CHAPTER 2

CLASSICAL ENCRYPTION TECHNIQUES

SYMMETRIC ENCRYPTION

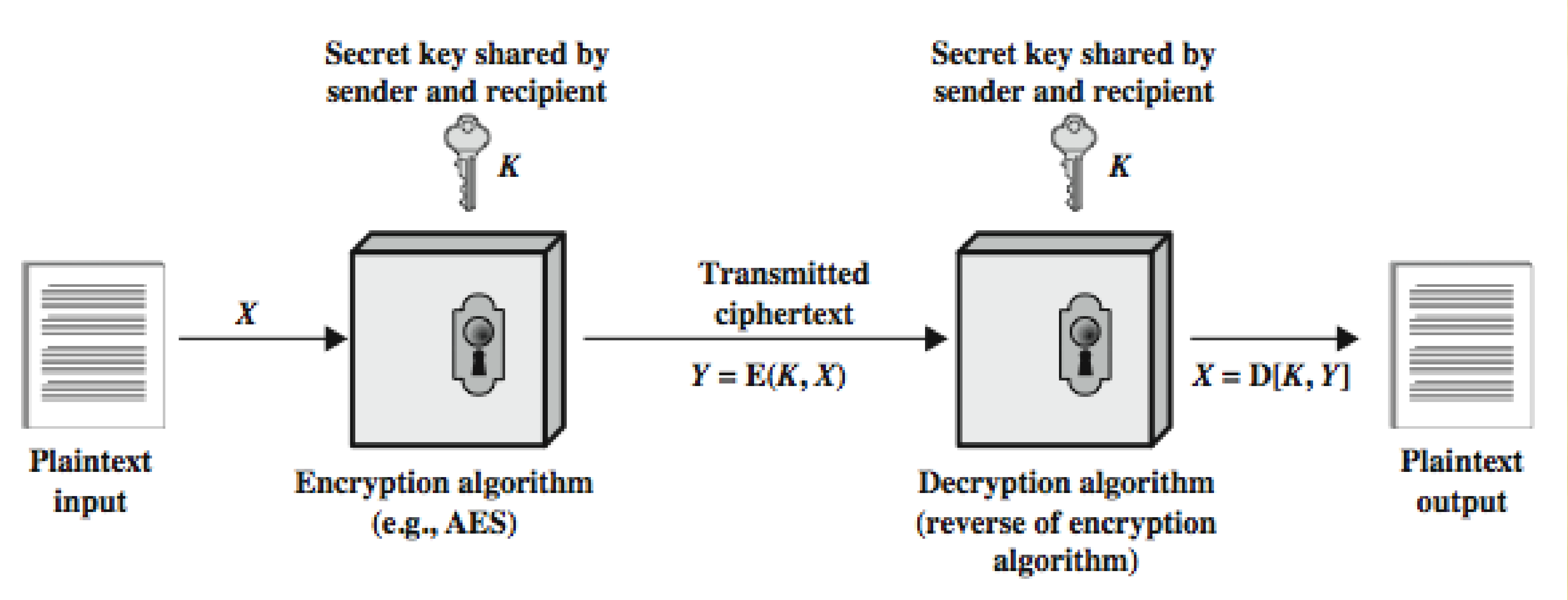
- or conventional / private-key / single-key
- sender and recipient share a common key

  Since both sender and receiver are equivalent, either can encrypt or decrypt messages using that common key.
- all classical encryption algorithms use a private-key

SOME BASIC TERMINOLOGY

- **plaintext** - original message
- **ciphertext** - coded message
- **cipher** - algorithm for transforming plaintext to ciphertext
- **key** - info used in cipher known only to sender/receiver
- **encipher (encrypt)** - converting plaintext to ciphertext
- **decipher (decrypt)** – converting ciphertext to plaintext
- **cryptography** - study of encryption principles / methods
- **cryptanalysis (codebreaking)** - study of deciphering ciphertext principles / methods *without* knowing key
- **cryptology** - field of both cryptography and cryptanalysis

SYMMETRIC CIPHER MODEL



The five components of the symmetric cipher model:

