# CS6111: Foundations of Cryptography

Assignment 3

S Srinivas Saurab

CS16B039

## Instructions

- Deadline: Wednesday, Oct 16.
- We encourage submissions by Latex. Paper is also accepted.

## References

- Introduction to Cryptography - Delfs and Knebl
- A Graduate Course in Applied Cryptography - Boneh and Shoup (link)
- Introduction to Modern Cryptography - Katz and Lindell

# 1 Number Theory

Let $G$ be some group and let $g \in G$ be an element of prime order $n$. That is, the set of elements generated by $g$ is a cyclic group of prime order, denoted as follows: $G_n = \langle g \rangle = \{g^i \mid i \geq 0\}$. We have that $\langle g \rangle^* = \{g^x : x \in \mathbb{Z}_n^*\}$ is the set of all generators of $G_n$. Clearly, $|\langle g \rangle^*| = \phi(n)$, which is known to be the number of generators of any cyclic group of order $n$.

1. (8 points) Show that for any $h \in \langle g \rangle$, the following conditions are equivalent:

   (1) $h \in \langle g \rangle^*$

   (2) $ord(h) = n$

   (3) $\langle h \rangle = \langle g \rangle$

   (4) $\langle h \rangle^* = \langle g \rangle^*$

   To show equivalency, you must prove the following:

   (a) (2 points) (1) $\implies$ (2)

   (b) (2 points) (2) $\implies$ (3)

   (c) (2 points) (3) $\implies$ (4)

   (d) (2 points) (4) $\implies$ (1)

**Solution:** Across all parts, we have the following common information:

$$g^n = e$$

$$\langle g \rangle^* = \{g^x : x \in \{1, 2, ....., n-1\}\}$$

This is because, actual definition has $x \in \mathbb{Z}_n^*$ and n is prime. Therefore, every number less than n is co-prime to n and thus

$$\mathbb{Z}_n^* = \{1, 2, 3, ....., n-1\}$$

(a) We need to prove that $h \in \langle g \rangle^* \implies ord(h) = n$.

If $h \in \langle g \rangle^*$, we have x in $\{1, 2, 3, ....., n-1\}$ such that

$$h = g^x$$

Since $h \in \langle g \rangle$, it belongs to the finite group as g itself has a finite order n. This implies h has a finite order as otherwise $\{h^1, h^2, h^3, ......\}$ will become a infinite group. This is a contradiction because each of these terms can be also expressed as powers of g as $h = g^x$. This is a contradiction as g is of finite order and so it cannot have infinite distinct number of powers.

Therefore,

$$\exists a; h^a = e$$

We have shown that h has a finite order. Now we will show that this order must be equal to n.

**If $a < n$:**

In this case, we have

$$e = h^a = g^{ax}$$

As $a, x < n$, we have: $g^{ax} = e$. As inverse exists in $\langle g \rangle$, consider $g^{ax-kn}$ such that ax-kn < n. In this case, we have a $y = ax - kn < n$ such that

$$g^y = e$$

This is a contradiction as we know that g's order is n. Therefore $a \geq n$.

**If $a > n$:** This cannot be the case because $h^n = g^{nx} = e$. Therefore, a has to be less than n.

Hence we have proved that a = ord(h) = n.

(b) We need to prove that $ord(h) = n \implies \langle h \rangle = \langle g \rangle$

We have ord(h) is n. Firstly, note that as $h \in \langle g \rangle$, we have $\exists a$ such that $h = g^a$.

Now, returning to the question, we need to basically prove that:

$$\langle h \rangle = \{h, h^2, h^3, .......\} = \{g, g^2, g^3.......\} = \langle g \rangle$$

As ord(h) = n = ord(g), this is same as showing:

$$\{h, h^2, h^3......, h^n\} = \{g, g^2, g^3....., g^n\}$$

$$\text{As n is prime, all the elements in } \{h, h^2, ..., h^n\} \text{ are different}$$

This is because if $h^i = h^j$, then $h^{i-j} = 1$ but $i - j < n$. Also, observe that any power of h is a power of g as $h = g^a$. As g's power is n, there are only 'n' distinct powers of g possible (which are simply $\{g, g^2, ...., g^n\}$). Thus from the above observations,

$$\langle h \rangle = \{h, h^2, ...., h^n\} = \{g, g^2, ..., g^n\} = \langle g \rangle$$

$$\text{Hence, } \langle h \rangle = \langle g \rangle$$

(c) We need to prove that $\langle h \rangle = \langle g \rangle \implies \langle h \rangle^* = \langle g \rangle^*$
Since n is prime, all the numbers in $\{1, 2, ..., n-1\}$ are co-prime to it. We need to show that if

$$\{h, h^2, h^3, ...\} = \{g, g^2, g^3 ....\}$$

then

$$\{h, h^2, ...., h^{n-1}\} = \{g, g^2, ..., g^{n-1}\}$$

From the common information, we are given that $h \in \langle g \rangle$. Therefore, for some $a \in \mathbb{N}, h = g^a$.

From this, we can arrive at the proof by considering three facts:
1) All of $\{h, h^2, ..., h^{n-1}\}$ are all different because n is prime.
2) Each power of h is also a power of g.
3) There are at most n distinct powers of g.
The first one is true because $\langle h \rangle = \langle g \rangle \implies h \neq 1$ and g's order is n. Other two claims are also easy to prove. From the above three claims, following similar proof as above, we get

$$\{h, h^2, ...., h^{n-1}\} = \{g, g^2, ..., g^{n-1}\}$$

(d) We need to prove that $\langle h \rangle^* = \langle g \rangle^* \implies h \in \langle g \rangle^*$
$\langle h \rangle^*$ has all generators of $\langle h \rangle$. Hence, obviously $h \in \langle h \rangle^*$. As $\langle h \rangle^* = \langle g \rangle^*$, then we directly have $h \in \langle g \rangle^*$ as $h \in \langle h \rangle^*$.

2. (2 points) Show that $a^{\log_h b} = b^{\log_h a}$ for any $a, b \in \langle g \rangle$ and $h \in \langle g \rangle^*$.

