COMS W4261: Introduction to Cryptography.
Instructor: Prof. Tal Malkin

Number Theory Recitation Handout
Edo Roth

Definitions

A **group** is a set G with binary operation * that satisfies the following 4 properties:

- Closure: If $a, b \in G$, then $a * b \in G$
- Associativity: $(a * b) * c = a * (b * c) \ \ \forall \ \ a, b, c \in G$
- Identity: $\exists e \in G$ s.t. $a * e = e * a \ \ \forall \ \ a \in G$
- Inverse: $\forall a \in G \ \ \exists b \in G$ s.t. $a * b = b * a = e$

Note that commutativity does not generally hold! (commutative groups are called abelian)

The **order** of a group is the number of elements in the group.

Let $+$ = addition as normally defined, $\times$ = multiplication as normally defined.
Examples: $(\mathbb{Z}, +)$, (Even $\mathbb{Z}, +$)
Non-Examples: $(\mathbb{Z}, \times)$, (Odd $\mathbb{Z}, +$), (Non-negative $\mathbb{Z}, +$)

$\mathbb{Z}_n = \{0, 1, 2, \ldots, n-1\}$ with the operation $+ \pmod n$ is always a group.
$\mathbb{Z}_n^* = \{a \in [1, N-1] | gcd(a, n) = 1\}$ with the operation $\times \pmod n$ is always a group.

In particular, we often work with $\mathbb{Z}_p^* = \{1, 2, \ldots, p-1\}$.

G' is a **subgroup** of G if:

- $G' \subseteq G$
- G' and G share an operation
- G' is a group

Example: Subgroups of $\mathbb{Z}_5^* = (\{1, 2, 3, 4\}, \times)$: $\{1, 2, 3, 4\}, \{1, 4\}, \{1\}$

Greatest Common Divisor (GCD): $gcd(a, b)$ is defined as the largest integer d s.t. $d|a$ and $d|b$. The Euclidean algorithm gives us an efficient way to calculate the gcd.

Thm: For all a, b, $\exists x, n \in \mathbb{Z}$ s.t $xa + nb = gcd(a, b)$
$\Rightarrow$ If a and b are relatively prime ($gcd(a, b) = 1$), then $xa + nb = 1 \Rightarrow a \equiv x^{-1} (mod \ \ n)$
$\Rightarrow a$ is an inverse to x modulo n!

Euclidean Algorithm Example: Find the inverse of 8 mod 11.

$$\begin{aligned}
11 &= 8 + 3 \qquad & 3 &= 11 - 8 \\
8 &= 2 \cdot 3 + 2 \qquad & 2 &= 8 - 2 \cdot 3 \\
3 &= 2 + 1 \qquad & 1 &= 3 - 2 \\
2 &= 2 \cdot 1 \\
&\Rightarrow \\
1 &= 3 - 2 \cdot 1 \\
&= 3 - (8 - 2 \cdot 3) = 3 \cdot 3 - 8 \\
&= 3(11 - 8) - 8 \\
&= 3 \cdot 11 - 4 \cdot 8
\end{aligned}$$

$\Rightarrow -4$ (or 7) is inverse of 8 modulo 11 – confirm that $8 \cdot 7 \equiv 56 \equiv 1 (\text{mod} 11)$.

Modular Arithmetic: $x \equiv y \pmod{n}$ iff $n|x - y$.

$$\begin{aligned}
x \equiv y \pmod{n} \Rightarrow \; & x + a \equiv y + a \pmod{n} \\
& x - a \equiv y - a \pmod{n} \\
& x \cdot a \equiv y \cdot a \pmod{n}
\end{aligned}$$

Note that this does not (in general) hold for division!
(e.g. $7 \cdot 3 = 4 \cdot 3 (\text{mod} 9)$ but $7 \not\equiv 4 (\text{mod} 9)$)

Euler number: $\phi(n) = |\mathbb{Z}_n^*|$. We see this implies $\phi(p) = p - 1$ for all primes p, and that $\phi(p^\alpha) = p^{\alpha-1}(p-1)$ (since all numbers besides multiples of p are relatively prime to p^α).

We also know that $\phi(mn) = \phi(m)\phi(n)$ if $gcd(m, n) = 1$.

