Math Rules Cyberspace

Travis H.

Stanford University, 13-14 Nov 2010

Intro

- Play clip 1 Cryptografie, Cryptografia intro clip (2:18)

Web Sites

Castles on the Internet

- Have a practical purpose so can never be perfect
- More valuable the data, the stronger it must be
- Always under attack

Web Sites

Castles on the Internet

- Have a practical purpose so can never be perfect
- More valuable the data, the stronger it must be
- Always under attack

Web Sites

Castles on the Internet

- Have a practical purpose so can never be perfect
- More valuable the data, the stronger it must be
- Always under attack

Web Sites

Castles on the Internet

- Have a practical purpose so can never be perfect
- More valuable the data, the stronger it must be
- Always under attack

Firewalls

Keeping Bad Guys Out

- But a real firewall has to let *something* in or out
- Every castle must have one door
- Otherwise there's no point

What Really *Is* a Hacker?

- Most people only see superficial details
- A hacker wants to *understand* the Matrix
- Not necessarily malicious

Hackers

- Play clip 2 from the movie *Hackers*, typical 90s stereotype (3:57)

System Crackers are Malicious Hackers

The Internet Ninjas

- Powers of invisibility
- Like to wear black
- Strike without warning
- Leave no trace
- Make most people uncomfortable

System Crackers are Malicious Hackers

The Internet Ninjas

- Powers of invisibility
- Like to wear black
- Strike without warning
- Leave no trace
- Make most people uncomfortable

System Crackers are Malicious Hackers

The Internet Ninjas

- Powers of invisibility
- Like to wear black
- Strike without warning
- Leave no trace
- Make most people uncomfortable

System Crackers are Malicious Hackers

The Internet Ninjas

- Powers of invisibility
- Like to wear black
- Strike without warning
- Leave no trace
- Make most people uncomfortable

System Crackers are Malicious Hackers

The Internet Ninjas

- Powers of invisibility
- Like to wear black
- Strike without warning
- Leave no trace
- Make most people uncomfortable

System Crackers are Malicious Hackers

The Internet Ninjas

- Powers of invisibility
- Like to wear black
- Strike without warning
- Leave no trace
- Make most people uncomfortable

White Hat Hackers

- Play clip 2, news on White Hat Hackers (2:33)

Security Experts

Internet Jedi

- No effective law enforcement on Internet, like Wild West
- No regulation of software industry
- Nobody to protect people from vendors and crackers

Security Experts

Internet Jedi

- No effective law enforcement on Internet, like Wild West
- No regulation of software industry
- Nobody to protect people from vendors and crackers

Security Experts
Internet Jedi

- No effective law enforcement on Internet, like Wild West
- No regulation of software industry
- Nobody to protect people from vendors and crackers

There Are Temptations

Cryptologists

Modern Wizards

- Start off very weak
- Require many years to develop their powers
- Pore over dusty tomes to find the information they need
- Books are incomprehensible to others
- Full of weird symbols and obscure incantations

Cryptologists

Modern Wizards

- Start off very weak
- Require many years to develop their powers
- Pore over dusty tomes to find the information they need
- Books are incomprehensible to others
- Full of weird symbols and obscure incantations

Cryptologists
Modern Wizards

- Start off very weak
- Require many years to develop their powers
- Pore over dusty tomes to find the information they need
- Books are incomprehensible to others
- Full of weird symbols and obscure incantations

Cryptologists
Modern Wizards

- Start off very weak
- Require many years to develop their powers
- Pore over dusty tomes to find the information they need
- Books are incomprehensible to others
- Full of weird symbols and obscure incantations

Cryptologists

Modern Wizards

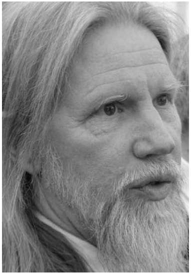
- Start off very weak
- Require many years to develop their powers
- Pore over dusty tomes to find the information they need
- Books are incomprehensible to others
- Full of weird symbols and obscure incantations

A Powerful Wizard

- Gandalf the White
- Most powerful wizard in Gondor

Coincidence? I think not.