**Solution:** As h generates $\langle g \rangle$, we have i,j such that $a = h^i$ and $b = h^j$.

$$a^{\log_h b} = a^j = h^{ij}$$

$$b^{\log_h a} = b^j = h^{ij}$$

Hence $a^{\log_h b} = b^{\log_h a}$.

## 2 Probability Theory

Let $X, Y$ be two random variables and $V$ denote the set of possible values for $X$ and $Y$. $\Delta(X;Y)$ represents the statistical distance between $X$ and $Y$.

1. (2 points) Prove the following proposition: $\Delta(X;Y) = 1 - \sum_{v \in V} \min(\Pr[X=v], \Pr[Y=v])$.

**Solution:** "The" statistical distance refers to Total Variation Distance. In simple terms, it is defined as the largest difference between the probabilities assigned by two distributions to an event, i.e., a subset in the probability space.

**Claim:** Largest differentiating subset is the set of a's such that Pr(Y=a) > Pr(X=a). i.e.,

$$S_{max} = \{a | Pr(Y = a) > Pr(X = a)\}$$

**Proof:** To begin with, if any set maximizing this probability difference has at least one element such that Pr(X=a) > Pr(Y=a), then simply removing such elements would increase the probability difference. Therefore, any maximizing set should not have such elements.

Moreover, if a set does not contain at least one element that satisfies Pr(Y=a) > Pr(X=a), then simply adding such an element would increase the probability. Therefore, all such elements must be present in such a maximizing set.

Hence we have

$$S_{max} = \{v | Pr(Y = v) > Pr(X = v)\}$$

(Note: it can be symmetrically opposite if we consider Pr(Y=v)<Pr(X=v).)

Therefore, we know that:

$$\Delta(X;Y) = \sum_{v \in S_{max}} (Pr(X = v) - Pr(Y = v))$$

Considering the complement event of $S_{max}$ which elicits the same difference in probabilities, we have:

$$\Delta(X;Y) = \sum_{v \in S_{max}} (Pr(X = v) - Pr(Y = v)) = \sum_{v \notin S_{max}} (Pr(Y = v) - Pr(X = v))$$

From the two equalities above, we have:

$$\Delta(X;Y) = \sum_{v \in V} \frac{1}{2} |(Pr(X = v) - Pr(Y = v))|$$

$$\Delta(X;Y) = \sum_{v \in V} \frac{1}{2} \{max(Pr(X = v), Pr(Y = v)) - min(Pr(X = v), Pr(Y = v))\}$$

We know that $\sum_{v \in V} \{max(Pr(X = v), Pr(Y = v)) + min(Pr(X = v), Pr(Y = v))\}$ is 2 as the probabilities over all elements in probability space add to two. By substituting this and cancelling half, we have:

$$\Delta(X;Y) = 1 - \sum_{v \in V} \min(\Pr[X = v], \Pr[Y = v])$$

2. (2 points) Prove the triangle inequality: $\Delta(X;Z) \leq \Delta(X;Y) + \Delta(Y;Z)$

**Solution:** To prove this, we need to use a slightly different formula from that derived in Q2.1.

We use,

$$\Delta(X;Z) = \sum_{v \in V} \frac{1}{2}\{max(Pr(X = v), Pr(Z = v)) - min(Pr(X = v), Pr(Z = v))\}$$

Now, consider what happens when a new distribution **'Y'** comes into play. At a given $v \in V$ assume that $Pr(Z = v) \geq Pr(X = v)$, we arrive at three cases:

**Case1:** $Pr(Y = v) \geq Pr(Z = v) \geq Pr(X = v)$.

In this case,

$$(\Delta(X;Y) + \Delta(Y;Z))|_v \geq \Delta(X;Z)|_v$$

This is because,

$max(Pr(X = v), Pr(Y = v)) - min(Pr(X = v), Pr(Y = v)) = Pr(Y = v) - Pr(X = v)$

and its already $\geq$

$max(Pr(X = v), Pr(Z = v)) - min(Pr(X = v), Pr(Z = v)) = Pr(Z = v) - Pr(X = v)$

**Case2:** $Pr(Z = v) \geq Pr(Y = v) \geq Pr(X = v)$.

In this case,

$$(\Delta(X;Y) + \Delta(Y;Z))|_v = \Delta(X;Z)|_v$$

This is because,

$max(Pr(X = v), Pr(Y = v)) - min(Pr(X = v), Pr(Y = v)) = Pr(Y = v) - Pr(X = v)$

and $max(Pr(Z = v), Pr(Y = v)) - min(Pr(Z = v), Pr(Y = v)) = Pr(Z = v) - Pr(Y = v)$

and their sum is exactly same as

$max(Pr(X = v), Pr(Z = v)) - min(Pr(X = v), Pr(Z = v)) = Pr(Z = v) - Pr(X = v)$

**Case3:** $Pr(Z = v) \geq Pr(X = v) \geq Pr(Y = v)$

In this case,

$$(\Delta(X;Y) + \Delta(Y;Z))|_v \geq \Delta(X;Z)|_v$$

This is because,

$max(Pr(Z = v), Pr(Y = v)) - min(Pr(Z = v), Pr(Y = v)) = Pr(Z = v) - Pr(Y = v)$

and its already $\geq$

$max(Pr(X = v), Pr(Z = v)) - min(Pr(X = v), Pr(Z = v)) = Pr(Z = v) - Pr(X = v)$

We have shown that for all v such that $Pr(Z = v) \geq Pr(X = v)$, $(\Delta(X;Y) + \Delta(Y;Z))|_v \geq \Delta(X;Z)|_v$. By symmetry, we can see that this holds also if $Pr(Z = v) \leq Pr(X = v)$.

