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.

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

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

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

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

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

Here's a Caesar Cipher applet for you to try....

Exercises #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,

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

Captain Crunch era

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

Exercises #3, p. 155, shows how to decode one.

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.

Here's 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 that quantity with the determinant of the matrix.

DHSLH CUDSZ XAFOH XZDYL O

The code is deciphered by inverting the matrix:

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

Here's a Good website for trying it out.

More to come.... (RSA)