Théorie Algorithmique des Nombres

Basé sur le cours de Christophe Ritzenthaler
Université de Rennes 1
Mathématiques de l'Information, Cryptographie
Master 2

Elie Noumon Allini

Table des matières

Chapitre 1

Réseaux et réductions

1.1 Réseaux euclidiens

1.1.1 Définitions

Réseau

Un réseau (Λ, q) est un $\mathbb{Z}$-module libre Λ de rang fini avec une forme quadratique définie positive q sur $\Lambda \otimes_{\mathbb{Z}} \mathbb{R}$.
Si $(b_i)_{1\leqslant i\leqslant n}$ une famille libre de $\mathbb{R}^d$, avec $n \leqslant d$, alors on a :

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

En général, on considère $n = d$.

Norme d'un vecteur

Soit $(e_i)_{1\leqslant i\leqslant n}$ une base orthonormale de $\mathbb{R}^n$, alors pour tout $x = (x_i)_{1\leqslant i\leqslant n}$ dans Λ, on prendra :

$$q(x) = \sum_{i=1}^{n} x_i^2.$$

On définit alors la norme de x par :

$$\|x\| = \sqrt{q(x)} = \left(\sum_{i=1}^{n} x_i^2 \right)^{1/2}.$$

Ceci permet d'associer à la forme quadratique q un produit scalaire $\langle x, y \rangle$.

1.1.2 Orthogonalisation de Gram-Schmidt (GSO)

Soit $(b_i)_{1\leqslant i\leqslant n}$ une base de $\mathbb{R}^n$, alors pour en déduire une base orthogonale, on pourra procéder comme suit :

1) on pose $b_1^* = b_1$

2) pour $i > 1$, on pose $b_i^* = b_i - \sum_{j=1}^{i-1} \mu_{ij} b_j^*$, avec $\mu_{ij} = \frac{\langle b_i, b_j^* \rangle}{\langle b_j^*, b_j^* \rangle}$.

La famille $(b_i^*)_{1\leqslant i\leqslant n}$ ainsi construite est alors une base orthogonale de $\mathbb{R}^n$. Afin de prouver ce résultat, on procéra par récurrence.

- Par définition, la famille (b_1^*) est orthogonale.
- Supposons que pour i fixé, la famille $(b_r^*)_{1\leqslant r\leqslant i-1}$ soit orthogonale, alors

on a $\langle b_r^*, b_s^* \rangle = \delta_r^s \langle b_r^*, b_r^* \rangle$, pour $r, s \in [\![1, i-1]\!]$. Ainsi :

$$\begin{aligned}
\langle b_i^*, b_r^* \rangle &= \left\langle b_i - \sum_{j=1}^{i-1} \mu_{ij} b_j^*, b_r^* \right\rangle \\
&= \langle b_i, b_r^* \rangle - \sum_{j=1}^{i-1} \mu_{ij} \langle b_j^*, b_r^* \rangle \\
&= \langle b_i, b_r^* \rangle - \mu_{ir} \langle b_r^*, b_r^* \rangle \\
&= \langle b_i, b_r^* \rangle - \langle b_i, b_r^* \rangle, \quad \text{car } \mu_{ir} = \frac{\langle b_i, b_r^* \rangle}{\langle b_r^*, b_r^* \rangle} \\
&= 0.
\end{aligned}$$

Ceci prouve $b_i^* \perp b_r^*$, pour $r < i$, et donc que $(b_r^*)_{1 \leqslant r \leqslant i}$ soit orthogonale.

Remarquons qu'inversement si $b_i = c_i + \sum_{j<i} \alpha_{ij} c_i$ et que les c_i sont orthogonaux alors on a en fait que $c_i = b_i^*$ par construction.
Désignons par B (respectivement B^*) la matrice en ligne relativement à la base canonique de la famille $(b_i)_{1 \leqslant i \leqslant n}$ (respectivement $(b_i^*)_{1 \leqslant i \leqslant n}$) et posons :

$$A = \begin{pmatrix} 1 & \mu_{2,1} & \cdots & \mu_{n,1} \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mu_{n,n-1} \\ 0 & \cdots & 0 & 1 \end{pmatrix}$$

alors on a $B = B^* A$.

Il est à noter que les familles $(b_i)_{1 \leqslant i \leqslant r}$ et $(b_i^*)_{1 \leqslant i \leqslant r}$ engendrent le même sous-espace vectoriel.

On déduit alors que pour $k > i$, on a $b_k^* \perp \langle b_1, \cdots, b_i \rangle$.

Remarque

On évite l'orthonormalisation pour ne pas prendre les racines carrées. Ainsi, si les b_i sont à coefficients dans $\mathbb{Q}$, alors les b_i^* le sont aussi.

1.1.3 Déterminant d'un réseau

Matrice de Gram

On appelle matrice de Gram de la base $(b_i)_{1 \leqslant i \leqslant n}$, la matrice :

$$\text{Gram}(b_1, b_2, \cdots, b_n) = \left(\langle b_i, b_j \rangle \right)_{1 \leqslant i,j \leqslant n}.$$

On peut écrire :

$$\text{Gram}(b_1, b_2, \cdots, b_n) = {}^tBB.$$

On déduit alors que $\text{Gram}(b_1, b_2, \cdots, b_n)$ est symétrique définie positive. Ceci permet d'écrire :

$$\det(\Lambda) = \sqrt{\det(\text{Gram}(b_1, b_2, \cdots, b_n))}.$$

Proposition 1.1

Le réel $\det(\Lambda)$ est bien défini et est égale au produit des $\|b_i^*\|$.