- Gandalf the White
- Most powerful wizard in Gondor



- Whitfield Diffie
- Chief Security Officer at Sun Microsystems

Have You Ever Wished?

- ...you could walk through walls?
- How about **fire**walls?
- ...you had one of those invisibility cloaks?
- How about remaining undetected **on the Internet**?
- ...you could tame monsters?
- How about **botnets**?

Have You Ever Wished?

- ...you could walk through walls?
- How about **fire**walls?
- ...you had one of those invisibility cloaks?
- How about remaining undetected **on the Internet**?
- ...you could tame monsters?
- How about **botnets**?

Have You Ever Wished?

- ...you could walk through walls?
- How about **fire**walls?
- ...you had one of those invisibility cloaks?
- How about remaining undetected **on the Internet**?
- ...you could tame monsters?
- How about **botnets**?

Have You Ever Wished?

- ...you could walk through walls?
- How about **fire**walls?
- ...you had one of those invisibility cloaks?
- How about remaining undetected **on the Internet**?
- ...you could tame monsters?
- How about **botnets**?

Have You Ever Wished?

- ...you could walk through walls?
- How about **fire**walls?
- ...you had one of those invisibility cloaks?
- How about remaining undetected **on the Internet**?
- ...you could tame monsters?
- How about **botnets**?

Have You Ever Wished?

- ...you could walk through walls?
- How about **fire**walls?
- ...you had one of those invisibility cloaks?
- How about remaining undetected **on the Internet**?
- ...you could tame monsters?
- How about **botnets**?

Crypto Intro

- Play clip 4 “A Brief History of Cryptography” (6:11)
- Covers what cryptography is

What All the Words Mean

cryptography is encrypting your information so that other people can't read it

cryptanalysis is trying to read other people's encrypted messages

cryptology is the study of both

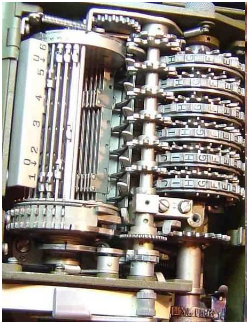
Roman Times

The Skytale

Jefferson Cylinder - 1790

WWII - Enigma

Korean War

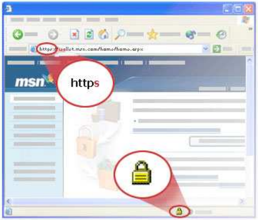


Late 20th Century

Russian Fialka Machine

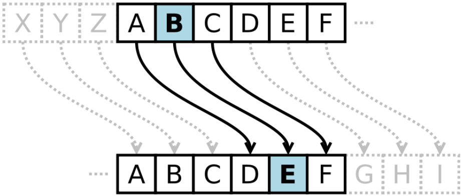
Modern Encryption Machines

Almost Everything on the Internet



- If you see these icons, your computer is doing encryption

Caesar Cipher



- Replace input (B) with letter three to the right (E)
- The number three is called the *key* to the cipher
- Wraps around
- To decrypt we do the reverse
- *ANT* becomes *DQW*

Caesar Cipher Example

substitution table

plaintext	ABCDEFGHIJKLMNOPQRSTUVWXYZ
ciphertext	DEFGHIJKLMNOPQRSTUVWXYZABC

example

plaintext	THE QUICK BROWN FOX JUMPS OVER THE LAZY DOG
ciphertext	WKH TXLFN EURZQ IRA MXPSV RYHU WKH ODCB GRJ

Changing Symbols

- We use 26 symbols (A-Z); this is called our *alphabet*.
- It's irrelevant to cryptographers.
- For example, we could change to numbers:

A	B	C	D	E	F	G	H	I	J	K	L	M
00	01	02	03	04	05	06	07	08	09	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

- These are called *ordinal numbers*.