Plaintext - original message
Encryption algorithm – performs substitutions/transformations on plaintext
Secret key – control exact substitutions/transformations used in encryption algorithm
Ciphertext - scrambled message
Decryption algorithm – inverse of encryption algorithm

REQUIREMENTS

- Two requirements for secure use of symmetric encryption:
  - a strong encryption algorithm
  - a secret key known only to sender / receiver
- mathematically have:

plaintext X, ciphertext Y, key K, encryption algorithm E, decryption algorithm D

$$Y = E(K, X)$$

$$X = D(K, Y)$$

- Assume encryption algorithm is known
- Implies a secure channel to distribute key

CRYPTOGRAPHY

- Cryptographic system can be characterized by:

1. The type of operations used for transforming plaintext to ciphertext :

All encryption algorithms are based on two general principles :

- **substitution:** in which each element in the plaintext (bit, letter, group of bits or letters) is mapped into another element, and
- **Transposition:** in which elements in the plaintext are rearranged.

Most systems, referred to as product systems, involve multiple stages of substitutions and transpositions.

2. The number of keys used

- **Symmetric, single-key, secret-key, or conventional encryption:** If both sender and receiver use the same key.
- **Asymmetric, two-key, or public-key encryption:** If the sender and receiver use different keys.

3. The way in which the plaintext is processed

- **A block cipher** processes the input one block of elements at a time, producing an output block for each input block.
- **A stream cipher** processes the input elements continuously, producing output one element at a time, as it goes along.

CRYPTANALYSIS

- To recover key not just message, there are two general approaches:
  - **Cryptanalysis:** relies on the nature of the algorithm plus some knowledge of the general characteristics of the plaintext or even some sample plaintext-ciphertext pairs. This type of attack exploits the characteristics of the algorithm to attempt to deduce a specific plaintext or the key being used.
  - **Brute-force attacks** try every possible key on a piece of ciphertext until an intelligible translation into plaintext is obtained. On average, half of all possible keys must be tried to achieve success.

THE VARIOUS TYPES OF CRYPTANALYTIC ATTACKS BASED ON THE AMOUNT OF INFORMATION KNOWN TO THE CRYPTANALYST

Type of Attack	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

DEFINITIONS

- **unconditional security**
  - If the ciphertext generated by the scheme does not contain enough information to determine uniquely the corresponding plaintext, no matter how much ciphertext is available.
- **computational security**
  - If either the cost of breaking the cipher exceeds the value of the encrypted information, or the time required to break the cipher exceeds the useful lifetime of the information.

- **For all reasonable encryption algorithms, we have to assume computational security where it either takes too long, or is too expensive, to bother breaking the cipher.**

BRUTE FORCE SEARCH

- always possible to simply try every key
- most basic attack, proportional to key size
- assume either know / recognise plaintext

Key Size (bits)	Number of Alternative Keys	Time required at 1 decryption/µs	Time required at 10^6 decryptions/µs
32	$2^{32} = 4.3 \times 10^9$	2^{31} µs = 35.8 minutes	2.15 milliseconds
56	$2^{56} = 7.2 \times 10^{16}$	2^{55} µs = 1142 years	10.01 hours
128	$2^{128} = 3.4 \times 10^{38}$	2^{127} µs = 5.4×10^{24} years	5.4×10^{18} years
168	$2^{168} = 3.7 \times 10^{50}$	2^{167} µs = 5.9×10^{36} years	5.9×10^{30} years
26 characters (permutation)	$26! = 4 \times 10^{26}$	2×10^{26} µs = 6.4×10^{12} years	6.4×10^6 years

A brute-force attack involves trying every possible key until an intelligible translation of the ciphertext into plaintext is obtained. On average, half of all possible keys must be tried to achieve success.