Preuve :

- Soient $(b_i)_{1 \leqslant i \leqslant n}$ et $(b_i')_{1 \leqslant i \leqslant n}$ deux bases de Λ dont les matrices respectives sont B et B', alors il existe une matrice U telle que $B' = UB$ et $\det(U) = \pm 1$ [1].
  On a alors :
  $${}^tB'B' = {}^t(UB)(UB) = {}^tB\,{}^tUUB.$$
  On a donc $\det B' = \det B$.
- Puisque nous avons $B = AB^*$, on peut écrire ${}^tB = {}^tB^*\,{}^tA$. On déduit alors :
  $$\text{Gram}(b_1, b_2, \cdots, b_n) = B\,{}^tB = AB^*\,{}^tB^*\,{}^tA.$$
  Comme $\det(A) = 1$, on a :
  $$\det(\text{Gram}(b_1, b_2, \cdots, b_n)) = \det(B^*\,{}^tB^*) = \det\left(\langle b_i^*, b_j^* \rangle \right)_{i,j}.$$

1. Voir l'exercice 1 de la section (**??**).

Puisque $(b_i^*)_{1\leqslant i\leqslant n}$ est une base orthogonale, on a $\langle b_i^*, b_j^*\rangle = \|b_i^*\|^2\delta_{i,j}$, et donc :

$$\big(\langle b_i^*, b_j^*\rangle\big)_{i,j} = \operatorname{diag}\big(\|b_1^*\|^2, \|b_2^*\|^2, \cdots, \|b_n^*\|^2\big).$$

On a alors :

$$\det(\operatorname{Gram}(b_1, b_2, \cdots, b_n)) = \prod_{i=1}^{n}\|b_i^*\|^2,$$

d'où : $\det(\Lambda) = \prod_{i=1}^{n}\|b_i^*\|$. ■

Remarque

Puisque $\det(\Lambda)$ est bien défini, on peut le considérer comme le volume d'un domaine fondamental : $\left\{\sum_i \lambda_i b_i,\ 0\leqslant\lambda_i<1\right\}$.

Corollaire 1.1 (Inégalité d'Hadamard)

On a : $|\det B| \leqslant \prod_{i=1}^{n}\|b_i\|$.

Preuve : D'après la proposition (1.1), on peut écrire $\det(\Lambda)^2 = \prod_{i=1}^{n}\|b_i^*\|^2$. Comme nous avons $b_i^* = b_i - \sum_{j=1}^{i-1}\mu_{ij}b_j^*$, on peut écrire $b_i = b_i^* + \sum_{j<i}\mu_{ij}b_j^*$, et déduire alors :

$$\|b_i\|^2 = \|b_i^*\|^2 + \sum_{j<i}|\mu_{ij}|^2\|b_j^*\|^2 \geqslant \|b_i^*\|^2.$$

On a donc :

$$|\det B| = \prod_{i=1}^{n}\|b_i^*\| \leqslant \prod_{i=1}^{n}\|b_i\|.$$

■

1.1.4 Minimum d'un réseau

On appelle minimum de Λ, et on note $\lambda(\Lambda)$, la longueur minimale des vecteurs non nul de Λ.

Etant donnée une base (b_i), trouver b tel que $\|b\| = \lambda(\Lambda)$ est un problème NP-complet.

Soient $x, y \in \mathbb{R}^n$, on écrira $x \equiv y \mod \Lambda$ si, et seulement si, il existe $b \in \Lambda$ tel que $x = y + b$.

Lemme 1.1

Soit C une partie mesurable de $\mathbb{R}^n$ telle que $\operatorname{vol}(C) > \det(\Lambda)$, alors il existe $x, y \in C$ distincts tels que $x \equiv y \mod \Lambda$.

Preuve : Soit (b_i) est une base de Λ et $\mathscr{F} = \left\{\sum_i \lambda_i b_i,\ 0\leqslant\lambda_i<0\right\}$ un domaine fondamental tels que $\operatorname{vol}(\Lambda) = \operatorname{vol}(\mathscr{F})$. Pour $x \in \Lambda$, on pose $C_x = (C-x)\cap\mathscr{F}$. Notons que $C\cap(\mathscr{F}+x)$ est le translaté de $(C-x)\cap\mathscr{F}$ [2] par x, alors on a :

$$\operatorname{vol}(C_x) = \operatorname{vol}(C\cap(\mathscr{F}+x))\ [3].$$

- Supposons les C_x disjoints, alors on a $\operatorname{vol}\left(\bigcup_{x\in\Lambda} C_x\right) = \sum_{x\in\Lambda}\operatorname{vol}(C_x)$, donc :

$$\operatorname{vol}\left(\bigcup_{x\in\Lambda} C_x\right) = \sum_{x\in\Lambda}\operatorname{vol}(C_x) = \sum_{x\in\Lambda}\operatorname{vol}(C\cap(\mathscr{F}+x)).$$

Puisque les $x\in\Lambda$ les $\mathscr{F}+x$ forment une partition de $\mathbb{R}^n$. Alors :

$$\operatorname{vol}\left(\bigcup_{x\in\Lambda} C_x\right) = \operatorname{vol}\left(\bigcup_{x\in\Lambda} C\cap(\mathscr{F}+x)\right) = \operatorname{vol}(C\cap\mathbb{R}^n) = \operatorname{vol}(C).$$

En outre, $\mathscr{F}$ contient tous les C_x, on peut alors écrire :

$$\operatorname{vol}\left(\bigcup_{x\in\Lambda} C_x\right) \leqslant \operatorname{vol}(\mathscr{F}) = \det(\Lambda).$$

On déduit alors $\operatorname{vol}(C)\leqslant\det(\Lambda)$, ce qui contredit $\operatorname{vol}(C)>\det(\Lambda)$.

2. On a :

$$y\in C\cap(\mathscr{F}+x) \iff \begin{cases} y\in C \\ y = \sum_i \lambda_i b_i + x \end{cases} \iff \begin{cases} y-x\in C-x \\ y-x\in\mathscr{F}\end{cases} \iff y\in C\cap\mathscr{F}+x.$$

3. Les translations sont des isométries.

- Ainsi les C_x ne sont pas disjoints, alors il existe $x, y \in \Lambda$ tels que $C_x \cap C_y \neq \emptyset$. Dans ce cas, il est possible de trouver $c_1, c_2 \in C$ tels que $c_1 - x = c_2 - y$, ce qui donne $c_1 = c_2 + x - y$. Comme on a $x, y \in \Lambda$, on déduit $c_1 \equiv c_2 \mod \Lambda$. ■

Théorème 1.1 (Minkowski)

Soit C un convexe symétrique de $\mathbb{R}^n$ tel que $\text{vol}(C) > 2^n \det(\Lambda)$, alors il existe un vecteur non nul de Λ dans C.