Encrypting Using Numbers

- 1. Replace *A* with *0*, *B* with *1*, *C* with *2*, ... *Z* with *25*
  2. Add the *key* (3) to each number
  3. Replace *0* with *A*, *1* with *B*, *2* with *C*, ... *25* with *Z*
- - But wait, what if we went over 25?
  - In that case we subtract 26 from the result
  - So 24 plus 3 is 27, but that's too high, so 27 - 26 = 1
  - This is called *modular addition*.
- - For decryption, we subtract the key *k*
  - If we go under zero, then we add 26

Encrypting Using Numbers

- 1. Replace *A* with *0*, *B* with *1*, *C* with *2*, ... *Z* with *25*
  2. Add the *key* (3) to each number
  3. Replace *0* with *A*, *1* with *B*, *2* with *C*, ... *25* with *Z*
- - But wait, what if we went over 25?
  - In that case we subtract 26 from the result
  - So 24 plus 3 is 27, but that's too high, so 27 - 26 = 1
  - This is called *modular addition*.
- - For decryption, we subtract the key *k*
  - If we go under zero, then we add 26

Encrypting Using Numbers

- 1. Replace *A* with *0*, *B* with *1*, *C* with *2*, ... *Z* with *25*
  2. Add the *key* (3) to each number
  3. Replace *0* with *A*, *1* with *B*, *2* with *C*, ... *25* with *Z*
- - But wait, what if we went over 25?
  - In that case we subtract 26 from the result
  - So 24 plus 3 is 27, but that's too high, so 27 - 26 = 1
  - This is called *modular addition*.
- - For decryption, we subtract the key *k*
  - If we go under zero, then we add 26

How Is This Math?

If x is the plaintext and y is the ciphertext, the equation we're using is:

$$y = (x + 3) \ mod \ 26$$

Or more generally, for a key k and an alphabet of n symbols:

$$y = (x + k) \ mod \ n$$

Decryption is similar:

$$x = (y - k) \ mod \ n$$

Cryptanalysis of Caesar Cipher

encrypted message

Nwcz akwzm ivl amdmv gmiza iow wcz nibpmza jzwcopb nwzbp wv bpqa kwvbqvmvb, i vme vibqwv, kwvkmqdml qv tqjmzbg, ivl lmlqkibml bw bpm xzwxwaqbqwv bpib itt umv izm kzmibml mycit.

- Q: How can we read such a message without knowing the key?

Trivial Cryptanlysis

encrypted message

Nwcz akwzm ivl amdmv gmiza iow wcz nibpmza jzwcopb nwzbp wv bpqa kwvbqvmvb, **i** vme vibqwv, kwvkmqdml qv tqjmzbg, ivl lmlqkibml bw bpm xzwxwaqbqwv bpib itt umv izm kzmibml mycit.

- If we know the input is English, there's only a few one-letter words.
- Since the rest of the input is gibberish, and it's not capitalized, "i" must be ciphertext for a.

Classical Countermeasures

- Use all one case (lower or upper)
- Remove punctuation and spaces
- You'll sometimes see them in five-letter groups; that's easier to read and was normal for telegraphs
- But I'm lazy and didn't do that in this talk
- Q: Why doesn't modern cryptography do this?

