Sampling algorithm from a discrete normal distribution over lattices

Morgane Guerreau

UVSQ

2021

Sommaire

1. Theoretical explanations
   - Introduction to lattices
   - Falcon Signature Scheme
   - Fast Fourier Nearest Plane
   - Why we need trapdoor samplers
2. Program demonstration

What is a lattice ?

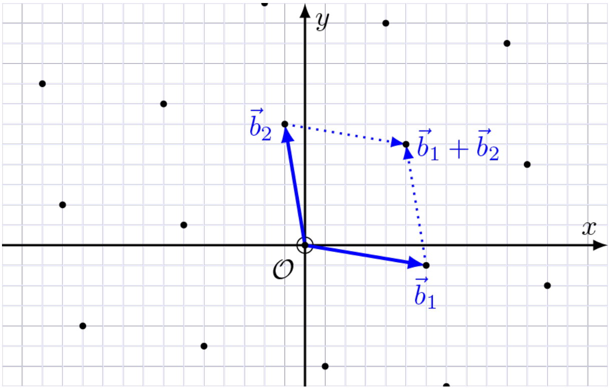
Definition

Let $H = \mathbb{R}^m$. A lattice is a discrete subgroup of H. For a basis $B = \{b_1, ..., b_n\} \in H^n$, we note $L(B)$ and call lattice generated by B the set of vectors

$$\left\{ \sum_{i=1}^{n} x_i b_i | x_i \in \mathbb{Z} \right\}$$

A lattice will be noted Λ or $L(B)$ when the basis B will be provided.

What is a lattice ?



Lattices problem

Definition

The i-th successive minimum λ_i of a lattice Λ is defined as the minimum radius $r \in \mathbb{R}$ of a n-dimensional sphere B with center 0 that contains i linearly independent lattice vectors:

$$\lambda_i(\Lambda) = min\{r | \dim(span(\Lambda \cap B_{r,0})) \geq i\}$$

Lattices problem

Definition (SVP - Shortest Vector Problem)

Given a n-dimensional lattice Λ, find a lattice vector v such that $||v|| = \lambda_1(\Lambda)$.

Definition (CVP - Closest Vector Problem)

Given a n-dimensional lattice Λ and a point $c \in H$, find a lattice vector v such that $||c - v|| = dist(c, \lambda) = \min_{z \in \Lambda} ||c - z||$.

Definition ($\text{SIS}_{n,m,q,\beta}$ - Shortest Integer Solution)

Let n and m,q = poly(n) be integers. Given a uniformly random matrix $A \in \mathbb{Z}_q^{n \times m}$, find a non-zero vector z such that $Az = 0 \mod q$ and $||z|| \leq \beta$.

Falcon Signature Scheme

- Lattice-based signature scheme
- Post-Quantum Cryptography (NIST finalist)
- Problem SIS
- GPV Framework
    - Map the message to a point of the space
    - Find the closest vector in the lattice

