Cryptography and Network Security Chapter 2

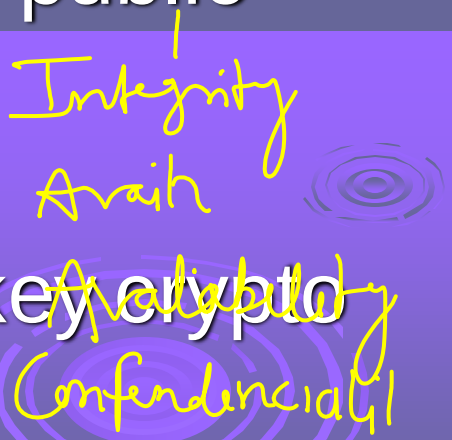
Fifth Edition

by William Stallings

Lecture slides by Lawrie Brown

Symmetric Encryption

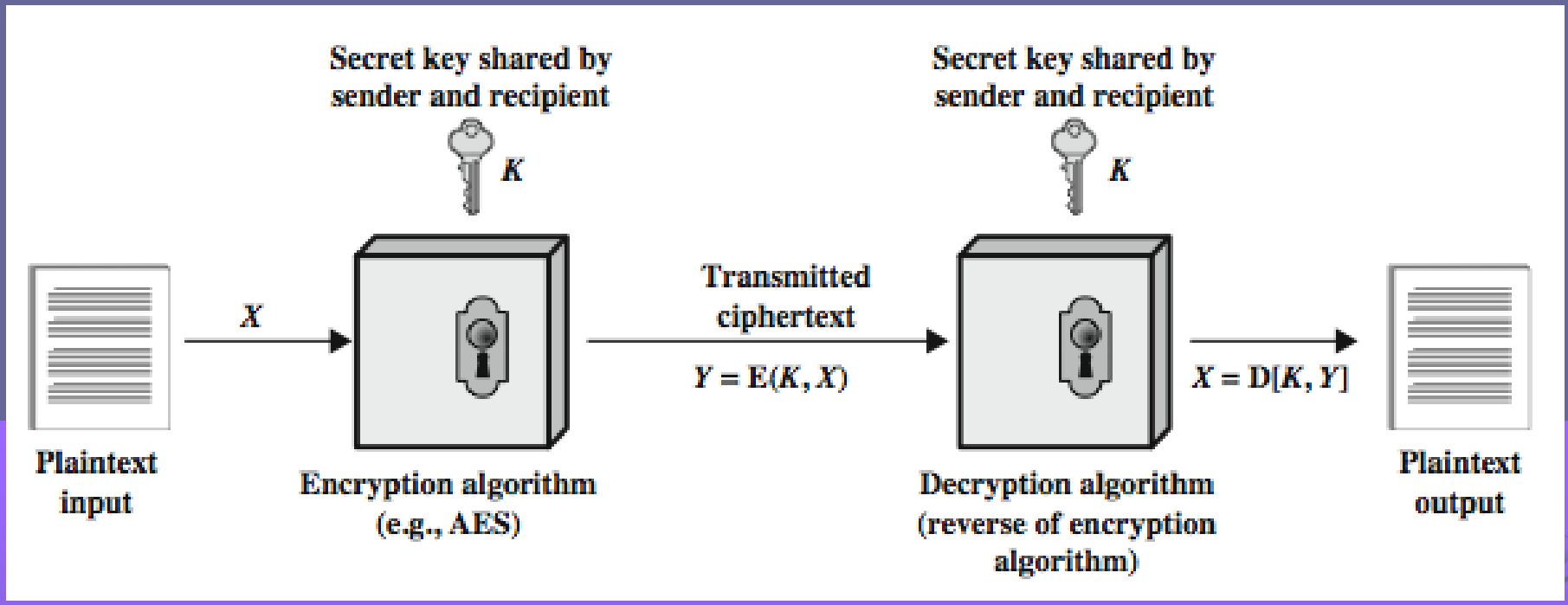
- or conventional / private-key / single-key
- sender and recipient share a common key
- all classical encryption algorithms are private-key
- was only type prior to invention of public-key in 1970's
- and by far most widely used (still)
- is significantly faster than public-key crypto



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)** - recovering plaintext from ciphertext
- **cryptography** - study of encryption principles/methods
- **cryptanalysis (codebreaking)** - study of principles/ methods of deciphering ciphertext *without* knowing key
- **cryptology** - field of both cryptography and cryptanalysis

Symmetric Cipher Model



Requirements

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

  $Y = E(K, X) = E_K(X) = \{X\}_K$

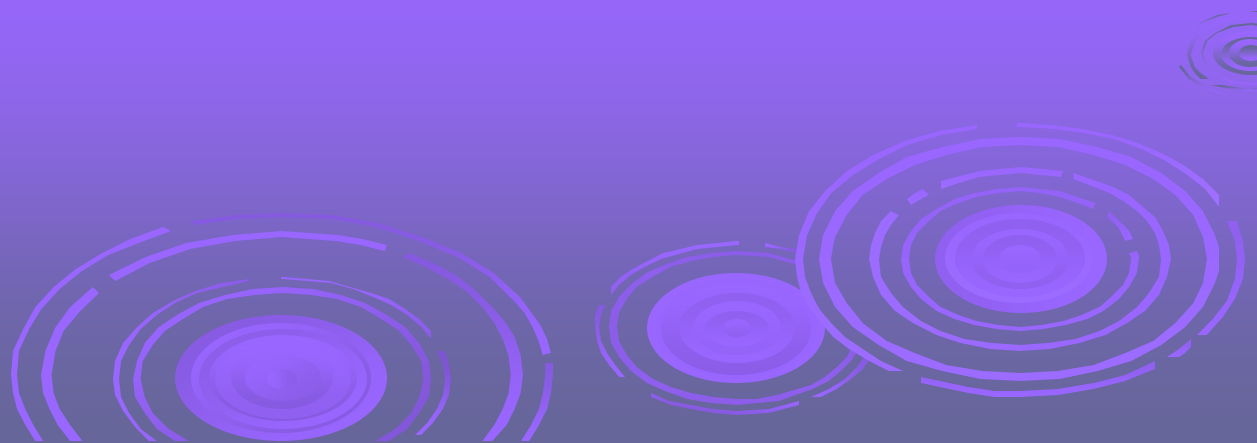
  $X = D(K, Y) = D_K(Y)$
- assume encryption algorithm is known
  - Kerckhoff's Principle: security in secrecy of key alone, not in obscurity of the encryption algorithm
- implies a secure channel to distribute key
  - Central problem in symmetric cryptography

Cryptography

- can characterize cryptographic system by:
  - type of encryption operations used
    - substitution
    - transposition
    - product
  - number of keys used
    - single-key or private
    - two-key or public
  - way in which plaintext is processed
    - block
    - stream

Cryptanalysis

- objective to recover key not just message
- general approaches:
  - cryptanalytic attack
  - brute-force attack
- if either succeed all key use compromised



