FINAL EXAM

3h
documents are not allowed
calculators are allowed
answer the two parts on two separate sheets
answers can be written in French or in English

Part I: lattices

Recall that by convention, in this part, we consider column vectors. In other words, when we say that a matrix B generates a lattice $\mathcal{L}$, we mean that the columns of B generate the lattice $\mathcal{L}$.

1 Course exercise

1. Let B and C be two matrices in $\mathrm{GL}_n(\mathbb{R})$. Under which condition do they generate the same lattice, i.e., $\mathcal{L}(B) = \mathcal{L}(C)$? (Give the condition, no justification needed)

2. Let $B = \begin{pmatrix} 2 & 0 \\ 0 & -3 \end{pmatrix}$ and $C = \begin{pmatrix} 10 & 14 \\ 6 & 9 \end{pmatrix}$, do they generate the same lattice?

3. Same question for $B = \begin{pmatrix} 2 & 0 \\ 0 & -3 \end{pmatrix}$ and $C = \begin{pmatrix} 2 & 4 \\ -3 & 0 \end{pmatrix}$.

4. Recall that the SIS problem with parameters m, n, q and β (with $m \geq n$ integers, $q \geq 2$ integer and $\beta \geq 1$ real number) asks, given as input a uniformly random matrix $A \leftarrow \mathrm{Uniform}(\mathbb{Z}_q^{m\times n})$ to output (if it exists) $x \in \mathbb{Z}^m$ such that $x^T A = 0 \bmod q$ and $\|x\| \leq \beta$. Describe a reduction from the SIS problem (for any choice of parameters m, n, q and β) to the exact Shortest Vector Problem (SVP_γ with $\gamma = 1$). (In other words, assuming that we have a polynomial time algorithm $\mathcal{B}$ solving the SVP_γ problem for $\gamma = 1$ in any lattice, describe a polynomial time algorithm $\mathcal{A}$ (using algorithm $\mathcal{B}$) solving the SIS problem.)

5. Using Minkowski's inequality, give an upper bound on $\lambda_1(\mathcal{L})$, where $\mathcal{L} = \mathcal{L}(B)$ is the lattice generated by the basis $B = \begin{pmatrix} -1 & 0 \\ -5 & 8 \end{pmatrix}$.

6. Let $K = \mathbb{Q}[X]/(X^4+1)$ and $\mathcal{O}_K = \mathbb{Z}[X]/(X^4+1)$. Let $\mathcal{M} \subseteq \mathcal{O}_K$ be the rank-1 module (i.e., the ideal) generated by the ($\mathcal{O}_K$)-basis $B = (a) \in {\mathcal{O}_K}^{1\times 1}$, with $a = 2 - X + 3X^2 + X^3 \in \mathcal{O}_K$. What is the ($\mathbb{Z}$-)rank of the module lattice $\Sigma(\mathcal{M})$ associated to $\mathcal{M}$? Give a ($\mathbb{Z}$-)basis of this module lattice $\Sigma(\mathcal{M})$.

2 Problem: NTRU

In this exercise, we will study the NTRU problem over $\mathbb{Z}$ (recall that NTRU is usually defined over the ring of integers $\mathcal{O}_K$ of a number field K; here we will consider the case where $\mathcal{O}_K = \mathbb{Z}$).

The NTRU problem with parameters q and B is defined as follows: given as input $h \in \mathbb{Z}_q$, find, if it exists, $f, g \in \mathbb{Z}$ with g invertible modulo q and $\|(f,g)\| \leq B$ such that $h = fg^{-1} \bmod q$. When a solution (f,g) exists, we say that h is an NTRU instance and that (f,g) is a trapdoor for h.

In all this exercise, we will assume that q is prime and that $B < q$.

1. Let h be an NTRU instance and (f, g) be a trapdoor for h. Is this trapdoor unique? I.e., does there exist no other pair $(f', g') \neq (f, g)$ with $h = f' \cdot (g')^{-1} \bmod q$ and $\|(f', g')\| \leq B$? If yes, prove it. If no, find a counter-example.

