYETEFETEOAAT

Row transposition cipher

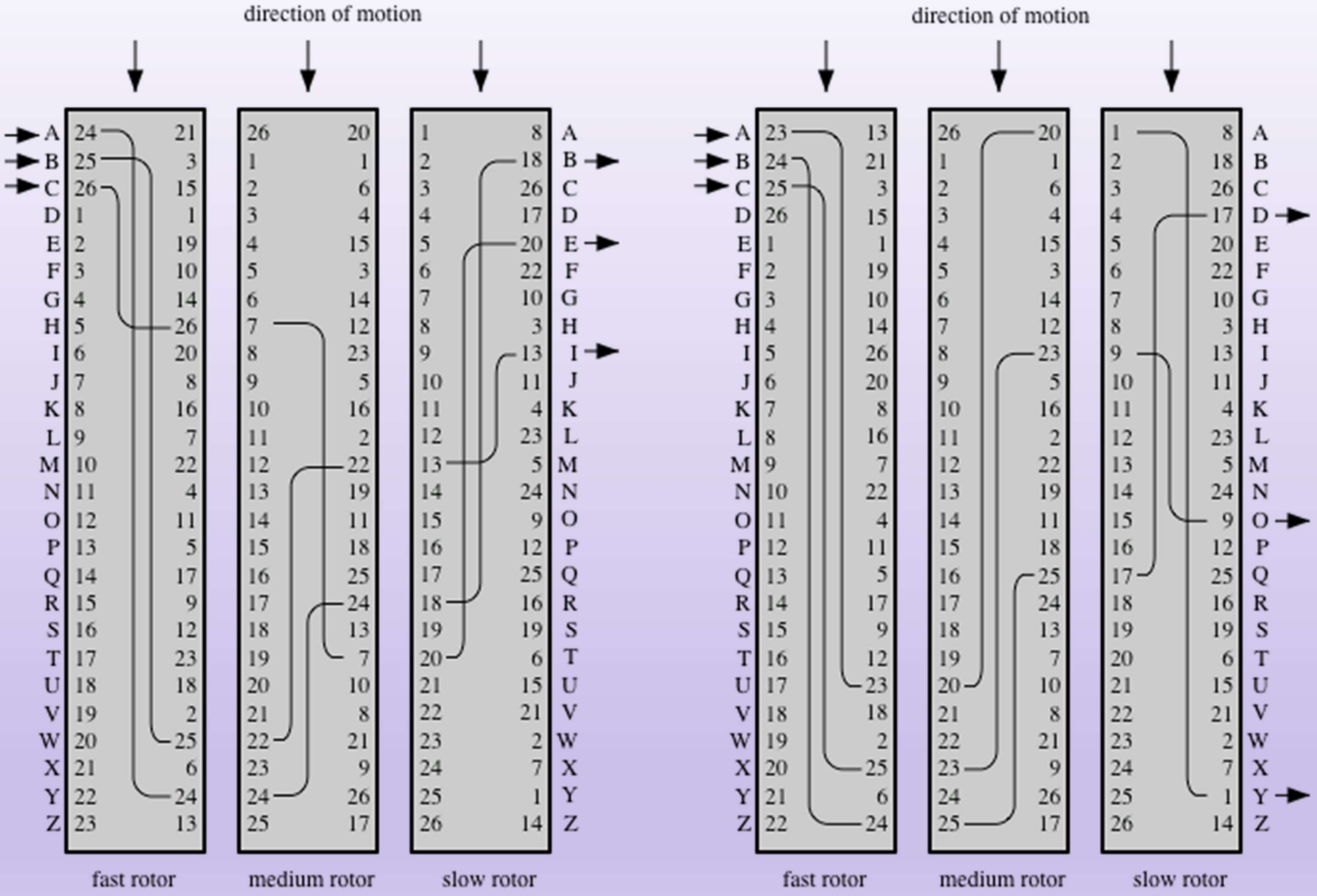
- Write message in a rectangle, row by row,
- Read the message off, column by column
- But, permute the order of the columns with key
- Example

