FACTORING POLYNOMIALS OVER NUMBER FIELDS

ANDREA MUNARO

Graduate Seminar

Contents

1. Introduction

The purpose of these notes is to give a substantially self-contained introduction to the factorization of polynomials over number fields. In particular, we present Zassenhaus' algorithm and a factoring algorithm using lattice reduction, which were, respectively, the best in practice and in theory, before 2002. We give references for the van Hoeij-Novocin algorithm, currently the best both in practice and in theory. The next section is devoted to introduce lattices, which are relevant for the algorithms.

2. Lattices

We consider $\mathbb{R}^n$ as a metric space with the Euclidean metric. Then, as in any topological space, we have the notion of discretness. We can reformulate it as follows.

Definition 1. A subset D of $\mathbb{R}^n$ is called *discrete* if it has no limit points, that is, for all $x \in D$, there exists $\rho > 0$ such that $\mathcal{B}(x, \rho) \cap D = \{x\}$.

Example 1. $\mathbb{Z}^n$ is discrete (take $\rho = 1/2$), while $\mathbb{Q}^n$ and $\mathbb{R}^n$ are not. The set $\{1/n : n \in \mathbb{N}^*\}$ is discrete but the set $\{0\} \cup \{1/n : n \in \mathbb{N}^*\}$ is not.

Definition 2. A *lattice* L in $\mathbb{R}^n$ is a discrete subgroup of the additive group $\mathbb{R}^n$. A set of independent generators of L is called a (*lattice*) *basis*. The *dimension* or *rank* of L, denoted by $\dim(L)$, is the cardinality of a lattice basis.

Remark 1. Let A be an additive subgroup of $\mathbb{R}^n$ and $0 \neq x \in A$. Suppose 0 is not a limit point and that there exists $y_\rho \neq x$ in $A \cap \mathcal{B}(x, \rho)$ for every $\rho > 0$. Then $x - y_\rho \in A$ and $\|x - y_\rho\| < \rho$. Hence $0 \neq x - y_\rho \in A \cap \mathcal{B}(0, \rho)$ for every $\rho > 0$, contradiction. Then we have that an additive subgroup is discrete if and only if 0 is not a limit point. In other words a lattice is any nonempty set $L \subseteq \mathbb{R}^n$ stable by subtraction and such that $L \cap \mathcal{B}(0, \rho) = \{0\}$ for some $\rho > 0$.

Example 2. By the above Remark it is clear that $\{0\}$ and $\mathbb{Z}^n$ are lattices. Furthermore, any subgroup of a lattice is a lattice.

Let $\mathbf{v}_1, \ldots, \mathbf{v}_r \in \mathbb{R}^n$ and let $\mathcal{L}(\mathbf{v}_1, \ldots, \mathbf{v}_r) = \left\{\sum_{i=1}^r n_i \mathbf{v}_i : n_i \in \mathbb{Z}\right\}$ be the set of all integral linear combinations of the $\mathbf{v}_i$'s. This is a subgroup of $\mathbb{R}^n$ but not necessarily a lattice, since discretness can fail to hold. On the other hand, if the $\mathbf{v}_i$'s are linearly independent, a positive answer is given by the following Theorem.

Theorem 1. *Let A be an additive subgroup of $\mathbb{R}^n$. Then A is a lattice in $\mathbb{R}^n$ if and only if there exist $\mathbb{R}$-linearly independent elements $\mathbf{v}_1, \dots, \mathbf{v}_r \in A$ such that $A = \mathcal{L}(\mathbf{v}_1, \dots, \mathbf{v}_r)$.*

Proof. Suppose $A = \mathcal{L}(\mathbf{v}_1, \dots, \mathbf{v}_r)$ with $\mathbf{v}_1, \dots, \mathbf{v}_r$ linearly independent over $\mathbb{R}$ and consider the parallelepiped $\mathcal{P} = \left\{\sum_{i=1}^r x_i \mathbf{v}_i : |x_i| < 1\right\}$. By linear independence we have $A \cap \mathcal{P} = \{0\}$. On the other hand, there exists $\rho > 0$ such that $\mathcal{B}(0, \rho) \subseteq \mathcal{P}$. Then 0 is not a limit point of A. Conversely, suppose that A is a lattice. Let r be the dimension of the $\mathbb{R}$-linear span of A and choose r linearly independent elements $\mathbf{w}_1, \dots, \mathbf{w}_r$ of A over $\mathbb{R}$. Consider the set

$$F = \{\mathbf{x} \in A : \mathbf{x} = x_1 \mathbf{w}_1 + \cdots + x_r \mathbf{w}_r,\ 0 \leq x_i \leq 1\} = A \cap \{x_1 \mathbf{w}_1 + \cdots + x_r \mathbf{w}_r,\ 0 \leq x_i \leq 1\}.$$

F is closed and bounded, thus compact. But being also discrete, it must be finite. Then for each $1 \leq i \leq r$ we can choose $\mathbf{v}_i \in F$ such that $\mathbf{v}_i = x_i \mathbf{w}_i + \cdots + x_r \mathbf{w}_r$ with $x_i > 0$ and minimal (note that $\mathbf{w}_i \in F$). The $\mathbf{v}_i$'s are clearly linearly independent. Let $\mathbf{v} \in A$ and write $\mathbf{v} = \sum_{i=1}^r \lambda_i \mathbf{v}_i$. Then $\mathbf{v}' := \sum_{i=1}^r \left(\lambda_i - \lfloor \lambda_i \rfloor\right) \mathbf{v}_i \in A$. Suppose there exists j such that $\lambda_j - \lfloor \lambda_j \rfloor \neq 0$, and consider the minimal for which $\lambda_j - \lfloor \lambda_j \rfloor > 0$. Then $\mathbf{v}' = x_j(\lambda_j - \lfloor \lambda_j \rfloor)\mathbf{w}_j + \cdots$ is an element of F which contradicts the minimality of $\mathbf{v}_j$. Then $\lambda_i \in \mathbb{Z}$ for any i, and so $A = \mathcal{L}(\mathbf{v}_1, \dots, \mathbf{v}_r)$. □

Lemma 1. *Let $\mathbf{v}_1, \dots, \mathbf{v}_r$ and $\mathbf{w}_1, \dots, \mathbf{w}_r$ be any two bases of a lattice L. Let A and B be the $n \times r$ matrices with columns $\mathbf{v}_1, \dots, \mathbf{v}_r$ and $\mathbf{w}_1, \dots, \mathbf{w}_r$, respectively. Then there exists an $r \times r$ integral matrix M with $\det(M) = \pm 1$ such that $B^t = MA^t$.*

Proof. By definition of basis, there exist integers m_{ij} and n_{ij}, for $1 \leq i, j \leq r$, such that $\mathbf{w}_j = \sum_{i=1}^r m_{ij} \mathbf{v}_i$ and $\mathbf{v}_j = \sum_{i=1}^r n_{ij} \mathbf{w}_i$. Then

$$\mathbf{v}_j = \sum_{k=1}^r n_{kj} \mathbf{w}_k = \sum_{k=1}^r n_{kj} \sum_{i=1}^r m_{ik} \mathbf{v}_i = \sum_{i=1}^r \left(\sum_{k=1}^r m_{ik} n_{kj}\right) \mathbf{v}_i.$$

By linear independence we have $\sum_{k=1}^r m_{ik} n_{kj} = \delta_{ij}$. This means that, by letting $M = (m_{ij})_{1 \leq i,j \leq r}$ and $N = (n_{ij})_{1 \leq i,j \leq r}$, we have $MN = I_r$. Therefore $\det(M)\det(N) = 1$. But both determinants are integers and so $\det(M) = \det(N) = \pm 1$. □

Definition 3. Let $\mathbf{v}_1, \dots, \mathbf{v}_r \in \mathbb{R}^n$. For $1 \leq m \leq r$, the $m \times m$ matrix $G_m := (\mathbf{v}_i \cdot \mathbf{v}_j)_{1 \leq i,j \leq m}$ is called the *m-th Gramian matrix* of $\mathbf{v}_1, \dots, \mathbf{v}_r$.

Note that $G_m = (\mathbf{v}_i \cdot \mathbf{v}_j)_{1 \leq i,j \leq m} = A^t A$, where A is the $n \times m$ real matrix with columns $\mathbf{v}_1, \dots, \mathbf{v}_m$. We have that $\mathbf{y}^t (A^t A) \mathbf{y} = (A\mathbf{y})^t (A\mathbf{y}) \geq 0$, for any $\mathbf{y} \in \mathbb{R}^r$. Then the m-th Gramian matrix is positive semidefinite and it is positive definite if and only if $\mathbf{v}_1, \dots, \mathbf{v}_m$ are linearly independent. In particular its determinant is positive if and only if $\mathbf{v}_1, \dots, \mathbf{v}_m$ are linearly independent.

Definition 4. Let L be a lattice in $\mathbb{R}^n$ with basis $\mathbf{v}_1, \dots, \mathbf{v}_r$. Consider the r-th Gramian matrix $(\mathbf{v}_i \cdot \mathbf{v}_j)_{1 \leq i,j \leq r}$. Then we define the *determinant* or *volume* of L by $d(L) = \sqrt{\det(\mathbf{v}_i \cdot \mathbf{v}_j)}$.

Lemma 2. *Let L be a lattice in $\mathbb{R}^n$. Then the determinant is independent of the choice of the basis. Furthermore, if L is a full-rank lattice, i.e. any basis has rank n, then $d(L) = |\det(\mathbf{v}_1, \dots, \mathbf{v}_n)|$ for any basis $\mathbf{v}_1, \dots, \mathbf{v}_n$ of L.*

Proof. Let $\mathbf{v}_1, \dots, \mathbf{v}_r$ and $\mathbf{w}_1, \dots, \mathbf{w}_r$ be two bases of L and let A and B be the matrices with columns $\mathbf{v}_1, \dots, \mathbf{v}_r$ and $\mathbf{w}_1, \dots, \mathbf{w}_r$, respectively. By Lemma 1 we have $B^t = MA^t$ for an integral matrix M with $\det(M) = \pm 1$. Then

$$d(L) = \sqrt{\det(B^t B)} = \sqrt{\det(MA^t AM^t)} = \sqrt{\det(A^t A)}.$$

The second statement is immediate by definition. □

Note 1. The determinant corresponds to the r-dimensional volume of the parallelepiped spanned by any basis of L.

Remark 2. Let L be a lattice of dimension ≥ 1 in $\mathbb{R}^n$. There exists a nonzero vector $\mathbf{v} \in L$. We have that $L \cap \mathcal{B}$ is a finite set, where $\mathcal{B}$ is the closed ball of radius $\|\mathbf{v}\|$ centered at $\mathbf{0}$. Since $\mathbf{v} \in L \cap \mathcal{B}$, there is a shortest nonzero vector in the intersection. Note that we will always refer to short vectors with respect to the Euclidean norm.

Definition 5. Let L be a lattice in $\mathbb{R}^n$. The Euclidean norm of a shortest nonzero vector in L (note that a shortest vector is not unique) is called the *first minimum* of L and it is denoted by $\lambda_1(L)$. We generalize the above definition as follows. For all $1 \leq i \leq \dim(L)$, the i-th minimum $\lambda_i(L)$ is defined as the minimum of $\max_{1\leq j\leq i} \|\mathbf{v}_j\|$ over all i linearly independent lattice vectors $\mathbf{v}_1, \ldots, \mathbf{v}_i \in L$, i.e. it is the smallest real number $\lambda_i(L)$ such that L contains i linearly independent vectors of length at most $\lambda_i(L)$.

Clearly, the minima are incresing: $\lambda_1(L) \leq \lambda_2(L) \leq \cdots \leq \lambda_{\dim(L)}(L)$. Our goal is not only to find short vectors, but also to construct bases of lattices with short vectors. The product of the norms of a basis is closely related to the determinant of the lattice. Indeed, we will see that this product is at least the determinant of the lattice, a result known as Hadamard's inequality. The following Theorem gives an upper bound for the product depending again on the determinant.

Theorem 2 (Hermite's inequality). *Let L be a d-dimensional lattice in $\mathbb{R}^n$. There exists a basis $(\mathbf{v}_1, \ldots, \mathbf{v}_d)$ of L such that*

$$\prod_{j=1}^{d} \|\mathbf{v}_j\| \leq \left(\frac{4}{3}\right)^{\frac{d(d-1)}{4}} d(L).$$

Let L be a lattice in $\mathbb{R}^n$. The orthogonal projection of L over a subspace of $\mathbb{R}^n$ is clearly a subgroup of $\mathbb{R}^n$, but it is not necessarily discrete. However, for "nice" choices of the subspace we can ensure discreteness. Therefore, we are going to prove Hermite's inequality by induction on the lattice dimension using projected lattices. We will denote simply by $\mathbf{v}^{\perp}$ the orthogonal complement of the linear span of $\mathbf{v}$.

Lemma 3. *Let L be a d-dimensional lattice in $\mathbb{R}^n$ and let $\mathbf{v}$ be a nonzero element of L. The orthogonal projection L' of L on $\mathbf{v}^{\perp}$ is a $(d-1)$-dimensional lattice in $\mathbb{R}^n$. Furthermore, $d(L) = \|\mathbf{v}\| \cdot d(L')$.*

Proof of Lemma 3. First, we need to recall an "extension of a basis result" for lattices:

Lemma 4. *Let L be a d-dimensional lattice in $\mathbb{R}^n$ and let $\mathbf{v}_1, \ldots, \mathbf{v}_{d'}$ $(d' < d)$ be linearly independent vectors in L. Then $\mathbf{v}_1, \ldots, \mathbf{v}_{d'}$ can be completed to a basis of L (i.e. there exist other $d - d'$ vectors $\mathbf{v}_{d'+1}, \ldots, \mathbf{v}_d \in L$ such that $\mathbf{v}_1, \ldots, \mathbf{v}_d$ is a basis of L) if and only if every vector in L which is a real linear combination of $\mathbf{v}_1, \ldots, \mathbf{v}_{d'}$ is in fact an integral linear combination.*

Proof. See [5]. □

Back to the proof. By the above Lemma we can extend $\mathbf{v}$ to a basis of L: there exist $\mathbf{v}_1, \ldots, \mathbf{v}_{d-1} \in L$ such that $(\mathbf{v}, \mathbf{v}_1, \ldots, \mathbf{v}_{d-1})$ is a basis of L. Denote by $\mathbf{x}'$ the orthogonal projection of $\mathbf{x} \in L$ on $\mathbf{v}^{\perp}$. Then $L' = \mathcal{L}(\mathbf{v}'_1, \ldots, \mathbf{v}'_{d-1})$. But $\mathbf{v}'_1, \ldots, \mathbf{v}'_{d-1}$ are linearly independent, and so, by Theorem 1, L' is a $(d-1)$-dimensional lattice. The second assertion is an easy exercise using Definition 4. □

If we can find a short vector in the orthogonal projection L', then we can find a reasonably short vector in L. This is made precise by the following statement.

