The goal of this homework is to cover the definition of cryptographic signatures, and there construction from lattice problems. Informally, a signature scheme allows to authenticate the source of a message. Authentication means that one can verify with the public key that a message was indeed signed of the corresponding secret key. More formally

DEFINITION 1 *A signature scheme is a triplet of poly-time algorithms* $(\textsc{KeyGen}, \textsc{Sign}, \textsc{Verif})$ *where:*

- $\textsc{KeyGen}(1^n) \to (pk, sk) \in \mathcal{PK} \times \mathcal{SK}$*: Takes a security parameter* n *as input, and outputs a key-pair*
- $\textsc{Sign}(sk \in \mathcal{SK}, \mu \in \mathcal{M}) \to s \in \mathcal{S}$*: Takes a secret-key and a message as inputs, and outputs a signature*
- $\textsc{Verif}(pk \in \mathcal{PK}, \mu \in \mathcal{M}, s \in \mathcal{S}) \to b \in \{0,1\}$*: Tests the validity of the signature* s *for the message* m *under public key* pk*.*

To avoid consideration related to the so called *random oracle model*[1], we will simply assume that all the messages to be signed are chosen uniformly at random.

DEFINITION 2 *A signature scheme is said correct, if, with overwhelming probability over the choice of* $m \leftarrow \mathcal{M}$*,* $(pk, sk) \leftarrow \textsc{KeyGen}(1^n)$*, and* $s \leftarrow \textsc{Sign}(sk, m)$ *it holds that:*

$$\textsc{Verif}(pk, m, s) = 1.$$

The simplest security notion for a signature scheme follows.

DEFINITION 3 *A signature scheme is said Unforgeable under Key-Only Attack (UF-KOA), if, for any PPT adversary* $\mathcal{A}$*, it holds with overwhelming probability that*

$$\textsc{Verif}(pk, \mu^*, s^*) = 0$$

over the random choices of $(pk, sk) \leftarrow \textsc{KeyGen}(1^n)$*,* $\mu^* \leftarrow \mathcal{M}$ *and* $s^* \leftarrow \mathcal{A}(pk, \mu^*)$*.*

The above notion is nevertheless too weak to be interesting. Indeed, we would also want to prevent an attacker to sign new messages after he sees (polynomially many) other signed messages.

DEFINITION 4 *A signature scheme is said Unforgeable under Random Message Attacks (UF-RMA), if, for any* $t = \text{poly}(n)$*, and any PPT adversary* $\mathcal{A}$*, it holds with overwhelming probability that*

$$\textsc{Verif}(pk, \mu^*, s^*) = 0$$

over the random choices of $(pk, sk) \leftarrow \textsc{KeyGen}(1^n)$*,* $\mu^* \leftarrow \mathcal{M}$*,* $\forall i \leq t, \mu_i \leftarrow \mathcal{M}$*,* $s_i \leftarrow \textsc{Sign}(sk, \mu_i)$*, and* $s^* \leftarrow \mathcal{A}(pk, \mu^*, [(\mu_i, s_i)]_{i=1\ldots t})$*. The list of signed messages* $[(\mu_i, s_i)]_{i=1\ldots t}$ *is referred to as the transcript.*

The homework contains three exercises.

Ex 1 covers the construction of trapdoored SIS instances; namely how to generate random SIS instance together with a short basis of the underlying lattice. This short basis will then be used as a secret key in our signature scheme.

[1] `https://en.wikipedia.org/wiki/Random_oracle`

Ex 2 proposes a tentative construction of a signature scheme. Although we will be able to prove that it is secure under key-only attack, we will also overview why it is not secure when the adversary has seen several message-signature pairs.

Ex 3 proposes to fix the above scheme. This will be achieved by resorting to Discrete Gaussian Sampling (DGS).

Scoring:

- Questions marked with a $*$ are bonus questions: a perfect score can be achieved without them.
- Questions marked with $^-$ have been covered in class, but are left to maintain the integrity of the problem.

