Number Theory Section Summary: 7.5
An application to Cryptography

1. Summary

Cryptography is that application which would have driven G. H. Hardy crazy: he was in love with Number Theory because of its purity, because it didn't have application. Well, roll over in your grave Godfrey Hardy!

2. Definitions

- *Cipher* – the code
- *Plaintext* – the message to be encrypted
- *Ciphertext* – the encrypted message
- *Frequency Analysis* – using the known distribution of letters (or words) to break a code.

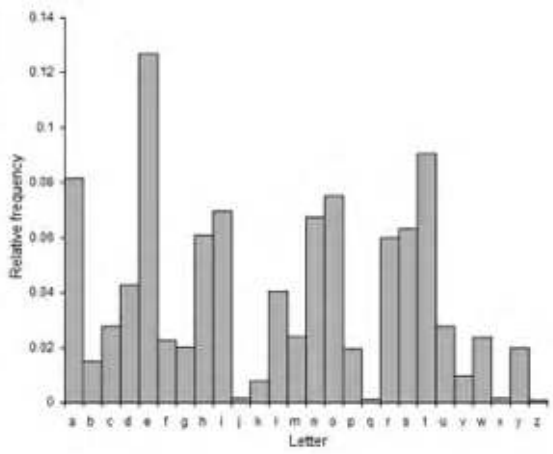

Figure 1: Frequency Analysis for the English language

Caesar Cypher (circa 50 B.C.) – Julius Caesar used this cipher to encode messages to Marcos Cicero: e.g.

$$C \equiv P + 3(\text{mod}26)$$

(any shift – other than multiples of 26! – will do). It's easy to decode:

$$P \equiv C - 3(\text{mod}26)$$

This system is *monoalphabetic*: each letter is always represented using the cipher letter, so it's vulnerable to frequency analysis attacks.

http://www.shodor.org/interactivate/activities/caesar/ provides a good Caesar Cipher applet for you to try....

Example: #1, 2, p. 155

A generalization of the Caesar cipher would be to choose a linear transformation with a slope other than one: in other words,

$$C \equiv aP + b(\text{mod}26)$$

with $\gcd(a, 26) = 1$.

Example: #3, p. 155, shows how to decode one. It's part of your homework!

Example: #4, p. 155

"aE + b = Q" $\quad a \cdot 4 + b \equiv 16 \pmod{26}$

"aT + b = J" $\quad a \cdot 19 + b \equiv 9 \pmod{26}$

$-15a \equiv 7 \pmod{26}$

$a \equiv 3 \pmod{26}$

$a = 3$

$b = 4$

Vigenère Cypher (1586): a one-time key sequence is used, repeated below the message, and the addition is performed character by character on the two strings.

$$C_i \equiv P_i + b_i (\text{mod} 26)$$

It's easy to decode, in blocks of length n, where n is the length of the key:

$$P_i \equiv C_i - b_i (\text{mod} 26)$$

This system is *polyalphabetic*: a letter is generally represented by multiple ciphertext letters, so it's less vulnerable to frequency analysis attacks. However, once the length n of the key is discovered, it becomes n copies of a monoalphabetic cipher, and is vulnerable again.

http://math.ucsd.edu/~crypto/java/EARLYCIPHERS/Vigenere.html is a good website for trying it out.

Of course, the choice of 26 is simply a convenience since we're dealing with the English language. There's nothing particularly special about 26.

Hill's cipher (1929): encrypts **blocks** of letters, rather than letter by letter. Basically, a block is transformed using linear algebra and linear congruences. Recall from section 4.4:

Theorem 4.9: The system of linear congruences

$$C_1 \equiv aP_1 + bP_2 (\text{mod n})$$
$$C_2 \equiv cP_1 + dP_2 (\text{mod n})$$

has a unique solution whenever $\gcd(ad - bc, n) = 1$. The quantity $ad - bc$ is the determinant of the matrix. We can work with larger n x n systems, replacing the quantity $ad - bc$ with the determinant of the n x n matrix.

The code is deciphered by inverting the matrix (just as in linear algebra):

$$P_1 \equiv dC_1 - bC_2 (\text{mod n})$$
$$P_2 \equiv -cC_1 + aC_2 (\text{mod n})$$

(works for a^{-1} if $ad-bc=1$)

http://www.louisville.edu/~ahdeso01/applets/Hill.html is a good website for trying it out.

More to come....

If $ad - bc \neq 1$:

$$\begin{bmatrix} e & f \\ g & h \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} \equiv \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} (\text{mod } 26)$$

M^{-1} M

$$\begin{bmatrix} ea+fc & eb+fd \\ ga+hc & gb+hd \end{bmatrix} \equiv \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} (\text{mod } 26)$$

① $ea + fc \equiv 1 \pmod{26}$

② $eb + fd \equiv 0 \pmod{26}$

$d \cdot (1) - c(2) \Rightarrow \quad e(ad-bc) \equiv d \pmod{26}$

$a \cdot (2) - b(1) \Rightarrow \quad f(ad-bc) \equiv -b \pmod{26}$

Similarly for g + h.

$$(ad-bc)\underbrace{\begin{bmatrix} e & f \\ g & h \end{bmatrix}} \equiv \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} \pmod{26}$$

Solve for the matrix, M^{-1}