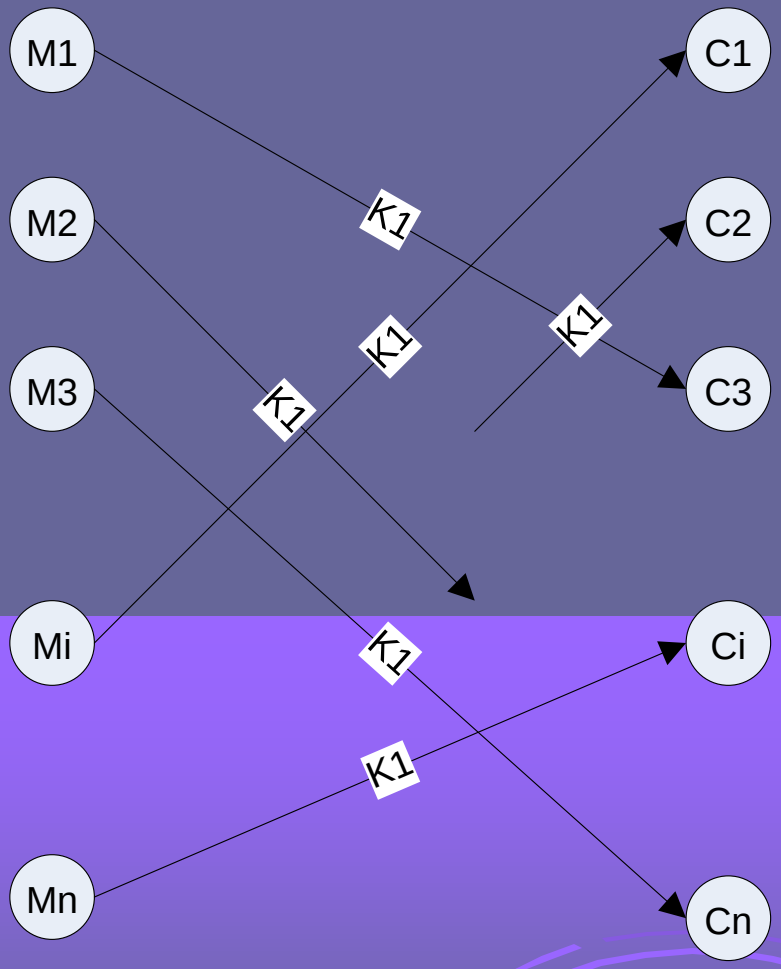
Cryptanalytic Attacks

- **ciphertext only**
  - only know algorithm & ciphertext, is statistical, can identify plaintext
- **known plaintext**
  - know/suspect plaintext & ciphertext
- **chosen plaintext**
  - select plaintext and obtain ciphertext
- **chosen ciphertext**
  - select ciphertext and obtain plaintext
- **chosen text**
  - select plaintext or ciphertext to en/decrypt

Cipher Strength

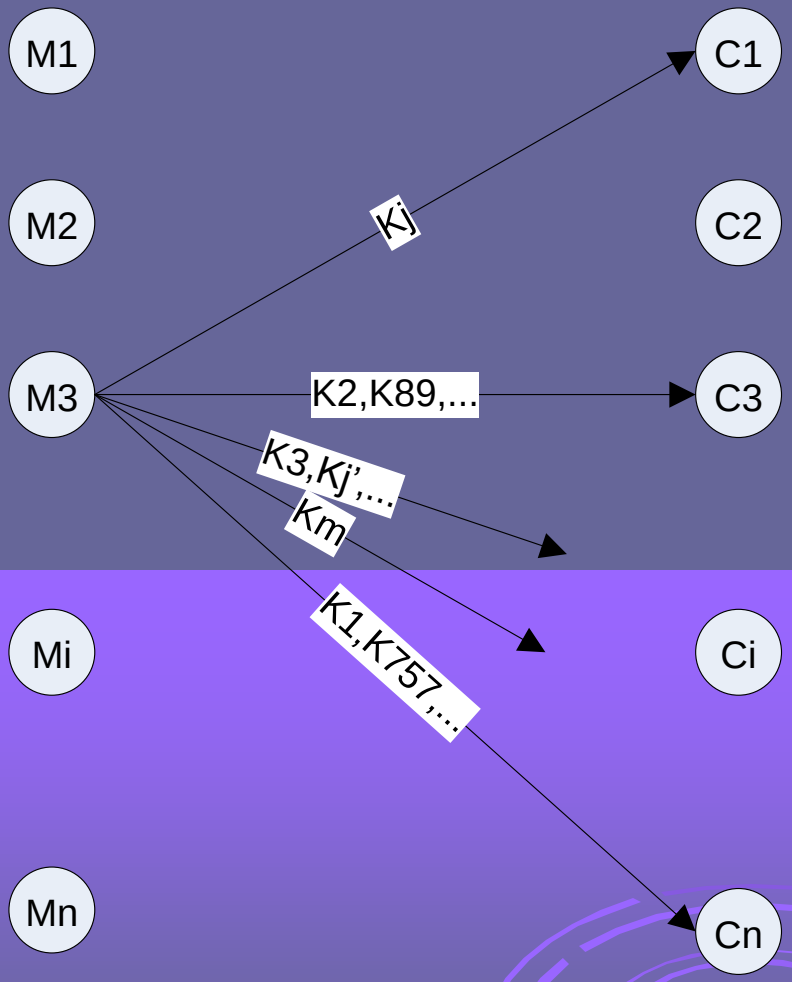
- **unconditional security**
  - no matter how much computer power or time is available, the cipher cannot be broken since the ciphertext provides insufficient information to uniquely determine the corresponding plaintext
- **computational security**
  - given limited computing resources (e.g. time needed for calculations is greater than age of universe), the cipher cannot be broken

Encryption Mappings



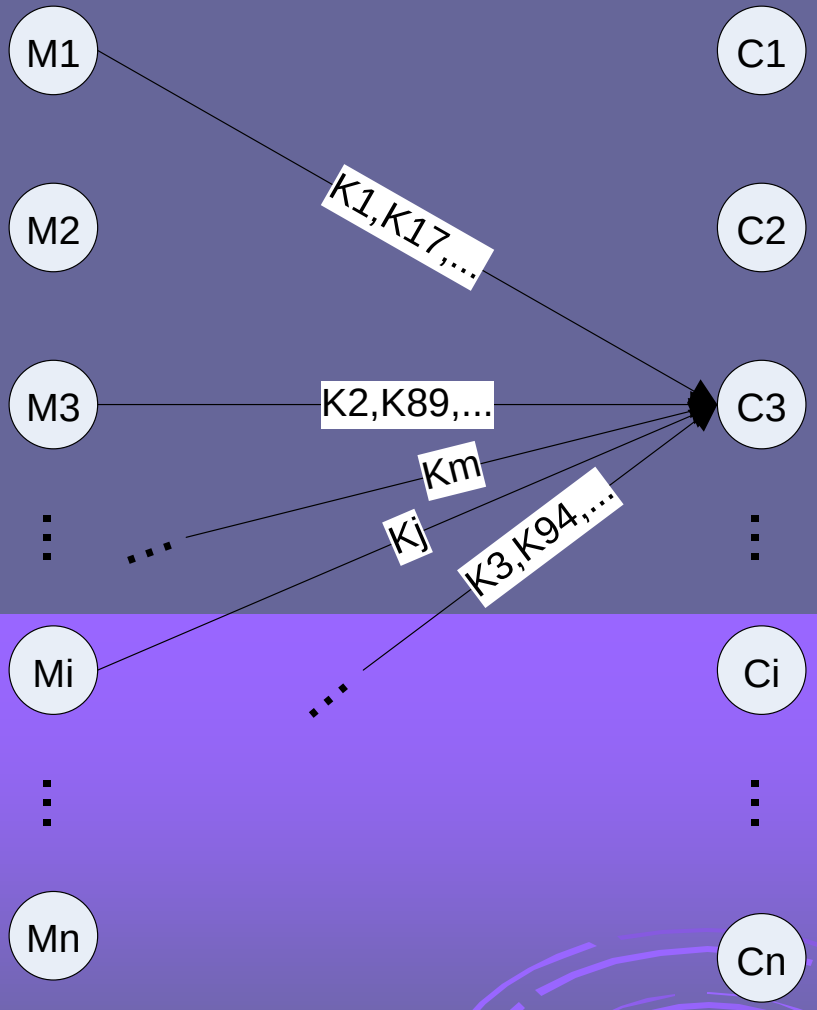
- **A given key (k)**
  - Maps any message Mi to some ciphertext E(k,Mi)
  - Ciphertext image of Mi is unique to Mi under k
  - Plaintext pre-image of Ci is unique to Ci under k
- **Notation**
  - ∀ key k and ∀ Mi in M, ∃! Cj in C such that E(k,Mi) = Cj
  - ∀ key k and ∀ ciphertext Ci in C, ∃! Mj in M such that E(k,Mj) = Ci
  - $E_k(.)$ is “one-to-one” (injective)
  - If |M|=|C| it is also “onto” (surjective), and hence bijective.