Exercise 1 (Generation of Trapdoored SIS instances) The goal of this exercise is to generate an almost uniform matrix $\mathbf{A} \in \mathbb{Z}_q^{n\times m}$ together with a trapdoor, that is, a short basis of the lattice $\Lambda_q^\perp(\mathbf{A})$.

Let $q = 2^{k+1}$, $m = 2\ell$, where $k = O(\log(n))$, $\ell = kn$ and let the *gadget* matrix $\mathbf{G}$ be

$$\mathbf{G} = \begin{bmatrix} 1 & 2 & \cdots & 2^k & & & & & & \\ & & & & 1 & 2 & \cdots & 2^k & & \\ & & & & & & & & \ddots & \\ & & & & & & & & 1 & 2 & \cdots & 2^k \end{bmatrix} \in \mathbb{Z}_q^{n\times\ell}$$

where empty entries denotes 0.

1. Construct a short basis $\mathbf{S_G} \in \mathbb{Z}_q^{\ell\times\ell}$ for the lattice $\Lambda_q^\perp(\mathbf{G})$, more precisely, each vector of $\mathbf{S_G}$ should have norm 2 or $\sqrt{5}$.
2. Give an poly-time algorithm (denoted $\mathbf{G}^{-1}(\mathbf{x})$) that for given $\mathbf{x} \in \mathbb{Z}_q^n$, outputs a binary solution $\mathbf{y} \in \{0,1\}^m$ to $\mathbf{Gy} = \mathbf{x}$. (Hint: There is a solution much Simpler than Running Babai's Nearest Plane Algorithm.)
3. Let $\mathbf{A}_1 \in \mathbb{Z}_q^{n\times\ell}$ be a uniform random matrix. Give an upper bound η_ε on $\eta_\varepsilon(\Lambda_q^\perp(\mathbf{A}_1))$ that holds with overwhelming probability over the choice of $\mathbf{A}_1$.
4. Set $\sigma = \Theta(n)$. Show that, for $\mathbf{R} \leftarrow D_{\mathbb{Z},\sigma}^{\ell\times\ell}$ the matrix $[\mathbf{A}_1, \mathbf{A}_1 \cdot \mathbf{R}] \in \mathbb{Z}_q^{n\times m}$ is statistically instinduishable from uniform. Conclude that $\mathbf{A} = [\mathbf{A}_1, \mathbf{G} - \mathbf{A}_1 \cdot \mathbf{R}] \in \mathbb{Z}_q^{n\times m}$ is statistically instinduishable from uniform.
6. Choosing the matrix $\mathbf{Y} \in \mathbb{Z}^{\ell\times\ell}$ appropriately (and using a poly-time algorithm), construct a short basis of $\Lambda_q^\perp(\mathbf{A})$ of the form

$$\mathbf{S_A} = \begin{bmatrix} \mathrm{Id}_\ell & \mathbf{R} \\ 0 & \mathrm{Id}_\ell \end{bmatrix} \cdot \begin{bmatrix} 0 & \mathrm{Id}_\ell \\ \mathbf{S_G} & \mathbf{Y} \end{bmatrix} = \begin{bmatrix} \mathbf{RS_G} & \mathrm{Id}_\ell + \mathbf{RY} \\ \mathbf{S_G} & \mathbf{Y} \end{bmatrix} \in \mathbb{Z}_q^{m\times m}$$

More precisely, ensure that $\|\mathbf{S_A}\| \le O(n^4)$ with overwhelming probability. (Hint: Ensure that $\mathbf{A} \cdot \mathbf{S_A} = 0 \bmod q$, and then compare determinants to conclude that it is indeed a basis.)