2. For $h \in \mathbb{Z}$, define the lattice $\mathcal{L}_h \subseteq \mathbb{Z}^2$ generated by the basis $B_h := \begin{pmatrix} 1 & 0 \\ h & q \end{pmatrix}$ (in columns). Show that if $h = h' \bmod q$, then $\mathcal{L}_h = \mathcal{L}_{h'}$.

Thanks to the previous question, we define $\mathcal{L}_h$ for any $h \in \mathbb{Z}_q$ as the lattice $\mathcal{L}_{\bar{h}}$ where $\bar{h}$ is any representative of h in $\mathbb{Z}$.

3. Let $h \in \mathbb{Z}_q$ be an NTRU instance with trapdoor (f, g). Show that $v := \begin{pmatrix} g \\ f \end{pmatrix}$ is in $\mathcal{L}_h$.

4. Let $h \in \mathbb{Z}_q$, show that if $v := \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \in \mathcal{L}_h$ with $v \neq 0$ and $\|v\| < q$, then $h = v_2 \cdot (v_1)^{-1} \bmod q$. (Don't forget that q is prime).

5. Combining the previous two questions, show that if h is an NTRU instance, then a shortest non-zero vector $v = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$ of $\mathcal{L}_h$ provides a trapdoor $(f, g) := (v_2, v_1)$ to the NTRU instance h. (Don't forget that q is prime and that $B < q$).

6. Using the previous questions, describe a polynomial time reduction from the NTRU problem to the exact Shortest Vector Problem (SVP$_\gamma$ with $\gamma = 1$) in lattices of dimension 2.

7. [Application] Let $q = 127$, $B = 5$ and $h = 62 \bmod q$. Is h an NTRU instance? If yes, compute a trapdoor (f, g) for h. (You may want to transform the problem into an SVP instance using the previous question, and then compute a shortest vector of the rank-2 lattice obtained using, e.g., Lagrange-Gauss algorithm.)

Part II: codes

3 Finding codewords of small weight

Let C be a code of even length n defined by a parity-check matrix $\mathbf{H}$ with $n/2$ rows. Let $\mathbf{h}_1, \mathbf{h}_2, \dots, \mathbf{h}_n$ be the columns of $\mathbf{H}$. Let $J \subset [1, n]$ be a subset of indices of cardinality $|J| = n/2$. We suppose that $\mathbf{H}$ is in systematic form, meaning that the submatrix $\mathbf{H}_J$ is the $n/2 \times n/2$ identity matrix, where $\mathbf{H}_J$ denotes the submatrix of $\mathbf{H}$ made up of the columns $\mathbf{h}_j$, $j \in J$. Assume that the index set J has been chosen randomly. Let $\mathbf{x}$ be a codeword of C of weight d.

1. What is the probability that $|\operatorname{supp}(\mathbf{x}) \cap J| = d - 1$? You may give an approximate value corresponding to d fixed and n tending to infinity. How do we recognise we are in the situation where $|\operatorname{supp}(\mathbf{x}) \cap J| = d - 1$?

2. What is the approximate cost of finding a codeword of weight d in this way?

3. Assuming that for any given J, computing the associated parity-check matrix $\mathbf{H}$ costs n^3 binary operations, is it more or less advantageous to look for a subset J for which $|\operatorname{supp}(\mathbf{x}) \cap J| = d-2$? And for $|\operatorname{supp}(\mathbf{x}) \cap J| = d-3$?

4 A cryptosystem

Let n be a multiple of 4. Let $\mathbf{E}$ be a binary $n/4 \times n$ matrix, where every column has weight $w = o(n)$. The matrix $\mathbf{E}$ is chosen randomly and uniformly under this constraint. Note that the average row weight of $\mathbf{E}$ is therefore $4w$. We define the code $C = \{\mathbf{x} \in \mathbb{F}_2^n,\ \mathbf{E}\mathbf{x}^\mathsf{T} = 0\}$.