CLASSICAL SUBSTITUTION CIPHERS

- Where letters of plaintext are replaced by other letters or by numbers or symbols, or

- If plaintext is viewed as a sequence of bits, then substitution involves replacing plaintext bit patterns with ciphertext bit patterns

CAESAR CIPHER

- Earliest known substitution cipher. The core idea is to replace one basic unit (letter/byte) with another.
- By Julius Caesar
- First attested use in military affairs
- Replaces each letter by 3rd letter on
- example:

```
plain:  meet me after the toga party
cipher: PHHW PH DIWHU WKH WRJD SDUWB
```

CAESAR CIPHER

- can define transformation as:

plain:	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:	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

- mathematically give each letter a number

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

n	o	p	q	r	s	t	u	v	w	x	y	z
13	14	15	16	17	18	19	20	21	22	23	24	25

- then have Caesar cipher as:

$c = \mathrm{E}(k, p) = (p + k) \bmod (26)$

$p = \mathrm{D}(k, c) = (c - k) \bmod (26)$

CRYPTANALYSIS OF CAESAR CIPHER

- only have 26 possible ciphers
  - A maps to A,B,..Z
- could simply try each in turn
- a **brute force search**
  - given ciphertext, just try all shifts of letters to recognize when have plaintext
  - ex., break ciphertext "GCUA VQ DTGCM“ gives "easy to break", with a shift of 2 (key C).

BREAK CIPHERTEXT "PHHW PH DIWHU WKH KRJD SDUWB"

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

MONOALPHABETIC CIPHER

- Rather than just shifting the alphabet, could shuffle (jumble) the letters arbitrarily
- Each plaintext letter maps to a different random ciphertext letter, hence key is 26 letters long