Brute Force Attack

encrypted message

nwczakwzmivlamdmvgmizaiowwcznibpmzajzwcopbnwzbpwv
bpqakwvbqvmvbivmevibqwvkwvkmqdmlqvtqjmzbgivllmlqki
bmlbwbpmxzwxwaqbqwvbpibittumvizmkzmibmlmycit

- Brute force attack tries all 26 possible keys (k=0. . . 25)
- One of them will yield a readable message
- Rest will still look encrypted

Frequency Analysis

- We know that e is the most common letter in English
- Count which is the most common letter in the message
- That's probably the letter e in the original

Frequency Analysis Example

- There's 18 occurences of the letter *m*

encrypted message

nwczakwz**m**ivla**m**d**m**vg**m**izaiowwcznibp**m**zajzwcopbnwzbpwv
bpqakwvbqv**m**vbiv**m**evibqwvkwvk**m**qd**m**lqvtqj**m**zbgivll**m**lqki
b**m**lbw bp**m**xzwxwaqbqwvbpibittu**m**viz**m**kz**m**ib**m**l**m**ycit

Let $y = ord(m) = 12$, $x = ord(e) = 4$, and remember:

$$x = (y - k) \ (mod \ n)$$

$$4 = (12 - k)$$

$$k = (12 - 4) = 8$$

Frequency Analysis Solution

encrypted message

nwczakwzmivlamdmvgmizaiowwcznibpmzajzwcopbnwzbpwv
bpqakwvbqvmvbivmevibqwvkwvkmqdmlqvTqjmzbgivllmlqki
bmlbwbpmxzwxwaqbqwvbpibittumvizmkzmibmlmycit

decrypted message (spaces and punctuation added)

Four score and seven years ago our fathers brought forth on this continent, a new nation, conceived in Liberty, and dedicated to the proposition that all men are created equal.

How Do We Improve the Cipher?

- How do we improve this cipher?
- First, we need to identify the problems.
- Q: What was the problem making brute force possible?

How Do We Improve the Cipher?

- How do we improve this cipher?
- First, we need to identify the problems.
- Q: What was the problem making brute force possible?

How Do We Improve the Cipher?

- How do we improve this cipher?
- First, we need to identify the problems.
- Q: What was the problem making brute force possible?

Caesar Cipher Review

substitution table

plaintext	ABCDEFGHIJKLMNOPQRSTUVWXYZ
ciphertext	DEFGHIJKLMNOPQRSTUVWXYZABC

- Note that the ciphertext alphabet is just the plaintext alphabet slid over
- All we need is one input and one output to figure out the *key* (the amount of the rotation)
- Q: How many keys are there?

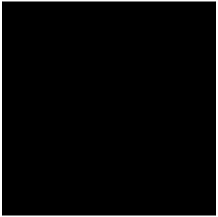
Substitution Cipher

- A *substitution cipher* maps from one alphabet to another
- Can map from and to same alphabet, but scrambled

substitution table

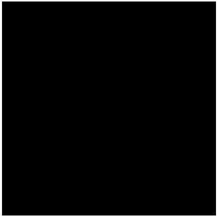
plaintext	ABCDEFGHIJKLMNOPQRSTUVWXYZ
ciphertext	THEQUICKBROWNFXJMPDVRLAZYG

Partial Function



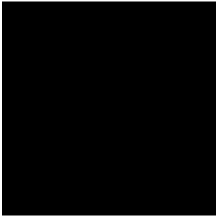
- With a *partial function*,
- some plaintext letters (elements of domain X) don't have known ciphertext letters (elements of co-domain Y)
- This isn't usually the case, and isn't the case here.

Total Function



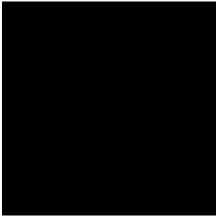
- In our case, every input symbol has an output symbol
- This is called a *total function*, or usually just a *function*

Injection



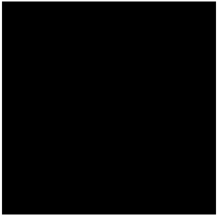
- Every input (in X) has *at most* one output (in Y)
- This makes it an *injective function*, or *one-to-one*

Surjection



- Every output (in Y) has *at least* one input (in X)
- This makes it a *surjective function*, or *onto*

Bijection



- A function that is both is known as *bijective*, or *one-to-one correspondence*
- This means that every output has exactly one input

Permutation

- A total bijective function whose outputs (codomain) are from the same set as its inputs (domain) is a *permutation*
- Basically, a permutation is just a scrambling of the elements

In Other Words

classical cryptographer	computer scientist	mathematician
plaintext	input	domain
ciphertext	output	codomain
letter	symbol	element
alphabet	alphabet	set

Does This Solve Our Problem?