Encryption Mappings (2)



- **A given plaintext (Mi)**
  - Mi is mapped to *some* ciphertext E(K,Mi) by every key k
  - Different keys may map Mi to the same ciphertext
  - There may be some ciphertexts to which Mi is never mapped by any key
- **Notation**
  - $\forall$ key k and $\forall$ Mi in M, $\exists!$ ciphertext Cj in C such that E(k,Mi) = Cj
  - It is possible that there are keys k and k' such that E(k,Mi) = E(k',Mi)
  - There may be some ciphertext Cj for which $\nexists$ key k such that E(k,Mi) = Cj

Encryption Mappings (3)



- **A ciphertext (Ci)**
  - Has a unique plaintext pre-image under each k
  - May have two keys that map the same plaintext to it
  - There may be some plaintext Mj such that no key maps Mj to Ci
- **Notation**
  - $\forall$ key k and $\forall$ ciphertext Ci in C, $\exists!$ Mj in M such that E(k,Mj) = Ci
  - There may exist keys k, k' and plaintext Mj such that E(k,Mj) = E(k',Mj) = Ci
  - There may exist plaintext Mj such that $\nexists$ key k such that E(k,Mj) = Ci

Encryption Mappings (4)

- Under what conditions will there always be some key that maps some plaintext to a given ciphertext?
- If for an intercepted ciphertext c_j, there is some plaintext m_i for which there does not exist any key k that maps m_i to c_j, then the attacker has learned something
- If the attacker has ciphertext c_j and known plaintext m_i, then many keys may be eliminated

Brute Force Search

- always possible to simply try every key
- most basic attack, exponential in key length
- 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

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
- by Julius Caesar
- first attested use in military affairs
- replaces each letter by 3rd letter on
- example:

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

Caesar Cipher

- can define transformation as:

```
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 =
  IN
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 =
  OUT
```

- mathematically give each letter a number

```
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
```

- then have Caesar (rotation) cipher as:

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

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

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
- do need to recognize when have plaintext
- eg. break ciphertext "GCUA VQ DTGCM"

Affine Cipher

- broaden to include multiplication
- can define affine transformation as:

  $c = E(k, p) = (ap + b) \bmod (26)$

  $p = D(k, c) = (a^{-1}(c - b)) \bmod (26)$
- key k=(a,b)
- a must be relatively prime to 26
  - so there exists unique inverse a^{-1}

Affine Cipher - Example

- example k=(17,3):

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

- example:

```
meet me after the toga party
ZTTO ZT DKOTG OST OHBD YDGOV
```

- Now how many keys are there?
  - 12 x 26 = 312
- Still can be brute force attacked!