Preuve : Puisque nous avons $\text{vol}(C) > 2^n \det(\Lambda)$, on a :

$$\text{vol}\left(\frac{C}{2}\right) = \frac{\text{vol}(C)}{2^n} > \det(\Lambda).$$

D'après le lemme (1.1), il existe alors c_1, c_2 distincts dans C et $\lambda \in \Lambda \setminus \{0\}$ tels que : $\frac{c_1}{2} = \frac{c_2}{2} + \lambda$, ce qui donne :

$$\lambda = \frac{1}{2}(c_1 - c_2).$$

Comme C est symétrique, on a $-c_2 \in C$, ainsi que $\frac{1}{2}(c_1 - c_2) \in C$ par convexité. On déduit donc que λ est un élément non nul de Λ dans C. ■

Corollaire 1.2 (Variante de Minkowski)

Soit C une partie convexe, compact et symétrique de $\mathbb{R}^n$ tel que $\text{vol}(C) \geqslant 2^n \det(\Lambda)$, alors il existe un vecteur non nul de Λ dans C.

Preuve : Désignons par C_k l'image de C par l'homothétie de rapport $k \in \mathbb{N} \setminus \{0\}$, alors C_k est un convexe symétrique et on a :

$$\text{vol}\left(C_{1+1/k}\right) = \text{vol}\left(\frac{C}{1+\frac{1}{k}}\right) = \frac{\text{vol}\, C}{\left(1+\frac{1}{k}\right)^n} > \frac{2^n}{\left(1+\frac{1}{k}\right)^n} \det(\Lambda) \geqslant \det(\Lambda),$$

car $2 \geqslant 1 + \frac{1}{k}$.

- D'après le lemme (1.1), il existe $c_1, c_2 \in C_{1+1/k}$ distincts et $x, y \in \Lambda$ tels que $c_1 - x = c_2 - y$, ce qui donne $c_1 - c_2 = x - y$. Puisque $C_{1+1/k}$ est un convexe symétrique, on a $c_1 - c_2 \in C_{1+1/k}$. On a également $x - y \in \Lambda$, car il est un groupe additif, on déduit donc $c_1 - c_2 \in \Lambda$. Vu que c_1 et c_2 sont distincts, on conclut $c_1 - c_2$ est non nul, et donc qu'il existe existe un élélement non nul α_k de Λ dans $C_{1+1/k}$.
- Puisque pour chaque k, on peut trouver un α_k dans $\Lambda \cap C_{1+1/k}$, on construit ainsi une suite (α_k) à valeurs dans $C_2 \cap (\Lambda \setminus \{0\})$. [4]

Comme $C_2 \cap (\Lambda \setminus \{0\})$ est un compact discret, il est fini. Ceci implique que la suite (α_k) admet une sous-suite constante, et cette constante α est à la fois dans $\Lambda \setminus \{0\}$ et dans $\overline{C}$. Mais comme C est un compact de $\mathbb{R}^n$, il est fermé et donc $\alpha \in C$. ■

Corollaire 1.3

Soit Λ un réseau de $\mathbb{R}^n$, alors on a : $\lambda(\Lambda) \leqslant \sqrt{n}(\det \Lambda)^{1/n}$.

Preuve : Désignons par C le parallélépipède de côté ℓ dans $\mathbb{R}^n$ centré en O, alors on a $\text{vol}(C) = \ell^n$. De plus C est un convexe, compact et symétrique de $\mathbb{R}^n$. En prenant $\ell = 2(\det \Lambda)^{1/n}$, alors on a $\ell = 2^n \det(\Lambda)$, et donc d'après la variante Minkowski il existe x non nul dans $C \cap \Lambda$. Puisque C est centré en O, pour chaque i on a $|x_i| \leqslant \frac{\ell}{2}$, on a donc :

$$\|x\|^2 = \sum |x_i|^2 \leqslant n\ell^2/4 \leqslant n(\det \Lambda)^{2/n}.$$

Puisque nous avons $\lambda(\Lambda) \leqslant \|x\|$, on déduit $\lambda(\Lambda) \leqslant \sqrt{n}(\det \Lambda)^{1/n}$. ■

Il est naturel d'étudier $\gamma_n = \sup_{\Lambda \,/\, \dim \Lambda = n} \frac{\lambda(\Lambda)^2}{(\det(\Lambda))^{2/n}}$ On ne connaît pas la valeur exacte de γ_n, pour tout n, mais on a le tableau suivant :

4. Si α et β sont deux entiers tels que $\alpha < \beta$, alors on a $0 \leqslant \frac{\alpha}{\beta} < 1$. Considérons un convexe C, alors on a :

$$x \in \alpha C \Longrightarrow x = \alpha c \Longrightarrow x = \beta \times \frac{\alpha}{\beta} c \Longrightarrow x = \beta c' \Longrightarrow x \in \beta C.$$

Comme pour tout k, on a $1 + \frac{1}{k} < 2$, on déduit $C_{1+1/k} \subset C_2$.

n	1	2	3	4	5	6	7	8	24
γ_n^n	1	4/3	2	4	8	64/3	64	256	4^{24}

Pour tout $n>0$, on a $\sqrt{\frac{n}{2\pi e}} \lesssim \gamma_n \lesssim \sqrt{\frac{n}{\pi e}}$.

1.1.5 Bases réduites

Base propre

Une base $(b_i)_{1\leqslant i\leqslant n}$ de $\mathbb{R}^n$ est dite propre si les coefficients de la matrice $(\mu_{ij})_{1\leqslant i,j\leqslant n}$ obtenue après la GSO vérifient $|\mu_{ij}|\leqslant \frac{1}{2}$, pour $i>j$.

Algorithme de proprification

Entrée une base $(b_i)_{1\leqslant i\leqslant n}$ de Λ
Sortie une base $(c_i)_{1\leqslant i\leqslant n}$ propre de Λ telle que $c_i = b_i + \sum_{j<i} x_j b_j$

1. Calculer la GSO de (b_i) et on en tire les μ_{ij}
2. Pour $i=1$ jusqu'à n :
3. pour $j=i-1$ jusqu'à 1 :
4. $x_j = \lfloor \mu_{ij} \rceil$ [5]
5. $b_i = b_i - x_j b_j$ [6]
6. $\mu_{ij} = \mu_{ij} - x_j$
7. pour $k=1$ jusqu'à $j-1$:
8. $\mu_{ik} = \mu_{ik} - x_j \mu_{jk}$
9. $c_i = b_i - \sum_{j<i} x_j b_j$

Preuve : Il s'agit de montrer que la base de sortie est propre, pour cela on procédera par récurrence.