Exercise 2 (A Naive Signature Scheme and an Attack) Keeping the notations from above, we set $\mathcal{PK} = \mathbb{Z}^{n\times m}$, $\mathcal{SK} = \mathbb{Z}^{m\times m}$, $\mathcal{M} = \mathbb{Z}_q^n$, $\mathcal{S} = \mathbb{Z}^m$, and define a signature scheme as follows:

- KeyGen(1^n): Generates a pair $(pk := \mathbf{A}, sk := \mathbf{S_A})$ as in Exercise 1.
- Sign$(\mathbf{S_A}, \mathbf{x})$: Uses Babai's Nearest Plane Algorithm on the short basis $\mathbf{S_A}$ to find a short solution $\mathbf{s}$ to $\mathbf{As} = \mathbf{x} \bmod q$ (say $\|\mathbf{s}\| \leq B$)
- Verif$(\mathbf{A}, \mathbf{x}, \mathbf{s})$: Verify that $\mathbf{As} = \mathbf{x} \bmod q$ and that $\|\mathbf{s}\| \leq B$.

1. Describe explicitly how to use Babai's Nearest-Plane algorithm the Signing procedure, and provide the value for the upper bound B.
2. For appropriate parameters (to be determined), and under the assumption that SIS is hard, prove that the above signature scheme is unforgeable under Key-Only Attack. (Hint: The SIS instance should have dimension $m + 1$.)

We now consider what information can be gained from signatures, and how to use it to recover the secret key.

3. For $\mathbf{s} \leftarrow$ Sign$(\mathbf{S_A}, \mathbf{x})$ where $\mathbf{x} \leftarrow \mathbb{Z}_q^n$ what is the distribution of $\mathbf{s}$ (as a function of $\widetilde{\mathbf{S_A}}$)?
4. What is the value of covariance matrix $\mathbf{C} = \mathbb{E}_{\mathbf{z}\leftarrow\mathcal{P}}(\mathbf{z}\mathbf{z}^t) \in \mathbb{R}^{\ell\times\ell}$ of a uniform variable $\mathbf{z}$ over some centered parallelepiped $\mathcal{P} = \mathcal{P}_{sym}(\mathbf{B})$? Given $\mathbf{C}$, how to compute in poly-time a linear transformation $\mathbf{L}$ such that $\mathbf{H} = \mathbf{L} \cdot \mathbf{B}$ satisfies $\mathbf{H}\mathbf{H}^t = \mathrm{Id}$?

5*. The k^{th} moment of a distribution $\mathcal{D}$ over $\mathbb{R}^\ell$ in direction[2] $\mathbf{v} \in \mathcal{S}^{\ell-1}$ is defined by:

$$M_{k,\mathcal{D}}(\mathbf{v}) = \mathbb{E}_{\mathbf{y}\leftarrow\mathcal{D}}\left[\langle \mathbf{v}, \mathbf{y}\rangle^k\right].$$

If $\mathcal{D}$ is a uniform distribution over a centered hypercube $\mathcal{H} = \mathcal{P}_{sym}(\mathbf{H})$ ($\mathbf{H}\mathbf{H}^t = \mathrm{Id}$), where are the local minima of $M_{k,\mathcal{D}}$ for $k = 4$? (Hint: First consider the case $\mathbf{H} = \mathrm{Id}$.)

6*. Propose a strategy to recover an approximation of GSO $\widetilde{\mathbf{S_A}}$ of the secret key given many signatures of random messages. A full formal proof that the attack works is not required.[3] A discussion on the difficulties toward a full proof of the attack is welcomed.

7. Would you use that scheme if your life depended on its security ? Why ?

Exercise 3 (A Provably Secure Signature Scheme) The security issue above can be viewed as a "statistical leak": each signatures reveals a little bit of information about the secret key. In this exercise, we will provably seal this leak by making the transcript statistically independent from the secret key $\mathbf{S_A}$.

Let us first assume that, given the short basis $\mathbf{S_A}$ that one can sample in poly-time from discrete Gaussian distribution $D_{\Lambda_q^{\perp\mathbf{x}}(\mathbf{A}),\sigma}$ of parameter σ over the lattice coset

$$\Lambda_q^{\perp\mathbf{x}}(\mathbf{A}) = \{\mathbf{y} \text{ s.t. } \mathbf{Ay} = \mathbf{x} \mod q\}.$$

[2] A direction in $\mathbb{R}^\ell$ is a unit vector of $\mathbb{R}^\ell$, i.e. a vector on the sphere $S^{\ell-1}$.

