Quick Review of Operations in Modular Arithmetic

$$\begin{aligned}
\Big[(a \bmod n) + (b \bmod n)\Big] \bmod n &= (a+b) \bmod n \\
\Big[(a \bmod n) - (b \bmod n)\Big] \bmod n &= (a-b) \bmod n \\
\Big[(a \bmod n)(b \bmod n)\Big] \bmod n &= (ab) \bmod n \\
(x^a \bmod n)^b \bmod n &= x^{ab} \bmod n
\end{aligned}$$

$$\begin{aligned}
x^{a+b+c+\cdots+k} \bmod n &= (x^a x^b \cdots x^k) \bmod n \\
&= \Big[(x^a \bmod n)(x^b \bmod n)\cdots(x^k \bmod n)\Big] \bmod n
\end{aligned}$$

Residues and Primitive Elements

- Start with a ring $\mathbf{Z}_n$.

Definition: $\mathbf{Z}_n^*$ is the set of residues modulo n that are relatively prime to n (prime residue group). That is, $\mathbf{Z}_n^* = \{a : (a,n) = 1, a \in \mathbf{Z}_n\}$.

- Note: $\mathbf{Z}_n^*$ does not contain zero.

- $\mathbf{Z}_n^*$ forms an abelian group under multiplication. (Proof to follow)

- For $n = p$, p a prime, $|\mathbf{Z}_p^*| = p - 1$ and $\mathbf{Z}_p^*$ is cyclic. (Proof to follow)

Claim: $\mathbf{Z}_n^*$ is an abelian group under multiplication.

Proof: That $\mathbf{Z}_n^*$ is abelian follows from the commutativity of the integers under multiplication. We now prove that $\mathbf{Z}_n^*$ is a group under multiplication.

1. Closure – for $a, b \in \mathbf{Z}_n^*$, $(a,n) = (b,n) = 1$. Clearly, $(ab, n) = 1$ and $ab \in \mathbf{Z}_n^*$.

2. Associativity – follows from the integers under multiplication.

3. Identity – We know that 1 is the multiplicative identity for the integers, and $1 \in \mathbf{Z}_n^*$, since $(1, n) = 1$.

4. Inverses – This requires the proof of the following theorem:

   Theorem 1.1: The equivalence $ax = b \bmod n$ has a unique solution $x \in \mathbf{Z}_n$ for every $b \in \mathbf{Z}_n$ if and only if $(a, n) = 1$.

   Proof: First assume $(a, n) = 1$.

   We first show existence of a solution x:

   - Fix $b \in \mathbf{Z}_n$
   - Apply the following theorem:

     Theorem 1.2: For integers a, n, $(a, n) = 1$ if and only if $\exists\ r, s \in \mathbf{Z}$ such that $ar + ns = 1$. (We will not prove this).
   - Since $(a, n) = 1$, we can write $ar + ns = 1$ by this theorem.
   - Thus $arb + nsb = b$, and $(arb + nsb) \bmod n = b \bmod n$
   - This implies that $(arb \bmod n) + (nsb \bmod n) = b \bmod n$ and $arb = b \bmod n$.
   - Thus, $x = rb$ is a solution of $ax = b \bmod n$.

   Now suppose $ax = b \bmod n$ has more than one solution (contradiction).

   - Thus, $\exists$ distinct $x_1, x_2 \in \mathbf{Z}_n$ with $ax_1 = b \bmod n$ and $ax_2 = b \bmod n$.
   - So $\exists\ r, s \in \mathbf{Z}$ with $ax_1 = rn + b$ and $ax_2 = sn + b$.

$$\begin{aligned} ax_1 - rn &= ax_2 - sn \\ (r - s)n &= (x_1 - x_2)a \end{aligned}$$

- Since $(a, n) = 1$, all the factors of n must be found in $x_1 - x_2$, which implies that n divides $(x_1 - x_2)$.
- So $x_1 - x_2 = 0 \bmod n \implies x_1 = x_2 \bmod n$ (contradiction)
- Thus, there is a unique solution to $ax = b \bmod n$.

Now assume $ax = b \bmod n$ has a unique solution.

- Write $ax = ns + b$ for some $s \in \mathbf{Z}$. Thus $ax - ns = b$.
- Consider the case $b = 1$, and $ax - ns = 1$. We apply the following theorem:

  Theorem 1.2: For integers a, n, $(a, n) = 1$ if and only if $\exists\ r, s \in \mathbf{Z}$ such that $ar + ns = 1$.
- Since x, s are integers, we conclude that $(a, n) = 1$.