5. $\lfloor \mu_{ij} \rceil$ est l'entier le plus proche de μ_{ij}.
6. Cette étape est là à titre d'information, on lui préferera l'étape 9.

- Supposons que pour un certain $i\in[\![1,n-1]\!]$, les vecteurs $b_1,b_2,\cdots,b_{i-1}$ soient propres, alors pour tout $k<j\leqslant i-1$ on a $|\mu_{jk}|\leqslant\frac{1}{2}$. Supposons aussi par récurrence que $|\mu_{ik}|\leqslant\frac{1}{2}$, pour $k\in\{j+1,\cdots,i-1\}$, et montrons $|\mu_{ik}|\leqslant\frac{1}{2}$ pour $k=j$. Comme on a :

$$b_i = b_i^* + \sum_{k<i}\mu_{ik}b_k^* \quad \text{et} \quad b_j = b_j^* + \sum_{k<j}\mu_{jk}b_k^*,$$

on peut écrire :

$$\begin{aligned} b_i - \lfloor\mu_{ij}\rceil b_j &= \left(b_i^* + \sum_{k<i}\mu_{ik}b_k^*\right) - \lfloor\mu_{ij}\rceil\left(b_j^* + \sum_{k<j}\mu_{jk}b_k^*\right) \\ &= b_i^* + \sum_{k<i}\mu_{ik}b_k^* - \lfloor\mu_{ij}\rceil b_j^* - \lfloor\mu_{ij}\rceil\sum_{k<j}\mu_{jk}b_k^* \\ &= b_i^* + \sum_{k=j+1}^{i-1}\mu_{ik}b_k^* + (\mu_{ij}-\lfloor\mu_{ij}\rceil)\,b_j^* + \sum_{k<j}(\mu_{ik}-\lfloor\mu_{ij}\rceil\mu_{jk})\,b_k^*. \end{aligned}$$

La nouvelle valeur de μ_{ij} est alors $\mu_{ij}-\lfloor\mu_{ij}\rceil$, on a donc bien $|\mu_{ij}|\leqslant\frac{1}{2}$. ■

Remarque

Si les $\|b_j^*\|$ sont petites, alors les $\|c_j\|$ le sont aussi. Comme le produit des $\|b_j^*\|$ est constant (et égal à $\det(\Lambda)$), il en est de même du produit des $\|c_j\|$. Ainsi , dans le cas où l'un des $\|c_i\|$ serait trop petite, alors une autre sera trop grande.

Base Siegel réduite

Une $(b_i)_{1\leqslant i\leqslant n}$ une base de Λ, est dite Siegel réduite si on a :

$$\|b_i^*\|^2 \leqslant 2\|b_{i+1}^*\|^2,$$

pour tout $i<n$.

Remarques

- Dans LLL, il y a une condition plus générale avec un paramètre $\delta \in]\frac{1}{4}, 1[$, qui est :

$$\delta \|b_{i-1}^*\|^2 \leqslant \|b_i^* + \mu_{i,i-1} b_{i-1}^*\|^2.$$

  Elle est appelée condition de Lovasz. Dans ce cours on prend $\delta = \frac{3}{4}$.

- Pour une base Siegel réduite, on a :

$$\|b_1\|^2 = \|b_1^*\|^2 \leqslant 2\|b_2^*\|^2 \leqslant \cdots \leqslant 2^{n-1}\|b_n^*\|^2,$$

  d'où : $\|b_1\|^n \leqslant \left(\prod_{i=1}^{n} \|b_i^*\|\right)(\sqrt{2})^{\frac{n(n-1)}{2}} = (\sqrt{2})^{\frac{n(n-1)}{2}} \det \Lambda.$

Base réduite

On appellera base réduite toute base Siegel réduite et propre.

Lemme 1.2

Soit $(b_i)_{1\leqslant i\leqslant n}$ une base Siegel réduite et propre, alors on a :

$$1 \leqslant \frac{\|b_i\|^2}{\|b_i^*\|^2} \leqslant 2^{i-1}.$$

Preuve : Puisque nous avons $b_i = b_i^* + \sum_{j<i} \mu_{ij} b_j^*$ on peut écrire :

$$\frac{\|b_i\|^2}{\|b_i^*\|^2} = 1 + \sum_{j<i} \mu_{ij}^2 \frac{\|b_j^*\|^2}{\|b_i^*\|^2}.$$

Comme la base est Siegel réduite, on a $\|b_j^*\|^2 \leqslant 2^{i-j}\|b_i^*\|^2$. D'autre part, on a aussi $|\mu_{ij}| \leqslant \frac{1}{2}$. On déduit alors :

$$\frac{\|b_i\|^2}{\|b_i^*\|^2} \leqslant 1 + \sum_{j=1}^{i-1} \frac{1}{4} 2^{i-j} = 2^{i-2} + \frac{1}{2} \leqslant 2^{i-2} + 2^{i-2} = 2^{i-1}.$$

D'où le résultat. ■

Théorème 1.2

Soit $x \in \Lambda \setminus \{0\}$ et $(b_i)_{1\leqslant i\leqslant n}$ une base réduite, alors

1) on a : $\|x\| \geqslant \min \|b_i^*\|$;
2) on a : $\|b_1\| \leqslant (\sqrt{2})^{n-1}\|x\|$;
3) si $x_1, x_2, \cdots, x_t$ sont des éléments linéairement indépendants de Λ, on a : $\|b_t\| \leqslant (\sqrt{2})^{n-1} \max_{i\leqslant t} \|x_i\|$.