Lemma 5. *Let L be a lattice in $\mathbb{R}^n$ and let $\mathbf{v}_1$ be a nonzero shortest vector of L. Then every element x' of L', the orthogonal projection of L on $\mathbf{v}_1^{\perp}$, is the orthogonal projection of some $x \in L$ such that $\|x\|^2 \leq (4/3)\|x'\|^2$.*

Proof. We may assume that $x' \neq 0$. Let x_0 be any element of L that projects on x'. Then $x_0 = x' - \alpha \mathbf{v}_1$ for some $\alpha \in \mathbb{R}$. The vectors $x = x_0 + m\mathbf{v}_1 = x' + (m - \alpha)\mathbf{v}_1$, for $m \in \mathbb{Z}$, projects on x'. For these x's we have $\|x\|^2 = \|x'\|^2 + (m-\alpha)^2 \|\mathbf{v}_1\|^2$. Choosing m as the nearest integer to α, i.e. $\lceil \alpha - 1/2 \rceil$, we have

$$\|x\|^2 \leq \|x'\|^2 + \frac{1}{4}\|\mathbf{v}_1\|^2 \leq \|x'\|^2 + \frac{1}{4}\|x\|^2.$$

□

Now we have all the ingredients for the proof of Hermite's inequality.

Proof of Hermite's inequality. We use induction on d. The case $d = 1$ is trivial. Let $d \geq 2$ and assume the result true up to $d - 1$. Let $\mathbf{v}_1$ be a shortest vector of L. By induction and Lemma 3, there exists a basis $(\mathbf{v}'_2, \ldots, \mathbf{v}'_d)$ of L' such that

$$\prod_{j=2}^{d} \|\mathbf{v}'_j\| \leq \left(\frac{4}{3}\right)^{\frac{(d-1)(d-2)}{4}} d(L').$$

By Lemma 5, for $j \geq 2$, each $\mathbf{v}'_j$ is the orthogonal projection of some $\mathbf{v}_j \in L$ such that

$$\|\mathbf{v}_j\|^2 \leq (4/3) \left\|\mathbf{v}'_j\right\|^2.$$

Therefore, since $d(L) = \|\mathbf{v}_1\| \, d(L')$, we have

$$\begin{aligned}
\prod_{j=1}^{d} \|\mathbf{v}_j\|^2 &\leq \|\mathbf{v}_1\|^2 \left(\frac{4}{3}\right)^{d-1} \prod_{j=2}^{d} \|\mathbf{v}'_j\|^2 \\
&\leq \|\mathbf{v}_1\|^2 \left(\frac{4}{3}\right)^{d-1} \left(\frac{4}{3}\right)^{\frac{(d-1)(d-2)}{2}} d(L')^2 \\
&\leq \left(\frac{4}{3}\right)^{\frac{d(d-1)}{2}} d(L)^2
\end{aligned}$$

□

We have the following immediate corollary, which gives an upper bound for the quantity $\lambda_1(L)/d(L)^{1/d}$.

Corollary 1. *Let L be a d-dimensional lattice in $\mathbb{R}^n$. Then*

$$\frac{\lambda_1(L)}{d(L)^{1/d}} \leq \left(\frac{4}{3}\right)^{\frac{d-1}{4}}.$$

Thanks to the above Corollary, the following definition makes sense. The square factor in the formula is because of historical reasons.

Definition 6. The supremum of $\lambda_1(L)^2/d(L)^{2/d}$ over all d-dimensional lattices L is called *Hermite's constant of dimension d* and it is denoted by γ_d.

Remark 3. Corollary 1 gives an upper bound on Hermite's constant which is exponential in the dimension d. Using Minkowski's first theorem one can obtain a linear bound. Namely, $\gamma_d \leq 1+d/4$ for all $d \geq 1$.

Note 2. It is known that, for all $d \geq 1$, there exists a d-dimensional lattice L such that $\gamma_d = \lambda_1(L)^2/d(L)^{2/d}$ but the exact value of γ_d is known only for $1 \leq d \leq 8$ and $d = 24$.

Example 3. As an easy example, let us calculate γ_2. By Corollary 1, we have $\gamma_2 \leq \sqrt{4/3}$. On the other hand, consider the hexagonal lattice spanned by $\mathbf{v}_1, \mathbf{v}_2$ such that $\|\mathbf{v}_1\| = \|\mathbf{v}_2\|$ and $\mathbf{v}_1 \cdot \mathbf{v}_2 = \|\mathbf{v}_1\|^2/2$. It gives $\gamma_2 = \sqrt{4/3}$. Then Hermite's inequality can be rephrased as $\gamma_d \leq \gamma_2^{d-1}$, for all $d \geq 2$.

It is natural to ask if we can find a lattice basis achieving the successive minima. In the next subsection we will see that this is the case for 2-dimensional lattices. However, this is no longer true as soon as the dimension is at least 5. Indeed, it can be seen that the lattice spanned by the columns of the following matrix

$$\begin{pmatrix} 2 & 0 & 0 & 0 & 1 \\ 0 & 2 & 0 & 0 & 1 \\ 0 & 0 & 2 & 0 & 1 \\ 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix},$$

has no basis reaching all the minima.

2.1. **Two-dimensional lattice reduction.** The Lattice Reduction Problem consists in finding, given a lattice basis as input, a basis with short, nearly orthogonal vectors.

Definition 7. Let $(\mathbf{v}_1, \mathbf{v}_2)$ be a basis of a two-dimensional lattice L in $\mathbb{R}^n$. The basis is called *Lagrange-reduced* if $\|\mathbf{v}_1\| \le \|\mathbf{v}_2\|$ and $|\mathbf{v}_1 \cdot \mathbf{v}_2| \le \|\mathbf{v}_1\|^2/2$.

Theorem 3. *If a basis $(\mathbf{v}_1, \mathbf{v}_2)$ of a two-dimensional lattice L in $\mathbb{R}^n$ is Lagrange-reduced, then $\|\mathbf{v}_1\| = \lambda_1(L)$ and $\|\mathbf{v}_2\| = \lambda_2(L)$, i.e. it reaches the succesive minima.*

Proof. Let $\mathbf{v}$ be any nonzero vector in L. Then $\mathbf{v} = a\mathbf{v}_1 + b\mathbf{v}_2$ for some $a, b \in \mathbb{Z}$, not both zero. By assumption we have $\|\mathbf{v}_1\| \le \|\mathbf{v}_2\|$ and $|\mathbf{v}_1 \cdot \mathbf{v}_2| \le \|\mathbf{v}_1\|^2/2$. Then

$$\begin{aligned}
\|\mathbf{v}\|^2 &= (a\mathbf{v}_1 + b\mathbf{v}_2) \cdot (a\mathbf{v}_1 + b\mathbf{v}_2) \\
&= a^2 \|\mathbf{v}_1\|^2 + 2ab(\mathbf{v}_1 \cdot \mathbf{v}_2) + b^2 \|\mathbf{v}_2\|^2 \\
&\ge a^2 \|\mathbf{v}_1\|^2 - |ab| \|\mathbf{v}_1\|^2 + b^2 \|\mathbf{v}_2\|^2 \\
&\ge a^2 \|\mathbf{v}_1\|^2 - |ab| \|\mathbf{v}_1\|^2 + b^2 \|\mathbf{v}_1\|^2 \\
&= (a^2 - |ab| + b^2) \|\mathbf{v}_1\|^2 \\
&\ge \|\mathbf{v}_1\|^2 .
\end{aligned}$$

Then $\mathbf{v}_1$ is a shortest vector of L.
Now suppose that $\mathbf{v} = a\mathbf{v}_1 + b\mathbf{v}_2$ and $\mathbf{v}_1$ are linearly independent, i.e. $b \ne 0$. We have

$$\begin{aligned}
\|\mathbf{v}\|^2 &\ge a^2 \|\mathbf{v}_1\|^2 - |ab| \|\mathbf{v}_1\|^2 + b^2 \|\mathbf{v}_2\|^2 \\
&= a^2 \|\mathbf{v}_1\|^2 - |ab| \|\mathbf{v}_1\|^2 + \frac{1}{4} b^2 \|\mathbf{v}_2\|^2 + \frac{3}{4} b^2 \|\mathbf{v}_2\|^2 \\
&= a^2 \|\mathbf{v}_1\|^2 - |ab| \|\mathbf{v}_1\|^2 + \frac{1}{4} b^2 \|\mathbf{v}_1\|^2 + \frac{3}{4} b^2 \|\mathbf{v}_2\|^2 \\
&= \left(|a| - \frac{1}{2}|b|\right)^2 \|\mathbf{v}_1\|^2 + \frac{3}{4} b^2 \|\mathbf{v}_2\|^2 .
\end{aligned}$$

If $|b| \ne 1$ then clearly $\|\mathbf{v}\| \ge \|\mathbf{v}_2\|$. So suppose $b = \pm 1$. But then we have

$$\|\mathbf{v}\|^2 \ge a^2 \|\mathbf{v}_1\|^2 - |a| \|\mathbf{v}_1\|^2 + \|\mathbf{v}_2\|^2 = |a| (|a| - 1) \|\mathbf{v}_1\|^2 + \|\mathbf{v}_2\|^2 \ge \|\mathbf{v}_2\|^2 .$$

□

Remark 4. The converse holds in the Theorem above.

In the following denote by $\lfloor x \rceil := \lceil x - 1/2 \rceil$ the nearest integer to $x \in \mathbb{R}$. Recall the following variant of the well-known Euclidean algorithm for computing the greatest common divisor of two integers.

Algorithm 1 Centered Euclidean algorithm

```
Input: $(n, m) \in \mathbb{Z}^2$.
Output: $\gcd(n, m)$.
if $|n| \le |m|$ then
    swap $n$ and $m$.
end if
while $m \ne 0$ do
    $r \leftarrow n - qm$ where $q = \lfloor \frac{n}{m} \rceil$.
    $n \leftarrow m$.
    $m \leftarrow r$.
end while
return $|n|$
```

The only difference with the classical Euclidean algorithm is that one takes for q the nearest integer to n/m rather than its integral part. The above algorithm can be reformulated in the language of lattices. We regard n and m as vectors in $\mathbb{R}$. We know that $n\mathbb{Z} + m\mathbb{Z} = \gcd(n,m)\mathbb{Z}$ and so $\gcd(n,m)$ is a shortest vector in $\mathcal{L}(n,m)$.

The following algorithm is a natural generalization in dimension two. The arithmetic operation of division with remainder in the Euclidean algorithm is replaced by the geometric operation of (rounded) orthogonal projection.

Algorithm 2 Lagrange's reduction algorithm

```
Input: a basis $(\mathbf{v}_1, \mathbf{v}_2)$ of a two-dimensional lattice $L$.
Output: a Lagrange-reduced basis of $L$.
1: if $\|\mathbf{v}_1\| < \|\mathbf{v}_2\|$ then
2:     swap $\mathbf{v}_1$ and $\mathbf{v}_2$.
3: end if
4: repeat
5:     $\mathbf{r} \leftarrow \mathbf{v}_1 - q\mathbf{v}_2$ where $q = \left\lfloor \frac{\mathbf{v}_1 \cdot \mathbf{v}_2}{\mathbf{v}_2 \cdot \mathbf{v}_2} \right\rceil$.
6:     $\mathbf{v}_1 \leftarrow \mathbf{v}_2$.
7:     $\mathbf{v}_2 \leftarrow \mathbf{r}$.
8: until $\|\mathbf{v}_1\| \leq \|\mathbf{v}_2\|$
9: return $(\mathbf{v}_1, \mathbf{v}_2)$
```

Lemma 6. *After each execution of the loop in Algorithm 2 we have*

$$\left|\mathbf{v}_1' \cdot \mathbf{v}_2'\right| \leq \frac{1}{2} \left\|\mathbf{v}_1'\right\|^2,$$

where $\mathbf{v}_1'$ and $\mathbf{v}_2'$ are the new vectors.

Proof. By definition of q we have

$$q(\mathbf{v}_2 \cdot \mathbf{v}_2) - \frac{1}{2}(\mathbf{v}_2 \cdot \mathbf{v}_2) < \mathbf{v}_1 \cdot \mathbf{v}_2 \leq q(\mathbf{v}_2 \cdot \mathbf{v}_2) + \frac{1}{2}(\mathbf{v}_2 \cdot \mathbf{v}_2),$$

from which

$$-\frac{1}{2}(\mathbf{v}_2 \cdot \mathbf{v}_2) < (\mathbf{v}_1 - q\mathbf{v}_2) \cdot \mathbf{v}_2 \leq \frac{1}{2}(\mathbf{v}_2 \cdot \mathbf{v}_2).$$

Therefore

$$\left|\mathbf{v}_1' \cdot \mathbf{v}_2'\right| = \left|\mathbf{v}_2 \cdot (\mathbf{v}_1 - q\mathbf{v}_2)\right| \leq \frac{1}{2}\left\|\mathbf{v}_2\right\|^2 = \frac{1}{2}\left\|\mathbf{v}_1'\right\|^2.$$

□

Theorem 4. *The Lagrange's algorithm works correctly.*

Proof. After each execution of Step 5 we get vectors $\mathbf{v}_1', \mathbf{v}_2'$ satisfying the following

$$\begin{pmatrix} \mathbf{v}_1' \\ \mathbf{v}_2' \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & -q \end{pmatrix} \begin{pmatrix} \mathbf{v}_1 \\ \mathbf{v}_2 \end{pmatrix}.$$

Since the matrix has determinant 1, we have that $(\mathbf{v}_1', \mathbf{v}_2')$ is a basis of the lattice. By Lemma 6, the output is a Lagrange-reduced basis.