Falcon Signature Scheme

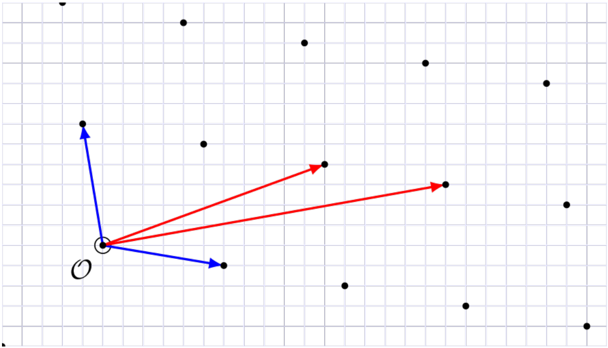


Figure: "Good" basis in blue, "bad" basis in red

Falcon Signature Scheme

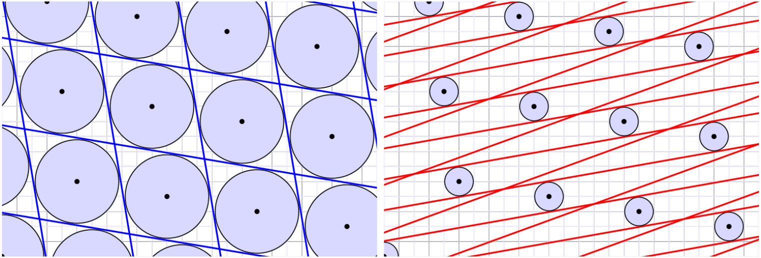


Figure: "Good" basis in blue, "bad" basis in red

NTRU lattices

- Lattice : discrete subgroup of a ring $R = \mathbb{Z}[x]/(\phi)$ with $\phi = x^n + 1$ and n being a power of two, $q \in \mathbb{N}^*$
- $F = \sum_{i=0}^{n-1} F_i x^i = (F_0, F_1, ..., F_{n-1})$
- Private key: $\begin{bmatrix} f & g \\ F & G \end{bmatrix}$ with $fG - gF = q \mod \phi$
- Public key: $\begin{bmatrix} 1 & h \\ 0 & q \end{bmatrix}$ with $h = g \cdot f^{-1} \mod q$

GPV Framework

- Public basis: $A = \begin{bmatrix} 1 & h^* \end{bmatrix}$
- Private basis: $B = \begin{bmatrix} g & -f \\ G & -F \end{bmatrix}$
- $B \times A^* = 0 \mod q$
- $\Lambda_q = L(A)$ and $\Lambda_q^{\perp} = L(B)$, with $\Lambda_q^{\perp}$ orthogonal to Λ_q
- Trapdoor sampler: takes in input a matrix A and a target c, finds a short vector s such that $s^t A = c \mod q$

Gram-Schmidt Orthogonalization

Lemma

Let $H = \mathbb{R}^m$ and $B = \{b_1, ..., b_n\} \in H^n$ be a basis. For any $k \in [1, n]$, we note $V_k = Span(B_k)$. There is a unique basis $\tilde{B} = \tilde{b}_1, ..., \tilde{b}_n \in H^n$ verifying any of these equivalent properties:

1. $\forall k \in [1, n], \tilde{b}_k = b_k - Proj(b_k, V_{k-1})$
2. $\forall k \in [1, n], \tilde{b}_k = b_k - \sum_{j=1}^{k-1} \frac{\langle b_k, \tilde{b}_j \rangle}{\langle \tilde{b}_j, \tilde{b}_j \rangle} \tilde{b}_j$
3. $\forall k \in [1, n], \tilde{b}_k \perp V_{k-1}$ *and* $(b_k - \tilde{b}_k) \in V_{k-1}$

Proposition

Let $B \in R^{n \times m}$ be a full-rank matrix. B can be uniquely decomposed as $B = L \cdot \tilde{B}$ where L is unit lower triangular, and the rows of $\tilde{B}$ are pairwise orthogonal.

Babai Nearest Plane

Definition

Let $B = \{b_1, ..., b_n\}$ be a real basis. We call fundamental parallelepiped generated by B and note P(B) the set $\sum_{i \leq j \leq n} [-\frac{1}{2}, \frac{1}{2}]\ b_j = [\frac{1}{2}, \frac{1}{2}]^n \cdot B$.

Algorithm 1 NearestPlane$_R(t, L)$

1: $z \leftarrow 0$
2: **for** j from n to 1 **do**
3: $\quad \bar{t}_j \leftarrow t_j + \sum_{i>j}(t_i - z_i)L_{ij}$
4: $\quad z_j \leftarrow \lfloor \bar{t}_j \rceil$
5: **end for**
6: **return** z

Factorization of L

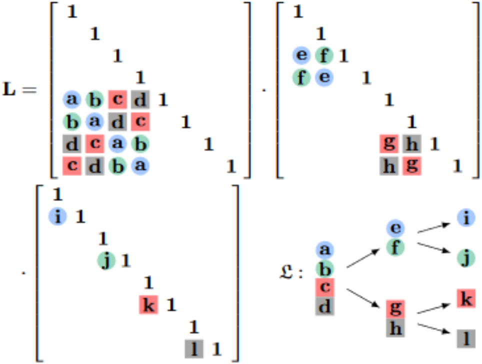


Figure: Factorization of L with L being the lower triangular matrix of the LDL decomposition

Coefficient vector and coefficient matrix

Definition

For any $d \in \mathbb{N}^*$, let R_d denote the ring $\mathbb{R}[x]/(x^d - 1)$, also known as circular convolution ring, or simply convolution ring.