Key	4	3	1	2	5	7	6
Plaintext	a	t	t	a	c	k	p
	o	s	t	p	o	n	e
	d	u	n	t	i	l	t
	w	o	a	m	x	y	z

- Ciphertext: TTNA APTM TSUO AODW COIX PETZ KNLY

Enigma

- German military ciphers in WW II
- A polyalphabetic cipher with intricate design for practical use
- A rotor is a monoalphabetic substitution cipher
- Concatenation of many rotors
  - If simply concatenated, equivalent to a monoalphabetic cipher
  - One stroke of input rotates one position in the first cylinder
  - One complete rotation of the first cylinder
    → one rotate position in the second cylinder, and so on
- If using 3 motors, there are $26 \times 26 \times 26 = 17576$ possible substitutions for an alphabet
  - Type 'a' continuously. The output sequence has a period of 17576
- A letter is substituted according to its position



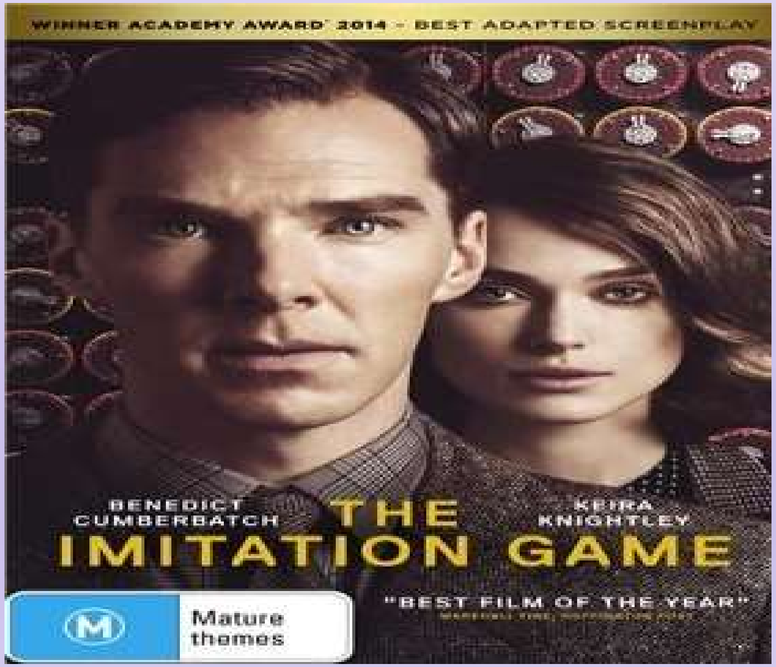
(a) Initial setting

(b) Setting after one keystroke

Enigma: machine photos

Enigma and Allen Turing

- Enigma was broken by a team led by Allen Turing
- This helped the Allies win WW II
- A movie in 2014: The imitation game



Steganography

- What information is carried in this letter?