Hence taking $\sum_{v \in V}$, we have:

$$\sum_{v \in V} (\Delta(X;Y) + \Delta(Y;Z))|_v \geq \Delta(X;Z)|_v$$

Such addition is justified because we are simply adding the (max-min of distributions) at each element in the system of events. Therefore, we have:

$$\Delta(X;Y) + \Delta(Y;Z) \geq \Delta(X;Z)$$

3. (2 points) Prove that $\Delta(f(X); f(Y)) \leq \Delta(X; Y)$ for any function $f$ defined on $V$.

**Solution:** Suppose the function $f$ maps the space of events in $V$ to the space of events in $W$, i.e., f:V→W. Now, consider W, the co-domain and let W' be the range of f. Hence $W' \subseteq W$

The contribution to variational distance will come only from the elements in W'. This is because both f(X) and f(Y) will have zero probability in $W \backslash W'$.

If f is one-one, then there is an isomorphism from V to W' (bijection) hence, there will be no difference in $\Delta(f(X); f(Y))$ and $\Delta(X; Y)$. To precisely prove this, we can consider $Pr(f(X) = a) = Pr(X = f^{-1}(a))$ as inverse can be properly defined.

Hence, f is one-one $\implies \Delta(f(X); f(Y)) = \Delta(X; Y)$

In the case that f is not a one-one function, there is at least one element in W' that is mapped to two elements in the domain.

**Claim:** If f is not one-one, then $\Delta(f(X); f(Y)) \leq \Delta(X; Y)$.

**Proof:** From Q2.1, we have

$$\Delta(X; Y) = 1 - \sum_{v \in V} \min(\Pr[X = v], \Pr[Y = v])$$

For the element y that is mapped has two pre-images, say a,b, we have:

$Pr(f(X) = y) = Pr(X = a) + Pr(X = b)$ and $Pr(f(Y) = y) = Pr(Y = a) + Pr(Y = b)$.

Therefore, we have:

$$\min(\Pr[f(X) = y], \Pr[f(Y) = y]) > \min(\Pr[X = a], \Pr[Y = a])$$

Summing over all y ∈ W', we have:

$$\sum_{y \in W'} [\min(\Pr[f(X) = y], \Pr[f(Y) = y]) > \sum_{a \in V} \min(\Pr[X = a], \Pr[Y = a])$$

Note that 'a' cannot obviously have two images as f is a function and hence the above summations cover all elements in V in the LHS.

$$\Delta(f(X); f(Y)) \leq \Delta(X; Y)$$

We have shown that whether f is one-one or not, $\Delta(f(X); f(Y)) \leq \Delta(X; Y)$. Therefore, we can state absolutely that $\Delta(f(X); f(Y)) \leq \Delta(X; Y)$.

4. (4 points) For $n \geq 1$, let $X \in_R \mathbb{Z}_n$ and $Y \in_R \mathbb{Z}_n^*$.

   (a) (2 points) Determine $\Delta(X; Y)$.

**Solution:** $\mathbb{Z}_n = \{0, 1, 2, ......., n-1\}$ and $\mathbb{Z}_n^* = \{1, ....., n-1\}$ ($\mathbb{Z}_n^*$ considers only co-primes).

Now, using result from Q2.1, we have:

$$\Delta(X;Y) = 1 - \sum_{n\in\mathbb{N}} min(Pr(X=n), Pr(Y=n))$$

$$\implies \Delta(X;Y) = 1 - \sum_{n\in\{0,1,2,....,n-1\}} min(Pr(X=n), Pr(Y=n))$$

$$\implies \Delta(X;Y) = 1 - \sum_{n\in\mathbb{Z}_n^*} min(Pr(X=n), Pr(Y=n)) - \sum_{n\in\mathbb{Z}_n\setminus\mathbb{Z}_n^*} min(Pr(X=n), Pr(Y=n))$$

(Dividing them into disjoint sets)
As Pr(Y=n) = 0 for $n \in \mathbb{Z}_n\setminus\mathbb{Z}_n^*$, we have the right most term as zero.

$$\implies \Delta(X;Y) = 1 - \sum_{n\in\mathbb{Z}_n^*} min(Pr(X=n), Pr(Y=n))$$

Now, we observe that within $\mathbb{Z}_n^*$, Pr(X=n) is lesser that Pr(Y=n), i.e.,

$$\forall n \in \mathbb{Z}_n^*, Pr(X=n) < Pr(Y=n)$$

Substituting this result and considering that $|\mathbb{Z}_n^*| = \varphi(n)$, we get:

$$\Delta(X;Y) = 1 - \frac{\varphi(n)}{n}$$

(b) (2 points) Show that $\Delta(X+Y;XY) = 0$, where addition and multiplication are done modulo $n$.

**Solution:** To show this, we need to rely on a interesting observation about $\mathbb{Z}_n^*$.

$$\forall x \in \mathbb{Z}_n^*\ , \forall y \in \mathbb{N}\ n|xy \implies n|y$$

**Claim1:** For an $x \in \mathbb{Z}_n^*$, $\{x, 2x, 3x, 4x, ....., nx\}$ must all represent different numbers.
**Proof:** This is true because if any of $ax = bx$, then $(a-b)x = 0$. As $n|0$, and $n \not| x$, $n$ must divide $a-b$. However, as $a - b < n$ and $a, b$ are distinct, n cannot divide $a - b$. Hence a contradiction.

From here, the proof of $\Delta(X+Y;XY) = 0$ straightforward. Consider any element 'a' in $\mathbb{Z}_n$. In the modulo 'n' world, we see that it is displaced by a unique amount w.r.t any element in $\mathbb{Z}_n^*$, i.e.,

$$\forall a \in \mathbb{Z}^n\ \forall n \in \mathbb{Z}_n^*\ \exists! d \in \mathbb{Z}_n \text{ such that } a = n + d$$

Now, from claim1, we have that all of $\{x, 2x, 3x, ...., nx\}$ must be different in modulo-n world. This means that this set should cover all of $\{0, 1, 2, ...., n-1\}$ in the modulo world as the cardinalities of $\{0, 1, 2, ...., n-1\}$ and $\{x, 2x, ..., nx\}$ are equal. From the above two arguments, we have:

$$\forall a \in \mathbb{Z}^n\ \forall n \in \mathbb{Z}_n^*\ \exists! d \in \mathbb{Z}_n \text{ such that } a = n * d$$

Note that we don't use the term 'nx', instead we use '0*x' if we need to arrive at 0 in $\mathbb{Z}^n$.