Preuve :

1) Comme x est un élément non nul de Λ, on peut trouver des $\lambda_i \in \mathbb{Z}$ non tous nuls tels que $x = \sum \lambda_i b_i$. Soit k le plus grand entier tel que $\lambda_k \neq 0$, alors on a :

$$x = \sum_{i\leqslant k} \lambda_i \left(b_i^* + \sum_{j<i} \mu_{ij} b_j^*\right) = \lambda_k b_k^* + \sum_{j<k} \nu_j b_j^*,$$

   avec $\nu_j \in \mathbb{R}$. On a alors :

$$\|x\|^2 = \lambda_k^2 \|b_k^*\|^2 + \sum_{j<k} \nu_j^2 \|b_j^*\|^2 \geqslant \lambda^2 \|b_k^*\|^2 \geqslant (\|b_k^*\|^2)^{7} \geqslant \min \|b_i^*\|^2.$$

2) Désignons par k l'entier tel que $\|b_k^*\| = \min \|b_i^*\|$. Comme la base $(b_i)_{1\leqslant i\leqslant n}$ est Siegel réduite, on a :

$$\|b_1\|^2 = \|b_1^*\|^2 \leqslant 2^{k-1}\|b_k^*\|^2.$$

   En utilisant la première assertion, on a $\|b_k^*\|^2 \leqslant \|x\|^2$, ce qui donne :

$$2^{k-1}\|b_k^*\|^2 \leqslant 2^{k-1}\|x\|^2 \leqslant 2^{n-1}\|x\|^2.$$

   On déduit alors :

$$\|b_1\|^2 \leqslant 2^{n-1}\|x\|^2.$$

7. Car λ est un entier naturel non nul, donc supérieur ou égal à 1.

3) Comme la base $(b_i)_{1\leqslant i\leqslant n}$ est base Siegel-réduite, on a :

$$\|b_j^*\|^2 \leqslant 2\|b_{j+1}^*\|^2 \leqslant \cdots \leqslant 2^{i-j}\|b_i^*\|^2.$$

De plus, pour tout j on a $x_j = \sum_{i=1}^{n} r_{ij}b_j$, avec $r_{ij} \in \mathbb{Z}$. Pour un $j \leqslant t$ donné, notons $I(j)$ le plus grand i tel que $r_{ij} \neq 0$. En se référant à la preuve du **1**, on a $\|x_j\|^2 \geqslant \|b_{I(j)}^*\|^2$ [8]. Quitte à changer l'ordre des x_j, on peut toujours supposer $I(1) \leqslant \cdots \leqslant I(t)$. Montrons alors par récurrence $j \leqslant I(j)$.

- On a $1 \leqslant I(1)$ car $x_1 \neq 0$.
- Supposons pour un j fixé, que la relation $j-1 \leqslant I(j-1)$ est vrai. Comme on a $I(j-1) \leqslant I(j)$, on peut écrire $j-1 \leqslant I(j)$. Si on avait $j-1 = I(j)$, alors on aurait $\langle x_1,\cdots,x_j\rangle \subset \langle b_1,\cdots,b_{j-1}\rangle$ ce qui contredirait [9] le fait que les x_i soient indépendants. On a donc $j \leqslant I(j)$.

Ainsi, on a en utilisant le Lemme 1.2 :

$$\|b_j\|^2 \leqslant 2^{j-1}\|b_j^*\|^2 \leqslant 2^{j-1}2^{I(j)-j}\|b_{I(j)}^*\|^2 = 2^{I(j)-1}\|b_{I(j)}^*\|^2.$$

Puisque nous avons $\|b_{I(j)}^*\|^2 \leqslant \|x_j\|^2$ et $2^{I(j)-1} \leqslant 2^{n-1}$, on déduit :

$$\|b_j\|^2 \leqslant 2^{n-1}\|x_j\|^2,$$

d'où :

$$\|b_j\| \leqslant (\sqrt{2})^{n-1}\|x_j\| \leqslant (\sqrt{2})^{n-1}\max_{i\leqslant t}\|x_i\|.$$

L'inégalité précédente étant vraie pour tout j, en prenant $j = t$, on obtient :

$$\|b_t\|^2 \leqslant (\sqrt{2})^{n-1}\max_{i\leqslant t}\|x_i\|.$$

■

Remarque

La première assertion du théorème (1.2) reste valable pour tout type de base, mais sa structure aide pour les autres points, c'est pour cela qu'elle est ici. Remarquons aussi que la borne sur $\|b_1\|$ n'utilise que la propriété Siegel réduite. La propreté sert à maîtriser la tailler des données dans l'algorithme LLL.

8. Si k est le plus grand entier tel que $\lambda_k \neq 0$ pour $x = \sum \lambda_i b_i$, alors on a $\|x\| \geqslant \|b_k^*\|$.
9. Puisque les x_i sont indépendants, $\langle x_1,\cdots,x_j\rangle$ est un espace de dimension j, alors que $\langle b_1,\cdots,b_{j-1}\rangle$ est de dimension $j-1$.

Corollaire 1.4

Soit $x \in \Lambda \setminus \{0\}$ et $(b_i)_{1\leqslant i\leqslant n}$ une base réduite, alors on a :

$$\|b_1\| \leqslant (\sqrt{2})^{n-1}\lambda(\Lambda).$$

Preuve : D'après la deuxième assertion du théorème précédent, on a :

$$\|b_1\| \leqslant (\sqrt{2})^{n-1}\|x\|,$$

pour tout vecteur non nul $x \in \Lambda$. En particulier, ceci est vrai pour les vecteurs minimaux de Λ, on a donc $\|b_1\| \leqslant (\sqrt{2})^{n-1}\lambda(\Lambda)$. ■

1.2 L'algorithme LLL et quelques applications

1.2.1 Présentation de l'algorithme

Des initiales de Arjen Lenstra, Hendrik Lenstra et László Lovász, le LLL est un algorithme de réduction de réseau qui s'exécute en temps polynomial.