[3] In cryptanalysis, heuristic reasoning is allowed.

1^-. Under which constraint can one claim that the two following distribution are statistically indistinguishable ?

- $(\mathbf{x}, \mathbf{y})$ where $\mathbf{x} \leftarrow \mathbb{Z}_q^n$ and $\mathbf{y} \leftarrow D_{\Lambda_q^{\perp \mathbf{x}}(\mathbf{A}),\sigma}$.
- $(\mathbf{x}, \mathbf{y})$ where $\mathbf{y} \leftarrow D_{\mathbb{Z}^m,\sigma}$ and $\mathbf{x} := \mathbf{A}\mathbf{x} \bmod q$.

(Hint: This question has been covered in class. You can limit your answer to a reference and a summary of the proof.)

2. Modify the parameters and the SIGN algorithm from Exercise 2, and prove its correctness, and its Unforgeability under random-messages attacks using the SIS assumption. (Hint: Get a SIS instance $\mathbf{A}$, forward (part of) it as the public key, simulate the generation of the transcript without knowledge of a secret key $\mathbf{S}_\mathbf{A}$.)

We now move to designing an algorithm to sample discrete gaussian. In the following, one can ignore all issues related to numerical precision, that is, assume that standard operation on $\mathbb{R}$ $(\times, +, -, /, \exp, \log)$ can be done in time $O(1)$.

3. Propose a poly-time algorithm to sample from a distribution statistically indistinguishable from $D_{\mathbb{Z},\sigma,c}$ given $c \in \mathbb{R}$, $\sigma > 0$. (Hint: Ignore the tails, and use rejection sampling[4]. Bonus question: Do not ignore the tails, and sample from $D_{\mathbb{Z},\sigma,c}$ perfectly.)

Now, consider the randomized variant of the Nearest-Plane algorithm given as Algorithm 1.

Algorithm 1: Randomized Nearest-Plane algorithm

Input : A basis $\mathbf{B} = (\mathbf{b}_1, \ldots, \mathbf{b}_n)$ of a lattice Λ, a target $\mathbf{t} \in \mathrm{span}(\Lambda)$, a parameter $\sigma > 0$.
Output: $(\mathbf{v}, \mathbf{e})$ such that $\mathbf{v} + \mathbf{e} = \mathbf{t}$, $\mathbf{v} \in \Lambda$, and $\mathbf{e}$ is small.

$\mathbf{e} := \mathbf{t}$
$\mathbf{v} := 0$
for $i = n$ *down to* 1 **do**
  $c_i := \langle \mathbf{e}, \widetilde{\mathbf{b}}_i \rangle / \|\widetilde{\mathbf{b}}_i\|^2$
  $\sigma_i := \sigma / \|\widetilde{\mathbf{b}}_i\|$
  $k_i \leftarrow D_{\mathbb{Z},\sigma_i,c_i}$
  $\mathbf{e} := \mathbf{e} - k_i \mathbf{b}_i$
  $\mathbf{v} := \mathbf{v} + k_i \mathbf{b}_i$
end
return $(\mathbf{v}, \mathbf{e})$

4^-. Prove that the output $\mathbf{v}$ of Algorithm 1 is statistically indistinguishable from the discrete Gaussian $D_{\mathcal{L}(\mathbf{B}),\mathbf{t},\sigma}$ assuming $\sigma \geq \max_i \|\widetilde{\mathbf{b}}_i\| \cdot \eta_\varepsilon(\mathbb{Z})$. (Hint: This question has been covered in class. You can limit your answer to a reference and a summary of the proof.)

5. Explicit how to use Algorithm 1 for the signing procedure $\mathbf{s} \leftarrow D_{\Lambda_q^{\perp \mathbf{x}}(\mathbf{A}),\sigma}$ of a message $\mathbf{x}$. Give an upper bound for $\mathbf{s}$ that holds with overwhelming probability.

6. Conclude: summarize all the required constraint for the constructed signature scheme to be poly-time, correct, and secure assuming the hardness of SIS.

[4] `https://en.wikipedia.org/wiki/Rejection_sampling`