**Claim2:** $X+Y$ and $XY$ define two probability distributions over $0, 1, ...., n-1$ and they are exactly equal.

**Proof:** We have seen that there are unique elements $d_1, d_2$ in $\mathbb{Z}^n$ for any element n in $\mathbb{Z}_n^*$ to reach any element a in $\mathbb{Z}_n$ through modulo addition and multiplication respectively.

Therefore,

$$Pr(X+Y=a) = \sum_{n\in\mathbb{Z}_n^*} Pr(X=n) * Pr(Y=d_1) = \sum_{n\in\mathbb{Z}_n^*} Pr(X=n) * \frac{1}{n} = \sum_{n\in\mathbb{Z}_n^*} \frac{1}{\varphi(n)} * \frac{1}{n} = \frac{1}{n}$$

Also,

$$Pr(XY=a) = \sum_{n\in\mathbb{Z}_n^*} Pr(X=n) * Pr(Y=d_2) = \sum_{n\in\mathbb{Z}_n^*} Pr(X=n) * \frac{1}{n} = \sum_{n\in\mathbb{Z}_n^*} \frac{1}{\varphi(n)} * \frac{1}{n} = \frac{1}{n}$$

Here, we have used that the variables X,Y are randomly chosen from their respective domains. Therefore, as our choice of 'a' has been arbitrary, we can conclude that $X+Y$ and $XY$ define the same probability distribution on all $a \in \{0, 1, ..., n-1\}$. Hence, from claim-2, $\Delta(X+Y; XY) = 0$

5. (2 points) For $n \geq d \geq 1$, let random variable $X$ take on values in $\{0, \cdots, d-1\}$, and let $U \in_R \{0, \cdots, n-1\}$. Show that $\Delta(U; X+U) \leq (d-1)/n$, and that this bound is tight.

**Solution:** We can see that the random variable $X+Y$ ranges from 0 to $n+d-2$. As U is in $\{0, 1, ..., n-1\}$, only these values contribute to the Total variational distance $\Delta(U; X+U)$.

Hence, we can determine $\Delta(U; X+U)$ by considering the distribution of $X+U$ only in this range of $\{0, 1, ..., n-1\}$.

Let us determine this probability distribution:

$Pr(X+U=a) = \begin{cases} \frac{1}{n} Pr(X \leq a) \text{ if } 0 \leq a \leq d-1 \\ \frac{1}{n} \text{ if } d \leq a \leq n-1 \end{cases}$ This is because, if $a \geq d$, there is always a 'b' such that $a+b=n$, hence the probability is simply $Pr(U=b) = \frac{1}{n}$. However, if the case that $a \leq d-1$, there is a possibility that $X > a$ and hence there would be no $U$ such that $X+U=a$. Therefore, we have the term $Pr(X \leq a)$.

Now, from Q2.1, we have:

$$\Delta(X;Y) = 1 - \sum_{v\in V} \min(\Pr[X=v], \Pr[Y=v])$$

On substituting Y by $X+U$ and considering the appropriate range, we have:

$$\Delta(X; X+U) = 1 - \sum_{a\in\{0,1,..,n-1\}} \min(\Pr[U=a], \Pr[X+U=a])$$

Considering RHS, we can split it as:

$$\Delta(X;X+U) = 1 - \sum_{a\in\{0,1..,d-1\}} \min(\Pr[U=a],\Pr[X+U=a]) - \sum_{a\in\{d,d+1..,n-1\}} \min(\Pr[U=a],\Pr[X+U=a])$$

Considering the distribution of $X+U$ as defined above, we have the last term as :

$$\Delta(X;X+U) = 1 - \sum_{a\in\{0,1,..,d-1\}} \min(\Pr[U=a],\Pr[X+U=a]) - \frac{1}{n} * (n-d)$$

Now, expanding the middle term, we have:

$$\Delta(X;X+U) = 1 - \sum_{a\in\{0,1,..,d-1\}} \frac{1}{n}(Pr(X \le a)) - \frac{1}{n} * (n-d)$$

$$\implies \Delta(X;X+U) = \frac{d}{n} - \frac{1}{n} * \sum_{a\in\{0,1,..,d-1\}} (Pr(X \le a))$$

Now, for a fixed 'd', we know that the right term in RHS is greater than 1. This is because, when $a = d-1$, the probability $Pr(X \le a) = 1$ for any distribution of 'X' and the other probabilities in the summation are at least zero. Hence, we have, **over all distributions of X**:

$$\Delta(X;X+U) \le \frac{(d-1)}{n}$$

The argument above gives us the idea on why the bound must be tight. It is tight when $Pr(X \le a) = 1$ when $a = d-1$ and is zero when $a < d-1$.
We can **construct a distribution where** $Pr(X = d-1) = 1$ **and is 0 otherwise.** For this, from the formula in Q2.1, we have:

$$\Delta(X;X+U) = 1 - 0*(d-1) - \frac{1}{n} * ((n-1)-(d-1)+1)$$

$$\Delta(X;X+U) = \frac{(d-1)}{n}$$

Hence the bound is tight.

6. (4 points) For $n \ge 1$, consider distributions $X, Y, Z$ given by

$$X = \{u : u \in_R \{0,\cdots,n-1\}\},$$
$$Y = \{2u : u \in_R \{0,\cdots,n-1\}\},$$
$$Z = \{2u+1 : u \in_R \{0,\cdots,n-1\}\}.$$

Clearly, $\Delta(Y;Z) = 1$.