Definition

For any $a = \sum_{i \in \mathbb{Z}_d} a_i x^i \in R_d$ where each $a_i \in \mathbb{R}$:

1. The coefficient vector of a is $c(a) = (a_0, ..., a_{d-1})$.
2. The circulant matrix of a is

$$C(a) = \begin{bmatrix} a_0 & a_1 & \dots & a_{d-1} \\ a_{d-1} & a_0 & \dots & a_{d-2} \\ \vdots & \vdots & \ddots & \vdots \\ a_1 & a_1 & \dots & a_0 \end{bmatrix} = \begin{bmatrix} c(a) \\ c(xa) \\ \vdots \\ c(x^{d-1}a) \end{bmatrix} \in \mathbb{R}^{d \times d}$$

Vectorization and Matrixification

Definition

Let d, d'$\in \mathbb{N}^*$ such that $d|d'$. We define the vectorization $V_{d/d'} : R_d^{n\times m} \to R_d^{x\times m(d/d')}$ inductively as follows:
Let k = d/gpd(d). For d' = gpd(d) and a single element $a \in R_d$, $a = \sum_{0\leq i\leq k_d} x^i a_i(x^k)$ where $a_i \in R_{d'}$ for each i. Then

$$V_{d/d'}(a) = (a_0, \ldots, a_{k-1}) \in R_{d'}^k$$

In other words, $V_{d/d'}(a)$ is the row vector whose coefficients are the $(a_i)_{0\leq i\leq k_d}$.

Vectorization and Matrixification

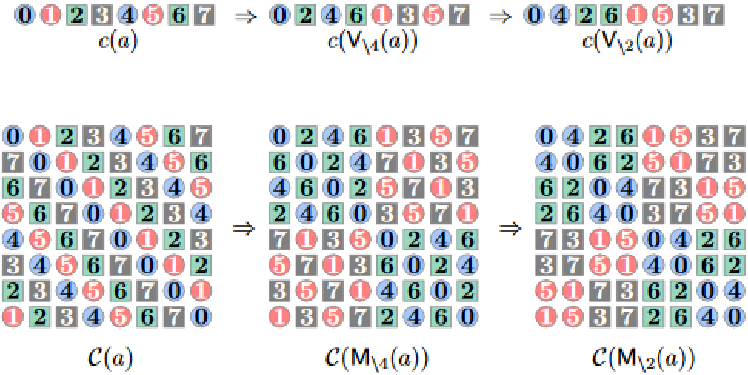
Definition

Following the notations of previous definition, we define the matrixification $M_{d/d'} : R_d^{n\times m} \to R_d^{n(d/d')\times m(d/d')}$ as follows:
Let k = d/gpd(d). For d' = gpd(d) and a single element $a \in R_d$, $a = \sum_{0\leq i\leq k_d} x^i a_i(x^k)$ where $a_i \in R_{d'}$ for each i. Then

$$M_{d/d'}(a) = \begin{bmatrix} a_0 & a_1 & \dots & a_{k-1} \\ xa_{k-1} & a_0 & \dots & a_{k-2} \\ \vdots & \vdots & \ddots & \vdots \\ xa_1 & xa_2 & \dots & a_0 \end{bmatrix} = \begin{bmatrix} V_{d/d'}(a) \\ V_{d/d'}(x^k a) \\ \vdots \\ V_{d/d'}(x^{(d'-1)k} a) \end{bmatrix} \in R_{d'}^{nk\times mk}$$

In particular, if d is prime, the $M_{d/1}(a) \in \mathbb{R}^{d\times d}$ is exactly the circulant matrix $C(a)$.