L'algorithme (1982)

Entrée : une base $(b_i)_{1\leqslant i\leqslant n}$ de $\Lambda \subset \mathbb{R}^n$
Sortie : une base réduite

1. $k = 2$
2. calculer les b_i^*
3. tant que $k \leqslant n$
4. proprification de b_k par rapport $(b_1, b_2, \cdots, b_{k-1})$ (dans l'algo de proprification, on fait seulement $i = k$)
5. si $k > 1$ et $\|b_{k-1}^*\|^2 > 2\|b_k^*\|^2$
6. échanger b_{k-1} et b_k
7. mettre à jour les b_i^*

8. $k = k-1$
9. sinon $k = k+1$
10. return (b_i)

1.2.2 Complexité de l'algorithme

Afin de pouvoir mesurer la taille des éléments plus facilement, on fera l'hypothèse bénigne que les coefficients des b_i sont dans $\mathbb{Z}$.

Théorème 1.3

Soit $A = \max_{i\leqslant n} \|b_i\|^2$, alors LLL :

1) finit en $O(n^4 \log A)$ opérations sur des entiers de taille $O(n \log A)$,
2) requiert $\tilde{O}(n^5 \log^2 A)$ opérations binaires,
3) requiert $O(n^3 \log A)$ bits d'espace.

Preuve : • Montrons que l'algo termine. Posons $\Lambda_i = \langle b_1, \cdots, b_i \rangle_{\mathbb{Z}}$, $D_i = \prod_{j\leqslant i} \|b_j^*\|^2 = \det {}^tB_i B_i \in \mathbb{Z}$ et $D = \prod_{i=1}^{n-1} D_i$. L'idée est de montrer que D change uniquement lors d'une transposition et devient inférieur à $\frac{3}{4}D$.

- Les b_i^* ne changent pas lors de la proprification : en effet, la construction de la GSO d'une base montre que si (b_i) et (b_i') sont deux bases reliées par une matrice de passage triangulaire supérieure avec des 1 sur la diagonale (comme c'est le cas pour une base et sa proprifiée) alors les GSO sont identiques. Donc D_i et D sont fixes. Par contre une transposition remplace (b_{k-1}^*, b_k^*) par (s,t) où $s = b_k^* + \mu_{k,k-1} b_{k-1}^*$. En effet : $b_k^* = b_k - \sum_{j<k} \mu_{kj} b_j^* = b_k - \sum_{j<k-1} \mu_{kj} b_j^* - \mu_{k,k-1} b_{k-1}^*$. Donc $\|s\|^2 \geqslant \|b_k^*\|^2$. Par ailleurs $\prod_{i=1}^n \|b_i^*\|^2 = \det(\Lambda)^2 (invariant)$ implique $\|s\|\|t\| = \|b_{k-1}^*\|\|b_k^*\|$ qui implique $\|t\| \leqslant \|b_{k-1}^*\|$.

  D_i reste invariant sauf pour D_{k-1} qui est multiplié par $\frac{\|s\|^2}{\|b_{k-1}^*\|^2}$.
  Or

$$\frac{\|s\|^2}{\|b_{k-1}^*\|^2} = \frac{\|b_k^*\|^2}{\|b_{k-1}^*\|^2} + \mu_{k,k-1}^2 \leqslant \frac{1}{2} + \frac{1}{4} \tag{1.1}$$

  en raison de la propreté et du passage dans le `if`.

- $0 < D \leqslant \|b_1\|^{2(n-1)} \cdots \|b_{n-1}\|^2 \leqslant A^{\frac{n(n-1)}{2}}$, donc le nombre de transpositions est au plus $\log_{3/4} A^{\frac{n(n-1)}{2}} = O(n^2 \log A)$. Chaque passage dans la boucle demande $O(n^2)$ opérations sur les b_i et μ_{ij} d'où un total de $O(n^4 \log A)$ opérations.
- Borne sur les dénominateurs

Lemme 1.3

À chaque étape k on a :

$$D_{k-1} b_k^* \in \mathbb{Z}^n \quad \text{et} \quad D_\ell \mu_{k,\ell} \in \mathbb{Z}$$

pour $\ell < k \leqslant n$.

Preuve : Écrivons $b_k^* = b_k - \sum_{\ell<k} \nu_{k,\ell} b_\ell$, pour des $\nu_{kl} \in \mathbb{Q}$. On a $\langle b_k^*, b_j \rangle = 0$ pour $j < k$ car $\langle b_1, \ldots, b_j \rangle = \langle b_1^*, \ldots, b_j^* \rangle$. En injectant dans l'égalité, on a

$$\sum_{\ell<k} \nu_{k,\ell} \langle b_\ell, b_j \rangle = \langle b_k, b_j \rangle \in \mathbb{Z},\ j < k$$

Le déterminant du système linéaire à inconnues les $\nu_{k,\ell}$ est D_{k-1} donc $D_{k-1}\nu_{k,\ell} \in \mathbb{Z}$ (en utilisant la résolution par la comatrice). Donc $D_{k-1} b_k^* = D_{k-1} b_k - \sum_{\ell<k} D_{k-1} \nu_{k,\ell} b_\ell \in \mathbb{Z}^n$.
Pour le second $D_\ell \mu_{k,\ell} = D_{\ell-1} \|b_\ell^*\|^2 \frac{\langle b_k, b_\ell * \rangle}{\|b_\ell^*\|^2} = \langle b_k, D_{\ell-1} b_\ell^* \rangle \in \mathbb{Z}$ ■

On a choisi A tel que $\|b_i^*\|^2 \leqslant \|b_i\|^2 \leqslant A$, pour tout i.

Lemme 1.4

Soit $(b_i)_{1\leqslant i\leqslant n}$ une base de Λ, alors LLL n'augmente pas $\max \|b_i^*\|$.

Preuve : les $\|b_i^*\|$ ne changent pas sauf dans les transpositions et les valeurs restent les mêmes sauf pour $\|s\|$ et $\|t\|$. Or $\|s\|^2 \leqslant \|b_{k-1}^*\|^2 \leqslant A$ d'après (1.1) et $\|t\|^2 \leqslant \|b_{k-1}^*\|^2 \leqslant A$. ■

Donc $D_i = \prod_{j<i} \|b_j^*\|^2 \leqslant A^i \leqslant A^n$.

Les lemmes impliquent que les dénominateurs des b_k^* et $\mu_{k,\ell}$ sont bornés par A^n et les dénominateurs des b_i sont 1. On déduit que tous les dénominateurs sont en $O(n\log A)$.