The lenght of $\mathbf{v}_1$ strictly decreases after each execution of Step 5. But $L \cap \mathcal{B}(\mathbf{0}, \rho)$ is finite for any real $\rho > 0$. Then the algorithm terminates after a finite number of executions of Step 5. □

Remark 5. The Lagrange algorithm is clearly polynomial time, and it gives as output a basis which achieve the first and second minimum and in which the two vectors are nearly orthogonal.

2.2. **Lattice reduction in any dimension.**

LLL algorithm takes as input a basis of a given lattice and returns an *LLL-reduced* basis, in which the vectors are short and nearly orthogonal. It does this by rearranging basis vectors such that latter vectors have "long" Gram-Schmidt lengths. The first vector in the output basis is an approximation of the shortest vector in the lattice. We know that, given a basis, we can obtain orthogonal vectors via Gram-Schmidt orthogonalization. But these vectors could not even belong to the lattice. Let us now review the Gram-Schmidt orthogonalization.

Definition 8. Let $\mathbf{v}_1, \dots, \mathbf{v}_r$ be vectors of $\mathbb{R}^n$. The *Gram-Schmidt ortogonalization (GSO)* of $\mathbf{v}_1, \dots, \mathbf{v}_r$ is the family $\mathbf{v}_1^*, \dots, \mathbf{v}_r^*$ defined as follows:

$$\begin{aligned}\mathbf{v}_1^* &= \mathbf{v}_1,\\ \mathbf{v}_i^* &= \mathbf{v}_i - \sum_{j=1}^{i-1} \mu_{ij}\mathbf{v}_j^* \ \ (2 \le i \le r), \quad \text{where} \quad \mu_{ij} = \frac{\mathbf{v}_i \cdot \mathbf{v}_j^*}{\mathbf{v}_j^* \cdot \mathbf{v}_j^*} \ \ (1 \le j < i \le r).\end{aligned}$$

Note that the vectors $\mathbf{v}_1^*, \dots, \mathbf{v}_r^*$ are not necessarily in $\mathcal{L}(\mathbf{v}_1, \dots, \mathbf{v}_n)$.

Theorem 5. *Let $(\mathbf{v}_1, \dots, \mathbf{v}_r)$ be vectors of $\mathbb{R}^n$ and let $(\mathbf{v}_1^*, \dots, \mathbf{v}_r^*)$ be its GSO. Denote by A, A^* the matrices with column vectors $\mathbf{v}_i$, $\mathbf{v}_i^*$, respectively. Then*

(1) $\mathbf{v}_i^* \cdot \mathbf{v}_j^* = 0$ *for* $1 \le i < j \le r$.

(2) $\operatorname{span}(\mathbf{v}_1^*, \dots, \mathbf{v}_k^*) = \operatorname{span}(\mathbf{v}_1, \dots, \mathbf{v}_k)$ *for* $1 \le k \le r$.

(3) *For* $1 \le k \le r$, *the vector* $\mathbf{v}_k^*$ *is the projection of* $\mathbf{v}_k$ *onto the orthogonal complement of* $\operatorname{span}(\mathbf{v}_1, \dots, \mathbf{v}_{k-1})$.

(4) $\|\mathbf{v}_k^*\| \le \|\mathbf{v}_k\|$ *for* $1 \le k \le r$.

(5) $\det A = \det A^*$.

Proof. We prove only (4) and (5). See [4] for the rest.

We have $\mathbf{v}_k = \mathbf{v}_k^* + \sum_{j=1}^{k-1} \mu_{kj}\mathbf{v}_j^*$. Then, by the orthogonality relation (1), we get

$$\|\mathbf{v}_k\|^2 = \|\mathbf{v}_k^*\|^2 + \sum_{j=1}^{k-1} \mu_{kj}^2 \left\|\mathbf{v}_j^*\right\|^2.$$

Since every term in the summand is nonnegative, (4) follows.

Set $\mu_{ij} = 1$ for $1 \le i = j \le r$, and $\mu_{ij} = 0$ for $1 \le i < j \le r$. Denote by M the lower triangular matrix $(\mu_{ij})_{1 \le i,j \le r}$. Then, by Definition 8, we have $A = A^*M^t$. But $\det M = 1$, and so we get (5). □

Corollary 2 (Hadamard's inequality). *Let $(\mathbf{v}_1, \dots, \mathbf{v}_r)$ be vectors of $\mathbb{R}^n$ and let $(\mathbf{v}_1^*, \dots, \mathbf{v}_r^*)$ be its GSO. For $1 \le k \le r$ we have*

$$\det G_k = \prod_{j=1}^{k} \left\|\mathbf{v}_j^*\right\|^2 \le \prod_{j=1}^{k} \|\mathbf{v}_j\|^2.$$

In particular, if $(\mathbf{v}_1, \dots, \mathbf{v}_r)$ is a basis of a lattice L in $\mathbb{R}^n$, we have

$$d(L) = \prod_{j=1}^{r} \left\|\mathbf{v}_j^*\right\| \le \prod_{j=1}^{r} \|\mathbf{v}_j\|.$$

Proof. Let A be the matrix with column vectors $\mathbf{v}_1, \dots, \mathbf{v}_k$. Recall that we have $A = A^*M^t$, where $M = (\mu_{ij})_{1 \le i,j \le k}$. Then

$$\det G_k = \det A^tA = \det(A^*M^t)^tA^*M^t = \det M(A^*)^tA^*M^t = \det(A^*)^tA^*.$$

But the columns of A^* are orthogonal, by Theorem 5. Then $(A^*)^tA^* = \operatorname{diag}(\|\mathbf{v}_1^*\|^2, \dots, \|\mathbf{v}_k^*\|^2)$. Finally, since $\|\mathbf{v}_k^*\| \le \|\mathbf{v}_k\|$ for $1 \le k \le r$, we get the first assertion.

The second assertion follows directly from the definition of volume. □

Definition 9. Let $1/4 < \delta \le 1$ be a real number. An ordered basis $(\mathbf{v}_1, \dots, \mathbf{v}_r)$ of a lattice L in $\mathbb{R}^n$ is called *size-reduced* if its GSO satisfies

$$|\mu_{ij}| \le 1/2, \qquad 1 \le j < i \le r.$$

It is called *LLL-reduced* with factor δ if it is size-reduced and

$$\left\|\mathbf{v}_{i+1}^* + \mu_{i+1,i}\mathbf{v}_i^*\right\|^2 \ge \delta \left\|\mathbf{v}_i^*\right\|^2, \qquad 1 < i < r.$$

Remark 6. The condition $|\mu_{ij}| \leq 1/2$ means that the projection $\mathbf{v}_i - \mathbf{v}_i^*$ of $\mathbf{v}_i$ over the linear span of $\mathbf{v}_1^*, \dots, \mathbf{v}_{i-1}^*$ is inside the parallelepiped $\mathcal{P} = \left\{\sum_{j=1}^{i-1} x_j \mathbf{v}_j^*,\ |x_j| \leq 1/2\right\}$, i.e. the vector $\mathbf{v}_i$ is "almost orthogonal" to the span of the previous vectors. Let us explain the condition $\left\|\mathbf{v}_{i+1}^* + \mu_{i+1,i}\mathbf{v}_i^*\right\|^2 \geq \delta \left\|\mathbf{v}_i^*\right\|^2$, called *Lovasz' condition*, which is equivalent to

$$\left\|\mathbf{v}_{i+1}^*\right\|^2 \geq (\delta - \mu_{i+1,i}^2) \left\|\mathbf{v}_i^*\right\|^2.$$

If we swap the vectors $\mathbf{v}_i$ and $\mathbf{v}_{i+1}$, then $\mathbf{v}_i^*$ and $\mathbf{v}_{i+1}^*$ could possibly change. The new $\mathbf{v}_i^*$ is exactly $\mathbf{v}_{i+1}^* + \mu_{i+1,i}\mathbf{v}_i^*$. This is the content of the following Lemma. Then Lovasz' condition guarantees that by swapping $\mathbf{v}_i$ and $\mathbf{v}_{i+1}$, the norm of $\mathbf{v}_i^*$ possibly decreases in a way quantified by δ, hence not too much. Notice also that $\mathbf{v}_{i+1}^* + \mu_{i+1,i}\mathbf{v}_i^*$ and $\mathbf{v}_i^*$ are the projections of $\mathbf{v}_{i+1}$ and $\mathbf{v}_i$ onto the orthogonal complement of $\operatorname{span}(\mathbf{v}_1, \dots, \mathbf{v}_{i-1})$, respectively. Therefore, if Lovasz' condition does not hold for some i, it does hold for the basis obtained by swapping $\mathbf{v}_{i+1}$ and $\mathbf{v}_i$.

Lemma 7. *Let $(\mathbf{v}_1, \dots, \mathbf{v}_r)$ be an ordered set of vectors in $\mathbb{R}^n$ and let $(\mathbf{v}_1^*, \dots, \mathbf{v}_r^*)$ be its GSO. Let $1 \leq j \leq r-1$ and let $\hat{\mathbf{v}}_1, \dots, \hat{\mathbf{v}}_r$ be the GSO obtained by swapping $\mathbf{v}_j$ and $\mathbf{v}_{j+1}$. Then $\hat{\mathbf{v}}_i^* = \mathbf{v}_i^*$ for $i \neq j, j+1$ and*

$$\hat{\mathbf{v}}_j^* = \mathbf{v}_{j+1}^* + \mu_{j+1,j}\mathbf{v}_j^*, \qquad \hat{\mathbf{v}}_{j+1}^* = \frac{\left\|\mathbf{v}_{j+1}^*\right\|^2}{\left\|\hat{\mathbf{v}}_j^*\right\|^2}\mathbf{v}_j^* - \mu_{j+1,j}\frac{\left\|\mathbf{v}_j^*\right\|^2}{\left\|\hat{\mathbf{v}}_j^*\right\|^2}\mathbf{v}_{j+1}^*.$$

Proof. The first assertion is clear using Theorem 5. For $\hat{\mathbf{v}}_j^*$ we have

$$\begin{aligned}
\hat{\mathbf{v}}_j^* &= \hat{\mathbf{v}}_j - \sum_{i=1}^{j-1} \frac{\hat{\mathbf{v}}_j \cdot \hat{\mathbf{v}}_i^*}{\hat{\mathbf{v}}_i^* \cdot \hat{\mathbf{v}}_i^*}\hat{\mathbf{v}}_i^* \\
&= \mathbf{v}_{j+1} - \sum_{i=1}^{j-1} \frac{\mathbf{v}_{j+1} \cdot \mathbf{v}_i^*}{\mathbf{v}_i^* \cdot \mathbf{v}_i^*}\mathbf{v}_i^* \\
&= \mathbf{v}_{j+1} - \sum_{i=1}^{j-1} \mu_{j+1,i}\mathbf{v}_i^* \\
&= \mathbf{v}_{j+1} - \sum_{i=1}^{j} \mu_{j+1,i}\mathbf{v}_i^* + \mu_{j+1,j}\mathbf{v}_j^* \\
&= \mathbf{v}_{j+1}^* + \mu_{j+1,j}\mathbf{v}_j^*.
\end{aligned}$$

We omit the formula for $\hat{\mathbf{v}}_{j+1}^*$. □

Theorem 6. *Let $(\mathbf{v}_1, \dots, \mathbf{v}_r)$ be an LLL-reduced basis with factor $1/4 < \delta \leq 1$ of a lattice L, and let $\alpha = 1/(\delta - 1/4)$. Then*

1. $\left\|\mathbf{v}_j\right\|^2 \leq \alpha^{i-1}\left\|\mathbf{v}_i^*\right\|^2$ *for $1 \leq j \leq i \leq r$,*
2. $d(L) \leq \prod_{j=1}^r \left\|\mathbf{v}_j\right\| \leq \alpha^{r(r-1)/4} d(L)$,
3. $\left\|\mathbf{v}_1\right\| \leq \alpha^{(r-1)/4} d(L)^{1/r}$,
4. *For any set of linearly independent vectors $\mathbf{y}_1, \dots, \mathbf{y}_m \in L$ we have*

$$\left\|\mathbf{v}_j\right\| \leq \alpha^{(r-1)/2} \max\left\{\left\|\mathbf{y}_1\right\|, \dots, \left\|\mathbf{y}_m\right\|\right\} \quad 1 \leq j \leq m.$$

Proof. Since $(\mathbf{v}_1, \dots, \mathbf{v}_r)$ is LLL-reduced with factor δ, we have

$$\left\|\mathbf{v}_i^*\right\|^2 \geq (\delta - \mu_{i,i-1}^2)\left\|\mathbf{v}_{i-1}^*\right\|^2 \geq \left(\delta - \frac{1}{4}\right)\left\|\mathbf{v}_{i-1}^*\right\|^2 = \frac{1}{\alpha}\left\|\mathbf{v}_{i-1}^*\right\|^2.$$

Therefore $\left\|\mathbf{v}_{i-1}^*\right\|^2 \leq \alpha\left\|\mathbf{v}_i^*\right\|^2$, and by induction we get

$$\left\|\mathbf{v}_j^*\right\|^2 \leq \alpha^{i-j}\left\|\mathbf{v}_i^*\right\|^2 \quad 1 \leq j \leq i \leq r. \tag{2.1}$$

Since $\mathbf{v}_1^*, \ldots, \mathbf{v}_r^*$ are orthogonal and using 2.1, we have

$$\begin{aligned}
\|\mathbf{v}_i\|^2 &= \|\mathbf{v}_i^*\|^2 + \sum_{j=1}^{i-1} \mu_{ij}^2 \left\|\mathbf{v}_j^*\right\|^2 \\
&\leq \|\mathbf{v}_i^*\|^2 + \sum_{j=1}^{i-1} \frac{1}{4}\alpha^{i-j} \left\|\mathbf{v}_i^*\right\|^2 \\
&= \left(1 + \frac{1}{4}\sum_{j=1}^{i-1} \alpha^{i-j}\right) \|\mathbf{v}_i^*\|^2 \\
&= \left(1 + \frac{1}{4}\frac{\alpha^i - \alpha}{\alpha - 1}\right) \|\mathbf{v}_i^*\|^2 .
\end{aligned}$$

An easy induction shows that

$$1 + \frac{1}{4}\frac{\alpha^i - \alpha}{\alpha - 1} \leq \alpha^{i-1}.$$

But then

$$\|\mathbf{v}_j\|^2 \leq \alpha^{j-1} \left\|\mathbf{v}_j^*\right\|^2 \leq \alpha^{i-1} \|\mathbf{v}_i^*\|^2 \quad 1 \leq j \leq i \leq r,$$

which gives (1).

