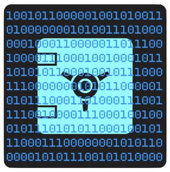
An Introduction to Mathematical Cryptography

James Guo

January 18th, 2024

What is Cryptography:



Cryptography is a subject under mathematics involving encrypting and decrypting codes. The goal of cryptography is to encode your message with someone so that others cannot understand.

A Story about Mary Queen of Scots: [1]

Figure of Queen Mary

Queen Mary was a queen of Scotland, and she was put on trial. From a future perspective, we know that she was a part of the conspirators, who wished to replace Queen Elizabeth (who is a Protestant) with Queen Mary (who is a fellow Catholic).

During this trial, Queen Mary was accused of treason. At the time, she had no secretaries to help her prepare her case. Queen Elizabeth was also contradictory on whether they should put Queen Mary to death due to her status, i.e., they needed more evidence.

What became the key was the correspondence between Queen Mary and the conspirators. They were all written in ciphers, meaning that they seemed like meaningless words unless someone could translate them into evidence that she committed treason.

Discussion:
Queen Mary's story ended with the result that Queen Mary's ciphers were cracked, announcing her death trail later. Meanwhile, this story also leads us towards more questions about cryptography: **Should cryptography (possibly part of privacy) be encouraged or better monitored to prevent it from being used for illegal things?**

Mathematical Foundations:

At this moment, you should probably be familiar with Fermat's Little Theorem, but now shall we introduce a stronger theorem, known as Euler's Theorem:

Euler's Theorem:
Let $m \in \mathbb{Z}^+$ and $a \in \mathbb{Z}$. If $\gcd(a, m) = 1$, then $a^{\phi(m)} \equiv 1 \pmod{m}$, where $\phi(m)$ is defined by the number of elements in $\{n \in \mathbb{Z} \colon 0 \leq n \leq m-1, \gcd(n, m) = 1\}$.

The **Fermat's Little Theorem** could be considered as a corollary from **Euler's Theorem**, in which the reader could figure this out by themselves (cf. Exercise 1).

Therefore, the proof for **Euler's Theorem** is provided, as follows:

[1] Source: *The Code Book - The Science of Secrecy from Ancient Egypt to Quantum Cryptography,* Simon Singh, 1999.

Proof. First, we may denote the elements in $\{n \in \mathbb{Z} \colon 0 \leq n \leq m-1, \gcd(n,m) = 1\}$ as $\{r_1, r_2, \cdots, r_{\phi(m)}\}$. We want to show that we can form a bijection between this set and $\{ar_1, ar_2, \cdots, ar_{\phi(m)}\}$, meaning they are one-to-one related by modulo m.
Clearly, we can conclude the following results:

- By $\gcd(r_i, m) = 1$ and $\gcd(a, m) = 1$, we know that $\gcd(ar_i, m) = 1$.
- $0 \leq r_i \leq m-1$ indicates that $r_i \not\equiv r_j \pmod{m}$. With $\gcd(a,m) = 1$, we know that $ar_i \not\equiv ar_j \pmod{m}$.
- For any $k \in \mathbb{Z}$ such that $\gcd(k,m) = 1$, with $\bar{a}$ be the inverse of a, we know that there exists some r_i such that $\bar{a}k \equiv r_i \pmod{m}$. This means that $k \equiv a\bar{a}k \equiv ar_i \pmod{m}$, meaning that any ar_i can be corresponded back to some $k \in \{r_1, r_2, \cdots, r_{\phi(m)}\}$.