Monoalphabetic Cipher

- rather than just shifting the alphabet
- could shuffle (permute) 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
```

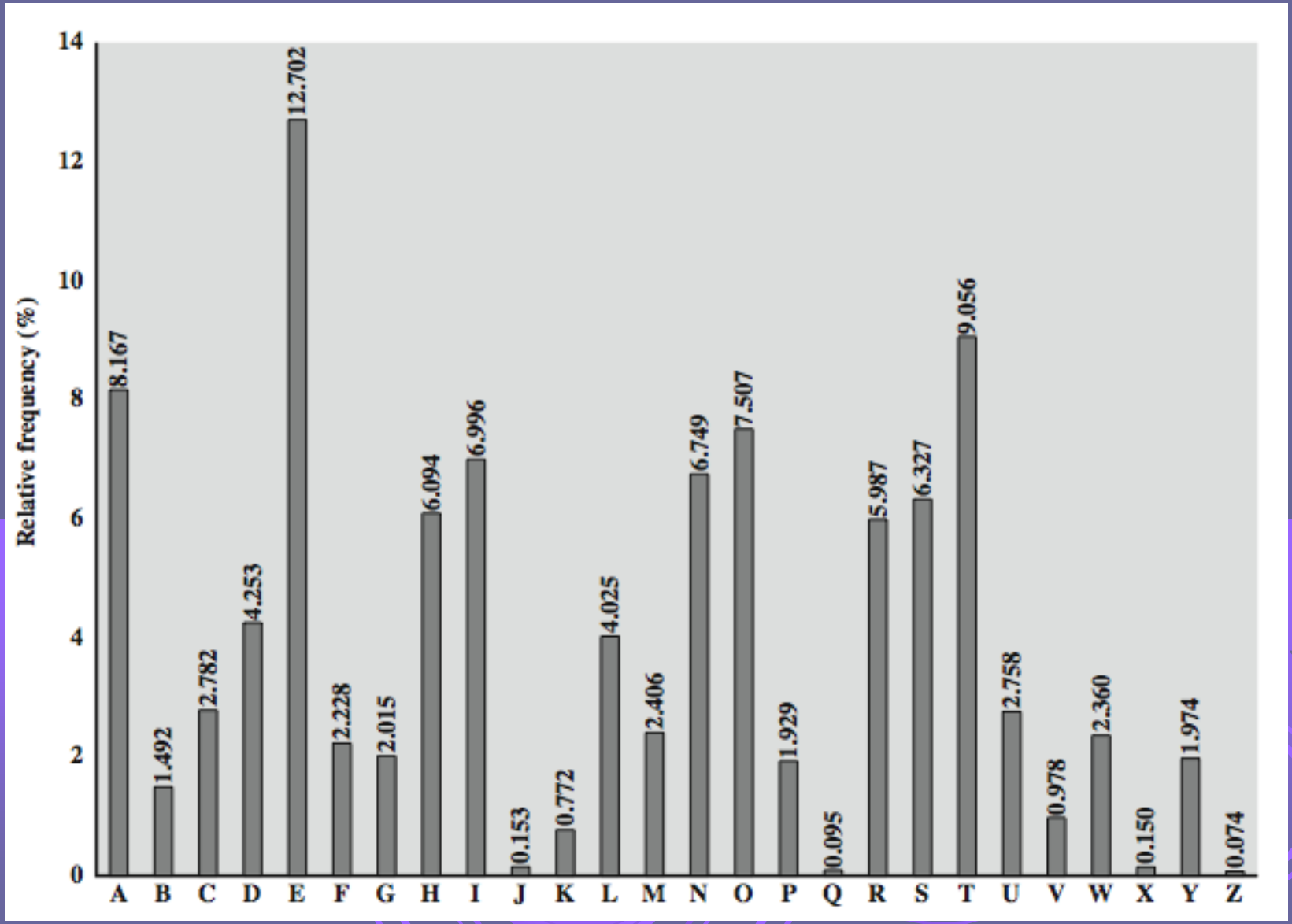
Monoalphabetic Cipher Security

- key size is now 25 characters…
- now have a total of 26! = 4 x 10^{26} keys
- with so many keys, might think is secure
- but would be **!!!WRONG!!!**
- problem is language characteristics

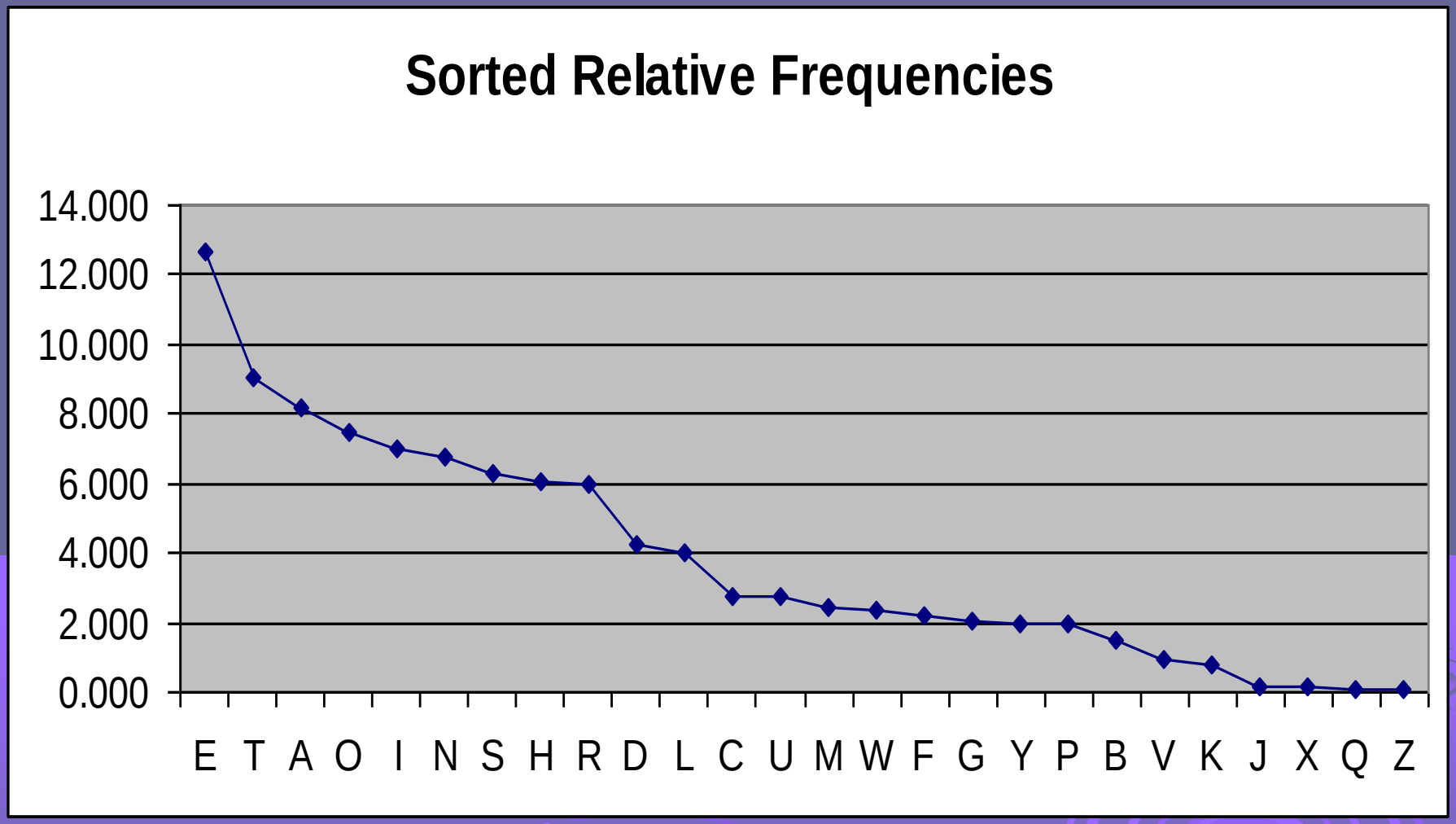
Language Redundancy and Cryptanalysis

- human languages are **redundant**
- e.g., "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

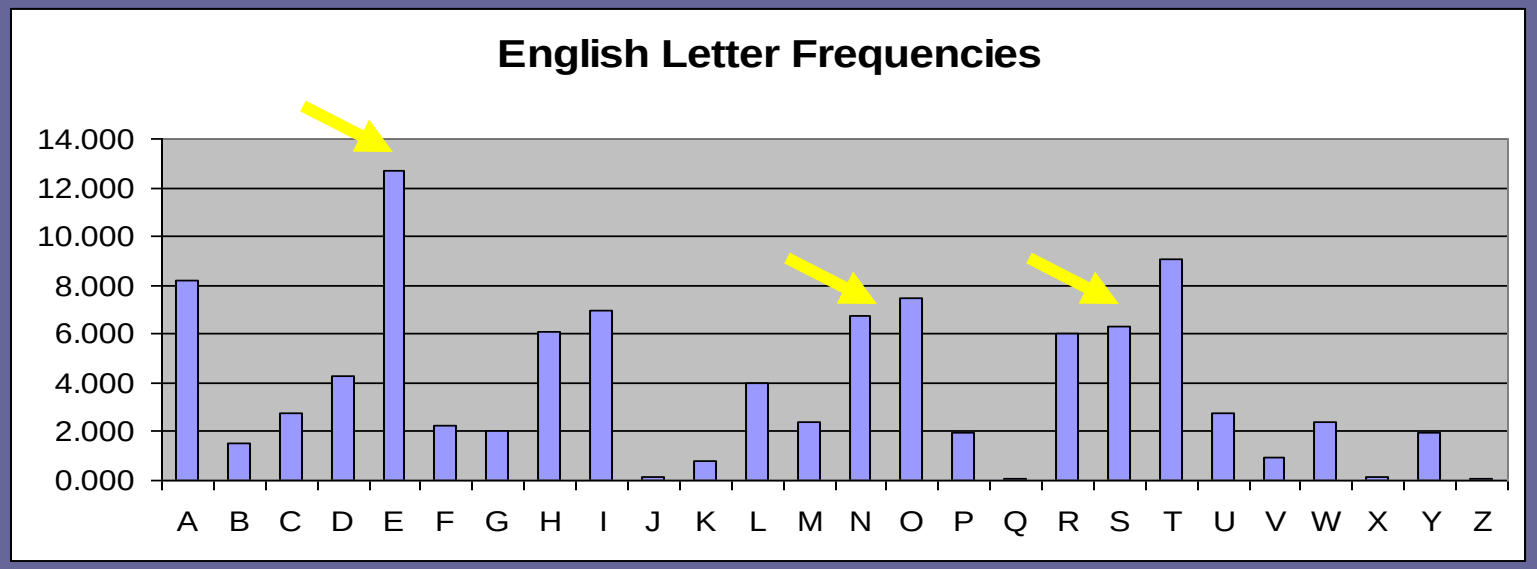
English Letter Frequencies

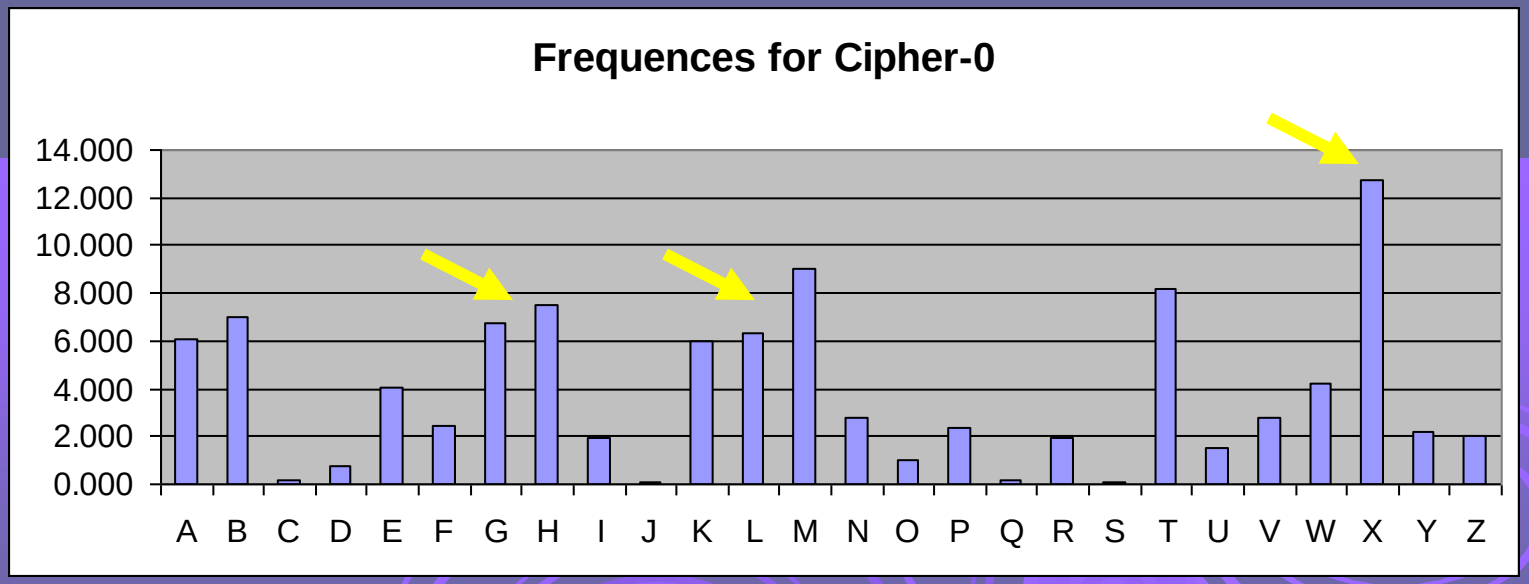


English Letter Frequencies

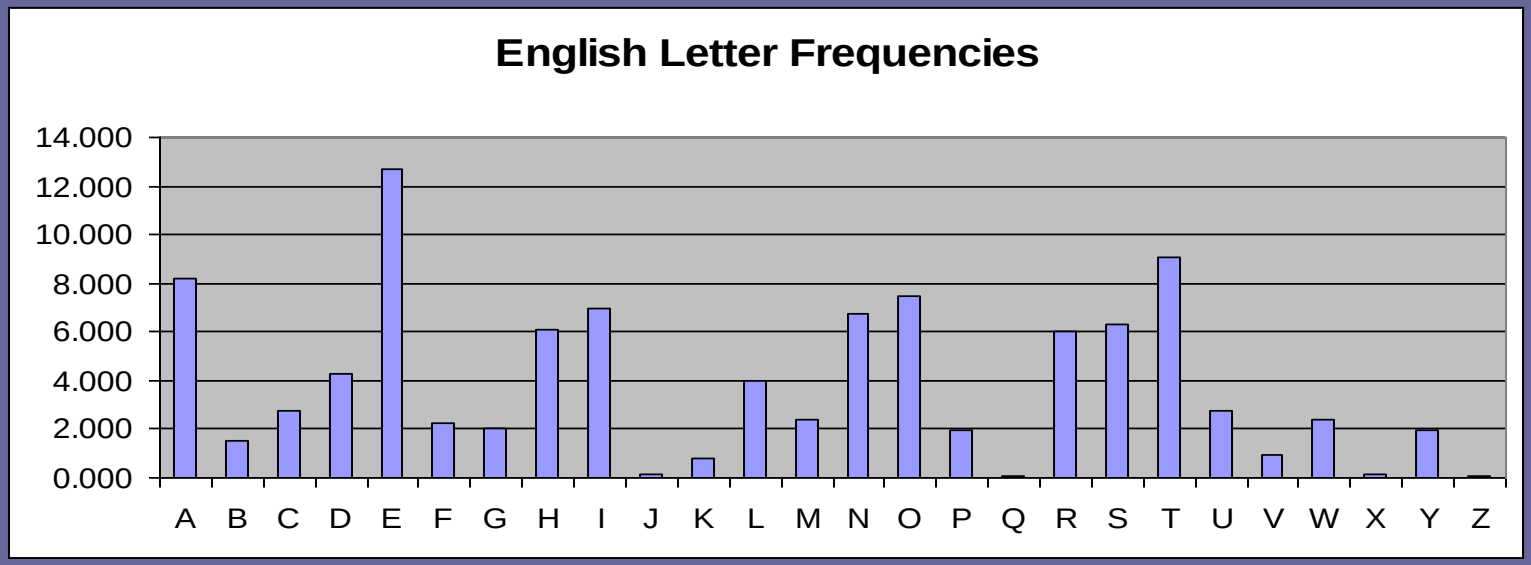


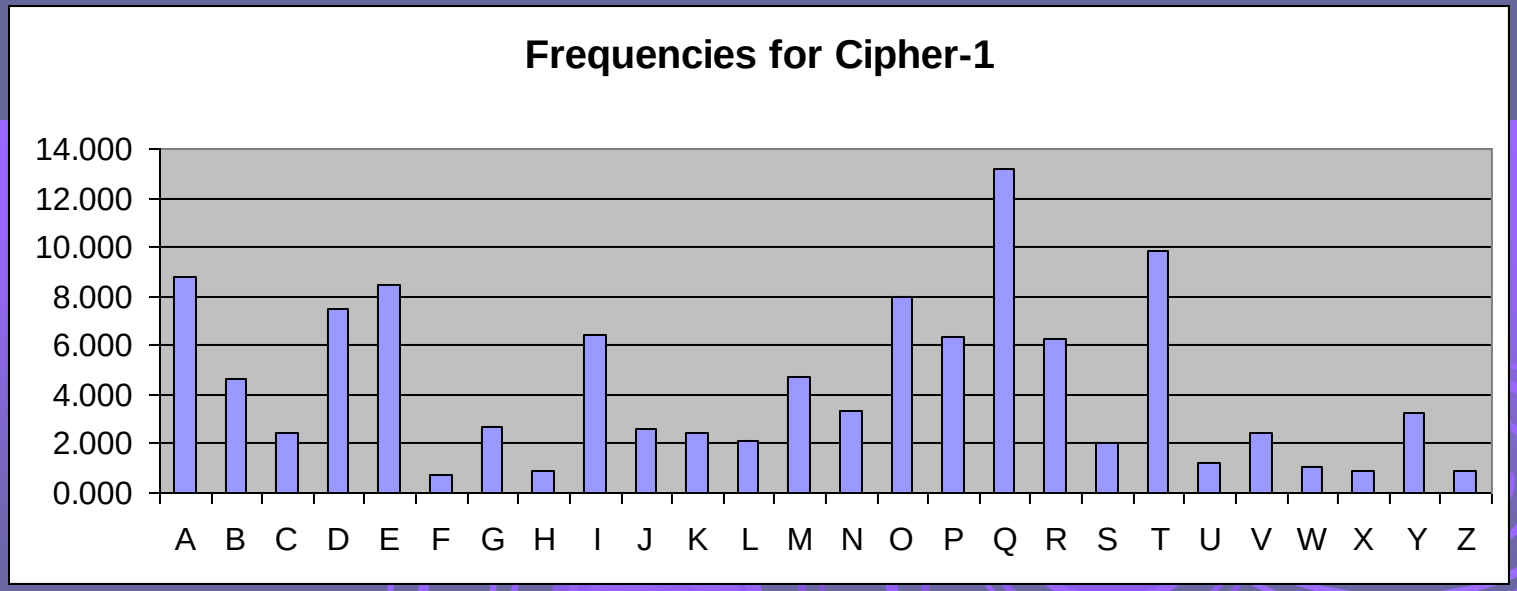
What kind of cipher is this?

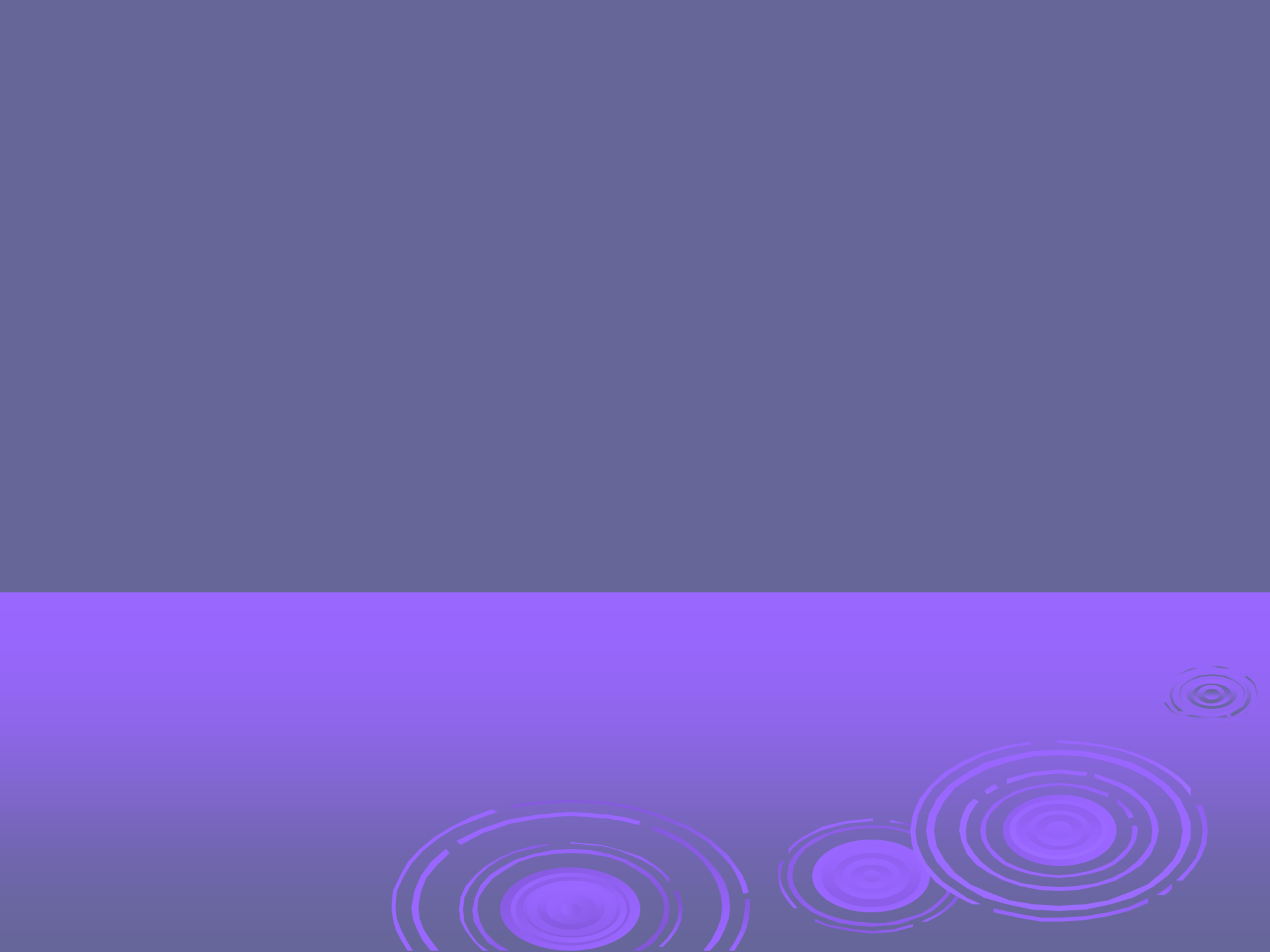




What kind of cipher is this?







Use in Cryptanalysis

- key concept - monoalphabetic substitution ciphers do not change relative letter frequencies
- discovered by Arabian scientists in 9th century
- calculate letter frequencies for ciphertext