- Bornes sur la valeur absolue des données
  - ∗ D'après l'inégalité de Cauchy-Schwarz, on a $|\mu_{ij}|^2 = \left(\frac{\langle b_i, b_j^*\rangle}{\|b_j^*\|^2}\right)^2 \leqslant \frac{\|b_i\|^2}{\|b_j^*\|^2}$
    Or $\|b_j^*\|^2 = \frac{D_j}{D_{j-1}} \geqslant \frac{1}{D_{j-1}}$ ce qui implique $|\mu_{ij}|^2 \leqslant D_{j-1}\|b_i\|^2$.
- Borne sur les $\|b_i\|$. L'idée est de montrer que $\|b_i\|^2 \leqslant nA$ sauf durant l'étape de proprification des b_i où $\|b_i\|^2 \leqslant n^2(4A)^{n+1}$.
  Au début, $\|b_i\|^2 \leqslant A \leqslant nA$ par définition. Le seul endroit où $\|b_i\|$ change c'est lorsque b_i est proprifié. On a $b_i = \mu_{ii}(=1)b_i^* + \mu_{i,i-1}b_{i-1}^* + \cdots + \mu_{i1}b_1^*$
  $\|b_i\|^2 = \sum_{j\leqslant i}|\mu_{ij}|^2\|b_j^*\|^2 \leqslant \sum_{j\leqslant i} m_i^2 A \leqslant nm_i^2 A$ où $m_i = \max(|\mu_{ij}|)$. A la fin de la proprification, on a $m_i = 1$ puisque $|\mu_{ij}| \leqslant \frac{1}{2}$ et $\mu_{ii} = 1$ et donc on a encore $\|b_i^*\|^2 \leqslant nA$.
  Pendant la proprification, on va montrer que les m_i n'augmentent pas trop. Au début $m_i^2 = \max_{1\leqslant j\leqslant i}\mu_{ij}^2 \leqslant \max D_{j-1}\|b_i\|^2 \leqslant A^{n-1}\max\|b_i\|^2 \leqslant A^n n$ en utilisant la borne sur les D_i. On regarde maintenant comment b_i change quand il est remplacé par $b_i - \lfloor\mu_{ij}\rceil b_j$. Après cette étape on a $|\mu_{i,l}| \leqslant 1/2$ pour $j \leqslant l < i$. Cela ne change donc pas la valeur de m_i. Pour $1 \leqslant l < j$ on a la nouvelle valeur de μ_{il}

$$|\mu_{il} - \lfloor\mu_{ij}\rceil\mu_{jl}| \leqslant |\mu_{il}| + |\lfloor\mu_{ij}\rceil||\mu_{jl}| \leqslant m_i + (m_i + 1/2)1/2 = 3/2m_i + 1/4 \leqslant 2m_i$$

  car $m_i \geqslant 1$. Durant toute la boucle, m_i ne peut donc augmenter que d'un facteur $2^{i-1} \leqslant 2^{n-1}$. Donc

$$m_i^2 \leqslant nA^n 2^{2n-2} \leqslant n(4A)^n$$

  et donc

$$\|b_i\|^2 \leqslant nm_i^2 A \leqslant n^2(4A)^n A \leqslant n^2(4A)^{n+1}.$$

- Borne finale sur les numérateurs $\|b_k\| \leqslant n(4A)^{\frac{n+1}{2}}$, $\|D_{k-1}b_k^*\| \leqslant A^n A^{1/2}$ car $\|b_k^*\|$ n'augmente pas dans l'algo et au départ $\|b_k^*\| \leqslant \|b_k\| \leqslant A^{1/2}$, $|D_l\mu_{k,\ell}| \leqslant A^n D_{\ell-1}^{1/2}\|b_k\| \leqslant A^n A^{n/2} n(4A)^{\frac{n+1}{2}}$ (la première inégalité est celle montrée avec Cauchy-Schwarz).
  Dons les numérateurs sont en en $\tilde{O}(n\log A)$. ■

1.2.3 Polynôme minimal

Théorème 1.4

Soit $z \in \mathbb{C}$, on suppose que :

- il existe un polynôme irréductible $P \in \mathbb{Z}[X]$ de degré n tel que :

$$P(z) = 0 \quad \text{et} \quad \|P\|_\infty = \max\{|coef|\} < A;$$

- on peut calculer $\tilde{z}(\varepsilon) \in \mathbb{Q}(i)$ tel que $|z - \tilde{z}| < \varepsilon$ pour tout $\varepsilon > 0$, avec $\log\frac{1}{\varepsilon} = O(n^2\log A)$

Alors on peut trouver P en temps polynomial en $\tilde{O}(n\log A)$.

Corollaire 1.5

On peut factoriser dans $\mathbb{Q}[X]$ en temps polynomial.

Preuve : On approxime une racine $z \in P$ par $\tilde{z}$ et on trouve son polynôme minimal en regardant les polynômes D de degré $\leqslant P$ (il est possible de donner une borne sur $\|D\|_\infty$ en fonction de P par la borne de Landau). Soit P est irréductible et on a finit, soit on recommence avec D et P/D. ■

Preuve : L'idée est de construire le réseau engendré par les colonnes de

$$\begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & & \vdots \\ 0 & 0 & \ddots & \\ 0 & \dots & \dots & 1 \\ CRe(\tilde{z}^n) & \dots & & CRe(\tilde{z}^0) \\ CIm(\tilde{z}^n) & \dots & & CIm(\tilde{z}^0) \end{bmatrix}$$

avec C un nombre réel très grand. Soit $v = (\lambda_n, \cdots, \lambda_0, CRe(\sum_{i=0}^n \lambda_i\tilde{z}^i), CIm(\sum_{i=0}^n \lambda_i\tilde{z}^i)$ le premier vecteur d'une base réduite. Si on veut qu'il soit court alors

$$|C||Re(\sum_{i=0}^n \lambda_i\tilde{z}^i)| \quad \text{et} \quad |C||Im(\sum_{i=0}^n \lambda_i\tilde{z}^i)|$$

doivent être petits, ce qui implique qu'ils sont nuls si $|C| >> 0$ et par suite $\sum_{i=0}^{n} \lambda_i \tilde{z}^i = 0$ et donc $Q = \sum_{i=0}^{n} \lambda_i X^i$ annule z.
Pour prouver le théorème, on va avoir besoin de trois lemmes.