Fermat's Little Theorem: $x^{\phi(n)} \equiv 1 \pmod{n}$ for $gcd(x, n) = 1$. This can be used to perform large exponential calculations very efficiently.
Example:

$$\begin{aligned}
1000^{77} (\text{mod} 13) \equiv 1000^{72+5} (\text{mod} 13) &\equiv 1000^5 (\text{mod} 13) \\
&\equiv 90^5 (\text{mod} 13) \\
&\equiv (-1)^5 (\text{mod} 13) \\
&\equiv -1 (\text{mod} 13) \\
&\equiv \mathbf{12 (mod 13)}
\end{aligned}$$

A **cyclic** group is a group that can be generated by a single element.
Example: $\mathbb{Z}_{13}^* = \{2, 4, 8, 3, 6, 12, 11, 9, 5, 10, 7, 1\}$ is generated by 2.

For *all* primes, in fact, $\mathbb{Z}_p^*$ is cyclic. We say that $x \in \mathbb{Z}_p^*$ is a quadratic residue (QR) if it has a square root in $\mathbb{Z}_p^*$. More specifically, we define:

$QR_n = \{x \in \mathbb{Z}_n^* | \exists w \text{ s.t. } x = w^2 \pmod{n}\}$
$QR_p = \langle g^2 \rangle$ for prime p

$\mathbb{Z}_p^*$ has a subgroup $\langle g^2 \rangle$, with $|\langle g^2 \rangle| = \frac{p-1}{2}$. If this size is prime, then p is known as a **safe prime**, and DDH believed to be hard.

Example of $\langle g^2 \rangle$: Look at $\mathbb{Z}_{13}^*$ again (not a safe prime!) - we see that $\langle 2^2 \rangle = \langle 4 \rangle = \{4, 3, 12, 9, 10, 1\}$. Note that $|\langle g^2 \rangle| = 6 = \frac{12}{2} = |\mathbb{Z}_{13}^*|/2$, as expected.

Assumptions

Discrete Log Assumption

- Choose cyclic group G of order q of bit length n, with a generator $g \in G$.
- Choose uniformly random exponent $a \in \mathbb{Z}_q$
- Output description of G, its order q, g, and $h = g^a$

Attacker receives this output (G, q, g, h), and outputs a' - wins iff $g^{a'} = h$.
Assumption: $\Pr[A \to a' \text{ s.t. } g^{a'} = h] \leq negl(n)$ for all PPT A

Computational Diffie-Hellman Assumption (CDH)

- Choose cyclic group G of order q of bit length n, with a generator $g \in G$.
- Choose uniformly random exponents $a, b \in \mathbb{Z}_q$
- Output description of group G, its order q, g, and $h_1 = g^a, h_2 = g^b$

Attacker receives this output (G, q, g, h_1, h_2), and outputs $z \in G$ - wins iff $z = g^{ab}$.
Assumption: $\Pr[A \to z \text{ s.t. } z = g^{ab}] \leq negl(n)$ for all PPT A

Decisional Diffie-Hellman Assumption (DDH)

- Choose cyclic group G of prime order q of bit length n, with a generator $g \in G$.
- Choose uniformly random exponents $a, b \in \mathbb{Z}_q$
- Output description of group G, order q, g, and $h_1 = g^a, h_2 = g^b$, as well as T, which is either g^{ab} or g^z, for uniformly random $z \in G$.

Attacker receives this output (G, q, g, h_1, h_2, T), and has to determine whether $T = g^{ab}$ or $T = g^z$.

Assumption: $|\Pr[A(G, q, g, h_1, h_2, g^{ab}) = 1] - \Pr[A(G, q, g, h_1, h_2, g^z) = 1]| \leq negl(n)$ for all PPT A

Note the relationship between these three assumptions. If we can break discrete log, then we can break CDH, and if we can break CDH then we can break DDH. Thus, DDH is the strongest assumption of these three (and discrete log is the weakest).

It is believed that discrete log and CDH hold for *any* $\mathbb{Z}_p^*$ for prime p, but DDH could only possibly be true over prime order groups, so not true for the group $\mathbb{Z}_p^*$ (which is of order $p-1$). If p is a safe prime $p = 2q + 1$ where q is also a prime, DDH is believed to be true for the order q subgroup of $\mathbb{Z}_p^*$ (namely over QR_p).