By Hadamard's inequality and part (1) we have

$$d(L) \leq \prod_{j=1}^{r} \|\mathbf{v}_j\| \leq \alpha^{\frac{0+1+\cdots+(r-1)}{2}} \prod_{j=1}^{r} \left\|\mathbf{v}_j^*\right\| = \alpha^{\frac{r(r-1)}{4}} d(L),$$

which gives part (2).

By setting $j = 1$ in part (1) and taking the product over $i = 1, \ldots, r$ we get

$$\|\mathbf{v}_1\|^{2r} \leq \alpha^{0+1+\cdots+(r-1)} \prod_{k=1}^{r} \|\mathbf{v}_k^*\|^2 = \alpha^{\frac{r(r-1)}{2}} d(L)^2.$$

Taking the $2r$-th roots gives part (3).

Let

$$\mathbf{y}_j = \sum_{i=1}^{r} a_{ij}\mathbf{v}_i \quad a_{ij} \in \mathbb{Z}, 1 \leq j \leq m,$$

and let $i(j)$ denote the largest index i for which $a_{ij} \neq 0$. Recalling that $\mu_{i,i} = 1$, we have

$$\mathbf{y}_j = \sum_{i=1}^{i(j)} a_{ij}\mathbf{v}_i = \sum_{i=1}^{i(j)} a_{ij} \sum_{k=1}^{i} \mu_{ik}\mathbf{v}_k^* = \sum_{i=1}^{i(j)} \sum_{k=1}^{i} a_{ij}\mu_{ik}\mathbf{v}_k^* \quad 1 \leq j \leq m.$$

Using the fact that $0 \neq a_{i(j),j} \in \mathbb{Z}$, we get

$$\|\mathbf{y}_j\|^2 \geq \left\|\mathbf{v}_{i(j)}^*\right\|^2 \quad 1 \leq j \leq m.$$

Assume wlog that $i(1) \leq \cdots \leq i(m)$. If $i(j) < j$ for some j, then $\mathbf{y}_1, \ldots, \mathbf{y}_j \in \mathcal{L}(\mathbf{v}_1, \ldots, \mathbf{v}_{i(j)})$, contradiction. Hence we have $j \leq i(j)$ for $1 \leq j \leq m$. Then, using part (1) with $i = i(j)$, we have

$$\begin{aligned}
\|\mathbf{v}_j\|^2 &\leq \alpha^{i(j)-1}\left\|\mathbf{v}_{i(j)}^*\right\|^2 \\
&\leq \alpha^{r-1}\left\|\mathbf{v}_{i(j)}^*\right\|^2 \\
&\leq \alpha^{r-1} \|\mathbf{y}_j\|^2 \\
&\leq \alpha^{r-1} \max\left\{\|\mathbf{y}_1\|^2, \ldots, \|\mathbf{y}_m\|^2\right\},
\end{aligned}$$

for $1 \leq j \leq m$. This completes the proof. □

Remark 7. (2) in the Theorem above tells us that an LLL-reduced basis with factor 1 satisfy Hermite's inequality. Moreover, $\delta = 1$ gives the best bounds but, as we will see later, we cannot guarantee the LLL algorithm to be polynomial time for such a δ.

Corollary 3. *Let $(\mathbf{v}_1, \dots, \mathbf{v}_r)$ be an LLL-reduced basis with factor $1/4 < \delta \leq 1$ of a lattice L, and let $\alpha = 1/(\delta - 1/4)$. Then, for every nonzero $\mathbf{y} \in L$, we have $\|\mathbf{v}_1\| \leq \alpha^{(r-1)/2} \|\mathbf{y}\|$.*

Remark 8. This gives an exponential upper bound for the length of the first vector in a LLL-reduced basis which applies uniformly to all lattices of dimension r. Furthermore, the first vector in an LLL-reduced basis is no longer than $\alpha^{(r-1)/2}$ times the shortest nonzero vector in the lattice. In particular, if $\delta = 3/4$ an LLL-reduced basis is a $2^{(r-1)/2}$-approximate solution to the shortest vector problem (SVP).

Corollary 4. *Let $(\mathbf{v}_1, \dots, \mathbf{v}_r)$ be an LLL-reduced basis of L and let $\mathbf{y}_1, \dots, \mathbf{y}_i \in L$ linearly independent vectors achieving the succesive minima. Then*

$$\alpha^{(1-i)/2} \max\{\|\mathbf{y}_1\|, \dots, \|\mathbf{y}_i\|\} \leq \|\mathbf{v}_i\| \leq \alpha^{(r-1)/2} \max\{\|\mathbf{y}_1\|, \dots, \|\mathbf{y}_i\|\}, \quad 1 \leq i \leq r.$$

Proof. By Theorem 6, we have $\|\mathbf{v}_j\|^2 \leq \alpha^{i-1} \|\mathbf{v}_i^*\|^2$, for $1 \leq j \leq i \leq r$, which, combined with Theorem 5, gives $\alpha^{1-i} \|\mathbf{v}_j\|^2 \leq \|\mathbf{v}_i^*\|^2 \leq \|\mathbf{v}_i\|^2$, for $1 \leq j \leq i \leq r$. Since $\mathbf{v}_1, \dots, \mathbf{v}_r$ are linearly independent, using Theorem 6 we get,

$$\alpha^{(1-i)/2} \max\{\|\mathbf{v}_1\|, \dots, \|\mathbf{v}_i\|\} \leq \|\mathbf{v}_i\| \leq \alpha^{(r-1)/2} \max\{\|\mathbf{v}_1\|, \dots, \|\mathbf{v}_i\|\}, \quad 1 \leq i \leq r.$$

Finally, by definition of $\mathbf{y}_i$ we get

$$\alpha^{(1-i)/2} \max\{\|\mathbf{y}_1\|, \dots, \|\mathbf{y}_i\|\} \leq \|\mathbf{v}_i\| \leq \alpha^{(r-1)/2} \max\{\|\mathbf{y}_1\|, \dots, \|\mathbf{y}_i\|\}, \quad 1 \leq i \leq r.$$

□

Remark 9. The Corollary above tells us that, just as the first vector of an LLL-reduced basis is an approximation of the shortest vector in the lattice, the other basis vectors provide approximations of the succesive minima.

The LLL algorithm can be viewed as an algorithmic generalization of the proof of Hermite's inequality $\gamma_d \leq \gamma_2^{d-1}$, for $d \geq 2$. Indeed, suppose we are given a basis $(\mathbf{v}_1, \dots, \mathbf{v}_d)$ of a lattice L and denote by π_i the orthogonal projection onto the orthogonal complement of $\mathrm{span}(\mathbf{v}_1, \dots, \mathbf{v}_{i-1})$. The algorithm makes sure that all the local bases $(\pi_i(\mathbf{v}_i), \pi_i(\mathbf{v}_{i+1}))$ are LLL-reduced and lift to a size-reduced basis. The following algorithm finds a size-reduced basis.

Algorithm 3 SIZE-REDUCE(k, l)

```
Input: a basis $(\mathbf{x}_1, \dots, \mathbf{x}_r)$ of an $r$-dimensional lattice $L$, and $l < k$.
Output: a basis $(\mathbf{v}_1, \dots, \mathbf{v}_r)$ of $L$ such that $|\mu_{kl}| \leq 1/2$.
for $i = 1$ to $r$ do
    $\mathbf{v}_i \leftarrow \mathbf{x}_i$.
end for
for $i = 1$ to $r$ do                                        ▷ compute the GSO
    $\mathbf{v}_i^* \leftarrow \mathbf{v}_i$.
    for $j = 1$ to $i - 1$ do
        $\mu_{ij} \leftarrow (\mathbf{v}_i \cdot \mathbf{v}_j^*)/(\mathbf{v}_j^* \cdot \mathbf{v}_j^*)$ and $\mathbf{v}_i^* \leftarrow \mathbf{v}_i^* - \mu_{ij}\mathbf{v}_j^*$.
    end for
end for
if $|\mu_{kl}| > 1/2$ then                          ▷ make $\mathbf{v}_k$ almost orthogonal to $\mathbf{v}_l$
    $\mathbf{v}_k \leftarrow \mathbf{v}_k - \lfloor \mu_{kl} \rceil \mathbf{v}_l$.
end if
for $j = 1$ to $l - 1$ do                                    ▷ update the GSO
    $\mu_{kj} \leftarrow \mu_{kj} - \lfloor \mu_{kl} \rceil \mu_{lj}$.
end for
$\mu_{kl} \leftarrow \mu_{kl} - \lfloor \mu_{kl} \rceil$.
return $(\mathbf{v}_1, \dots, \mathbf{v}_r)$.
```

Remark 10. (k, l)-reduction makes the vector $\mathbf{v}_k$ almost orthogonal to $\mathbf{v}_l$. It reduces $\mathbf{v}_k$ by subtracting the nearest integer to μ_{kl} times $\mathbf{v}_l$. This is the best possible reduction since $\mathbf{v}_k$ has to stay in the lattice. The algorithm then updates the GSO basis and coefficients.

Lemma 8. *The algorithm* `SIZE-REDUCE`(k,l) *is correct. Furthermore, the GSO of the output is equal to the GSO of the input.*

Proof. The first statement is an easy calculation. As for the second, let $A = A^*M^t$ and $B = B^*N^t$ be the matrix equations for the GSO before and after the algorithm and let $E = I_r - \lfloor \mu_{kl} \rceil E_{lk}$, where E_{lk} is the elementary (k,l)-matrix. Since B is obtained from A by subtracting $\lfloor \mu_{kl} \rceil$ times column l from column k, we have $B = AE$. But $\det E = 1$ and so the output is a lattice basis $(\mathbf{v}_1, \dots, \mathbf{v}_r)$ such that $|\mu_{kl}| \leq 1/2$. Since $l < k$, the span of the vectors $\mathbf{v}_1, \dots, \mathbf{v}_i$ does not change for all i, and so the orthogonal basis $\mathbf{v}_1^*, \dots, \mathbf{v}_r^*$ does not change by Theorem 5. □

Finally, we can state the LLL algorithm.

Algorithm 4 LLL algorithm

Input: a basis $(\mathbf{x}_1, \dots, \mathbf{x}_r)$ of an r-dimensional lattice L, and a real number $1/4 < \delta < 1$.
Output: an LLL-reduced basis $(\mathbf{v}_1, \dots, \mathbf{v}_r)$ of L.

```
for $i = 1$ to $r$ do
    $\mathbf{v}_i \leftarrow \mathbf{x}_i$.
end for
for $i = 1$ to $r$ do                                              ▷ compute the GSO
    $\mathbf{v}_i^* \leftarrow \mathbf{v}_i$.
    for $j = 1$ to $i - 1$ do
        $\mu_{ij} \leftarrow (\mathbf{v}_i \cdot \mathbf{v}_j^*)/\gamma_j^*$.
        $\mathbf{v}_i^* \leftarrow \mathbf{v}_i^* - \mu_{ij}\mathbf{v}_j^*$.
    end for
    $\gamma_i^* \leftarrow \mathbf{v}_i^* \cdot \mathbf{v}_i^*$.
end for
$k \leftarrow 2$.
while $k \leq r$ do
    SIZE-REDUCE$(k, k-1)$.
    if $\gamma_k^* \geq (\delta - \mu_{k,k-1}^2)\gamma_{k-1}^*$ then               ▷ check Lovasz' condition
        for $l = k - 2$ downto 1 do
            SIZE-REDUCE$(k, l)$.
        end for
        $k \leftarrow k + 1$.
    else
        $\mathbf{b} \leftarrow \mathbf{v}_{k-1}, \mathbf{v}_{k-1} \leftarrow \mathbf{v}_k, \mathbf{v}_k \leftarrow \mathbf{b}$.                  ▷ exchange $\mathbf{v}_{k-1}$ and $\mathbf{v}_k$
        $\nu \leftarrow \mu_{k,k-1}$.
        $\lambda \leftarrow \gamma_k^* + \nu^2\gamma_{k-1}^*$.
        $\mu_{k,k-1} \leftarrow \nu\gamma_{k-1}^*/\lambda$.
        $\gamma_k^* \leftarrow \gamma_k^*\gamma_{k-1}^*/\lambda$.
        $\gamma_{k-1}^* \leftarrow \lambda$.
        for $j = 1$ to $k - 2$ do
            $t \leftarrow \mu_{k-1,j}$, $\mu_{k-1,j} \leftarrow \mu_{kj}$, $\mu_{kj} \leftarrow t$              ▷ exchange $\mu_{k-1,j}$ and $\mu_{kj}$
        end for
        for $i = k + 1$ to $n$ do
            $\xi \leftarrow \mu_{ik}$.
            $\mu_{ik} \leftarrow \mu_{i,k-1} - \nu\mu_{ik}$.
            $\mu_{i,k-1} \leftarrow \mu_{k,k-1}\mu_{ik} + \xi$.
        end for
        if $k > 2$ then
            $k \leftarrow k - 1$.
        end if
    end if
end while
return $(\mathbf{v}_1, \dots, \mathbf{v}_r)$.
```

Note 3. At the start of each loop iteration $13-39$, the first $k-1$ vectors are already LLL-reduced. Then the k-th vector is size-reduced. If it does not satisfy Lovasz' condition, the vectors $\mathbf{v}_k$ and $\mathbf{v}_{k-1}$ are swapped and the counter k is decremented. Otherwise, k is incremented. The loop goes until k reaches the value r.