(a) (2 points) Show that $\Delta(X;Y) = \Delta(X;Z) = 1/2$ for even $n$

**Solution:** We have:

$$X \in_R \{0, 1, 2, ...., n-1\},$$
$$Y \in_R \{0, 2, 4...., 2n-2\},$$
$$Z \in_R \{1, 3, 5, ...., 2n-1\}.$$

Considering the distributions in range of X will be enough when we consider the formula in Q2.1. When n is even, say = 2k, then number of values of Y in {0,1,...,2k-1} are k exactly. Using Q2.1, we have:

$$\Delta(X;Y) = k*(0) + k*(\frac{1}{n})$$
$$\text{Hence, } \Delta(X;Y) = k*(\frac{1}{n}) = \frac{1}{2}$$

Similarly, then number of values of Z in {0,1,...,2k-1} are k exactly. Using Q2.1, we have:

$$\Delta(X;Z) = k*(0) + k*(\frac{1}{n})$$
$$\text{Hence, } \Delta(X;Z) = k*(\frac{1}{n}) = \frac{1}{2}$$

(b) (2 points) Determine $\Delta(X;Y)$ and $\Delta(X;Z)$ for odd $n$.

**Solution:** Similar to the above question, When n is odd, say 2k+1, then we have exactly k+1 even numbers in 0,1,...,2k and k odd numbers in 0,1,...,2k. Therefore, the corresponding statistical distances are:

$$\Delta(X;Y) = k*(0) + (k+1)*(\frac{1}{n}) = \frac{n+1}{2n}$$
$$\Delta(X;Z) = (k+1)*(0) + k*(\frac{1}{n}) = \frac{n-1}{2n}$$

7. (5 points) For $n$ prime, let $h$ and $M_0$ be arbitrary, fixed elements of $G_n = \langle g \rangle, h \neq 1$. Consider distributions $X, Y, Z$ given by

$$X = \{(A, B) : A \in_R \langle g \rangle, B \in_R \langle g \rangle\},$$
$$Y = \{(g^u, h^u M) : u \in_R \mathbb{Z}_n, M \in_R \langle g \rangle\},$$
$$Z = \{(g^u, h^u M_0) : u \in_R \mathbb{Z}_n\}.$$

**Solution:** Across, all parts of the question, as h and $M_0$ are fixed, let us consider them as $h = g^i$ and $M_0 = g^j$. Also, as $\langle g \rangle = G_n$, we have: $\langle g \rangle = \{g, g^2, g^3, ...., g^n\}$.

(a) (2 points) Show that $\Delta(X;Y) = 0$

**Solution:** The random variable $X$ is an ordered pair of two powers taken uniformly randomly from $\langle g \rangle$. The left coordinate of Y is also uniformly randomly chosen power of g. Hence the distributions are identical with respect to the left coordinate. We claim that the RHS's distributions are also identical.
**Claim:** $h^u M$ is equivalent to the uniform distribution as defined by B.
**Proof:** From our pre-assumptions, we have: $h = g^i$. Let us say $M = g^k$, then $h^u M = g^{iu+k}$ where u,k $\in_R \mathbb{Z}_n$. ($\mathbb{Z}_n$ because when k = n, we can consider k as 0 instead and the distributions wouldn't change.)
Analogous to 'Affine cipher' we will show that the randomness offered by 'k' wil be enough to make the distribution of '$iu + k$' uniformly random.
**Proof:** As order of g is 'n' all the powers can be considered modulo n. Therefore, $Pr(iu + k = a) = \sum_{y \in \mathbb{Z}_n} Pr(i = y) * Pr(k = (a - yu)\, modn)$. As $(a - yu)\, modn$ lies in $\mathbb{Z}_n$ for all y, this probability must be simply $\frac{1}{n}$. Therefore, we have:

$$Pr(iu + k = a) = \sum_{y \in \mathbb{Z}_n} Pr(i = y) * \frac{1}{n}$$

$$Pr(iu + k = a) = \frac{1}{n} * \sum_{y \in \mathbb{Z}_n} Pr(i = y)$$

As this summation is 1, we have:

$$Pr(iu + k = a) = \frac{1}{n}$$

Hence, the RHS is equivalent to uniform distribution. Therefore, as left and right coordinates of X,Y match in their respective ranges, the joint probabilities of their coordinates also match over the cross product of ranges. Hence the distributions are **exactly the same** and thus:

$$\Delta(X; Y) = 0$$

We did not even use $h \neq 1$ because that will not affect the summation of probabilities which will be 1 anyway.

(b) (2 points) Show that $\Delta(Y; Z) = 1 - 1/n$

**Solution:** Let us look at the Probability distributions:

$$Pr(X = (g^a, g^b)) = (\frac{1}{n} * \frac{1}{n})$$
$$Pr(Z = (g^a, g^b)) = \text{Let us calculate}$$

To calculate the probability distribution of Z, let us first look at the ranges of a,b,u,i,j:
As h $\neq 1$, $i \in \{1, 2, ..., n - 1\}$.
$a, b, j \in \{1, 2, ...., n\}$
$u \in \{0, 1, ..., n - 1\}$

Now, for $g^u = g^a \implies u = a$ as we have considered their ranges only in the modulo world. Also,

$h^u\mathbb{M}_0 = g^b \implies iu + j = b.$

For iu+j=b to be true,

$$u = i^{-1} * (b - j)$$

As n is prime and $i^{-1}$ is defined uniquely. Hence we have he condition that

$$a = i^{-1} * (b - j)$$

This happens only in n pairs of $(g^a, g^b)$. And for each of this pair, the probability that $Pr(Z = (g^a, g^b))$ is only $\frac{1}{n}$ as 'u' is chosen randomly. Therefore, from formula in Q2.1,

$$\Delta(Y; Z) = 1 - (n^2 - n) * 0 - n * (\frac{1}{n^2})$$

$$\Delta(Y; Z) = 1 - \frac{1}{n}$$

(c) (1 point) Show that $\Delta(X; Z) = 1 - 1/n$

**Solution:** We have shown that distributions of X,Y are identical at all points. Since statistical distance strictly depends only on the probability value at each point in the domain, we can infer that

$$\Delta(X; Z) = 1 - 1/n$$

## Omitted

8. (0 points) Prove the following proposition:

   1. $0 \leq \Delta(X; Y) \leq 1$, "nonnegativity" and "boundedness".
   2. $\Delta(X; Y) = 0$ if and only if $\forall_{v \in V} \Pr[X = v] = Pr[Y = v]$, "identical distributions".
   3. $\Delta(X; Y) = \Delta(Y; X)$, "symmetry".
   4. $\Delta(X; Z) \leq \Delta(X; Y) + \Delta(Y; Z)$. "triangle inequality".