Provided example (wrong!): This example was shown in the recitation, but as we will discuss, it is actually not secure, as we use a non-safe prime. (Reproduced from online notes at http://yums.org.uk/wp-content/uploads/2013/06/Diffie-Hellman-Key-Exchange-and-ElGamal.pdf)

Let $p = 9967$ and $g = 3$ (a generator). Alice secretly chooses $a = 34$, and computes $g^a = 3^34$. She can do this in polynomial time by repeatedly taking the exponent (we are working in mod 9967):

$$
\begin{aligned}
3^2 &= 9 \\
3^4 &= 9^2 = 81 \\
3^8 &= 81^2 = 6561 \\
3^{16} &= 6561^2 = 43046721 \equiv 9215 (\text{mod} 9967) \\
3^{32} &= 9215^2 (\text{mod} 9967) \equiv 84916225 \equiv 7352 (\text{mod} 9967) \\
3^{34} &= 3^{32} \times 3^2 \equiv 7352 \times 9 \equiv 6366 (\text{mod} 9967)
\end{aligned}
$$

Alice sends 6366 to Bob.

Bob, in the mean-time, picks $b = 37$, and we can use the same calculations above to see $3^{37} = 3^{32} \times 3^4 \times 3 \equiv 7452 \times 81 \times 3 \equiv 2443 (\text{mod} 9967)$.

Bob sends 2433 to Alice.
Thus, currently, $g^a = 6366$ and $g^b = 2433$ are public, while $a = 34$ and $b = 37$ are secret to Alice and Bob, respectively.

Alice can use this same fast exponentation to raise the public 2433 to the 34th power, obtaining $k = 2443^{34} \equiv 7782 (\text{mod} 9967)$.

Likewise, Bob computes $k = 6366^{37} \equiv 7782 (\text{mod} 9967)$.

Thus, the keys match! $g^{ab} = (g^a)^b = (g^b)^a$ as expected, and Bob and Alice now have a key that they can use in public key encryption schemes.

Explanation: In this example, we use $G = Z_{9967}^*$, with generator $g = 3$. The correctness does indeed goes through as shown – clearly Bob and Alice do generate the same key k. However, the only type of security we can get here, is that the adversary cannot completely reconstruct the key that Alice and Bob agree on (7782) [this holds under the CDH assumption – it's hard to find g^{ab} from seeing just g^a and g^b].

However, it is **not** true that the key they agree on is indistinguishable from a random element in the group (this we would get from DDH assumption, but DDH assumption is false in general in Z_p^*). In particular, if you look at the distribution of the agreed key g^{ab} when Alice chooses a at random and Bob chooses b at random, we can show that g^{ab} has probability

$\frac{1}{4}$ to be g^{even} and probability $\frac{3}{4}$ to be g^{odd} (while half $of Z_p^*$ is g^{even} and half is g^{odd}, so the probability distribution of the agreed key is skewed). This means the adversary, just by knowing that they are using Diffie-Helman Key Exchange (without even looking at communication), has some information on the key they agreed on.

A similar attack would work for any group of composite order. So for Diffie-Helman Key Exchange (as well as El Gamal Public Key Encryption, and any other example that requires DDH assumption for security), you must have a group of prime order.

If instead we are working in Z_p^* but with generator $g^2 \pmod p$, then all the elements generated will *all* be g^{even} (and in fact, we are working in the subgroup QR_p). In this example, if you use $3^2 = 9$ as a generator, you end up working in the subgroup QR_{9967}, which is a group of size $9966/2 = 4983$.

If we were lucky with this example and 4983 happened to be prime, then we believe DDH holds and this is indeed what Alice and Bob can use (ie., the same example, but with 9 as a generator, not 3).

Unfortunately, 4983 is not a prime (it's divisible by 3). So, while all the elements they will use are indeed in QR_{9967} (i.e. all are 3^{even}), the key is *NOT* distributed randomly over these elements, it's biased (as it always will be for any group of non-prime order – now it's more likely to be 9^{even} rather than 9^{odd}). So, the adversary learns information about the key (even though they can't completely reconstruct the key).

A correct and secure example would be to take something like this, but choose a safe prime p, and then work with g^2 rather than with g as a basis. See, for example, Problem 1 on HW 5.

Note that working in Z_p^* was the original suggestions of DH and those early years, before we developed a better understanding of security requirements.

From Edo: Apologies to the recitation-goers for not doing justice to this explanation. You guys should have called out my mistake! ☺