Lemma 9. *The LLL algorithm is correct.*

Proof. It follows from Remark 6 and the correctness of `SIZE-REDUCE`. □

Lemma 10. *Suppose Lovasz' condition is not satisfied for a given k, i.e. Steps $20-38$ are executed. Let $A = A^* M^t$ and $B = B^* N^t$ be the matrix equations for the GSO before and after the execution of Steps $20-38$. Then*

$$\mathbf{b}_i^* = \mathbf{v}_i^* \quad (i \neq k-1, k), \qquad \left\|\mathbf{b}_{k-1}^*\right\|^2 < \delta \left\|\mathbf{v}_{k-1}^*\right\|^2, \qquad \left\|\mathbf{b}_k^*\right\| \leq \left\|\mathbf{v}_{k-1}^*\right\|.$$

Proof. For all $i \neq k-1, k$ we have $\mathbf{v}_i = \mathbf{b}_i$. Then $\operatorname{span}(\mathbf{v}_1, \ldots, \mathbf{v}_{i-1}) = \operatorname{span}(\mathbf{b}_1, \ldots, \mathbf{b}_{i-1})$ and by Theorem 5 we have $\mathbf{b}_i^* = \mathbf{v}_i^*$ for all i different from $k-1$ and k. After exchanging, in Step 21, we have $\mathbf{b}_k = \mathbf{v}_{k-1}$ and $\mathbf{b}_{k-1} = \mathbf{v}_k$. Then the vector $\mathbf{b}_{k-1}^*$ is the component of $\mathbf{v}_k$ orthogonal to $\operatorname{span}(\mathbf{v}_1, \ldots, \mathbf{v}_{k-2})$. Recall that $\mathbf{v}_k = \mathbf{v}_k^* + \sum_{l=1}^{k-1} \mu_{kl} \mathbf{v}_l^*$. Then

$$\mathbf{b}_{k-1}^* = \mathbf{v}_k^* + \mu_{k,k-1} \mathbf{v}_{k-1}^* \quad \text{and} \quad \left\|\mathbf{b}_{k-1}^*\right\|^2 = \left\|\mathbf{v}_k^*\right\|^2 + \mu_{k,k-1}^2 \left\|\mathbf{v}_{k-1}^*\right\|^2.$$

Since the Lovasz' condition is not satisfied we have

$$\left\|\mathbf{b}_{k-1}^*\right\|^2 = \left\|\mathbf{v}_k^*\right\|^2 + \mu_{k,k-1}^2 \left\|\mathbf{v}_{k-1}^*\right\|^2 < \left(\delta - \mu_{k,k-1}^2\right) \left\|\mathbf{v}_{k-1}^*\right\|^2 + \mu_{k,k-1}^2 \left\|\mathbf{v}_{k-1}^*\right\|^2 = \delta \left\|\mathbf{v}_{k-1}^*\right\|^2.$$

We omit the proof of the last inequality. □

Note 4. The Lemma above tells us that the squared lengths of the orthogonal basis vectors decreases by a factor of at least δ. We will use this fact to show termination of the algorithm and it is also the reason we have restricted δ to be < 1. Indeed, for $\delta = 1$, it is unknown if an LLL-reduced basis can be computed in polynomial time.

Now we are going to prove that the LLL-algorithm is polynomial time.

Lemma 11. *Suppose that $\mathbf{v}_1, \ldots, \mathbf{v}_r$ is a basis of L. For $1 \leq k \leq r$, let d_k be the determinant of the k-th Gramian matrix G_k of $\mathbf{v}_1, \ldots, \mathbf{v}_r$. Then, after the execution of `SIZE-REDUCE(k,l)` the d_i's do not change. Furthermore, after the execution of Steps $20-38$ the d_i's do not change for $i \neq k-1$, but d_{k-1} changes to a new value $d'_{k-1} \leq \delta d_{k-1}$.*

Proof. By Hadamard's inequality we have $d_k = \prod_{l=1}^{k} \|\mathbf{v}_l^*\|^2$ and since, by Lemma 8, the orthogonal basis does not change, the first claim follows. For $i < k-1$, the execution of Steps $20-38$ does not change the Gramian matrix, and so d_i. For $i > k-1$, after the execution of Steps $20-38$ the new determinant d'_i is equal to $(-1)^2 d_i$, so again no effect. For $i = k-1$, we write $\mathbf{v}_l^*$ and $\mathbf{b}_l^*$ for the vectors before and after the execution of Steps $20-38$. We have

$$d'_{k-1} = \left\|\mathbf{b}_{k-1}^*\right\|^2 \prod_{l=1}^{k-2} \|\mathbf{b}_l^*\|^2 \leq \delta \left\|\mathbf{v}_{k-1}^*\right\|^2 \prod_{l=1}^{k-2} \|\mathbf{b}_l^*\|^2 = \delta \left\|\mathbf{v}_{k-1}^*\right\|^2 \prod_{l=1}^{k-2} \|\mathbf{v}_l^*\|^2 = \delta \prod_{l=1}^{k-1} \|\mathbf{v}_l^*\|^2 = \delta d_{k-1},$$

where the inequality and the third equality are given by Lemma 10. □

We associate to a basis of L the quantity $D = \prod_{k=1}^{r-1} d_k$ and denote by D_0 the value of D at the start of the LLL algorithm.

Lemma 12. *We have $D_0 \leq c^{r(r-1)}$, where $c = \max\{\|\mathbf{x}_1\|, \ldots, \|\mathbf{x}_r\|\}$.*

Proof. We have

$$D_0 = \prod_{k=1}^{r-1} d_k = \prod_{k=1}^{r-1} \prod_{l=1}^{k} \|\mathbf{x}_l^*\|^2 = \prod_{k=1}^{r-1} \|\mathbf{x}_k^*\|^{2(r-k)} \leq \prod_{k=1}^{r-1} \|\mathbf{x}_k\|^{2(r-k)} \leq \prod_{k=1}^{r-1} c^{2(r-k)} = \prod_{k=1}^{r-1} c^{2k} = c^{r(r-1)},$$

where the first inequality follows from Theorem 5. □

In the following, to simplify the investigation of the running time, we will assume that the input basis in the algorithm is integral.

Lemma 13. *Let E be the total number of executions of Steps $20 - 38$. Then*

$$E \leq -\frac{\log c}{\log \delta} r(r-1).$$

Proof. By Lemma 11, after each execution of Steps $20 - 38$, D is at most δD. Since D is a positive integer throughout the algorithm, we have $1 \leq \delta^E D_0$. Then, by Lemma 12, we have $\delta^{-E} \leq D_0 \leq c^{r(r-1)}$. The claim follows by taking logarithms. □

Note 5. The Lemma above implies termination of the algorithm. Indeed, the last execution of Steps $20 - 38$ is the last time the index k decreases. Since $2 \leq k \leq r$, there will be at most $r - 1$ more passes through the loop 13.

Lemma 14. *The total number of passes through the loop 13 is at most*

$$-\frac{2\log c}{\log \delta} r(r-1) + (r-1).$$

Proof. Let E' be the number of executions of Steps $16 - 19$. We need to evaluate $E + E'$. Every time E increases by 1, the index k decreases by 1 and every time E' increases by 1, the index k increases by 1. Then the quantity $k + E - E'$ remains constant throughout the algorithm. At the start we have $k = 2$ and $E = E' = 0$. At the end we have $k = r + 1$. Then $r + 1 + E - E' = 2$, from which $E + E' = 2E + r - 1$. We conclude using the Lemma above. □

We proved that the number of iterations is bounded by a polynomial in the input size. More specifically it is $O(r^2 \log c)$, where $c = \max\{\|\mathbf{x}_1\|, \ldots, \|\mathbf{x}_r\|\}$. To bound the running time we need to show that each iteration takes polynomial time. The number of arithmetic operation performed at each iteration is $O(r^4 \log c)$. Indeed, the computation of the GSO in Steps $4 - 11$ requires $O(r^3)$ arithmetic operations, Step 14 requires $O(n)$ operations, the $k - 2$ reductions in Steps $16 - 18$ require $O(r^2)$ operations, and the exchange requires $O(n)$ operations. Then, it is enough to show that the size of the numbers involved is polynomially bounded. This is given by the following Theorem.

Theorem 7. *The binary lengths of the integers arising in the algorithm are $O(r \log c)$.*

Proof. See [4]. □

3. Factoring over $\mathbb{Q}[X]$

Definition 10. Let R be a UFD. The *content* of a polynomial $P \in R[X]$ is the greatest common divisor of its coefficients, and it is denoted by $c(P)$. A polynomial with content 1 is called *primitive.*

Lemma 15 (Gauss' Lemma)**.** *Let R be a UFD. If $A, B \in R[X]$ are primitive, then their product AB is primitive too.*

Proof. Let $A = \sum_n a_n X^n$ and $B = \sum_n b_n X^n$. Suppose that AB is not primitive. Then there exists a prime $p \in R$ that divides all the coefficients of AB, but not all the coefficients of A and not all the coefficients of B. So we can define $h = \min\{n : p \nmid a_n\}$ and $k = \min\{n : p \nmid b_n\}$. The coefficient of X^{h+k} in AB is $c = a_{h+k}b_0 + \cdots + a_h b_k + \cdots + a_0 b_{h+k}$. p divides all terms of c but $a_h b_k$, and so p does not divide c, a contradiction. □

Remark 11. Let R be a UFD and F its field of fractions. Let $0 \neq A \in F[X]$. Then $rA \in R[X]$ for some $0 \neq r \in R$. Hence we can write $rA = sA_0$, with $s \in R$ and $A_0 \in R[X]$ primitive. Then A can be written as cA_0, with $0 \neq F$ and $A_0 \in R[X]$ primitive.

Corollary 5. *Let R be a UFD and F its field of fractions. If $A \in R[X]$ has positive degree and is reducible in $F[X]$, then $A = BC$ with $B, C \in R[X]$ having positive degree.*

Proof. Wlog we may assume that A is primitive. Suppose $A = B'C'$ with $B', C' \in F[X]$ having positive degree. Multiplying B' by $s \in R$, the least common multiple of the denominators of its coefficients, we get a primitive polynomial $B \in R[X]$. In the same way $C = tC' \in R[X]$ is primitive, where $t \in R$ is the least common multiple of the denominators of the coefficients of C'.

On the other hand $stA = stB'C' = BC$ is primitive by Gauss' Lemma. Hence $st \in R$ is a unit and the thesis follows. □

Remark 12. The Corollary above tells us that factoring in $\mathbb{Z}[X]$ is equivalent to factoring in $\mathbb{Q}[X]$ plus factoring in $\mathbb{Z}$. Since the best known algorithms for factoring in $\mathbb{Z}$ are much less efficient than those for $\mathbb{Z}[X]$ we will assume that our polynomials are primitive. Moreover, it is easy to see that we may assume the polynomials to be squarefree. Note that we could also assume the polynomials to be monic. This would simplify a little bit our calculations. Nevertheless, as already pointed out, we assume only primitiveness.

3.1. **Zassenhaus' algorithm.**

The idea is to take a "small" prime $p \in \mathbb{Z}$ not dividing the leading coefficient of f and such that its reduction modulo p is squarefree. Now we factor f modulo p using, for example, Cantor-Zassenhaus algorithm. Note that, if p is "small", this can be done quickly. Then we lift the factorization modulo p to a factorization modulo p^l, via the so called Hensel's lift. On the other hand, if f factors as $f = f_1 \cdots f_k$ in $\mathbb{Z}[X]$, then $\overline{f} = \overline{f_1} \cdots \overline{f_k}$ in $\mathbb{F}_{p^l}[X]$, where the bar denotes reduction modulo p^l. But if f_i is irreducible, then $\overline{f_i}$ need not to be irreducible, and a priori we don't know which of the irreducible modular factors of the f_i's belong together. So if we had a bound on the size of the coefficients of any integral factor of f we could try an exhaustive search on the factorization modulo p^l, i.e. we simply try all possible factor combinations to recover the integral factors. The bound is given by results of Landau and Mignotte. We will see that this leads to an exponential algorithm in the worst case (Zassenhaus' algorithm), because of the exhaustive search step. In the next section we will show how to circumvent this possibility using the LLL algorithm.

Assume we have a factorization of our polynomial f modulo a certain prime p. As mentioned above, we want to find a factorization of sufficient precision, i.e. modulo p^l for some l. This is done via the Hensel's Lemma, which essentialy shows that a factorization modulo p leads to a p-adic factorization.

Let R be an integral domain. Let $f, g, h \in R[X]$ and $m \in R$ be such that $f \equiv gh \bmod m$. We want to "lift" this to a factorization $f \equiv \hat{g}\hat{h} \bmod m^2$. Assume there exist $s, t \in R[X]$ with $sg + th \equiv 1 \bmod m$. Note that if m is a prime element this condition is equivalent to g, h being coprime in $(R/(m))[X]$, and s, t can be found using the extended Euclidean algorithm in $(R/(m))[X]$. Now we calculate

$$e_0 = f - gh, \quad \hat{g} = g + te_0, \quad \hat{h} = h + se_0$$

and note that

$$\begin{aligned} f - \hat{g}\hat{h} &= f - gh - sge_0 - the_0 - ste_0^2 \\ &= (1 - (sg + th))e_0 - ste_0^2 \\ &\equiv 0 \bmod m^2, \end{aligned}$$

since $e_0 \equiv 0 \bmod m$ and $1 - sg - th \equiv 0 \bmod m$. Then $\hat{g}\hat{h}$ is a factorization of f modulo m^2. But $\deg \hat{g} > \deg g$ and $\deg \hat{h} > \deg h$. To get equality on the degrees we need to use division with remainder in $R[X]$. This is possible if we work with a monic polynomial, as stated in the following Lemma.

