# Lecture Notes 7

## 19 Strongly One-Way from Weakly One-Way Functions

We now complete the proof theorem 1 from lecture 6 by constructing a strongly one-way function $g$ from a weakly one-way function $f$.

Let $f$ be a weakly one-way function with associated polynomial $p(\cdot)$. Assume w.l.o.g. that $f$ is length-preserving. Let $t(n) = n \cdot p(n)$, and let $T = \{n \cdot t(n) \mid n \in \mathbb{N}\}$. Let $g$ be the function on length $n \cdot t(n)$ strings defined by $g(x_1, \ldots, x_{t(n)}) = (f(x_1), \ldots, f(x_{t(n)}))$, where $|x_1|, \ldots, |x_{t(n)}| = n$. That is, given a string $x$ of length $n \cdot t(n)$, $g$ splits it into $t(n)$ length-$n$ strings $x_1, \ldots, x_{t(n)}$, applies $f(\cdot)$ to each, and concatenates the $t(n)$ result strings so obtained.

**Lemma 1** *$g$ is strongly one-way on lengths in $T$.*

**Proof:** Assume $g$ is not strongly one-way on lengths in $T$. We proceed to derive a contradiction.

Since $g$ is assumed not strongly one-way, there exists a p.p.t. algorithm $B'$ and a polynomial $q(\cdot)$ such that for infinitely many $m \in T$,

$$\Pr[B' \text{ inverts } g(U_m)] > \frac{1}{q(m)}. \tag{1}$$

Let $M'$ be the infinite set of integers for which inequality 1 holds, and let $N' = \{n \mid n^2 p(n) \in M'\}$.

We describe a p.p.t. algorithm $A'$ for inverting $f$ on input $y$. First consider the procedure $I'$ for inverting $f$.

Procedure $I'(y)$:
- For $i = 1$ to $t(n)$ do:
  1. Choose $x_1, \ldots, x_{t(n)} \in \{0, 1\}^n$ uniformly and independently.
  2. Compute $(z_1, \ldots, z_{t(n)}) = B'(f(x_1), \ldots, f(x_{i-1}), y, f(x_{i+1}), \ldots, f(x_{t(n)}))$.
  3. If $f(z_i) = y$, then halt and output $z_i$ and declare "success".
- If $f(z_i) \neq y$ for all $i$, then halt and declare "failure".

Now, algorithm $A'(y)$ runs $I'(y)$ repeatedly a total of $a(n) = 2n^2 \cdot p(n) \cdot q(n^2 p(n))$ times. If any of the runs of $I'(y)$ succeed, then $A'$ succeeds and gives the output of the first successful $I'$; otherwise, $A'$ fails.

For $n \in N'$, we will show that

$$\Pr[A' \text{ inverts } f(U_n)] > 1 - \frac{1}{p(n)},$$

contradicting the assumption that $f$ is weakly one-way.

Let

$$S_n = \{x \mid \Pr[I' \text{ inverts } f(x)] > \frac{n}{a(n)}\}$$

be the set of *good* inputs of length $n$. Claim 1 shows that $S_n$ is the set of inputs on which $I'$ succeeds often enough so that $A'$ has an exponentially small failure probability.

**Claim 1** *For all $x \in S_n$, $\Pr[A'$ inverts $f(x)] > 1 - \frac{1}{2^n}$.*

**Proof:** Immediate since for $x \in S_n$,

$$\Pr[A' \text{ fails on } f(x)] < \left(1 - \frac{n}{a(n)}\right)^{a(n)} < \frac{1}{e^n} < \frac{1}{2^n}.$$

■

We now show in Claim 2 that almost all inputs $x$ are in $S_n$ for those lengths $n$ that correspond to values $m = n^2 p(n) \in M'$ on which $B'$ has a success probability greater than $1/q(m)$. (See inequality 1.)

**Claim 2** *For all $n \in N'$,*

$$\frac{|S_n|}{2^n} > 1 - \frac{1}{2p(n)}.$$

**Proof:** Assume to the contrary that

$$|S_n| \leq \left(1 - \frac{1}{2p(n)}\right) \cdot 2^n \tag{2}$$

and let $m = n^2 p(n)$. By inequality 1,

$$s(n) \stackrel{\text{df}}{=} \Pr[B' \text{ inverts } g(U_m)] > \frac{1}{q(m)}. \tag{3}$$

The random variable $U_m$ consists of $n \cdot p(n)$ independent $n$-bit blocks $U_n^{(1)}, \ldots, U_n^{(n \cdot p(n))}$. Define

$$s_1(n) = \Pr[(B' \text{ inverts } g(U_m)) \wedge (\exists i) U_n^{(i)} \notin S_n];$$

$$s_2(n) = \Pr[(B' \text{ inverts } g(U_m)) \wedge (\forall i) U_n^{(i)} \in S_n].$$

Clearly, $s(n) = s_1(n) + s_2(n)$.