- compare counts/plots against known values
- if caesar cipher look for common peaks/troughs
  - peaks at: A-E-I triple, N-O pair, R-S-T triple
  - troughs at: J-K, U-V-W-X-Y-Z
- for monoalphabetic must identify each letter
  - tables of common double/triple letters help (digrams and trigrams)
- amount of ciphertext is important – statistics!

Example Cryptanalysis

- given ciphertext:

  UZQSOVUOHXMOPVGPOZPEVSGZWSZOPFPESXUDBMETSXAIZ

  VUEPHZHMDZSHZOWSFPAPPDTSVPQUZWYMXUZUHSX

  EPYEPOPDZSZUFPOMBZWPFUPZHMDJUDTMOHMQ

- count relative letter frequencies (see text)

Example Cryptanalysis

- given ciphertext:

  UZQSOVUOHXMOPVGPOZPEVSGZWSZOPFPESXUDBMETSXAIZ
  VUEPHZHMDZSHZOWSFPAPPDTSVPQUZWYMXUZUHSX
  EPYEPOPDZSZUFPOMBZWPFUPZHMDJUDTMOHMQ

- 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

Playfair Example

- Message = Move forward
- Plaintext = mo ve fo rw ar dx
- Here x is just a filler, message is padded and segmented
- mo -> ON; ve -> UF; fo -> PH, etc.
- Ciphertext = ON UF PH NZ RM BZ

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

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 (versus 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
```