For the case $b = 1$, $ax = 1 \bmod n$ implies $\exists!\ x \in \mathbf{Z}_n$ where x is the inverse of a modulo n. It now remains to show that the inverse of a, denoted x, is in $\mathbf{Z}_n^*$.

- Consider $ax = 1 \bmod n$

$$\begin{aligned} ax &= 1 \bmod n \\ ax &= ns + 1, \text{ for some integer } s \\ ax - ns &= 1 \end{aligned}$$

- Apply Theorem 1.2 to conclude that $(x, n) = 1$.
- Thus $x \in \mathbf{Z}_n^*$ by definition.

Therefore, a has a multiplicative inverse in $\mathbf{Z}_n^*$.

Thus, $\mathbf{Z}_n^*$ is a group under multiplication.

Consider the group $\mathbf{Z}_p^*$ for a prime p. Clearly, $\mathbf{Z}_p^* = \{1, 2, \ldots, p-1\}$. We now prove $\mathbf{Z}_p^*$ is cyclic.

Case 1: Suppose $|Z_p^*| = p - 1 = q^s$ for some prime q.

- Since Z_p^* is a commutative ring with unity and is also a group under multiplication, we know that every element in Z_p^* is a unit. This implies that Z_p^* is a field.

- Consider the element $x^n - 1 \in Z_p[x]$. By unique factorization in fields, we know that $x^n - 1$ can be factored into at most n linear factors. Thus, there are at most n solutions to $x^n = 1$.

- Consider an element $a \in Z_p^*$ with largest possible order q^r.

- The elements $1, a, a^2, \ldots, a^{q^r-1}$ are all distinct, and each element solves the equation $x^{q^r} = 1$.

- Since there are at most q^r solutions to $x^{q^r} = 1$, and there are q^r such solutions in the set $\{1, a, a^2, \ldots, a^{q^r-1}\}$, there are no other solutions to $x^{q^r} = 1$.

- Let $b \in Z_p^*$ be arbitrary with $|b| = q^t$ where $t \leq r$. We see that $b^{q^r} = (b^{q^t})^{q^{r-t}} = (1)^{q^{r-t}} = 1$.

- Since b solves the equation $x^{q^r} = 1$, $b = a^i$ for some i.

- Thus, Z_p^* is cyclic.

Case 2: Let $|Z_p^*| = q_1^{s_1} \cdots q_k^{s_k}$ where the q_i are distinct primes.

- Since the group Z_p^* is abelian, all its Sylow subgroups $S_{q_1}, \ldots, S_{q_k}$ are normal. Also, since the q_i are distinct primes, $S_{q_i} \cap S_{q_j} = \{e\}$ for $i \neq j$.

- Since the Sylow subgroups are normal and have trivial intersection, we can write $Z_p^* \approx S_{q_1} \times \cdots \times S_{q_k}$.

- From Case (1), each S_{q_i} is cyclic; thus, $Z_p^* \approx Z_{q_1} \times \cdots \times Z_{q_k}$.

- Since all the q_i are relatively prime, the group $S_{q_1} \times \cdots \times S_{q_k}$ is cyclic, with generator $b = (b_1, b_2, \ldots, b_k)$ for generators b_i of S_{q_i}.

- Thus, the group Z_p^* is cyclic.

Since Z_p^* is cyclic, we can now define the following:

Definition: An element α that generates $\mathbf{Z}_p^*$ is a **primitive element** (root) modulo p.

Discrete Logarithms

- Let α be a primitive element for a prime p. Thus, α generates $\mathbf{Z}_p^*$.
- Let $b \in \mathbf{Z}$. Then we can write $b = r \bmod p$ for $0 \leq r \leq p - 1$.
- Since $r \in \mathbf{Z}_p^*$, there exists a unique i such that $b = \alpha^i \bmod p$ for $0 \leq i \leq p - 2$.

Definition: We define i to be the index of b for the base $\alpha \pmod{p}$.

- The index is denoted as $i = \mathrm{ind}_{\alpha,p}(b)$ or $i = \log_\alpha(b)$.
- The indices are often referred to as discrete logarithms, since they resemble logarithms both in definition and in operation.

Diffie-Hellman Key Exchange

Background

Diffie-Hellman Key Exchange was developed by Whitfield Diffie and Martin Hellman at Stanford in 1976. Their paper describing the idea was the first published public-key technique.