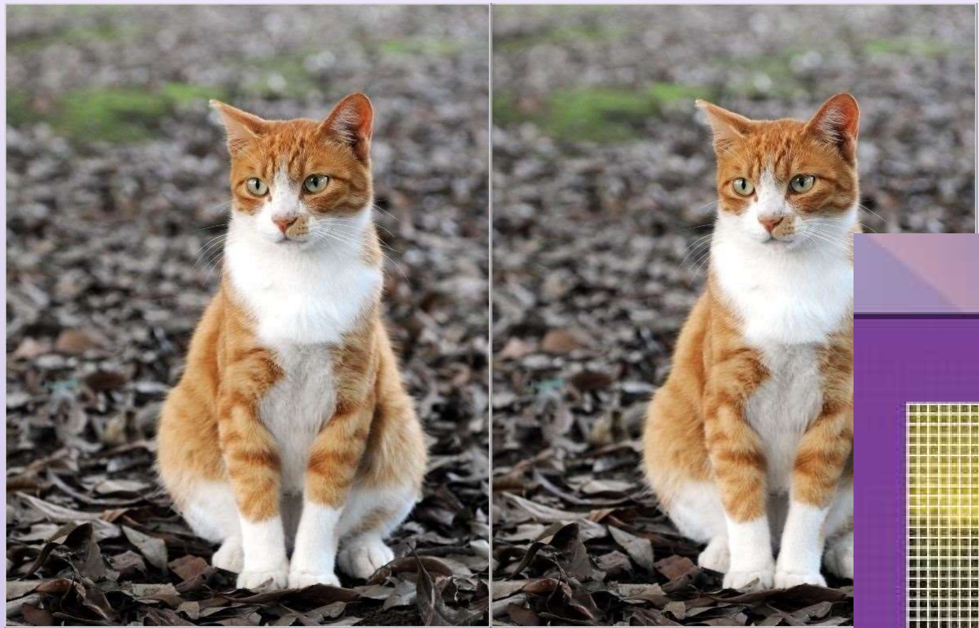
3rd March

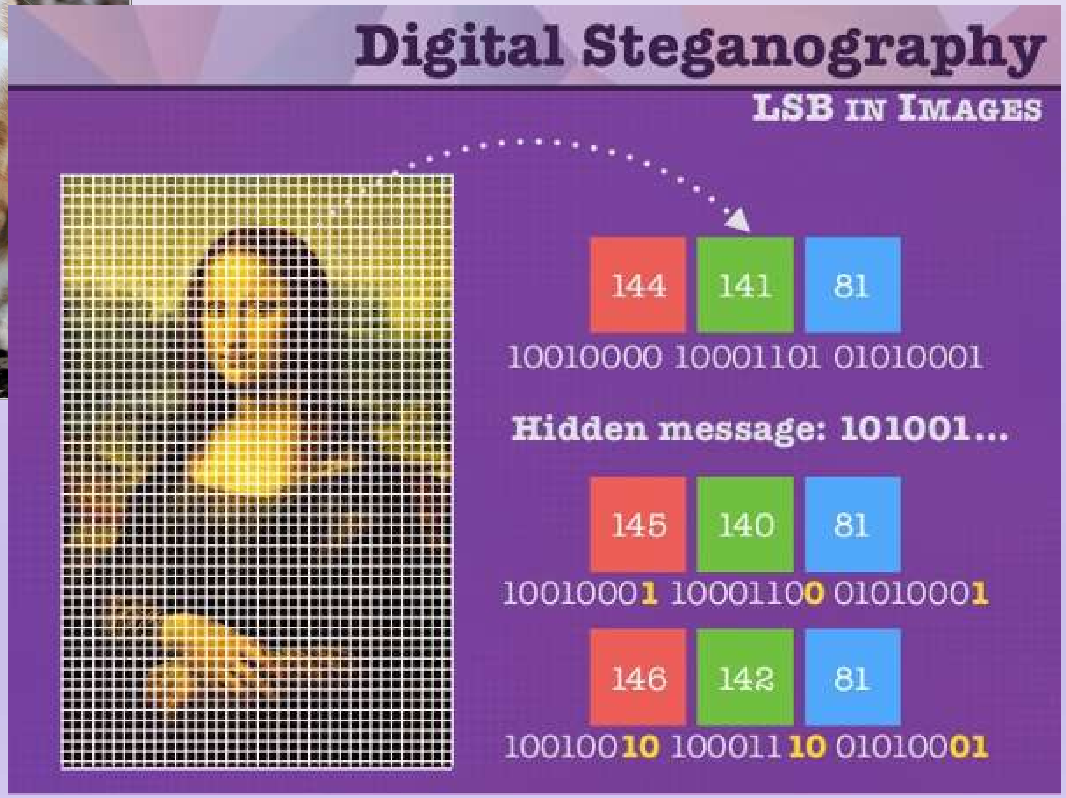
Dear George,

Greetings to all at Oxford. Many thanks for your
letter and for the Summer examination package.
All Entry Forms and Fees Forms should be ready
for final despatch to the Syndicate by Friday
20th or at the very latest, I'm told, by the 21st.
Admin has improved here, though there's room
for improvement still; just give us all two or three
more years and we'll really show you! Please
don't let these wretched 16+ proposals destroy
your basic O and A pattern. Certainly this
sort of change, if implemented immediately,
would bring chaos.

Sincerely yours,

Hide data in images





National Cryptologic Museum, US