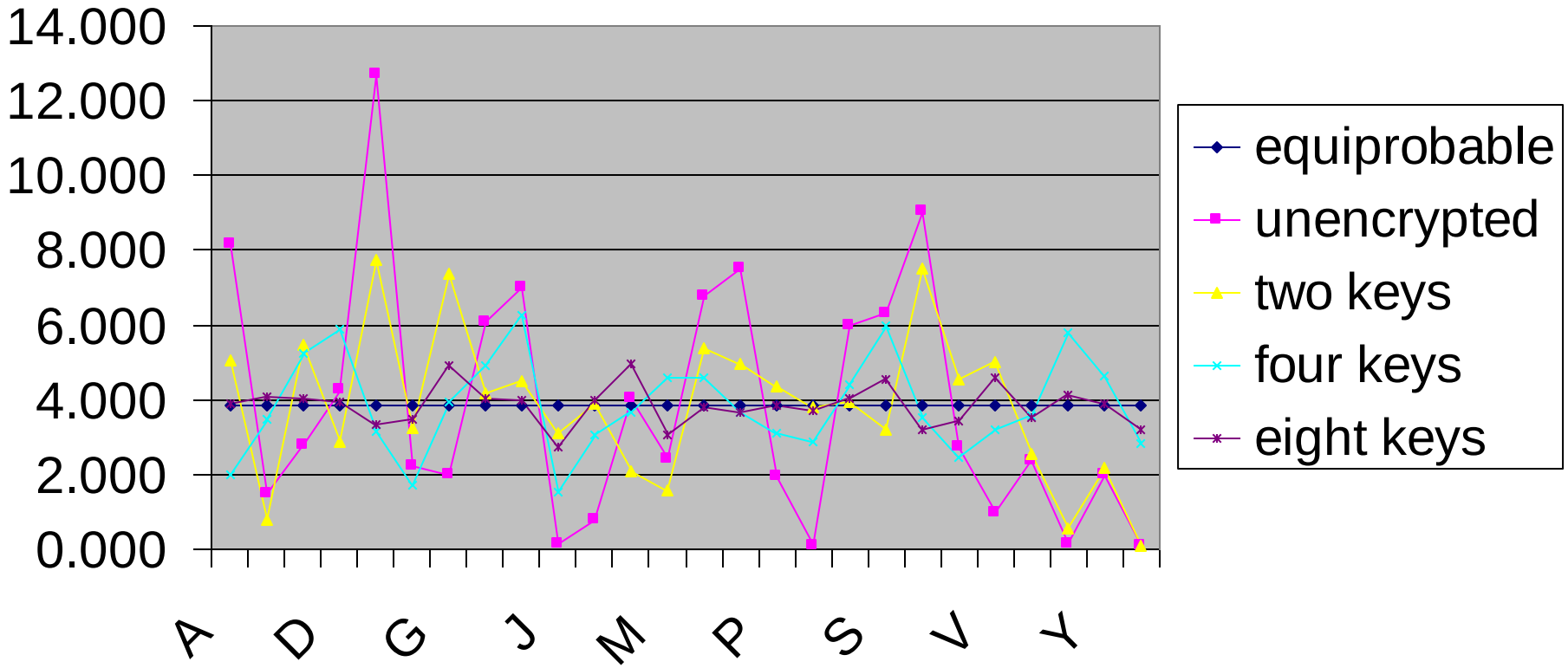
Aids

- simple aids can assist with en/decryption
- a **Saint-Cyr Slide** is a simple manual aid
  - a slide with repeated alphabet
  - line up plaintext 'A' with key letter, eg 'C'
  - then read off any mapping for key letter
- can bend round into a **cipher disk**
- or expand into a **Vigenère Tableau**