Lemma 16. *Let $f, g \in R[X]$, with g nonzero and monic. Then there exist unique polynomials $q, r \in R[X]$ with $f = qg + r$ and $\deg r < \deg g$. Furthermore, if f, g, q, r are as above and $f \equiv 0 \bmod m$ for some $m \in R$, then $q \equiv r \equiv 0 \bmod m$.*

Proof. The proof of the first assertion is given by the usual division algorithm for $F[X]$, where F is a field. The second assertion follows easily. □

In the following we assume $h \in R[X]$ to be a monic polynomial. Our process consists of two steps:

- Compute
$$e \equiv f - gh \bmod m^2.$$

By the Lemma above, we can perform division with remainder of se by h in $R[X]$ to get

$$se \equiv qh + r \bmod m^2, \quad \deg r < \deg h.$$

Since $e \equiv 0 \bmod m$, we have $se \equiv 0 \bmod m$ and, using the Lemma above, we get $q \equiv r \equiv 0 \bmod m$. Now define

$$g^* \equiv g + te + qg \bmod m^2, \quad h^* \equiv h + r \bmod m^2.$$

Clearly $g^* \equiv g \bmod m$ and $h^* \equiv h \bmod m$. Furthermore, h^* is monic and $\deg h^* = \deg h$, since $\deg r < \deg h$. We have

$$\begin{aligned} f - g^*h^* &\equiv f - (g + te + qg)(h + se - qh) \\ &= f - gh - (sg + th)e - ste^2 - (sg - th)qe + ghq^2 \\ &\equiv (1 - sg - th)e - ste^2 - (sg - th)qe + ghq^2 \\ &\equiv 0 \bmod m^2, \end{aligned}$$

since $1-sg-th \equiv e \equiv 0 \bmod m$ and $q \equiv 0 \bmod m$. Finally, if the leading coefficient of f is not a zero divisor modulo m, we have that $\deg g^* = \deg f - \deg h^* = \deg f - \deg h = \deg g$.

- Now compute

  $$e^* \equiv sg^* + th^* - 1 \bmod m^2.$$

  As before, we can perform division with remainder of se^* by h^* in $R[X]$ to get

  $$se^* \equiv q^*h^* + r^* \bmod m^2, \quad \deg r^* < \deg h^*.$$

  Since $e^* \equiv 0 \bmod m$, we have $se^* \equiv 0 \bmod m$ and, using the Lemma above, we get $q^* \equiv r^* \equiv 0 \bmod m$. Now define

  $$s^* \equiv s - r^* \bmod m^2, \quad t^* \equiv t - te^* - q^*g^* \bmod m^2.$$

  Clearly $s^* \equiv s \bmod m$ and $t^* \equiv t \bmod m$. Furthermore, it is not difficult to see that $\deg s^* < \deg h^*$, $\deg t^* < \deg g^*$ and $s^*g^* + t^*h^* \equiv 1 \bmod m^2$.

The two steps above constitutes an algorithmic proof of the following result. Moreover, note that the algorithm above is essentially linear in the input size.

Lemma 17 (Quadratic Hensel's Lemma). *Let R be an integral domain, $0 \neq p \in R$ not a zero divisor, and $n \in \mathbb{N}_{>0}$. Let $f, g, h, s, t \in R[X]$ be polynomials such that*

$$f \equiv gh \bmod p, \quad sg + th \equiv 1 \bmod p,$$

the leading coefficient of f is not a zero divisor modulo p, h is monic, $\deg f = \deg g + \deg h$, $\deg s < \deg h$ and $\deg t < \deg g$.

Then there exist polynomials $g^*, h^*, s^*, t^* \in R[X]$ such that

$$f \equiv g^*h^* \bmod p^{2^n}, \quad s^*g^* + t^*h^* \equiv 1 \bmod p^{2^n},$$

h^* is monic, $g^* \equiv g \bmod p$, $h^* \equiv h \bmod p$, $s^* \equiv s \bmod p$, $t^* \equiv t \bmod p$, $\deg g^* = \deg g$, $\deg h^* = \deg h$, $\deg s^* < \deg h^*$, and $\deg t^* < \deg g^*$. Moreover, the polynomials g^* and h^* are unique modulo p^{2^n}.

Remark 13. As mentioned above, if $p \in R$ is prime, the existence of s, t satisfying $sg+th \equiv 1 \bmod p$ is equivalent to the fact that g and h are coprime in $(R/(p))[X]$.

Proof. After n iterations of the algorithm given above we clearly get the result.

It remains to prove uniqueness. Suppose that $g_1, h_1, g_2, h_2 \in R[X]$ are as in the thesis but $g_1 \not\equiv g_2 \bmod p^{2^n}$ or $h_1 \not\equiv h_2 \bmod p^{2^n}$. Let $1 \leq i < n$ be maximal such that p^{2^i} divides both $g_1 - g_2$ and $h_1 - h_2$, i.e. $g_1 - g_2 = up^{2^i}$ and $h_1 - h_2 = vp^{2^i}$ for some $u, v \in R[X]$ such that $p \nmid u$ or $p \nmid v$. Assume wlog that $p \nmid u$. Then

$$0 \equiv g_1h_1 - g_2h_2 = g_1(h_1 - h_2) + h_2(g_1 - g_2) = (g_1v + h_2u)p^{2^i} \bmod p^{2^n}.$$

Since p is not a zero divisor, we have $p \mid p^{2^{n-i}} \mid (g_1v + h_2u)$. Denoting by a bar the reduction modulo p, we have $\overline{sg_2} + \overline{th_2} = 1$, $\overline{g_1} = \overline{g_2}$, and $\overline{g_1v} + \overline{h_2u} = 0$. Then

$$0 = \bar{t}(\overline{g_1v} + \overline{h_2u}) = \overline{tg_2v} + (1 - \overline{sg_2})\overline{u} = (\overline{tv} - \overline{su})\overline{g_2} + \overline{u},$$

from which $\overline{g_2} \mid \overline{u}$. We assumed $g_1 - g_2 = up^{2^i}$, with $p \nmid u$. But g_1 and g_2 have the same leading coefficient and the same degree. Hence $\deg \overline{u} < \deg \overline{g_2}$. Therefore $\overline{u}$ is the zero polynomial, a contradiction. □

Hensel's Lemma is true in much more generality. Indeed, let K be a field complete with respect to a nonarchimedean valuation $|\ |$. Let $\mathcal{O} = \{x \in K : |x| \leq 1\}$ be the corresponding valuation ring, with maximal ideal $\mathfrak{p} = \{x \in K : |x| < 1\}$ and residue class field $k = \mathcal{O}/\mathfrak{p}$.

Theorem 8 (Hensel's Lemma)**.** *Let $f \in \mathcal{O}[X]$ be a polynomial such that $f \not\equiv 0 \bmod \mathfrak{p}$. If f admits modulo $\mathfrak{p}$ a factorization*

$$f \equiv \overline{g}\overline{h} \bmod \mathfrak{p}$$

into coprime polynomials $\overline{g}, \overline{h} \in k[X]$, then f admits a factorization

$$f = gh$$

into polynomials $g, h \in \mathcal{O}[X]$ such that $\deg g = \deg \overline{g}$ and

$$g \equiv \overline{g} \bmod \mathfrak{p}, \quad h \equiv \overline{h} \bmod \mathfrak{p}.$$

Proof. See [13]. □

Remark 14. Let $K = \mathbb{Q}_p$ be the completion of $\mathbb{Q}$ with respect to the nonarchimedean p-adic absolute value $|\ |_p : \mathbb{Q} \longrightarrow \mathbb{R}$. Then the corresponding valuation ring is $\mathbb{Z}_p$, the ring of p-adic integers, with maximal ideal $p\mathbb{Z}_p$. It is not difficult to see that the residue class field is isomorphic to $\mathbb{Z}/p\mathbb{Z}$. Now, if we think about the elements of $\mathbb{Z}_p$ as formal series

$$\sum_{i=0}^{\infty} a_i p^i, \quad 0 \leq a_i < p,$$

(more formally, we are identifying $\mathbb{Z}_p$ with the projective limit $\varprojlim \mathbb{Z}/p^n\mathbb{Z}$), it is clear that the Theorem above is indeed a generalization of Lemma 17. Moreover, it is worth noting that one of the equivalent formulations of Hensel's Lemma is in fact an adaptation of Newton's method in calculus for finding roots of polynomials. For further discussions on this topic we refer to the excellent presentations in [13] and [11].

At this point let us recall our strategy to factor $f \in \mathbb{Z}[X]$:

(1) make f primitive and squarefree

(2) pick a suitable prime p

(3) factor f in $\mathbb{F}_p[X]$

(4) lift to a factorization modulo a large enough power p^k

(5) recover the true factors in $\mathbb{Z}[X]$.

In the following we will see that the right value of k to be chosen in step (4) depends on how large the coefficients of an irreducible factor of f can be. Moreover, a bound for the coefficients is needed. Indeed, if we had no bound, we would have to continue lifting indefinitely, and our method would fail to terminate on some inputs. So let us look for a bound on the coefficients of the integral factors of f.

Definition 11. Let $f = \sum_{0 \leq i \leq n} f_i X^i \in \mathbb{C}[X]$ and $s > 0$. The *s-norm* of f is

$$\|f\|_s = \left(\sum_{i=0}^{n} |f_i|^s \right)^{1/s},$$

where $|a|$ is the absolute value of $a \in \mathbb{C}$.

Recall that the 2-norm of f is nothing but the Euclidean norm of the vector $(f_0, \dots, f_n) \in \mathbb{C}^{n+1}$.

We want a bound for the norm of factors of f in terms of $\|f\|_2$. At a first glance one might hope that $\|h\|_2 \leq \|f\|_2$ for any factor $h \in \mathbb{Z}[X]$ of f. This is not the case. Indeed, consider $f = X^n - 1$ and as h take the n-th cyclotomic polynomial. The following Theorem gives an upper bound involving the Mahler's measure of f.

Theorem 9. *Let $h = \sum_{0\leq i\leq m} h_i X^i \in \mathbb{C}[X]$ be a factor of*

$$f = \sum_{0\leq i\leq n} f_i X^i = f_n \prod_{1\leq i\leq n} (X - z_i) \in \mathbb{C}[X].$$

Then

$$|h_i| \leq \left|\frac{h_m}{f_n}\right| \binom{m}{i} M(f), \quad 0 \leq i \leq m$$

where $M(f) := |f_n| \prod_{1\leq i\leq n} \max\{1, |z_i|\}$ is called the Mahler's measure of f.

Proof. Write $h = h_m \prod_{1\leq i\leq m}(X - u_i)$, with $u_i \in \mathbb{C}$. By Vieta's formulas, we have

$$h_i = (-1)^{m-i} h_m \sum_{\substack{S\subseteq\{1,\dots,m\}\\ |S|=m-i}} \prod_{j\in S} u_j.$$

Hence

$$|h_i| \leq |h_m| \sum_S \prod_{j\in S} |u_j| \leq \binom{m}{i} |h_m| \prod_{1\leq j\leq m} \max\{1, |u_j|\} = \binom{m}{i} M(h),$$

for $0 \leq i \leq m$. Since $M(h)/|h_m| \leq M(f)/|f_n|$, the thesis follows. □

Remark 15. The Mahler's measure of an algebraic number is defined as the Mahler's measure of its minimal polynomial. The proof of Theorem 9 gives us a proof of a special case of the Northcott's Theorem, which says that there are only finitely many algebraic numbers of bounded degree and bounded Mahler's measure. The Mahler's measure is related to the so called *height* of an algebraic number, which is a function required to satisfy an analogous property to the one in Northcott's Theorem, and it is commonly used to prove finiteness results. Examples of this can be found in Diophantine geometry (Mordell-Weil theorem), Diophantine approximation (subspace theorem) and arithmetic dynamics (finiteness of preperiodic points). See [3] and [10].

Theorem 10 (Northcott's Theorem, special case)**.** *For every $d \in \mathbb{N}$ and $B \in \mathbb{R}$ the set*

$$\mathcal{M} := \left\{\alpha \in \overline{\mathbb{Q}} : [\mathbb{Q}(\alpha) : \mathbb{Q}] = d,\ M(\alpha) \leq B\right\}$$

is finite.

Proof. Suppose $\alpha \in \mathcal{M}$ has minimal polynomial

$$M_\alpha = a_0 + \cdots + a_d X^d = a_d(X - \alpha_1)\cdots(X - \alpha_d) \in \mathbb{Z}[X].$$

Then, as in the proof of Theorem 9, we have

$$|a_i| \leq \binom{d}{i} |a_d| \prod_{j=1}^{d} \max\{1, |\alpha_j|\} \leq 2^d B.$$

Therefore, there are at most $\left(2^{d+1}B + 1\right)^{d+1}$ distinct minimal polynomials of some $\alpha \in \mathcal{M}$ in $\mathbb{Z}[X]$. Any of these has at most d roots, hence $|\mathcal{M}| < d\Big(\left(2^{d+1}B + 1\right)^{d+1}\Big) < \infty$. □

Corollary 6 (Kronecker's theorem)**.** *Let α be a nonzero algebraic number. Then $M(\alpha) = 1$ if and only if α is a root of unity.*

Proof. One implication is trivial. Then suppose that $M(\alpha) = 1$ for $\alpha \neq 0$ and $[\mathbb{Q}(\alpha) : \mathbb{Q}] = d$. By Northcott's Theorem, the set

$$\mathcal{M} = \left\{\beta \in \overline{\mathbb{Q}} : [\mathbb{Q}(\beta) : \mathbb{Q}] \leq d,\ M(\beta) \leq 1\right\}$$

is finite. Moreover, it is not difficult to see that $M(\alpha^2) \leq M(\alpha)^2$. Then $\alpha^2, \alpha^4, \dots$ are all contained in $\mathcal{M}$. Hence there exist $i < j$ such that $\alpha^{2^i} = \alpha^{2^j}$, from which $\alpha^{2^j - 2^i} = 1$. □

Remark 16. For $f \in \mathbb{Z}[X]$, clearly $M(f) \geq 1$, and Kronecker's theorem tells us that $M(f) = 1$ precisely when f is a product of cyclotomic polynomials and a power of X. *Lehmer's problem* asks if there exists a polynomial f satisfying $1 < M(f) < 1+\varepsilon$, for every $\varepsilon > 0$. In 1933 Lehmer noted that the polynomial $l = X^{10} + X^9 - X^7 - X^6 - X^5 - X^4 - X^3 + X + 1$ has $M(l) = 1.176280\ldots$, and this value remains the smallest known. Lehmer's problem has been solved only for several special classes of polynomials. See [3] and [10] for further discussions on this topic.