Vectorization and Matrixification



Factorization of L

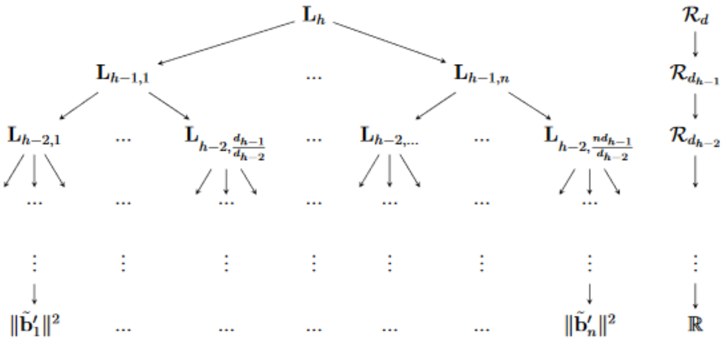
Theorem

Let $d \in \mathbb{N}$ and $1 = d_0 | d_1 | \ldots | d_h = d$ be a tower of proper divisors of d. Let $b \in R_d^m$ such that $M_{d/1}(b)$ is full-rank. There exists a GSO of $M_{d/1}$ as follows:

$$M_{d/i}(b) = \left(\prod_{i=0}^{h-1} M_{d_i/1}(L_i) \right) \cdot \tilde{B}_0$$

where $\tilde{B}_0 \in \mathbb{R}^{d \times dm}$ is orthogonal, and each $L_i \in R_{d_i}^{(d/d_i)\times(d/d_i)}$ is a block-diagonal matrix with unit lower triangular matrices of $R_{d_i}^{(d_{i+1}/d_i)\times(d_{i+1}/d_i)}$ as its d/d_{i+1} diagonal blocks.

Factorization of L



Factorization of L

Algorithm 2 $\mathrm{ffLDL}_{R_d}(G)$

1: **if** d = 1 **then**
2: **return** (G, [ ])
3: **end if**
4: $(L, D) \leftarrow \mathrm{LDL}_{R_d}(G)$
5: **for** i from 1 to n **do**
6: $T_i \leftarrow \mathrm{ffLDL}_{R_{\mathrm{gpd}(d)}}(M_{d/\mathrm{gpd}(d)}(D_{ii}))$
7: **end for**
8: **return** $(L, (T_i)_{1 \leq i \leq n})$

Fast Fourier Nearest Plane

Algorithm 3 $\text{ffNearestPlane}_{R_d}(t, T)$

if t is a 1-dimensional vector in $\mathbb{R}$ **then**
    return $\lfloor t \rceil$
end if
$\text{L} \leftarrow \text{T.Node()}$
for j from n to 1 **do**
    $\bar{t}_j \leftarrow t_j + \sum_{i>j}(t_i - z_i)L_{ij}$
    $z_j \leftarrow V^{-1}_{d/\text{gpd(d)}} \left[\text{ffNearestPlane}_{R_{\text{gpd(d)}}}(V_{d/\text{gpd(d)}}(\bar{t}_j), \text{T.Child(j)}) \right]$
end for
return z

The hidden parallelepiped problem

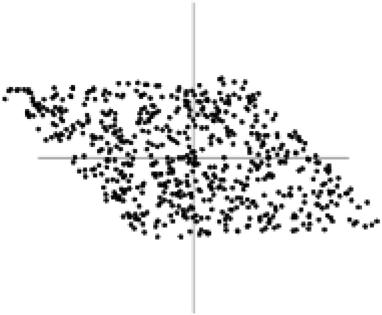
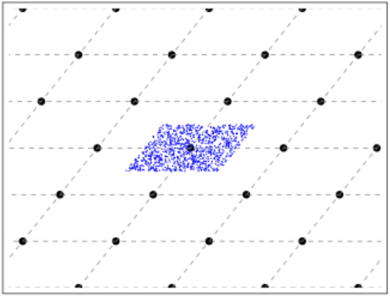
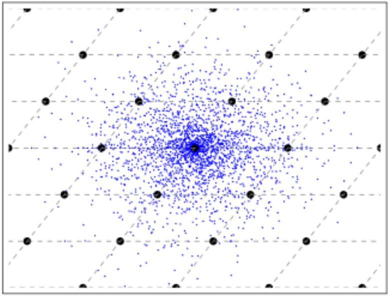


Figure: The hidden parallelepiped problem

The hidden parallelepiped problem



GGH [GGH97], NTRUSign[HHGP+03]



"GPV" framework [GPV08]

Program demonstration

```
make TEST=true
./ffsampling
valgrind --leak-check=full ./ffsampling
make clean && make
./ffsampling [dim] [-s|-v] [message]
```