9. (0 points) Prove the following proposition:

   1. $\Delta(X; Y) = \sum_{v \in V^+}(\Pr[X = v] - \Pr[Y = v])$, with $V^+ = \{v \in V : \Pr[X = v] > \Pr[Y = v]\}$.
   2. $\Delta(X; Y) = \sum_{v \in V}(\Pr[X = v] \dot{-} \Pr[Y = v])$, with $x \dot{-} y = \max(x - y, 0)$ ("x monus y").
   3. $\Delta(X; Y) = 1 - \sum_{v \in V} \min(\Pr[X = v], \Pr[Y = v])$.
   4. $\Delta(X; Y) = \max_{W \subseteq V} |\Pr[X \in W] - \Pr[Y \in W]|$.

10. (0 points) For $n, d \geq 1$, consider distributions $X$ and $Y$ given by

$$X = \{u : u \in_R \{0, \cdots, n-1\}\},$$
$$Y = \{u + d : u \in_R \{0, \cdots, n-1\}\}.$$

Determine $\Delta(X; Y)$, assuming $d \leq n$. Also, what is $\Delta(X; Y)$ if $d > n$?

# 3 One Way Functions/ Permutations

1. (2 points) Let $g_1 : \{0,1\}^n \rightarrow \{0,1\}^n$ and $g_2 : \{0,1\}^n \rightarrow \{0,1\}^n$ be two length preserving one-way functions. Define $f(x) = g_1(x)||g_2(x)$. Show that $f$ is not necessarily a one way function.

**Solution:**
It will be enough if we show one counterexample wherein f(x) will not be a OWF.
**Construction:**
Consider 'g' to be an arbitrary OWF from $\{0,1\}^{\frac{n}{2}}$to$\{0,1\}^{\frac{n}{2}}$. Now, select $g_1(x_1 \circ x_2) = x_1 \circ g(x_2)$ and $g_2(x_1 \circ x_2) = g(x_1) \circ x_2$.

**Claim:** $g_1$ and $g_2$ are both OWFs.
**Proof:** We will first show that $g_1$ is a OWF. By symmetry, $g_2$ will also be a OWF.
Assume that $g_1$ is not one-way. Then we have a PPT Adversary A for $g_1$ which can invert with non-negligible probability. From here, we have:

$$g(A(g_1(x_1 \circ x_2))_{[\frac{n}{2}.......n]}) = g(x_2)$$

whenever A succeeds. Therefore, by taking the right half of the A's inversion, we can always succeed in inverting 'g' when A succeeds in inverting $g_1$.
Hence, a polynomial blackbox reduction exists between the adversary A of $g_1$ to the shown adversary of $g$ (say B). Also we have shown that whenever A inverts $g_1$, B inverts g. From our assumption that A inverts with a probability : non-negl($\frac{n}{2}$), we can tell that B must invert with the same non-negl probability. But, non-negl($\frac{n}{2}$) is also non-negl(n). This is a contradiction that $g$ is a OWF.
Hence, $g_1$ is a OWF. And, by symmetry, we have $g_2$ is also a OWF.

From here, it is easy to see why $f(x) = g_1(x)||g_2(x)$ is not a one-way function. We can rewrite the equation, by plugging in definitions as: $f(x_1 \circ x_2) = x_1 \circ g(x_2)||g(x_1) \circ x_2$. For a given y, $f^{-1}(y)$ can simply be found by considering first and last quaritles of the y and invert with a probability 1, i.e.,

$$Pr(f(y_{[0,1,\ldots,\frac{n}{2}]}||y_{[\frac{3n}{2},\ldots,2n-1]}) = y) = 1$$

Hence, there exists a perfect inversion for 'f' by just looking at the output in linear time. Hence **f is not a one-way function, although $g_1$ and $g_2$ are one-way**.

2. (4 points) Let $f_1$ and $f_2$ be one way functions, where $\exists\ |x_1| = |x_2| \implies |f(x_1)| = |f(x_2)|$ (same sized output for same sized inputs). Let $f(x) = f_1(x_1) \oplus f_2(x_2)$ where $x = x_1||x_2$ and $|x_1| = |x_2|$ (assume even inputs).

(a) Give an example where $f_1$, $f_2$ and $f$ are one way functions.

**Solution:** To begin with, we need to assume that OWFs exist. Let $g$ be one such OWF. It is easy to see that

$$f_1(x) = 0^n||g(x)$$

and

$$f_2(x) = g(x)||0^n$$

are OWFs.

**Claim:** $f(x)$=$f_1(x_1) \oplus f_2(x_2)$ must be a OWF.

**Proof:** We have

$$f(x_1 \circ x_2) = g(x_2)||g(x_1)$$

To prove that this $f$ is a OWF, we can follow the proof strategies as done in assignment-2. To begin with, assume that 'f' is not a OWF. Then there exists a PPT non-negliglble inverter say A. Let us use A to construct B (a PPT, non-negl inverter for g).