- Caesar cipher had only 26 possible keys
- Q: How many does a substitution cipher have?

How Many Ways to Scramble 26 Letters?

- First letter may map to any of the 26 letters
- Second letter may map to 25 remaining letters
- Third letter may map to any of 24 remaining
- Do you see a pattern?

It's a Factorial!

$$26 \star (25 \star (24 ...)) = 26!$$

$$26! = 403291461126605635584000000$$

You've now discovered how to find out how many permutations are possible for a set.
For n elements, it's n factorial!
This is far too many to try them all (brute force)

It's a Factorial!

$$26 \star (25 \star (24 \ldots)) = 26!$$

$$26! = 403291461126605635584000000$$

You've now discovered how to find out how many permutations are possible for a set.

For n elements, it's n factorial!

This is far too many to try them all (brute force)

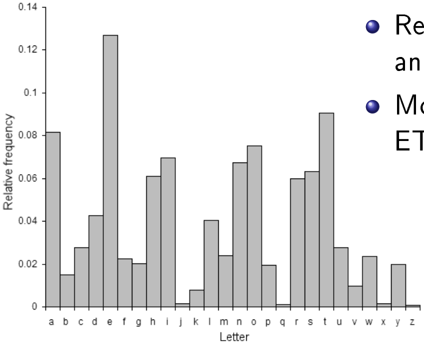
Sample Substitution

encrypted message

```
LIVITCSWPIYVEWHEVSRIQMXLEYVEOIEWHRXEXIPFEMVEWHKVSTYLXZIXLIKIIXPIJVSZEYPER
RGERIMWQLMGLMXQERIWGPSRIHMXQEREKIETXMJTPRGEVEKEITREWHEXXLEXXMZITWAWSQWXSW
EXTVEPMRXRSJGSTVRIEYVIEXCVMUIMWERGMIWXMJMGCSMWXSJOMIQXLIVIQIVIXQSVSTWHKPE
GARCSXRWIEVSWIIBXVIZMXFSJXLIKEGAEWHEPSWYSWIWIEVXLISXLIVXLIRGEPIRQIVIIBGII
HMWYPFLEVHEWHYPSRRFQMXLEPPXLIECCIEVEWGISJKTVWMRLIHYSPHXLIQIMYLXSJXLIMWRIG
XQEROIVFVIZEVAEKPIEWHXEAMWYEPPXLMWYRMWXSGSWRMHIVEXMSWMGSTPHLEVHPFKPEZINTC
MXIVJSVLMRSCMWMSWVIRCIGXMWYMX
```

- So how would you *cryptanalyze* this?
- Taking the spaces out is a common trick and *usually* doesn't hurt readability of plaintext

English Letter Frequency Distribution



- Remember frequency analysis?
- Most common letters are: ETAOINSHRDLCU...

English Bigram Distribution

- *bigrams* are pairs of letters
- most common is "th", followed by "he", and others

English Trigram Distribution

- *trigrams* are three letters in a row
- most common is “the”, followed by “and”, “tha”, etc.