Lemma 18. *Let $f \in \mathbb{C}[X]$ and $z \in \mathbb{C}$. Then $\left\|(X-z)f\right\|_2 = \left\|(\overline{z}X - 1)f\right\|_2$.*

Proof. Write $f = \sum_{0\leq i\leq n} f_i X^i$ and let $f_{-1} = f_{n+1} = 0$. Then

$$\begin{aligned}
\left\|(X-z)f\right\|_2^2 &= \sum_{0\leq i\leq n+1} \left|f_{i-1} - zf_i\right|^2 \\
&= \sum_{0\leq i\leq n+1} \left(f_{i-1} - zf_i\right)\left(\overline{f_{i-1}} - \overline{z}\overline{f_i}\right) \\
&= \|f\|_2^2\left(1+|z|^2\right) - \sum_{0<i<n+1}\left(z\overline{f_{i-1}}f_i + \overline{z}f_{i-1}\overline{f_i}\right) \\
&= \sum_{0\leq i\leq n+1}\left(\overline{z}f_{i-1} - f_i\right)\left(z\overline{f_{i-1}} - \overline{f_i}\right) \\
&= \sum_{0\leq i\leq n+1}\left|\overline{z}f_{i-1} - f_i\right|^2 \\
&= \left\|(\overline{z}X - 1)f\right\|_2^2.
\end{aligned}$$

□

Theorem 11 (Landau's inequality)**.** *For any $f = \sum_{0\leq i\leq n} f_i X^i = f_n \prod_{1\leq i\leq n}(X - z_i) \in \mathbb{C}[X]$, we have $M(f) \leq \|f\|_2$.*

Proof. We arrange the roots so that $|z_1|, \ldots, |z_k| > 1$ and $|z_{k+1}|, \ldots, |z_n| \leq 1$. Then $M(f) = |f_n \cdot z_1 \cdots z_k|$. Let

$$g = f_n \prod_{1\leq i\leq k}\left(\overline{z_i}X - 1\right)\prod_{k<i\leq n}(X - z_i) = g_n X^n + \cdots + g_0 \in \mathbb{C}[X].$$

Then, using repeatedly Lemma 18, we have

$$M(f)^2 = |f_n \cdot z_1 \cdots z_k|^2 = |g_n|^2 \leq \|g\|_2^2 = \left\|\frac{g}{(\overline{z_1}X - 1)\cdots(\overline{z_k}X - 1)}(X - z_1)\cdots(X - z_k)\right\|_2^2 = \|f\|_2^2.$$

□

Theorem 12 (Mignotte's bound)**.** *Suppose $f, g, h \in \mathbb{Z}[X]$ have degrees* $\deg f = n \geq 1$, $\deg g = m$, $\deg h = k$, *and that gh divides f in $\mathbb{Z}[X]$. Then*

$$\|g\|_\infty \|h\|_\infty \leq \|g\|_2 \|h\|_2 \leq \|g\|_1 \|h\|_1 \leq 2^{m+k}\|f\|_2 \leq (n+1)^{1/2} 2^{m+k}\|f\|_\infty.$$

In particular, letting $g = 1$, we have

$$\|h\|_\infty \leq \|h\|_2 \leq \|h\|_1 \leq 2^k \|f\|_2 \leq (n+1)^{1/2} 2^k \|f\|_\infty.$$

Proof. By the proof of Theorem 9 we have $|g_i| \leq \binom{m}{i} M(g)$ for $0 \leq i \leq m$. Then, summing over i, we have

$$\|g\|_1 = \sum_{0\leq i\leq m} |g_i| \leq 2^m M(g).$$

We get similar inequalities for h. Multiplying them, we obtain

$$\|g\|_1 \|h\|_1 \leq 2^{m+k} M(g) M(h) \leq 2^{m+k} M(f) \leq 2^{m+k}\|f\|_2.$$

The remaining inequalities are well-known. □

Let $f = f_0 + \cdots f_n X^n \in \mathbb{Z}[X]$ be a squarefree primitive polynomial of degree n and set

$$B = (n+1)^{1/2} 2^n \|f\|_\infty |f_n| .$$

We choose a prime p such that $f_n \not\equiv 0 \bmod p$ and such that $f \bmod p$ is squarefree. Now take a positive integer k such that $p^k > 2B$. Thanks to Hensel's Lemma we have a factorization of f modulo p^k:

$$f \equiv f_n g_1 \cdots g_r \bmod p^k, \quad g_1, \ldots, g_r \in \mathbb{Z}[X].$$

We assume that $g_1, \ldots, g_r$ are monic and, using symmetric representatives, that $\|g_i\|_\infty < p^k/2$ for all i. Now we choose $g, h \in \mathbb{Z}[X]$ such that

$$g \equiv f_n \prod_{i \in S \subsetneq \{1,\ldots,r\}} g_i \bmod p^k,$$

$$h \equiv f_n \prod_{i \in \{1,\ldots,r\} \setminus S} g_i \bmod p^k,$$

and

$$\|g\|_\infty , \|h\|_\infty < p^k/2.$$

Lemma 19. *With notation as above, $f_n f = gh$ in $\mathbb{Z}[X]$ if and only if $\|g\|_1 \|h\|_1 \leq B$. In words, the factorization with symmetric representatives modulo p^k is a factorization over $\mathbb{Z}$.*

Proof. Suppose $f_n f = gh$ in $\mathbb{Z}[X]$. Then, by Mignotte's bound, we have

$$\|g\|_1 \|h\|_1 \leq (n+1)^{1/2} 2^n \|f_n f\|_\infty = (n+1)^{1/2} 2^n \|f\|_\infty |f_n| = B.$$

Conversely, suppose that $\|g\|_1 \|h\|_1 \leq B$. By the well-known inequalities for the norms, we have

$$\max\big\{ \|g\|_\infty , \|h\|_\infty \big\} \leq \|gh\|_\infty \leq \|gh\|_1 \leq \|g\|_1 \|h\|_1 \leq B < \frac{p^k}{2}.$$

Therefore, $f_n f \equiv gh \bmod p^k$ is actually an equality in $\mathbb{Z}[X]$. □

Therefore, all what we have to do is to find an appropriate set S. This will be done by exhaustive search. Note that, in order to apply the Cantor-Zassenhaus algorithm, we need to find a "small" prime p such that $f \bmod p$ is squarefree. This can be done, for example, using the sieve of Eratosthenes. First, we need a bound on the "bad" primes. With this in mind, let us recall briefly the notion of resultant, which is based on the following Lemma.

Lemma 20. *Let F be a field and let $f, g \in F[X]$ be nonzero polynomials. Then* $\gcd(f,g) \neq 1$ *if and only if there exist $s, t \in F[X] \setminus \{0\}$ such that $sf + tg = 0$,* $\deg s < \deg g$, *and* $\deg t < \deg f$.

Proof. Let $d = \gcd(f,g)$. If $d \neq 1$, then $\deg h \geq 1$ and it is enough to take $s = -g/d$, $t = f/h$. Conversely, assume there exist s, t as in the statement. Suppose that f, g are coprime. Since $sf = -tg$, then $f \mid t$. This is a contradiction, assuming that $t \neq 0$ and $\deg f > \deg t$. □

Definition 12. Let R be a commutative ring and let $f = f_0 + \cdots + f_n X^n, g = g_0 + \cdots + g_m X^m$ be polynomials in $R[X]$. The **Sylvester matrix** of f and g is the following $(m+n) \times (m+n)$ matrix:

$$S(f,g) = \begin{pmatrix} f_n & & & g_m & & \\ f_{n-1} & \ddots & & g_{m-1} & \ddots & \\ \vdots & \ddots & f_n & \vdots & \ddots & g_m \\ f_0 & & f_{n-1} & g_0 & & g_{m-1} \\ & \ddots & \vdots & & \ddots & \vdots \\ & & f_0 & & & g_0 \end{pmatrix}.$$

The **resultant** of f and g is defined as $\operatorname{res}(f,g) = \det S$. The **discriminant** of f, denoted by $\operatorname{disc}(f)$, is the resultant of f and its formal derivative f'.

Remark 17. If R is a field, in the Definition above, then $S(f,g)$ is nothing but the matrix representing the linear map

$$\varphi_{f,g} : P_m \times P_n \longrightarrow P_{m+n}, \quad \varphi_{f,g}(s,t) = sf + tg,$$

where P_i is the i-dimensional F-vector space of polynomials with degree less than i, with respect to the canonical bases.

Corollary 7. *Let $f, g \in F[X]$ be nonzero polynomials of degrees n, m, respectively. Then* $\gcd(f,g) = 1$ *if and only if $res(f,g) \neq 0$*

Proof. By Lemma 20, $\deg \gcd(f,g) > 1$ if and only if there exists a nonzero $(s,t) \in P_m \times P_n$ such that $\varphi_{f,g}(s,t) = 0$. Since $P_m \times P_n$ and P_{m+n} have the same dimension, this happens if and only if $0 \neq \det S(f,g) = \operatorname{res}(f,g)$. □

Lemma 21. *Let $f, g \in \mathbb{Z}[X]$ with degrees n, m, respectively. Then*

$$|res(f,g)| \leq \|f\|_2^m \|g\|_2^n \leq (n+1)^{m/2}(m+1)^{n/2} \|f\|_\infty^m \|g\|_\infty^n .$$

Proof. The first inequality follows easily from Hadamard's inequality. The second is the well-known inequality between norms. □

Remark 18. If $g = f'$ in the Lemma above, then $m = n-1$ and $\|f'\|_\infty \leq n \|f\|_\infty$. Therefore

$$\begin{array}{rcl} |\operatorname{disc}(f)| & \leq & (n+1)^{(n-1)/2} n^{n/2} \|f\|_\infty^{n-1} \|f'\|_\infty^n \\ & \leq & (n+1)^{(n-1)/2} n^{3n/2} \|f\|_\infty^{2n-1} \\ & \leq & (n+1)^{2n} \|f\|_\infty^{2n-1} . \end{array}$$

Theorem 13. *Let $f \in \mathbb{Z}[X]$ be a nonzero squarefree polynomial and let p be a prime not dividing the leading coefficient of f. Then $f \bmod p$ is squarefree if and only if p does not divide disc(f).*

Proof. See [17]. □

Now, with the sieve of Eratosthenes, we compute the first γ primes, where

$$\gamma = \left\lceil 2 \log_2 \left((n+1)^{2n} \|f\|_\infty^{2n-1} \right) \right\rceil .$$

As a consequence of the Prime Number Theorem, each of them is less than $2\gamma \ln \gamma$ and, by our choice of γ, no more than half of these primes divide $\operatorname{disc}(f)$. If needed, see [8] for details. Finally, we can state a factorization algorithm. In the following denote by $\operatorname{lc}(f)$ the leading coefficient of the polynomial f.

Algorithm 5 Zassenhaus' algorithm

```
Input: a squarefree primitive nonconstant polynomial $f = f_0 + \cdots + f_n X^n \in \mathbb{Z}[X]$ with $f_n > 0$.
Output: the irreducible factors $\{f_1, \ldots, f_r\} \subseteq \mathbb{Z}[X]$ of $f$.
1: $n \leftarrow \deg f$.
2: if $n = 1$ then
3: return $\{f\}$.
4: else
5:     $b \leftarrow \mathrm{lc}(f)$.
6:     $C \leftarrow (n+1)^{2n} \|f\|_\infty^{2n-1}$.
7:     $r \leftarrow \lceil 2 \log_2 C \rceil$.
8:     $B \leftarrow (n+1)^{1/2} 2^n \|f\|_\infty b$.
9:     Using the sieve of Eratosthenes, find a prime $p < 2r \ln r$ such that $p \nmid b$ and $p \nmid \mathrm{disc}(f)$.
10:    $k \leftarrow \lceil \log_p(2B+1) \rceil$.
11:    Using Cantor-Zassenhaus, find noncostant monic irreducible polynomials $h_1, \ldots, h_r \in \mathbb{Z}[X]$ such that $f \equiv b h_1 \cdots h_r \bmod p$ and $\|h_j\|_\infty < p/2$ for all $j$.
12:    Using Quadratic Hensel's Lemma repeatedly, find noncostant monic irreducible polynomials $g_1, \ldots, g_r \in \mathbb{Z}[X]$ such that $f \equiv b g_1 \cdots g_r \bmod p^k$, $g_j \equiv h_j \bmod p$ and $\|g_j\|_\infty < p^k/2$ for all $j$.
13:    $T \leftarrow \{1, \ldots, r\}$.        ▷ initialize the index set $T$ of modular factors still to be treated
14:    $I \leftarrow \emptyset$.                ▷ initialize the set of factors found
15:    $F \leftarrow f$.                        ▷ initialize the polynomial still to be factored
16:    $s \leftarrow 1$.
17:    while $2s \leq |T|$ do                   ▷ factor combination
18:        for all subsets $S \subseteq T$ of cardinality $s$ do
19:            Find $g, h \in \mathbb{Z}[X]$ such that $g \equiv b \prod_{j \in S} g_j \bmod p^k$, $h \equiv b \prod_{j \in T \setminus S} g_j \bmod p^k$ and $\|g\|_\infty, \|h\|_\infty < p^k/2$.
20:            if $\|g\|_1 \|h\|_1 \leq B$ then
21:                $T \leftarrow T \setminus S$.
22:                $I \leftarrow I \cup \{g/c(g)\}$.
23:                $F \leftarrow h/c(h)$.
24:                $b \leftarrow \mathrm{lc}(F)$.
25:                break the loop 18 and goto Step 17.
26:            end if
27:        end for
28:        $s \leftarrow s + 1$.
29:    end while
30: end if
31: return $I \cup F$.
```

Theorem 14. *Zassenhaus' algorithm works correctly.*

Proof. For a factor $u \in \mathbb{Z}[X]$ of f, denote by $\mu(u)$ the number of monic irreducible factors which divide u modulo p. Since $\mathbb{F}_p$ is a UFD, these factors form a subset of $\{g_1, \ldots, g_r\}$. We claim that the following are invariants:

- $F \equiv b \prod_{j \in T} g_j \bmod p^k$
- $b = \mathrm{lc}(F)$
- $f = F \prod_{g^* \in I} g^*$
- each polynomial in I is irreducible
- F is primitive and each of its irreducible factors $u \in \mathbb{Z}[X]$ has $\mu(u) \geq s$

They hold initially. So assume that they hold before Step 19 and the condition in Step 20 is true for some $S \subseteq T$ of cardinality s. Then $gh = bF$, by Lemma 19, and $g/c(g)$ is a factor of $bF/c(bF) = F$. But $\mu(g) = s$ and $\mu(u) \geq s$ for each irreducible factor u of f. Therefore, $g/c(g)$ is an irreducible factor of f. Then it is trivial to see that all the points above are true after Step 20.