Setting Up the System

1. First, one user chooses two public elements: a prime p, and a primitive element of p, denoted α.

2. Alice determines two new elements using p and α:

   - She selects $X_A \in \mathbf{Z}$ with $X_A < p$.
   - She calculates Y_A by $Y_A = \alpha^{X_A} \bmod p$.
   - X_A remains private, while Y_A is public.

3. Similarly, Bob determines X_B and Y_B:

   - $X_B \in \mathbf{Z}$ with $X_B < p$
   - $Y_B = \alpha^{X_B} \bmod p$
   - X_B is private and Y_B is public.

4. They can then each generate the key, K:

   - Alice calculates $K = (Y_B)^{X_A} \bmod p$
   - Bob calculates $K = (Y_A)^{X_B} \bmod p$

5. They now each have the same key with which to encrypt messages.

Verification of the Equality of the Keys

Taking the key calculated by Alice and employing operations in modular arithmetic, we can write:

$$\begin{aligned}
K &= (Y_B)^{X_A} \bmod p \\
&= (\alpha^{X_B} \bmod p)^{X_A} \bmod p \\
&= \alpha^{X_B X_A} \bmod p \\
&= \alpha^{X_A X_B} \bmod p \\
&= (\alpha^{X_A} \bmod p)^{X_B} \bmod p \\
&= (Y_A)^{X_B} \bmod p
\end{aligned}$$

Thus, the keys generated by Alice and Bob are equal.

Simple Example

1. Suppose we choose $p = 17$ and $\alpha = 3$. We can verify that α is a primitive element modulo p by determining that α generates the group $\mathbf{Z}_p^*$.

2. Suppose Alice chooses $X_A = 7$ and calculates $Y_A = (3)^7 \bmod 17 = 11$.

3. Suppose Bob chooses $X_B = 12$ and calculates $Y_B = (3)^{12} \bmod 17 = 4$.

4. The elements $p = 17, \alpha = 3, Y_A = 11$ and $Y_B = 4$ are all public; thus, Alice obtains Bob's Y_B value and Bob obtains Alice's Y_A value.

5. To generate the key:

   - Alice calculates $K = 4^7 \bmod 17 = 13$.

   - Bob calculates $K = 11^{12} \bmod 17 = 13$.

6. They can now use the key 13 to send an encrypted message.

Why is this key generation hard to break?

- The only public elements an opponent has to work with are p, α, Y_A, Y_B. To determine the key, the opponent needs to calculate either X_A or X_B. Since $Y_A = \alpha^{X_A} \bmod p$, we have that $X_A = \text{ind}_{\alpha,p}(Y_A)$. Thus, the opponent must calculate a discrete logarithm, which is computationally difficult. Currently, the fastest known algorithm to do so is on the order of $e^{((\ln\ p)^{1/3}\ln(\ln\ p))^{2/3}}$, which becomes infeasible for large primes.

- The original algorithm published in 1976 is vulnerable only to a man-in-the-middle attack as follows:

  Suppose Alice was sending her Y_A value to Bob. An opponent, Eve, intercepts this message and substitutes her own value for Y_A, which she then sends on to Bob. Since Bob has no way of verifying that the message actually came from Alice, he unsuspectingly uses Eve's value to calculate his key. Diffie, along with several others, published a revised version of his original algorithm in 1992 which defeats this attack by introducing user authentication.

- Today, the algorithm is used in systems such as Virtual Private Networks as the first stage in a method of encryption. The public keys calculated by Diffie-Hellman are slow at encryption, whereas private keys encrypt large blocks of text quickly. Thus, two users would first calculate a public key using Diffie-Hellman, and then use that key to encrypt a private key which they would then exchange. This private key would then be used to encrypt large blocks of data.

The Euclidean Algorithms

Theorem 1.1: The equivalence $ax = b \bmod n$ has a unique solution $x \in \mathbf{Z}_n$ for every $b \in \mathbf{Z}_n$ if and only if $(a, n) = 1$. (previously proven)

Corollary: Following from the previous theorem, $(a, n) = 1$ if and only if $\exists\, a^{-1} \in \mathbf{Z}_n$ such that $a \cdot a^{-1} = 1 \bmod n$. That is, a has an inverse in $\mathbf{Z}_n$ if and only if $(a, n) = 1$.

- $a \in \mathbf{Z}_n^*$ has a multiplicative inverse, since $(a, n) = 1$.
- Also, for integers m, n, $\exists\, c, d \in \mathbf{Z}$ such that $(m, n) = cm + dn$.