Plain : abcdefghijklmnopqrstuvwxyz
Cipher : DKVQFIBJWPESCXHTMYAUOLRGZN

Plaintext : ifwewishtoreplaceletters
Ciphertext: WIRFRWAJUHYFTSDVFSFUUFYA

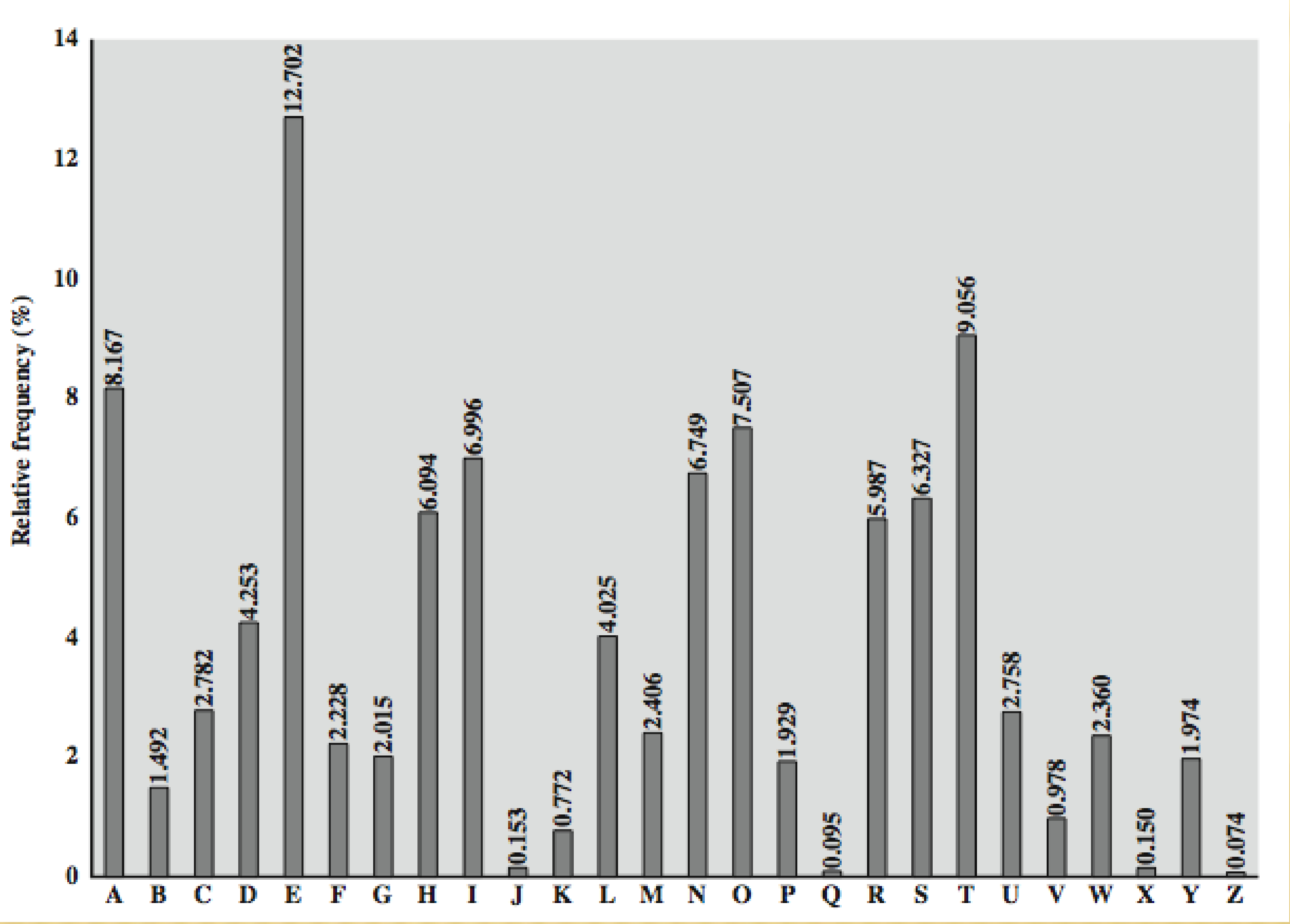
MONOALPHABETIC CIPHER SECURITY

- Have a total of 26! = 4×10^{26} keys
- with so many keys, might think is secure, but would be **!!!WRONG!!!**
- problem is language characteristics

LANGUAGE REDUNDANCY AND CRYPTANALYSIS

- human languages are **redundant**
- eg "th lrd s m shphrd shll nt wnt"
- letters are not equally commonly used, in English E is by far the most common letter
  - followed by T,R,N,I,O,A,S
- other letters like Z,J,K,Q,X are fairly rare
- have tables of single, double & triple letter frequencies for various languages

ENGLISH LETTER FREQUENCIES



USE IN CRYPTANALYSIS

- key concept - Monoalphabetic substitution ciphers do not change relative letter frequencies
- discovered by Arabian scientists in 9^{th} century
- calculate letter frequencies for Ciphertext
- compare counts/plots against known values
- if Caesar cipher looks for common peaks/troughs
  - peaks at: A-E-I triple, NO pair, RST triple
  - troughs at: JK, X-Z
- for Monoalphabetic must identify each letter
  - tables of common double/triple letters help

EXAMPLE CRYPTANALYSIS

- given ciphertext:

  UZQSOVUOHXMOPVGPOZPEVSGZWSZOPFPESXUDBMETSXAIZ
  VUEPHZHMDZSHZOWSFPAPPDTSVPQUZWYMXUZUHSX
  EPYEPOPDZSZUFPOMBZWPFUPZHMDJUDTMOHMQ

- count relative letter frequencies (see text)
- guess P & Z are e and t
- guess ZW is th and hence ZWP is the
- proceeding with trial and error finally get:

  it was disclosed yesterday that several informal but
  direct contacts have been made with political
  representatives of the viet cong in moscow

PLAYFAIR CIPHER

- Not even the large number of keys in a monoalphabetic cipher provides security
- one approach to improving security was to encrypt multiple letters
- The **Playfair Cipher** is an example
- invented by Charles Wheatstone in 1854, but named after his friend Baron Playfair

PLAYFAIR KEY MATRIX

- a 5X5 matrix of letters based on a keyword
- fill in letters of keyword (sans duplicates)
- fill rest of matrix with other letters
- eg. using the 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

ENCRYPTING AND DECRYPTING

- plaintext is encrypted two letters at a time
  1. if a pair is a repeated letter, insert filler like 'X’
  2. if both letters fall in the same row, replace each with letter to right (wrapping back to start from end)
  3. if both letters fall in the same column, replace each with the letter below it (wrapping to top from bottom)
  4. otherwise each letter is replaced by the letter in the same row and in the column of the other letter of the pair

SECURITY OF PLAYFAIR CIPHER

- Security much improved over monoalphabetic
- Since have 26 x 26 = 676 digrams
- Would need a 676 entry frequency table to analyse (verses 26 for a monoalphabetic) and correspondingly more ciphertext
- was widely used for many years
  - eg. by US & British military in WW1
- it **can** be broken, given a few hundred letters
- since still has much of plaintext structure