**Reduction from B to A:**

For a given y where $|y| = n$, we need to find a $x \in g^{-1}(y)$ to successfully invert 'g'. Suppose the adversary 'B' has the access to the blackbox encrypter of 'g' (considering Kirckhoff's security). Then he can simply consider an arbitrary string 'z' such that $|z| = n$ and encrypt it using 'g's blackbox. Now, B can append y and g(z) to get a '2n' string. He can simply feed this string to 'A' which will give him say 'r'.

$$\text{i.e., Say} A(g(z)\|y) = r$$

Now, we have, if $r = r_1 \circ r_2$, $f(r) = g(r_2)||g(r_1)$, in the scenarios where 'A' successful inverts 'f'. Also, as f(r) is originally $y||g(z)$, we have $g(r_2) = y$. This means that whenever A succeeds, B can succeed by considering the second half on input to invert g. As A is non-negl adversary in '2n', B is also a non-negl adversary in '2n' and this is non-negl(n). Hence a contradiction that g is OW. **Therefore, f must be a OWF.**

To put simply, B's algorithm is:

B(y):

Consider a random n-bit string z

Find g(z)

Give A the input g(z)$\circ y, say$r = A(g(z)$\circ y$

If r = $r_1 \circ r_2$, return r2

PS: For odd size inputs to $f_1$ or $f_2$, split them unevenly, i.e., use $f_1(x_1 \circ x_2) = 0^{n+1}||g(x_1)$ and $f_2(x_1 \circ x_2) = g(x_1)||0^{n+1}$ instead and ignore f[0] and f[-1].

(b) Give an example where $f_1, f_2$ are one way but not $f$.

**Solution:** Following the hint given, let us set x to be equal to $x_1||x_2||x_3||x_4$ and the output f(x) to be $y_1||y_2$. Now, let us consider the one-way functions $f_1$ and $f_2$ identical to those in question-1. i.e.,

$$f_1(x_1 \circ x_2) = x_1 \circ g(x_2)$$

and

$$f_2(x_3 \circ x_4) = g(x_3) \circ x_4$$

where 'g' is a original one-way function from $\{0,1\}^{\frac{n}{4}} \to \{0,1\}^{\frac{n}{4}}$.

From here, we can see that $y_1||y_2$ = f($x_1||x_2||x_3||x_4$) = $x_1 \oplus g(x_3) \;||\; x_4 \oplus g(x_2)$. This means that for any given $y_1||y_2$, we have the adversary A with access to g's encryption (considering Kirckhoff's security):

**Algorithm for A($y_1||y_2$):**

Consider random $\frac{n}{2}$ bit strings a,b

Calculate $y_a = g(a)$ and $y_b = g(b)$

Consider $c = y_1 \oplus y_b$ and $d = y_2 \oplus y_a$

Return $(c||a||b||d)$

Obviously, when we find f($c||a||b||d$), we see that it is equal to $c \oplus g(b)||g(a) \oplus d$, i.e.,

$$f(c||a||b||d) = c \oplus g(b)||g(a) \oplus d$$

$$f(c||a||b||d) = c \oplus y_b||y_a \oplus d$$

From our choices of c and d, this is same as:

$$f(c||a||b||d) = y_1||y_2$$

Therefore, the given adversary A always succeeds in inverting 'f' with probability of 1. Hence **f is not a one-way function although $f_1$ and $f_2$ are.**

3. (8 points) Let $g(x)$ be a length preserving one-way function. Let $x = x_1||x_2$ where $x_1 = x_2$ (assume even inputs). Which of following are one-way functions? Prove your answers.

(a) $f_a(x) = g(\bar{x})$, where $\bar{x}$ is the bitwise compliment of $x$

**Solution:** If there is a PPT, non-negligible inverter for $f_a$, say A, then we will try to construct a non-negligible PPT inverter B for g.

B(y):

z := A(y)

x := $\bar{z}$

**Claim:** This algorithm of B successfully inverts g whenever A inverts $f_a$ .

**Proof:** Say $A(g(\bar{x})) = y$. If A inverts $f_a$, i.e., $g(\bar{x})$ successfully

$$g(\bar{y}) = g(\bar{x})$$

Therefore, given an input encryption of $\bar{x}$, it found a $\bar{y}$ such that:

$$g(\bar{y}) = g(\bar{x})$$

Now, replace $a = \bar{x}$ and $b = \bar{y}$ in the above proposition without loss of generality.
Then we find get the new proposition: $A(g(a)) = \bar{b}$. If A inverts $f_a$, i.e., $g(a)$ successfully

$$g(b) = g(a)$$

$$\implies g(B(g(a)) = g(a)$$

Hence both A,B invert $f_a, g$ respectively the same probability. As A inverts with non-negl, so does B and this creates a contradiction that g is one-way. **Therefore, $f_a$ is a one-way function.**

(b) $f_b(x) = g(x_1 \oplus x_2)$

**Solution:** $f_b(x_1||x_2) = g(x_1 \oplus x_2)$. Suppose $f_b$ is not a OWF. A PPT non-negl adversary for $f_b$ say A, can take in $g(x_1 \oplus x_2)$ and give us $x' = x_1'||x_2'$ such that $g(x_1' \oplus x_2') = g(x_1 \oplus x_2)$. i.e.,

$$A(g(x_1 \oplus x_2)) = x' \text{ such that } g(x_1 \circ x_2) = g(x_1' \circ x_2')$$

with non-negl probability. Let us construct B, an adversary for g as:

Algorithm for B(g(x)):
Find x' := A(g(x))
Consider x' = $x_1'||x_2'$
Return $x_1' \oplus x_2'$

B will succeed whenever A does because of the definition of A's success that $A(g(x_1 \oplus x_2)) = x'$ such that $g(x_1 \circ x_2) = g(x_1' \circ x_2')$. Therefore, as A is non-negl(2n), PPT(2n) it is also non-negl(n) and also PPT(n). Hence, a contradiction as B will now succeed with non-negl probability over input space. Therefore, $f_b$ is a OWF.

A part $f_c(x) = \begin{cases} 0^{|x|} \text{ if exactly one bit of } x_1 \text{ is } 1 \\ 0^{|x_1|} \cdot g(x_2) \quad \text{otherwise} \end{cases}$

**Solution:** If x = $x_1 \circ x_2$, and $|x_1| = n = |x_2|$ the number of $x_1$'s that have exactly one bit as '1' are 'n'. We have already proven that $f(x) = 0^n||g(x_2)$ is a OWF.