Attacking It 1

encrypted message

LIVITCSWPIYVEWHEVSRIQMXLEYVEOIEWHRXEXIPFEMVEWHKVSTYLXZIXLIKIIXPIJVSZEYPER
RGERIMWQLMGLMXQERIWGPSRIHMXQEREKIETXMJTPRGEVEKEITREWHEXXLEXXMZITWAWSQWXSW
EXTVEPMRXRSJGSTVRIEYVIEXCVMUIMWERGMIWXMJMGCSMWXSJOMIQXLIVIQIVIXQSVSTWHKPE
GARCSXRWIEVSWIIBXVIZMXFSJXLIKEGAEWHEPSWYSWIWIEVXLISXLIVXLIRGEPIRQIVIIBGII
HMWYPFLEVHEWHYPSRRFQMXLEPPXLIECCIEVEWGISJKTVWMRLIHYSPHXLIQIMYLXSJXLIMWRIG
XQEROIVFVIZEVAEKPIEWHXEAMWYEPPXLMWYRMWXSGSWRMHIVEXMSWMGSTPHLEVHPFKPEZINTC
MXIVJSVLMRSCMWMSWVIRCIGXMWYMX

- I was most common letter, XL most common bigram, XLI most common trigram
- Guessed that XLI=the

Attacking It 2

encrypted message

```
heVeTCSWPeYVaWHaVSReQMthaYVaOeaWHRtatePFaMVaWHKVSTYhtZetheKeetPeJVSZaYPaR
RGaReMWQhMGhMtQaReWGPSReHMtQaRaKeaTtMJTPRGaVaKaeTRaWHatthattMZeTWAWSQWtSW
atTVaPMRtRSJGSTVReaYVeatCVMUeMWaRGMeWtMJMGCSMWtSJOMeQtheVeQeVetQSVSTWHKPa
GARCStRWeaVSWeeBtVeZMtFSJtheKaGAaWHaPSWYSWeWeaVtheStheVtheRGaPeRQeVeeBGee
HMWYPFhaVHaWHYPSRRFQMthaPPtheaCCeaVaWGeSJKTVWMRheHYSPHtheQeMYhtSJtheMWReG
tQaROeVFVeZaVAaKPeaWHtaAMWYaPPthMWYRMWtSGSWRMHeVatMSWMGSTPHhaVHPFKPaZeNTC
MteVJSVhMRSCMWMSWVeRCeGtMWYMt
```

- heVe = here, Rtate = state, atthattMZe = atthattime
- means V=r, R=s, M=i, Z=m

Attacking It 3

encrypted message

```
hereTCSWPeYraWHarSseQithaYraOeaWHstatePFairaWHKrSTYhtmetheKeetPeJrSmaYPas
sGaseiWQhiGhitQaseWGPSseHitQasaKeaTtiJTPsGaraKaeTsaWHatthattimeTWAWSQWtSW
atTraPistsSJGSTrseaYreatCriUeiWasGieWtiJiGCSiWtSJOieQthereQeretQSrSTWHKPa
GAsCStsWearSWeeBtremitFSJtheKaGAaWHaPSWYSWeWeartheStherthesGaPesQereeBGee
HiWYPFharHaWHYPSssFQithaPPtheaCCearaWGeSJKTrWisheHYSPHtheQeiYhtSJtheiWseG
tQasOerFremarAaKPeaWHtaAiWYaPPthiWYsiWtSGSWsiHeratiSWiGSTPHharHPFKPameNTC
iterJSrhisSCiWiSWresCeGtiWYit
```

- remarA = remark, and so on...

Done

decrypted message

```
hereuponlegrandarosewithagraveandstatelyairandbroughtmethebeetlefromaglas
scaseinwhichitwasencloseditwasabeautifulscarabaeusandatthattimeunknownton
aturalistsofcourseagreatprizeinascientificpointofviewtherewaretworoundbla
ckspotsnearoneextremityofthebackandalongoneneartheotherthescaleswereexcee
dinglyhardandglossywithalltheappearanceofburnishedgoldtheweightoftheinsec
twasveryremarkableandtakingallthingsintoconsiderationicouldhardlyblamejup
iterforhisopinionrespectingit
```

- Add spaces between words, and...