POLYALPHABETIC CIPHERS

- Polyalphabetic substitution ciphers
- Improve security using multiple cipher alphabets
- Make cryptanalysis harder with more alphabets to guess and flatter frequency distribution
- Use a key to select which alphabet is used for each letter of the message
- Use each alphabet in turn
- Repeat from start after end of key is reached

VIGENÈRE CIPHER

- simplest polyalphabetic substitution cipher
- effectively multiple caesar ciphers
- key is multiple letters long $K = k_1\ k_2\ \ldots\ k_d$
- i^{th} letter specifies i^{th} alphabet to use
- use each alphabet in turn
- repeat from start after d letters in message
- decryption simply works in reverse

EXAMPLE OF VIGENÈRE CIPHER

- write the plaintext out
- write the keyword repeated above it
- use each key letter as a caesar cipher key
- encrypt the corresponding plaintext letter
- eg using keyword *deceptive*

key: deceptivedeceptivedeceptive
plaintext: wearediscoveredsaveyourself
ciphertext:ZICVTWQNGRZGVTWAVZHCQYGLMGJ

SECURITY OF VIGENÈRE CIPHERS

- Have multiple ciphertext letters for each plaintext letter
- Hence letter frequencies are obscured, but not totally lost
- Start with letter frequencies
  - see if look monoalphabetic or not
- if not, then need to determine number of alphabets, since then can attach each

AUTOKEY CIPHER

- Ideally want a key as long as the message
- Vigenère proposed the **autokey** cipher
- With keyword is prefixed to message as key
- knowing keyword can recover the first few letters
- use these in turn on the rest of the message
- but still have frequency characteristics to attack
- eg. given key *deceptive*

key: deceptivewearediscoveredsav
plaintext: wearediscoveredsaveyourself
ciphertext: ZICVTWQNGKZEIIGASXSTSLVVWLA

ONE-TIME PAD

- If a truly random key as long as the message is used, the cipher will be secure
- Unbreakable since ciphertext bears no statistical relationship to the plaintext
- Since for **any plaintext** & **any ciphertext** there exists a key mapping one to other
- Can only use the key **once** though
- Problems in generation & safe distribution of key

TRANSPOSITION CIPHERS

- Consider classical **transposition** or **permutation** ciphers
- These hide the message by rearranging the letter order without altering the actual letters used
- Can recognise these since have the same frequency distribution as the original text

RAIL FENCE CIPHER

- Write message letters out diagonally over a number of rows , then read off cipher row by row
- e.g. Write message out as:

```
m e m a t r h t g p r y
 e t e f e t e o a a t
```

- Giving ciphertext

```
MEMATRHTGPRYETEFETEOAAT
```

ROW TRANSPOSITION CIPHERS

write letters of message out in rows over a specified number of columns

- then reorder the columns according to some key before reading off the rows

```
            1 2 3 4 5 6 7
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

Key:        3 4 2 1 5 6 7
Plaintext:  T A T A C K P
            T P S O O N E
            N T U D I L T
            A M O W X Y Z
Ciphertext: TTNAAPTMTSUOAODWCOIXKNLYPETZ
```