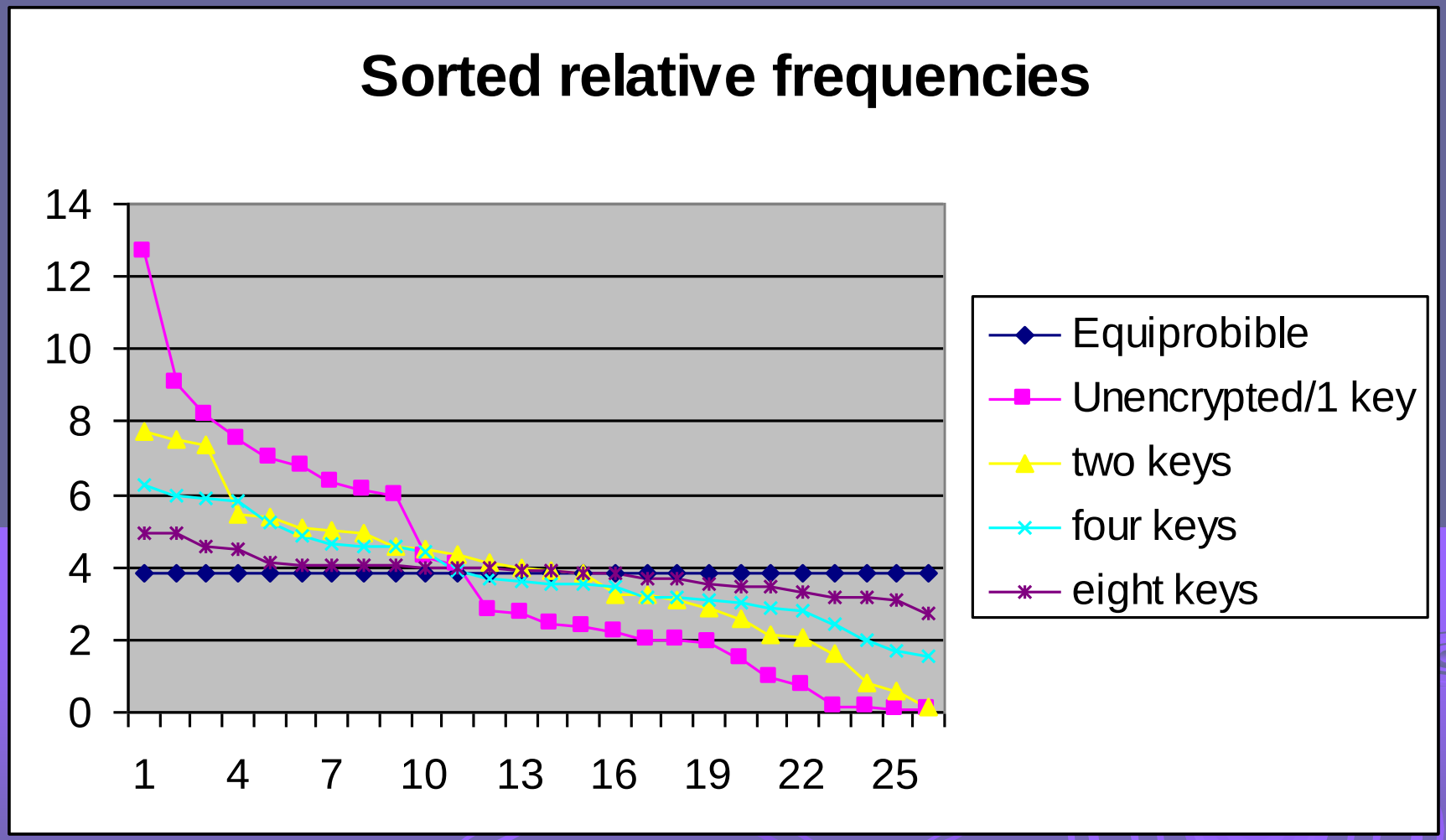
Security of Vigenère Ciphers

- have multiple ciphertext letters for each plaintext letter
- hence letter frequencies are obscured
- but not totally lost
- start with letter frequencies
  - see if it looks monoalphabetic or not
- if not, then need to determine number of alphabets, since then can attack each

Frequencies After Polyalphabetic Encryption



Frequencies After Polyalphabetic Encryption



Homework 1

- Due next class
- Question 1:

What is the best “flattening” effect you can achieve by carefully selecting two monoalphabetic substitutions? Explain and give an example. What about three monoalphabetic substitutions?

Kasiski Method

- method developed by Babbage / Kasiski
- repetitions in ciphertext give clues to period
- so find same plaintext a multiple of key length apart
- which results in the same ciphertext
- of course, could also be random fluke
- e.g. repeated “VTW” in previous example
- distance of 9 suggests key size of 3 or 9
- then attack each monoalphabetic cipher individually using same techniques as before

Example of Kasiski Attack

- Find repeated ciphertext trigrams (e.g., VTW)
- May be result of same key sequence and same plaintext sequence (or not)
- Find distance(s)
- Common factors are likely key lengths

```
key:        deceptivedeceptivedeceptive
plaintext:  wearediscoveredsaveyourself
ciphertext:ZICVTWQNGRZGVTWAVZHCQYGLMGJ
```

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
```

Homophone Cipher

- rather than combine multiple monoalphabetic ciphers, can assign multiple ciphertext characters to same plaintext character
- assign number of homophones according to frequency of plaintext character
- Gauss believed he made unbreakable cipher using homophones
- but still have digram/trigram frequency characteristics to attack
- e.g., have 58 ciphertext characters, with each plaintext character assigned to ceil(freq/2) ciphertext characters – so e has 7 homophones, t has 5, a has 4, j has 1, q has 1, etc.

Vernam Cipher

- ultimate defense is to use a key as long as the plaintext
- with no statistical relationship to it
- invented by AT&T engineer Gilbert Vernam in 1918
- specified in U.S. Patent 1,310,719, issued July 22, 1919
- originally proposed using a very long but eventually repeating key
- used electromechanical relays

One-Time Pad

- if a truly random key as long as the message is used, the cipher will be secure
- called a One-Time pad (OTP)
- is 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

- now 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
- eg. 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

- is a more complex transposition
- 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

```
Key: 4312567
Column Out 4 3 1 2 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
Ciphertext: TTNAAPTMTSUOAODWCOIXKNLYPETZ
```

Block Transposition Ciphers

- arbitrary block transposition may be used
- specify permutation on block
- repeat for each block of plaintext