We derive a contradiction by showing that both $s_1(n)$ and $s_2(n)$ are bounded from above by $n^2 \cdot p(n)/a(n)$.

Note that

$$\Pr[I' \text{ inverts } f(x)] \geq \Pr[B' \text{ inverts } g(U_m) \mid U_n^{(i)} = x] \tag{4}$$

This is because algorithm $I'$ succeeds on $y = f(x)$ whenever $B'$ succeeds on $g(U_m)$ and $U_n^{(i)} = x$.

Following the text, we have

$$\begin{aligned}
s_1(n) &= \Pr[(\exists i)((B' \text{ inverts } g(U_m)) \wedge U_n^{(i)} \notin S_n)] && (5)\\
&\leq \sum_{i=1}^{n \cdot p(n)} \Pr[(B' \text{ inverts } g(U_m)) \wedge U_n^{(i)} \notin S_n] && (6)\\
&\leq \sum_{i=1}^{n \cdot p(n)} \sum_{x \notin S_n} \Pr[(B' \text{ inverts } g(U_m)) \wedge U_n^{(i)} = x] && (7)\\
&= \sum_{i=1}^{n \cdot p(n)} \sum_{x \notin S_n} \Pr[U_n^{(i)} = x] \cdot \Pr[B' \text{ inverts } g(U_m) \mid U_n^{(i)} = x] && (8)
\end{aligned}$$

$$
\begin{aligned}
&\leq \sum_{i=1}^{n \cdot p(n)} \max_{x \notin S_n} \{\Pr[B' \text{ inverts } g(U_m) \mid U_n^{(i)} = x]\} && (9)\\
&\leq \sum_{i=1}^{n \cdot p(n)} \max_{x \notin S_n} \{\Pr[I' \text{ inverts } f(x)]\} && (10)\\
&\leq n \cdot p(n) \cdot \frac{n}{a(n)} = \frac{n^2 \cdot p(n)}{a(n)}. && (11)
\end{aligned}
$$

Step (10) follows from inequality (4), and step (11) follows from the definition of $S_n$ and the obvious fact that $\Pr[I' \text{ inverts } f(x)] \leq 1$ since all events have probability at most 1.

The following bound on $s_2(n)$ holds for all sufficiently large $n$.

$$
\begin{aligned}
s_2(n) &\leq \Pr[(\forall i) U_n^{(i)} \in S_n] && (12)\\
&\leq \left(1 - \frac{1}{2p(n)}\right)^{n \cdot p(n)} < \frac{1}{2^{n/2}} && (13)\\
&< \frac{n^2 \cdot p(n)}{a(n)} && (14)
\end{aligned}
$$

Here, step (12) follows from the definition of $s_2(n)$, step (13) is by the assumed inequality (2)), and step (14) holds because every positive rational function is greater than an inverse exponential for all sufficiently large $n$.

From (11) and (14), we have

$$
s(n) = s_1(n) + s_2(n) \leq \frac{2n^2 \cdot p(n)}{a(n)} = \frac{1}{q(n^2 p(n))} = \frac{1}{q(m)}.
$$

This contradicts (3), completing the proof of the claim. ■

To finish the proof of the lemma, we observe that

$$
\Pr[U_n \in S_n] \geq 1 - \frac{1}{2p(n)}
$$

follows immediately from claim 2, and

$$
\Pr[A' \text{ inverts } f(U_n) \mid U_n \in S_n] \geq 1 - \frac{1}{2^n}
$$

follows from claim 1 since the bound applies to every $x \in S_n$. Hence,

$$
\begin{aligned}
\Pr[A' \text{ inverts } f(U_n)] &\geq \Pr[(A' \text{ inverts } f(U_n)) \wedge U_n \in S_n]\\
&= \Pr[U_n \in S_n] \cdot \Pr[A' \text{ inverts } f(U_n) \mid U_n \in S_n]\\
&\geq \left(1 - \frac{1}{2p(n)}\right) \cdot \left(1 - \frac{1}{2^n}\right) > 1 - \frac{1}{p(n)}.
\end{aligned}
$$

This contradicts the assumption that $f$ is weakly one-way with associated polynomial $p(\cdot)$, concluding the proof of the lemma. ■