Now we will show that setting f(x) to $0^{2n}$ for $\frac{n}{2^n}$ fraction of the input does not affect the one-wayness of f. Assume that an arbitrary adversary A of f could not invert any of these x's successfully. Now the adversary of $f_c(x)$ say $A_c$ can invert this fraction successfully with Pr=1. This means that

$$Pr[x \leftarrow \{0,1\}^{2n}, A_c(1^{2n}, f_c(x)) \in f_c^{-1}(f_c(x))] \leq Pr[x \leftarrow \{0,1\}^{2n}, A(1^{2n}, f(x)) \in f^{-1}(f(x))] + \frac{n}{2^n} * 1$$

As A can be any adversary, let us take it equal to $A_c$ except over this fraction of input and wrong-inverting over this $\frac{n}{2^n}$ fraction of the input. In this case, f is a OWF, the RHS term is negligible. Hence the LHS cannot be non-negligible.

Hence $f_c(x)$ is a OWF.

(c) $\text{f}_d(x) = \begin{cases} 0^{|x|} \text{ if at least one bit of } x_1 \text{ is } 1 \\ 0^{|x_1|} \cdot g(x_2) \;\; \text{otherwise} \end{cases}$

**Solution:** Just invert everything to $0^{2n}$, we will be correct with a probability of 1-$\frac{1}{2^n}$ as only when $x_1 = 0^n$, we might be wrong. Therefore, $f_d$ cannot be a OWF.

4. (2 points) We know that $f$ may be one way but $f(f(x))$ may not be one way. What about $f(x)||f(f(x))$?

**Solution:** $g(x) = f(x)||f(f(x))$ must be a OWF. We can give a reduction where if we have a non negligible PPT inverter black box for $g$, we can invert $f$ also.

**Reduction using Black Box Inverter of $g$:**
Suppose g is not a OWF, it has an PPT non-negl inverting adversary. Let B be one such inverter of 'g'. Let us construct an adversary A for 'f'. On receiving an string 'y', A can use f's blackbox (considering Kirckhoff's security) to encrypt that string again. It can then append to get $y||f(y)$. Now, if this string is fed to B, we have, say :

$$B(y||f(y)) = x$$

Then, if B successfully inverter 'g', this means that:

$$y||f(y) = f(x)||f(f(x))$$

Hence, this means that f(x) must be y. This means that if 'A' simply returns whatever B returns, it has to follow: $f(B(y||f(y))) = y$ whenever B is successful.
Hence, A succeeds whenever B does. This is a contradiction as now f can be inverted with non-negl probability. **Hence g must be a OWF**.

5. (3 points) Given a strong one way function $f$, construct a weak one way function $g$ that is NOT a strong one way function.

**Solution:** Let us begin by defining two classes of One-Way Functions.

**Strong One-Way Functions:** A strong one-way function can be computed by a PPT and $\forall$ PPT A, $\exists$negl fuction$\epsilon$ and $\forall n \in N$ such that $Pr[x \leftarrow \{0,1\}^n, A(1^n, f(x)) \in f^{-1}(f(x))] \leq \epsilon(n)$

**Weak One-Way Functions:** A weak one-way function can be computed by a PPT and $\exists$ polynomial $q(x)$ $\forall$ PPT A, $\forall n \in N$ such that $Pr[x \leftarrow \{0,1\}^n, A(1^n, f(x)) \in f^{-1}(f(x))] \leq 1 - \frac{1}{q(x)}$

We need to construct a weak one-way function from a strong one.

**Idea:** We construct a function g that acts like identity function on $\frac{3}{4}^{th}$ of the input and acts like 'f' on the rest of $\frac{1}{4}^{th}$. This quartile is chosen tactically, as we will see. Now we show that any adversary that gets 'considerably' close to probability '1' of inversion, can invert 'f' with a non-negligible probability, thereby forming a contradiction.

**Construction:** Map $\frac{3}{4}^{th}$ of the input (I) to itself, i.e., g(x) = x and map $\frac{1}{4}^{th}$ of the input (I') to f, i.e., g(x) = f(x). **This subset of domain, the quartile I', is chosen to be non-polynomial**. Such quartile must exist from pigeon hole principle as OWFs (i.e f) cannot have a polynomial range.

**Claim:** G is a Weak One-Way Function.

**Proof:** We prove that it **is a** OWF and that it **is not a strong** OWF.

**It is weak and it is OW:**

We show that there cannot be an adversary that gets within inverse polynomial range to probability of 1. From the definition of weak one-way functions, we have g is a weak OWF if:

$$\exists \text{ polynomial } q(x), \forall PPTA,\ \forall n \in N \text{ such that } Pr[x \leftarrow \{0,1\}^n, A(1^n, f(x)) \in f^{-1}(f(x))] \leq 1 - \frac{1}{q(x)}$$

Suppose, this definition is not satisfied, i.e., we have an inverter A that can invert with a probability that is $> 1 - \frac{1}{q(x)}$ for each polynomial q(X) for at least one n$\in \mathbb{N}$. Then, it has to invert the total range leaving out the g(x) = x edges with a probability $> \frac{1}{4} - \frac{1}{q(x)}$ for all polynomials q(x). All these inversions find a preimage into I' as we did not consider g(x)=x images.

**Arriving at the contradiction:** Now, considering the original function 'f', the same inverter can invert successfully with a probability $> \frac{1}{4} - \frac{1}{q(x)}$. This is because the edges of I' are not modified but only edges from $I \backslash I'$ are put back. Clearly $> \forall q\ \ \frac{1}{4} - \frac{1}{q(x)}$ implies a non-negligible fraction. Therefore, this **contradicts the assumption that f is a OWF.** Hence, no inverter can invert the constructed f(I') with a probability $> \frac{1}{4} - \frac{1}{q(x)}$ and therefore, cannot invert 'g' with probability $> 1 - \frac{1}{q(x)}$. **Therefore, g is weak and g is OW**.

**It is not strong:**

This is the case because a trivial inverter that returns $A^{-1}(x) = x$ would give us a probability of at least $\frac{3}{4}$ which is not negligible. **Hence, g is not strong.**

**Thus, we have constructed g, a weak OWF that is not strong.**