  Suppose $(m, n) = 1$

$$\begin{aligned} 1 \bmod n &= (cm + dn) \bmod n \\ &= (cm \bmod n) + (dn \bmod n) \\ &= cm \bmod n \\ &= cm \end{aligned}$$

- Note: c is the multiplicative inverse of m modulo n.
- The **Euclidean Algorithm** calculates the gcd of any two integers using the Division Algorithm recursively.
- The **Extended Euclidean Algorithm** calculates the gcd, as well as the values of c and d as defined above for any two integers.

The Extended Euclidean Algorithm

Given $m, n \in \mathbf{Z}$, with $m > n$, we define the following elements recursively:

$$a_0 = m, \quad a_1 = n, \quad q_k = \lfloor \tfrac{a_{k-1}}{a_k} \rfloor$$

$$a_k = a_{k-2} - q_{k-1} \cdot a_{k-1}$$

$$x_0 = 1, \quad x_1 = 0, \quad y_0 = 0, \quad y_1 = 1$$

$$x_k = x_{k-2} - q_{k-1} \cdot x_{k-1}, \quad y_k = y_{k-2} - q_{k-1} \cdot y_{k-1}$$

Then, for every step k of the recursive process, $a_k = x_k a_0 + y_k a_1$.

- This process ends when $a_k = 0$ for some k.
- Then, $a_{k-1} = (m, n)$ (Euclidean Algorithm)
- Also, $(m, n) = x_{k-1}m + y_{k-1}n$ (Extended Euclidean Algorithm)

Proof of the Extended Euclidean Algorithm

We want to prove $a_k = x_k a_0 + y_k a_1$, and we do this by induction on k:

- Let $k = 0$. Then,

$$\begin{aligned} x_0 a_0 + y_0 a_1 &= 1 \cdot a_0 + 0 \cdot a_1 \\ &= a_0 \end{aligned}$$

  Thus, the statement holds for $k = 0$.

- Now assume the statement holds for all k with $0 \leq k \leq s$. Consider the following:

$$\begin{aligned} x_{s+1} a_0 + y_{s+1} a_1 &= (x_{s-1} - q_s \cdot x_s) a_0 + (y_{s-1} - q_s \cdot y_s) a_1, \\ &= x_{s-1} a_0 - q_s x_s a_0 + y_{s-1} a_1 - q_s y_s a_1, \\ &= (x_{s-1} a_0 + y_{s-1} a_1) - q_s (x_s a_0 + y_s a_1), \\ &= a_{s-1} - q_s a_s, \text{ by definition,} \\ &= a_{s+1} \text{ by definition.} \end{aligned}$$

Thus, the statement holds for $k = s + 1$.

By induction, we conclude that $a_k = x_k a_0 + y_k a_1$ for all k.

An Example Using the Extended Euclidean Algorithm

Suppose we wish to find the multiplicative inverse of 25 mod 48. Note that $(25, 48) = 1$, which allows us to find the inverse using the Extended Euclidean Algorithm.

- Let $a_0 = 48$ and $a_1 = 25$.
- Calculate $q_1 = \lfloor \frac{48}{25} \rfloor = 1$. Then $a_2 = a_0 - q_1 a_1 = 48 - 1(25) = 23$.
- The next step would then calculate $q_2 = \lfloor \frac{25}{23} \rfloor = 1$. Then $a_3 = 25 - 1(23) = 2$.
- The next step calculates $q_3 = \lfloor \frac{23}{2} \rfloor = 11$. Then $a_4 = 23 - 11(2) = 1$.
- Since we know that $(25, 48) = 1$, the process ends here. We now determine the x and y values.
- Define $x_0 = 1, x_1 = 0, y_0 = 0, y_1 = 1$.
- Calculate $x_2 = x_0 - q_1 x_1 = 1 - 1(0) = 1$ and $y_2 = y_0 - q_1 y_1 = 0 - 1(1) = -1$.
- The next step calculates $x_3 = 0 - 1(1) = -1$ and $y_3 = 1 - 1(-1) = 2$.
- The next step calculates $x_4 = 1 - 11(-1) = 12$ and $y_4 = -1 - 11(2) = 23$. We stop here.
- Now we can write $a_4 = x_4 a_0 + y_4 a_1$. Thus, $1 = 12(48) + (-23)(25)$.
- This implies that 1 mod 48 $= \big(12(48) \bmod 48\big) + \big((-23)(25) \bmod 48\big)$.
- Thus, 1 mod48 $= (-23)(25)$ mod 48, and -23, which is the same as 25 mod 48, is the multiplicative inverse of 25 modulo 48.