With these we formed our bijection, which means that for all ar_i, it is equivalent to one unique r_j, who is also uniquely equivalent to ar_i modulo m, allowing us to make the following equivalence under modulo m:

$$\prod_{i=1}^{\phi(m)} r_i \equiv \prod_{i=1}^{\phi(m)} (ar_i) \equiv a^{\phi(m)} \prod_{i=1}^{\phi(m)} r_i \pmod{m}.$$

Since $\gcd(r_i, m) = 1$ for all r_i, this means that $\gcd\left(\prod_{i=1}^{\phi(m)} r_i, m\right) = 1$, which indicates that $a^{\phi(m)} = 1$. □

Prepositions on Euler's ϕ Function:
Let p be a prime, n as an integer, then

$$\phi(p^n) = p^n - p^{n-1}.$$

Let m and n be integers such that $\gcd(m,n) = 1$, then

$$\phi(mn) = \phi(m)\phi(n).$$

Then, with these two functions, we would be able to calculate the Euler's ϕ function for all positive integers by the **Fundamental Theorem of Arithmetics**.

Given the limits of length, we hereby provide a sketch of proof:

Sketch of Proof. For the first equality, it is not hard to find $\{1, 2, \cdots, p^n\} \backslash \{p, 2p, \cdots, p^{n-1}p\}$ are all the integers fulfilling the requirements for the set.
For the second equality, with $\{r_1, r_2, \cdots, r_{\phi(m)}\}$ and $\{s_1, s_2, \cdots, s_{\phi(n)}\}$ being the two sets for all co-prime residues under modulo m and modulo m, we want to show that all $nr_i + ms_j$ for $1 \leq i \leq \phi(m)$ and $1 \leq j \leq \phi(n)$ forms a set for all co-prime residues under modulo mn, which fulfills:

1. $\gcd(nr_i + ms_j, mn) = 1$ for all $1 \leq i \leq \phi(m)$ and $1 \leq j \leq \phi(n)$;
2. $nr_i + ms_j \not\equiv nr_k + ms_l$ whenever $(i,j) \neq (k,l)$;
3. For any $t \in \mathbb{Z}$ such that $\gcd(t, mn) = 1$, there exists (r_i, s_j) such that $t \equiv nr_i + ms_j \pmod{mn}$.

The rest of the proofs for these prepositions are left as an exercise for the diligent readers (cf. Exercise 2).

You might have found the above proofs lengthy. With the developments of Modern Algebra (or Abstract Algebra), one can prove it in a more canonical approach. Even without sufficient context, some smart readers might have noticed that the multiplications forms a "cycle", noted as a **Cyclic group** in **Group theory**.

Group Theory (Brief Introduction):
One could consider a set G and a binary operator $\bullet_G : G \to G$ a **group** if it satisfies the following:

1. $\bullet_G$ is associative: $(\forall g, h, k \in G): \ (g \bullet_G h) \bullet_G k = g \bullet_G (h \bullet_G k)$;
2. There exists an identity element e_G: $(\exists e_G \in G)(\forall g \in G): \ g \bullet_G e_G = g = e_G \bullet_G g$;
3. Every element has an inverse with respect to $\bullet_G$: $(\forall g \in G)(\exists h \in G): \ g \bullet_G h = e_G = h \bullet_G g$.

Furthermore, **Group Theory** can be understood by **Category Theory**, which is more generalized.

Without additional contexts, we may consider G consists powers of a modulo m (or a class containing all integers of same remainder modulo m) with multiplications as operator, denoted $(G, \cdot)$, as follows:

$$\begin{tikzcd}
a \arrow[r, "\cdot a"] & a^2 \arrow[r, "\cdot a"] & \cdots \arrow[r, "\cdot a"] & a^{\phi(m)-1} \arrow[r, "\cdot a"] & a^{\phi(m)} \equiv 1 \arrow[llll, bend left, "\cdot a"]
\end{tikzcd}$$

Notation-wise, this can be considered as a **Cyclic group** generated by a single element a. There are more discussions on order and index of groups. This is a very brief example of **Group Theory** from Abstract Algebra. If you found this interesting, consider taking some college level courses on such topic.

Public RSA:

Back to cryptography, what we want the most is to find some algorithms where the code is securely encrypted while decrypting them is not too complicated. Furthermore, what **Public RSA** aims for is to provide the way to encrypt to the public while having a secret way to encrypt it.