Adding Spaces

decrypted message

```
Hereupon Legrand arose, with a grave and stately air, and brought me the
beetle from a glass case in which it was enclosed.  It was a beautiful
scarabaeus, and, at that time, unknown to naturalists-of course a great
prize in a scientific point of view.  There were two round black spots
near one extremity of the back, and a long one near the other.  The
scales were exceedingly hard and glossy, with all the appearance of
burnished gold.  The weight of the insect was very remarkable, and,
taking all things into consideration, I could hardly blame Jupiter for
his opinion respecting it.
```

- Abracadbra, we're done.
- How do we solve this? Well...

Cryptography 101

- Play clip 5, video which talks about transposition ciphers (9:59)

Polyalphabetic Substitution Ciphers

- Play clip 6 (20:28), skip to 15:34 for Vigenère cipher (4:56 net)
- Covers polyalphabetic substitution ciphers

Breaking Substitution Ciphers

- Breaking these is an advanced topic
- It's called Kasiski Examination
- But you have half the knowledge already

The Manhattan Project Cipher

- Cut out two 11x11 squares of graph paper
- Number them 0..9 along X and Y axes; this gives you a 10x10 grid
- Put letters these number of times in the same place on each grid
- A 8, B 1, C 3, D 4, E 13, F 2, G 2, H 6, I 7, J 1, K 1, L 4, M 2, N 6, O 7, P 2, Q 1, R 6, S 6, T 9, U 2, V 1, W 2, X 1, Y 2, Z 1
- Encrypt by picking a letter at random, then writing down the X, Y coordinates (commas are not necessary)

Steganography

- Play clip S to explain steganography (2:30)
- Steganography is *not* cryptography, but similar in some ways

Playfair

- Play clip P which rapidly covers playfair cipher (3:00)

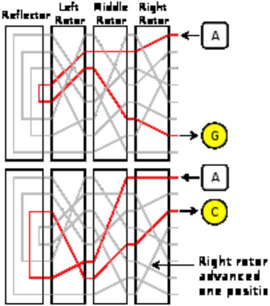
One-Time Pad (OTP)

- Play OTP video 1 explaining OTPs (2:42)
- Play OTP video 2 recording of a numbers station (1:56)
- Q: Notice anything about the groupings of numbers?

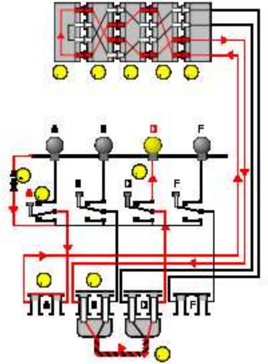
Enigma

- Play clip E1 “The Enigma Machine” (3:39)
- Talk about how rotors act as substitution ciphers
- Play clip E2 video “My Enigma” (3:44)

Rotor Assembly

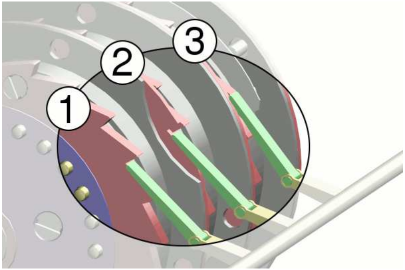


Wiring Diagram



- Red shows electrical current
- Battery (1) to key A (2) to plugboard (3) to fixed entry wheel (4) through rotors (5), to reflector (6)

But That's Not All



- With each keypress, the wheels rotated
- rotors 1-2-3 advance like seconds-minutes-hours on a clock

This Is Math Too

$$E = PRMLUL^{-1}M^{-1}R^{-1}P^{-1}$$

- Where E is encryption function, P is plugboard, R/M/L right/middle/left rotors, U is reflector

$$E = P(r^iRr^{-i})(r^jMr^{-j})(r^kLr^{-k})U(r^kL^{-1}r^{-k})(r^jM^{-1}r^{-j})(r^iR^{-1}r^{-i})P^{-1}$$

- r is the cyclic permutation operator
- Don't feel bad if that makes your brain hurt