Lemme 1.5 (Borne de Cauchy)

Soit $z \in \mathbb{C}$ et $P = \sum a_i X^i \in \mathbb{C}[X]$ such that $P(z) = 0$ and $a_n \neq 0$. Then

$$|z| \leqslant 2 \max_{0 \leqslant i < n} \left| \frac{a_i}{a_n} \right|^{1/(n-i)}$$

Preuve : Supposons par contradiction que

$$|z|^{n-i} > 2^{n-i} \left| \frac{a_i}{a_n} \right|$$

for all $i < n$. Then $2^{i-n}|a_n||z|^n \geqslant |a_i z^i|$ and

$$|a_n z^n| > |a_n z^n| \underbrace{\sum_{i<n} 2^{i-n}}_{<1} > \sum_{i<n} |a_i z^i|.$$

Or $P(z) = 0$ peut s'écrire $a_n z^n = -\sum_{i<n} a_i z^i$ et donc

$$|a_n z^n| \leqslant \sum_{i<n} |a_i z^i|$$

d'où la contradiction. ■

Lemme 1.6

Soit $P \in \mathbb{Z}[X]$ irréductible sur $\mathbb{Q}$ avec $\deg P = n$, $||P||_\infty \leqslant A$ et $z \in \mathbb{C}$ une racine de P. Alors pour tout $Q \in \mathbb{Z}[X]$ tel que $\deg Q \leqslant n$ et $||Q||_\infty \leqslant A$, on a soit $Q(z) = 0$ ou $|Q(z)| > \eta > 0$ où $\log(1/\eta) = \tilde{O}(n^2 \log A)$.

Preuve : On écrit $P = a_n \prod_{i=1}^{n} (X - z_i)$, $z_i \in \mathbb{C}$ avec$z_1 = z$.Par le Lemme 1.5, on a $|z_i| < D = 2A$ et on peut supposer que $D \geqslant 1$. Considérons

$$R = a_n^{\deg Q} \prod_{i=1}^{n} Q(z_i) = \mathrm{Res}_X(P, Q) \in \mathbb{Z}.$$

Rappelons que ce résultant est nul si et seulement si P et Q ont une racine en commun dans $\mathbb{C}$ ce qui implique si c'est le cas que $P|Q$ puisque P est irréductible et que $P = Q$ (à un facteur près) puisque $\deg Q \leqslant n$. Les racines de P et de Q sont donc les mêmes. Si on suppose que $Q(z) \neq 0$ alors $|R| \geqslant 1$. D'autre part on a, en écrivant $Q = \sum \lambda_i X^i$ que

$$|Q(z_i)| \leqslant \sum |\lambda_i||z_i|^i \leqslant A(n+1)(2A)^n$$

et que $|a_n| \leqslant A$ donc

$$1 \leqslant |A| \leqslant Q(z) A^n (A(n+1)(2A))^{n-1}$$

ce qui donne bien la borne recherchée. ■

Lemme 1.7

Soient $z, \tilde{z} \in \mathbb{C}$ tel que $|z - \tilde{z}| < \varepsilon < 1$, $|z| \leqslant A$ et $P \in \mathbb{C}[X]$ tel que $\deg P \leqslant n$ et $||P||_\infty \leqslant A$. Alors $|P(z) - P(\tilde{z}) < \varepsilon B$ où $\log B = \tilde{O}(n \log A)$.

Preuve : En utilisant les développements de Taylor à l'ordre 1, on a $P(z) - P(\tilde{z}) = (z - \tilde{z})P'(\zeta)$ avec $\zeta = z + t(\tilde{z} - z)$ pour $t \in [0,1]$. Ainsi $|\zeta| \leqslant A + \varepsilon \leqslant A + 1$ et $|P'(z)| \leqslant n^2 A (A+1)^{n-1}$ en bornant chacun des coefficients de P'. ■

On peut maintenant revenir au théorème. Supposons que $Q(z) \neq 0$, alors en utilisant Lemme 1.6, on a que $C|Q(z)| > C\eta$. Ainsi $C|Q(\tilde{z})| > C(\eta - \varepsilon B)$ par le Lemme 1.7. Remarquons que le résultat est vide si $\varepsilon B > \eta$.
Puisque v est le premier vecteur de la base réduite on a que $||v||^2 \leqslant 2^n \lambda(\Lambda)$. Or dans le réseau Λ il y a le vecteur $w = (a_n, \ldots, a_0, C\mathrm{Re}P(\tilde{z}), C\mathrm{Im}P(\tilde{z}))$ et $||w||^2 \leqslant (n+1)A^2 + C^2|P(\tilde{z})|^2$ puisque $(\mathrm{Re}P(\tilde{z}))^2 + (\mathrm{Im}P(\tilde{z})^2 = |P(\tilde{z})|^2$. En utilisant le Lemme 1.7 avec $P(z) = 0$ on a donc

$$||v||^2 \leqslant 2^n((n+1)A^2 + C^2\varepsilon^2 B^2).$$

D'autre part,

$$||v||^2 \geqslant C^2 |Q(\tilde{z}|^2 > C^2(\eta - \varepsilon B)^2$$

si $\eta - \varepsilon B \geqslant 0$. On obtient donc une contradiction dès lors que

$$C^2(\eta - \varepsilon B)^2 > 2^n((n+1)A^2 + C^2\varepsilon^2 B^2)$$

c'est-à-dire $\eta - \varepsilon B > 2^{(n+1)/2}\varepsilon B$ si l'on choisit $C\varepsilon B = \sqrt{n+1}A$. On prend pour cela $\varepsilon < \eta/(2B)$ tel qu'on a aussi $2^{(n+1)/2}\varepsilon B \leqslant \eta/2$. La valeur $\varepsilon = \eta/(B2^{n+3}/2)$ convient. On trouve donc

$$\log(1/\varepsilon) = \log(1/\eta) + \log B + O(n) = \tilde{O}(n^2 \log A)$$

$$\log C = O(\log(nA) - \log(\varepsilon B)) = \tilde{O}(n^2 \log A)$$

d'où le résultat final. ■