This must sound counter-intuitive for many of you, as we have learned a lot about inverse functions. For many elementary functions, there exists an inverse function that 'undo' the operation. While for some others, the 'undo' process is as hard for the decrypts even if people know everything about the operation.

Eventually, modular arithmetic came to the view of people wishing to secure their message.

Public RSA Algorithms:
Utilizing **Euler's Theorem**, we can develop the following algorithms with public keys (n, r) such that n is the product of two prime numbers p_1 and p_2, while $\gcd(\phi(n), r) = 1$. The original message could be an integer M. *Note that this integer M should not be greater than n, else would cause not one-to-one mapping.*

$$\begin{tikzcd}
\text{Message Encoder} & M \arrow[r] & M^r \pmod{n} \arrow[r] & E \arrow[ddr] & & \\
\text{Public} & (n,r) \arrow[ur] \arrow[dr] & & & & \\
\text{Message Decoder} & n = p_1 \times p_2 \arrow[u] & r^{-1} \pmod{\phi(n)} \arrow[r] & s \arrow[r] & E^s \pmod{n} \arrow[r] & M \\
& (p_1, p_2) \arrow[u] \arrow[r] & \phi(n) = (p_1 - 1)(p_2 - 1) \arrow[u] & & &
\end{tikzcd}$$

$*$ Also note that (mod n) is really a notation to find the *remainder* of the division, always represented by `%` in Computer Science syntax.

You might now be wondering why this is secure. The key is on factoring a composite number. Seemingly straightforward, there is not yet a sufficient algorithm to factor. In reality, banking services use this **Public RSA** systems. When the primes are thousands digits long, the complexity to crack is too significant.

Number of Primes:
People might ask: Are there infinitely many prime numbers? There is an exquisite proof by contradiction.

Proof. Assume that there exists a finite number of primes, denoted $\{2, 3, 5, \cdots, p_n\}$. We know that $2 \times 3 \times 5 \times \cdots \times p_n + 1$ is not divisible by any of the prime numbers, which is a contradiction. Hence, there does not exist the largest prime, meaning there are infinitely many of them. □

With such foundations, this is one of the best solutions for **Public RSA**. For simplicity, here is an example with small primes demonstration the encrypting and decrypting 30 using public key $(221, 101)$:

$$\begin{tikzcd}
 & \text{Message: } 30 \arrow[r] & 30^{101} \pmod{221} \equiv 166 \arrow[r] & 166 \arrow[dd] \\
 & \text{Key: } (221, 101) \arrow[ur] & & \\
\text{Primes: } (13, 17) \arrow[r] \arrow[dr] & 13 \times 17 = 221 \arrow[u] & 101^{-1} \pmod{192} \equiv 173 \arrow[r] & 166^{173} \pmod{221} = \boxed{30} \\
 & \phi(221) = 12 \times 16 = 192 \arrow[ur] & &
\end{tikzcd}$$

Future about Public RSA:
Currently, there are no efficient algorithms to factor prime. However, with potentials of quantum computing and newly developed algorithms, this version of **Public RSA** is not necessarily safe. It is important to develop new "one-way" functions to secure the message encryption in the future. Meanwhile, the advancements in solving these problems do push mathematics and other areas of science to newer generations.

Exercises:

1. Prove Fermat's Little Theorem by **Euler's Theorem**.

   Remark: Fermat's Little Theorem is that *if p is a prime and $a \in \mathbb{Z}$, then* $a^p \equiv a \pmod{p}$.

2. Prove the prepositions on **Euler's ϕ Function** using the sketch of proof.

3. A public key is $(239812014798221, 103)$, in which n can be factored as 15485863 and 15485867 that are two prime numbers. Given an encrypted message is 216642813890413, find the original message.

   Remark: You should use Wolfram Alpha, or other equivalent tools, for the scope of calculation.

4. Prove that there exist infinitely many primes congruent to 5 modulo 6 based off the proof that there are infinitely many primes.