Secret Codes, Ciphers and Cryptography

> Few persons can be made to believe that it is not quite an easy thing to invent a method of secret writing that shall baffle investigation. Yet it may be roundly asserted that human ingenuity cannot concoct a cipher which human ingenuity cannot resolve.[1]
>
> **Edgar Allan Poe** (American author and poet; 1809 - 1849)

The long history of secret codes was mentioned in the introduction. We have now mathematically arrived at a locale where we can begin to view the mathematics that underlies modern encryption schemes.

In the previous lesson you saw how the Chinese remainder theorem can be used to recover an unknown number from knowledge of specific residues. This gives rise to a **cipher**, an algorithm for encrypting secret information. If you wanted to share secret numbers with a friend, all you would need to do is secretly agree beforehand on the the *key* - a large number which is the product of unique primes. To share a secret number you could send your friend a list of the residues and they could reconstruct the secret number. Even if your list of residues were intercepted by a foe, it would be hard for your secret number to be reconstructed without the identity of the key.

This is a good first example see how a mathematical cipher works. There are practical limitations to this particular method. And there is the enormous limitation that has plagued all ciphers throughout history - the sender and receiver must both have prior knowledge of the key. In any arena where the goal is to keep information secret, how can keys be effectively shared in secret? It

[1] Quoted in "Cryptology: From Caesar Ciphers to Public-key Cryptosystems," by D. Luciano and G. Prichett, in *The College Mathematics Journal*, Vol. 18, No. 1, Jan. 1987, pp. 2-17.

was an enormous breakthrough in the 1970's when a host of computer scientists and mathematicians developed methods of **public key cryptography** where there are two distinct keys, a **private key** which is known by the receiver and a **public key** which is revealed to the world and allows anyone to encrypt messages which can only be decrypted by the holder of the private key.

One of the most important such public key crypotsystems is the *RSA algorithm* named after **Ron Rivest** (American computer scientist; 1947 -), **Adi Shamir** (Israeli computer scientist; 1952 -), and **Leonard Adleman** (American computer scientist and biologist; 1945 -). As you shall see, the underlying mechanism that drives this method is *Fermat's little theorem* - the result you found previously.

To employ the RSA algorithm the Receiver must build the keys. This is done as follows:

- Two very large primes p and q are chosen.
- The product $n = p \cdot q$ is computed. The number n is one part of the public key.
- An exponent e is chosen so that $1 < e < (p-1)(q-1)$ and e shares no common factors with $(p-1)(q-1)$. The number e is the other part of the public key.
- The equation $k \cdot e \equiv 1 \pmod{(p-1)(q-1)}$ is solved for k. The number k is the private key.
- The public key (n,e) is made public for all to see.

With the keys identified, but before we describe the workings of the cipher, it is important to understand the essential security matter - factoring large numbers into primes. The primes p and q are generally chosen to have about 200 digits and so n has about 400 digits. This number is part of the public key, made public for all of the world to see!

1. Suppose a foe with knowledge of the public key was able to factor n into its prime factors. Describe why this would enable the foe to determine the private key and break any message within this particular RSA scheme.

This is a prime example of why our understanding of primes is so important. And why the incredible abilities of The Twins might have helped uncover one of the great mysteries in all of mathematics.

2. A Sender needs to translate their message text into a number so it can be encrypted before it is sent. After decryption the Receiver will need to translate the number which is output by this algorithm back into the message text. Determine a way that you could represent any alphabetic, text based message as a single number in such a way that the message can be easily reconstructed from the number.

To send a secret message a Sender simply converts the text message into a single number, as in Investigation **2**. Denote this number by m for "message." The Sender then computes:
$c \equiv m^e \pmod{n}$.

The encrypted ciphertext, denoted by c, can then be sent as a *public message*, as it can only be decrypted by the Receiver who is the only person in possession of the private key. The Receiver decrypts the message by computing:

$$c^k \bmod n.$$

But why does this work? Why does this recover the secret message m?

3. Since k is defined to satisfy $k \cdot e \equiv 1 \pmod{(p-1)(q-1)}$, explain why we can write $k \cdot e - 1 = j(p-1)(q-1)$ for some integer j.

4. Explain why the number computed by the Receiver, $c^k \bmod n$, is equal to
$m^{e \cdot k} \bmod n$.

5. Explain why $m^{e \cdot k} = m^{e \cdot k - 1} \cdot m$.

We are now ready to invoke Fermat's little theorem, which will be done both mod$p$ and mod$q$ and then combined.

6. Explain why
$m^{e \cdot k} = \left(m^{p-1}\right)^{j(q-1)} \cdot m.$

7. Modulo p this means
$m^{e \cdot k} \equiv \left(m^{p-1}\right)^{j(q-1)} \cdot m \pmod{p}$.

Use Fermat's little theorem to explain why this expression on the right is congruent to m.

8. Explain why $m^{e \cdot k} - m$ is a multiple of p.

9. Explain why $m^{e \cdot k} - m$ must also be a multiple of q.

10. Since $m^{e \cdot k} - m$ is a multiple of both of the different primes p and q, explain why it must be a multiple of the product $p \cdot q$.

11. Determine the number $m^{e \cdot k} \bmod n$ and explain why this means that the Receiver has recovered the secret message.

So this is some slightly technical algebra. But consider the enormity of what you have just shown. You have just rediscovered the inner workings a special application of a 300 year-old, theoretical result about whole numbers - one that has revolutionized secrecy and that is fundamental to the information age. Unlike the cute application of algebra in "chocolate math," in RSA we have an algorithm that, together with related advances in public key encryption methods, is fundamental to *all* of e-commerce, computer security, weaponry codes, and secure communication. The linchpin to all of this? As yet we have found no efficient methods for factoring large numbers or finding the pattern to the primes - the key builders in encryption can stay far ahead of the numerical lock pickers.

For most of the history of cryptography, Poe's declaration that opened this section seemed valid. But human ingenuity combined with powerful mathematics has managed to create ciphers which cannot be resolved.