PRODUCT CIPHERS

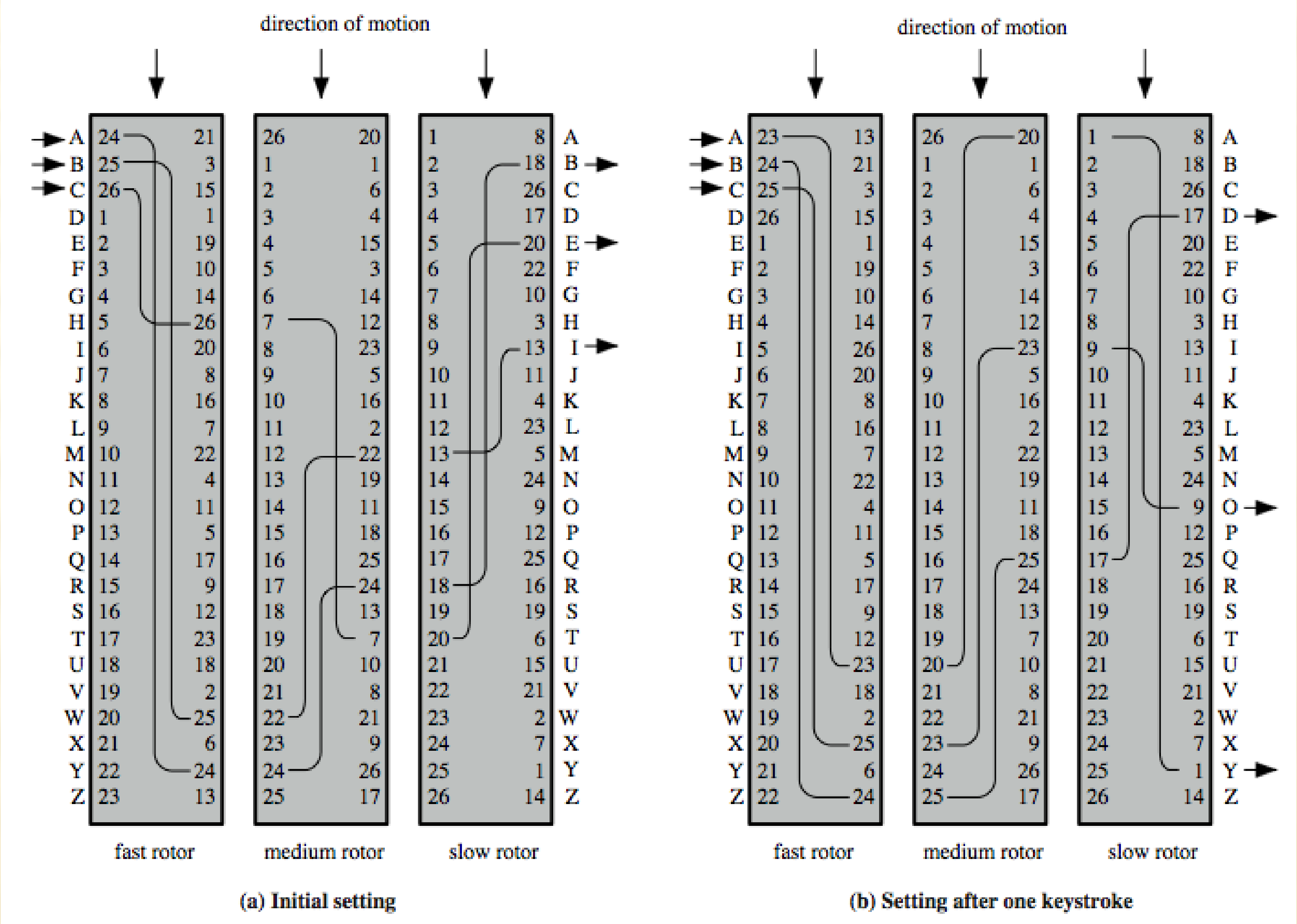
- Ciphers using substitutions or transpositions are not secure because of language characteristics
- Hence consider using several ciphers in succession to make harder, but:
  - two substitutions make a more complex substitution
  - two transpositions make more complex transposition
  - but a substitution followed by a transposition makes a new much harder cipher
- this is bridge from classical to modern ciphers

ROTOR MACHINES

- Before modern ciphers, rotor machines were most common complex ciphers in use
- Widely used in WW2
  - German Enigma, Allied Hagelin, Japanese Purple
- Implement a very complex, varying substitution cipher
- Use a series of cylinders, each giving one substitution, which rotated and changed after each letter was encrypted
- with 3 cylinders have 26^3=17576 alphabets

HAGELIN ROTOR MACHINE

ROTOR MACHINE PRINCIPLES



SUMMARY

- Classical cipher techniques and terminology
- Monoalphabetic substitution ciphers
- Cryptanalysis using letter frequencies
- Playfair cipher
- Polyalphabetic ciphers
- Transposition ciphers
- Product ciphers and rotor machines