```
Key: 4931285607
Plaintext:  attackpost poneduntil twoamxyzab
```

```
Ciphertext: CTATTSKPAO DLEONIDUPT MBAWOAXYTZ
```

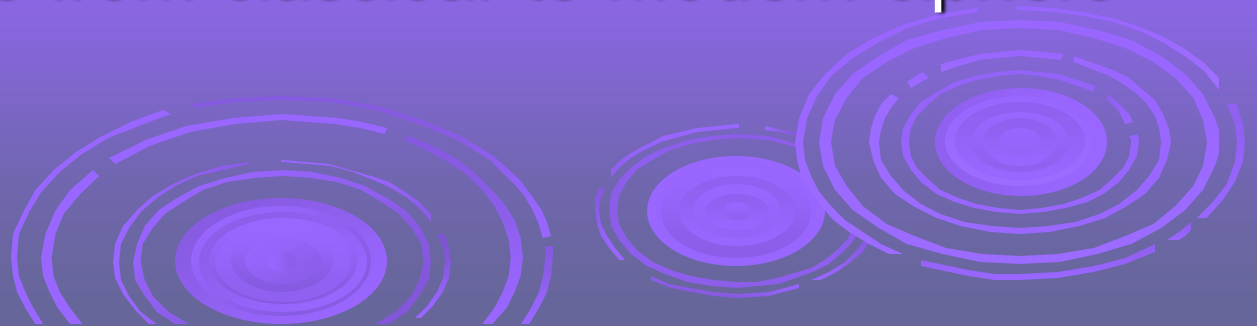
Homework 1

- Due next class
- Question 2:

Mathematically specify an arbitrary block transposition cipher with block length B and permutation π:[0..B-1] $\rightarrow$ [0..B-1] for plaintext $P=p_0p_1p_2p_3\ldots p_{N-1}$, where N is a multiple of B.

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



Homework 1

- Due next class
- Question 3: (be mathematical)

a. What is the result of the product of two rotational substitutions?

b. What is the result of the product of two affine substitutions?

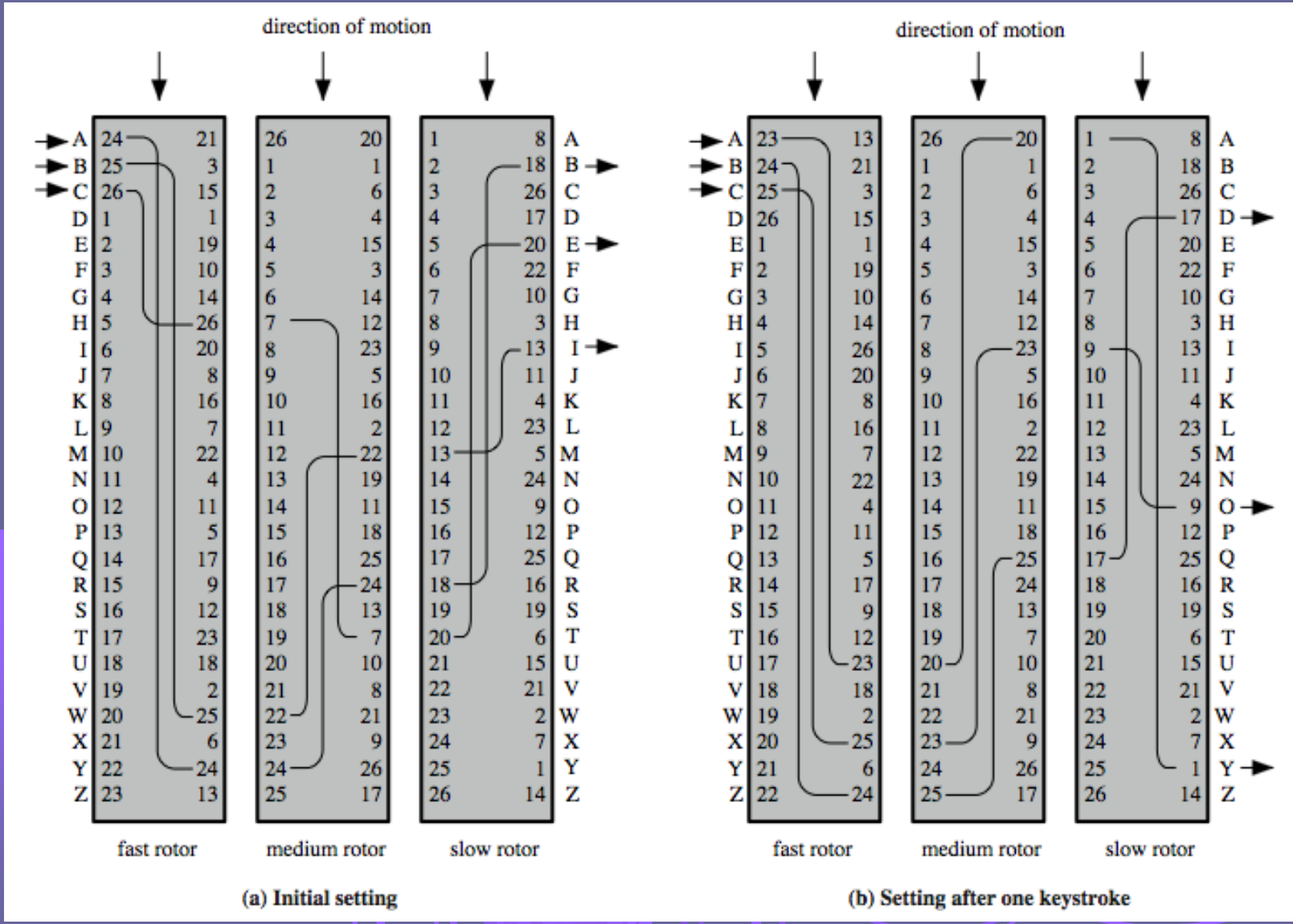
c. What is the result of the product of two block transpositions?

Rotor Machines

- before modern ciphers, rotor machines were most common complex ciphers in use
- widely used in WW2
  - German Enigma, Allied Hagelin, Japanese Purple
- implemented a very complex, varying substitution cipher
- used 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

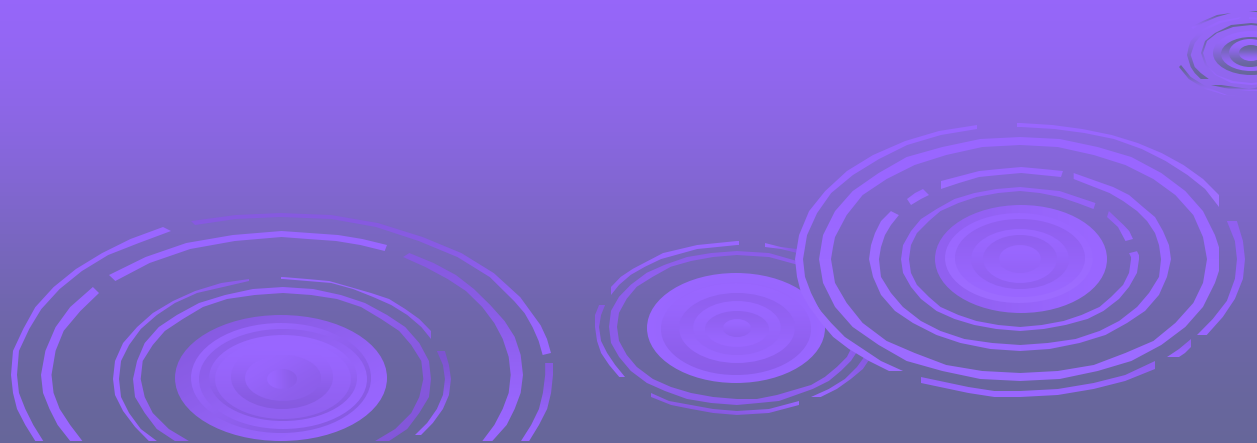
Rotor Machine Principles



Homework 1

- Due next class
- Question 4:

Give a mathematical description of a two-rotor cipher.



Rotor Ciphers

- Each rotor implements some permutation between its input and output contacts
- Rotors turn like an odometer on each key stroke (rotating input and output contacts)
- Key is the sequence of rotors and their initial positions

Steganography

- an alternative to encryption
- hides existence of message
  - using only a subset of letters/words in a longer message marked in some way
  - using invisible ink
  - hiding in LSB in graphic image or sound file
  - hide in “noise”
- has drawbacks
  - high overhead to hide relatively few info bits
- advantage is can obscure encryption use

Summary

- have considered:
  - classical cipher techniques and terminology
  - monoalphabetic substitution ciphers
  - cryptanalysis using letter frequencies
  - Playfair cipher
  - polyalphabetic ciphers
  - transposition ciphers
  - product ciphers and rotor machines
  - steganography