1. Let $\mathbf{G}$ be a fixed, randomly chosen, generator matrix of C. Let $k = 3n/4$. From the plaintext $\mathbf{m} \in \mathbb{F}_2^k$ we create a ciphertext through the correspondence
$$\mathbf{m} \mapsto \mathbf{y} = \mathbf{m}\mathbf{G} + \mathbf{e} \tag{1}$$
where $\mathbf{e} \in \mathbb{F}_2^n$ is a vector of small weight t. How can we decipher (recover $\mathbf{m}$) from $\mathbf{y}$ with the help of the secret key $\mathbf{E}$? For this to work, how should the parameters t and w be chosen? (Give approximate values for t and w).

2. Let $\mathbf{A}$ be a *uniform* random $n/4 \times n$ binary matrix. We now define the code C as $C = \{\mathbf{x} \in \mathbb{F}_2^n,\ \mathbf{E}\mathbf{x}^\mathsf{T} = 0 \text{ and } \mathbf{A}\mathbf{x}^\mathsf{T} = 0\}$, and let $\mathbf{G}$ be again a random generator matrix for C What is now the dimension of the message (plaintext) space if we continue to apply (1) to create a ciphertext?

3. We now suppose:

   - There does not exist an efficient algorithm that, given a random binary code of length n and dimension $n/2$, is able to recover a uniform random codeword $\mathbf{c}$ from $\mathbf{c} + \mathbf{e}$, where $\mathbf{e}$ is uniform random of weight t.
   - There does not exist an efficient algorithm $\mathcal{A}$ that
     - takes as input two $n/4 \times n$ matrices $\mathbf{A}$ and $\mathbf{B}$, where $\mathbf{A}$ is guaranteed to be uniformly random, and where $\mathbf{B}$ is guaranteed to be chosen
       (i) either uniformly random and independent of $\mathbf{A}$,
       (ii) or of the form $\mathbf{B} = \mathbf{S}\mathbf{A} + \mathbf{E}$, where $\mathbf{S}$ is uniform random of order $n/4 \times n/4$ and where $\mathbf{E}$ is constructed as before.
     - decides with a non-negligible advantage over a random coin-flip wether $\mathbf{B}$ has been created with distribution (i) or distribution (ii).

   Under this assumption, prove the security of the cryptosystem introduced in the previous question, assuming that the plaintext $\mathbf{m}$ is uniformly random.

4. We now drop the condition that $\mathbf{m}$ be uniformly random, and suppose that $\mathbf{m} \in \{\mathbf{m}_0, \mathbf{m}_1\}$, where $\mathbf{m}_0, \mathbf{m}_1$ are two fixed messages known to the adversary. How do you break the cryptosystem in this case?

5 Another cryptosystem

Let A be the ring $A = \mathbb{F}_2[X]/(X^n + 1)$ and let $\mathbf{h}$ be a element of A chosen randomly and uniformly in A and made public. We identify binary n-tuples $(a_0, \ldots, a_{n-1})$ with elements of A represented by the polynomial $a_0 + a_1X + \cdots + a_{n-1}X^{n-1}$.

1. If $\mathbf{a}$ and $\mathbf{b}$ are two elements of A of Hamming weight w_a and w_b, show that the element $\mathbf{ab}$ of A has Hamming weight at most $w_a w_b$.

2. We propose the following cryptosystem: the public key consists of P and $\mathbf{G}$, where
   - P is an element of A of the form $P = \mathbf{hb} + \boldsymbol{\beta}$, where $\mathbf{b}$ and $\boldsymbol{\beta}$ are elements of A of (small) weight t,
   - $\mathbf{G}$ is a $k \times n$ generator matrix of an error-correcting code C that comes with an efficient decoding algorithm. It can be a Goppa code for example.

Let $\mathbf{m} \in \mathbb{F}_2^k$ be the plaintext. The ciphertext is constructed as:

$$\mathcal{C}(\mathbf{m}) = (\mathbf{ha} + \alpha, \mathbf{mG} + \mathbf{a}P)$$

Show how to decipher with knowledge of the secret key $\mathbf{b}$, and with suitable assumptions on

- the parameter t,
- the number of errors that the decoding algorithm for C should be guaranteed to decode.

3. Does the decoding algorithm for the code C need to be secret?