TD 7: Collision-Resistant Hash Functions

Exercise 1.

Suppose $h_1 : \{0,1\}^{2n} \to \{0,1\}^n$ is a collision-resistant hash function.

1. Define $h_2 : \{0,1\}^{4n} \to \{0,1\}^n$ as follows: Write $x = x_1 \| x_2$ with $x_1, x_2 \in \{0,1\}^{2n}$; return the value $h_2(x) = h_1(h_1(x_1) \| h_1(x_2))$. Prove that h_2 is collision-resistant.

2. For $i \geq 2$, define $h_i : \{0,1\}^{2^i n} \to \{0,1\}^n$ as follows: Write $x = x_1 \| x_2$ with $x_1, x_2 \in \{0,1\}^{2^{i-1}n}$; return $h_i(x) = h_1(h_{i-1}(x_1) \| h_{i-1}(x_2))$. Prove that h_i is collision-resistant.

Exercise 2.

1. In the Merkle-Damgård transform, the message is split into consecutive blocks, and we add as a last block the binary representation of the length of this message. Suppose that we do not add this block: does this transform still lead to a collision-resistant hash function?

2. Before HMAC was invented, it was quite common to define a MAC by $\mathrm{Mac}_k(m) = H^s(k \parallel m)$ where H is a collision-resistant hash function. Show that this is not a secure MAC when H is constructed via the Merkle-Damgård transform.

Exercise 3.

Let $m \geq n \geq 2$, $q \geq 2$ and $B > 0$ such that $mB \leq q/4$, with q prime. Recall that the $\mathsf{LWE}_{m,n,q,B}$ hardness assumption states that the distribution $(\mathbf{A}, \mathbf{A}\mathbf{s} + \mathbf{e})$, where $\mathbf{A} \hookleftarrow U(\mathbb{Z}_q^{m \times n})$, $\mathbf{s} \hookleftarrow U(\mathbb{Z}_q^n)$ and $e \hookleftarrow U((-B,B]^m)$ is computationally indistinguishable from $U(\mathbb{Z}_q^{m \times n} \times \mathbb{Z}_q^m)$. Define the following hash function:

$$
\begin{aligned}
H_{\mathbf{A}} : & \{0,1\}^m \to \mathbb{Z}_q^n \\
& \mathbf{x} \mapsto \mathbf{x}^\top \cdot \mathbf{A} \bmod q
\end{aligned}
$$

1. **(a)** Recall the definition of the compression factor, and compute it for H.

   (b) Show how to break the $\mathsf{LWE}_{m,n,q,B}$ assumption given a vector $\mathbf{x} \in \{-1,0,1\}^m$ such that $\mathbf{x}^\top \mathbf{A} = \mathbf{0} \bmod q$ and $\mathbf{x} \neq \mathbf{0}$.

   (c) Conclude on the collision-resistance of H.

Exercise 4.

Pedersen's hash function is as follows:

- Given a security parameter n, algorithm `Gen` samples (G, g, p) where $G = \langle g \rangle$ is a cyclic group of known prime order p. It then sets $g_1 = g$ and samples g_i uniformly in G for all $i \in \{2, \ldots, k\}$, where $k \geq 2$ is some parameter. Finally, it returns $(G, p, g_1, \ldots, g_k)$.
- The hash of any message $M = (M_1, \ldots, M_k) \in (\mathbb{Z}/p\mathbb{Z})^k$ is $H(M) = \prod_{i=1}^k g_i^{M_i} \in G$.

1. Bound the cost of hashing, in terms of k and the number of multiplications in G.
2. Assume for this question that G is a subgroup of prime order p of $(\mathbb{Z}/q\mathbb{Z})^\times$, where $q = 2p+1$ is prime. What is the compression factor in terms of k and q? Which k would you choose? Justify your choice.
3. Assume for this question that $k = 2$. Show that Pedersen's hash function is collision-resistant, under the assumption that the Discrete Logarithm Problem (DLP) is hard for G.
4. Same question as the previous one, with $k \geq 2$ arbitrary.