Now assume that the condition in Step 20 is false for all subsets $S \subseteq T$ of cardinality s, but that F has an irreducible factor $\tilde{g} \in \mathbb{Z}[X]$ with $\mu(\tilde{g}) = s$. Let $\tilde{h} = F/\tilde{g}$ and let $\tilde{S} \subseteq T$ be such that

$$\mathrm{lc}(\tilde{h})\tilde{g} \equiv b \prod_{j \in \tilde{S}} h_j \bmod p, \quad \text{and} \quad \mathrm{lc}(\tilde{g})\tilde{h} \equiv b \prod_{j \in T \setminus \tilde{S}} h_j \bmod p.$$

Since $F = \tilde{g}\tilde{h}$, and using the invariants above, we have that

$$bF = \mathrm{lc}(\tilde{g})\mathrm{lc}(\tilde{h})\tilde{g}\tilde{h} \bmod p^k.$$

Moreover, let

$$\hat{g} \equiv b \prod_{j \in \tilde{S}} g_j \bmod p^k \quad \text{and} \quad \hat{h} \equiv b \prod_{j \in T \setminus \tilde{S}} g_j \bmod p^k,$$

as in Step 19. By the first invariant, we have that $bF \equiv \hat{g}\hat{h} \bmod p^k$. Uniqueness in Hensel's Lemma implies that $\mathrm{lc}(\tilde{h})\tilde{g} \equiv \hat{g} \bmod p^k$ and $\mathrm{lc}(\tilde{g})\tilde{h} \equiv \hat{h} \bmod p^k$. By Mignotte's bound the coefficients of $\mathrm{lc}(\tilde{h})\tilde{g}$ and $\mathrm{lc}(\tilde{g})\tilde{h}$ are at most $B < p^k/2$. On the other hand $\|\hat{g}\|_\infty, \|\hat{h}\|_\infty < p^k/2$. Therefore, $\mathrm{lc}(\tilde{h})\tilde{g} = \hat{g}$, $\mathrm{lc}(\tilde{g})\tilde{h} = \hat{h}$ and $\hat{g}\hat{h} = bF$. But this means, by Lemma 19, that the condition in Step 20 is true for $\tilde{S}$, a contradiction. Then F has no irreducible factor $\tilde{g}$ with $\mu(\tilde{g}) = s$, and Step 28 guarantees that the invariants hold again at the next pass through Step 17.

It remains to show that F is irreducible if $2s > |T|$ in Step 17. Let $g \in \mathbb{Z}[X]$ be an irreducible factor of F and $h = F/g$. By the fifth invariant, we have $s \leq \mu(g), \mu(h) \leq |T|$ if h is nonconstant. But $2s \leq \mu(g) + \mu(h) = |T|$. Therefore h is a constant and F is irreducible. □

Unfortunately, the number of subsets S to consider in Step 18 is a priori exponential in $n = \deg f$. Indeed, if f is irreducible, we have to consider $2^{n/2} - 1$ sets. Even worse, there exist "bad" polynomials for which this could happen no matter which prime p we choose. This is the content of the following Proposition.

Proposition 1. *Let p_n denote the n-th prime number and consider the field extension*

$$\mathbb{Q}(\sqrt{p_1}, \ldots, \sqrt{p_n})/\mathbb{Q}.$$

Let f_n be the minimal polynomial of $\alpha_n = \sqrt{p_1} + \cdots + \sqrt{p_n}$, also called the ***$n$-th Swinnerton-Dyer polynomial****. Then f_n factorizes modulo any prime p into factor of degree at most* 2.

Proof. It can be seen that the roots of f_n are $\varepsilon_1\sqrt{p_1} + \cdots + \varepsilon_n\sqrt{p_n}$, where $\varepsilon_i \in \{-1, 1\}$. Then $f_n = \prod(X \pm \sqrt{p_1} \pm \cdots \pm \sqrt{p_n})$. We claim that $f_n \in \mathbb{Z}[X]$. Indeed, since $\sqrt{p_i} \in \mathbb{Q}(\sqrt{p_1}, \ldots, \sqrt{p_n})$ is an algebraic integer, so is α_n. Therefore, its minimal polynomial is in $\mathbb{Z}[X]$. Now let p be a prime. Since $\mathbb{F}_{p^2}$ contains all the square roots $\sqrt{p_1} \bmod p, \ldots, \sqrt{p_n} \bmod p$, then the reduction of f_n modulo p splits into linear factors over $\mathbb{F}_{p^2}$ and so into factor of degree at most 2 over $\mathbb{F}_p$. □

On the other hand it is clear that f_n is an irreducible polynomial in $\mathbb{Q}[X]$ of degree 2^n. Then we have the following Corollary.

Corollary 8. *Zassenhaus' algorithm is exponential time in the worst case.*

Remark 19. There are other examples of "bad" polynomials. Indeed, the cyclotomic polynomials Φ_n split modulo every prime, for all n such that $(\mathbb{Z}/n\mathbb{Z})^*$ is not cyclic. The reason of this phenomena is Galois theoretic. Note also that $\mathrm{Gal}(\mathbb{Q}(\sqrt{p_1}, \ldots, \sqrt{p_n})/\mathbb{Q} \cong (\mathbb{Z}/2\mathbb{Z})^n$. See [13] for details.

3.2. **Factoring using LLL algorithm.** In this section we will provide a replacement for the factor combination in the Zassenhaus' algorithm, avoiding the exhaustive search. This will result in a deterministic polynomial time algorithm, which on contrast is not very practical. Roughly speaking, the idea is the following. We want to factorize f. For a suitable small prime p we find a p-adic irreducible factor u of f, to a certain precision. We look for a polynomial g which also has u as a p-adic factor to the same precision. This condition can be translated into membership to a certain lattice, and a short vector in this lattice will enable us to find a factor of f. Let us begin with a technical Lemma.

Lemma 22. *Let R be an integral domain and $f, g \in R[X]$ nonzero polynomial with* $\deg f + \deg g \geq 1$*. Then there exist nonzero $s, t \in R[X]$ such that $sf + tg = res(f, g)$, $\deg s < \deg g$, and $\deg t < \deg f$.*

Proof. Let F be the field of fractions of R. If $\operatorname{res}(f, g) = 0$, then the thesis follows easily from Remark 17. So assume the resultant to be nonzero. Then, by Corollary 7, f and g are coprime in $F[X]$, and so there exist nonzero $s^*, t^* \in F[X]$ satisfying the required degree bounds and such that $s^*f + t^*g = 1$. Again by Remark 17, s^*, t^* are the unique solution of a linear system with coefficient matrix $S(f, g)$. Therefore, by Cramer's rule, $s^* = s/\det S(f, g)$ and $t^* = t/\det S(f, g)$, where $s, t \in R[X]$ are the determinants of submatrices of $S(f, g)$. □

The following Lemma is the crucial observation which will lead to the algorithm. It says that if two polynomials in $\mathbb{Z}[X]$ have a nonconstant common divisor modulo an m which is larger than their resultant, then they have a nonconstant common factor in $\mathbb{Z}[X]$.

Lemma 23. *Let $f, g \in \mathbb{Z}[X]$ have positive degrees n, k, respectively. Let $u \in \mathbb{Z}[X]$ be a nonconstant, monic polynomial dividing both f and g modulo m for some $m > \|f\|_2^k \|g\|_2^n$. Then* $\gcd(f, g) \in \mathbb{Z}[X]$ *is nonconstant.*

Proof. Suppose that $\gcd(f, g) = 1$ in $\mathbb{Q}[X]$. By the Lemma above, there exist $s, t \in \mathbb{Z}[X]$ such that $sf + tg = \operatorname{res}(f, g)$. Since u divides both f and g modulo m, then it divides $\operatorname{res}(f, g)$ modulo m. But u is monic and nonconstant, and so $\operatorname{res}(f, g) \equiv 0 \bmod m$. On the other hand, by Lemma 21, we have $|\operatorname{res}(f, g)| \leq \|f\|_2^k \|g\|_2^n < m$. Then, $\operatorname{res}(f, g) = 0$, a contradiction with Corollary 7. Therefore, $\gcd(f, g)$ in $\mathbb{Q}[X]$ is nonconstant, and the same is true for $\gcd(f, g)$ in $\mathbb{Z}[X]$. □

The idea is the following. Let $f \in \mathbb{Z}[X]$ be the primitive squarefree polynomial of degree n to be factored. Suppose we have computed a monic polynomial $u \in \mathbb{Z}[X]$ that divides f modulo some m. Then we look for a polynomial $g \in \mathbb{Z}[X]$ which is also divisible by u modulo m and which is "short", i.e. $\|g\|_2^n < m \|f\|_2^{-\deg g}$. Then the Lemma above gives us a nontrivial factor of f in $\mathbb{Z}[X]$. Now consider the lattice $L \subseteq \mathbb{Z}^j$ with basis given by the coefficient vectors of

$$\left\{uX^i : 0 \leq i < j - \deg u\right\} \cup \left\{mX^i : 0 \leq i < \deg u\right\}.$$

If $g \in L$, then $g = qu + rm$ with $q, r \in \mathbb{Z}[X]$, $\deg q < j - \deg u$, $\deg r < \deg u$. Therefore, $\deg g < j$ and u divides g modulo m. Conversely, suppose that $g \in \mathbb{Z}[X]$ is of degree less than j and divisible by u modulo m. Then $g = q^*u + r^*m$ for some $q^*, r^* \in \mathbb{Z}[X]$. Division with remainder by the monic polynomial u yields $q^{**}, r^{**} \in \mathbb{Z}[X]$ such that $r^* = q^{**}u + r^{**}$ and $\deg r^{**} < \deg u$. Therefore, $g = q^*u + (q^{**}u + r^{**})m = (q^* + mq^{**})u + r^{**}m$ with

$$\deg\left(q^* + mq^{**}\right) \leq \max\left\{\deg q^*, \deg q^{**}\right\} \leq \max\left\{\deg g - \deg u, \deg r^* - \deg u\right\} < j - \deg u.$$

To summarize, $g \in L$ if and only if $\deg g < j$ and u divides g modulo m. Our task is to find a short vector in L corresponding to a polynomial $g \in \mathbb{Z}[X]$ satisfying the properties above and such that $\|g\|_2^n < m \|f\|_2^{-\deg g}$.

Algorithm 6 LLL factoring algorithm

```
Input: a squarefree primitive nonconstant polynomial $f = f_0 + \cdots + f_n X^n \in \mathbb{Z}[X]$ with $f_n > 0$.
Output: the irreducible factors $\{f_1, \dots, f_r\} \subseteq \mathbb{Z}[X]$ of $f$.
 1: $n \leftarrow \deg f$.
 2: if $n = 1$ then
 3: return $\{f\}$.
 4: else
 5:     $b \leftarrow \mathrm{lc}(f)$.
 6:     $C \leftarrow (n+1)^{2n} \|f\|_\infty^{2n-1}$.
 7:     $r \leftarrow \lceil 2 \log_2 C \rceil$.
 8:     $B \leftarrow (n+1)^{1/2} 2^n \|f\|_\infty$.
 9:     Using the sieve of Eratosthenes, find a prime $p < 2r \ln r$ such that $p \nmid b$ and $p \nmid \mathrm{disc}(f)$.
10:     $k \leftarrow \left\lceil \log_p(2^{n^2/2} B^{2n}) \right\rceil$.
11:     Using Cantor-Zassenhaus, find noncostant monic polynomials $h_1, \dots, h_r \in \mathbb{Z}[X]$ irreducible modulo $p$ such that $f \equiv b h_1 \cdots h_r \bmod p$ and $\|h_j\|_\infty < p/2$ for all $j$.
12:     Using Quadratic Hensel's Lemma repeatedly, find noncostant monic polynomials $g_1, \dots, g_r \in \mathbb{Z}[X]$ such that $f \equiv b g_1 \cdots g_r \bmod p^k$, $g_j \equiv h_j \bmod p$ and $\|g_j\|_\infty < p^k/2$ for all $j$.
13:     $T \leftarrow \{1, \dots, r\}$.        ▷ initialize the index set $T$ of modular factors still to be treated
14:     $I \leftarrow \emptyset$.        ▷ initialize the set of factors found
15:     $F \leftarrow f$.        ▷ initialize the polynomial still to be factored
16:     while $T \neq \emptyset$ do
17:         Choose $u$ among $\{g_t : t \in T\}$ of maximal degree.
18:         $d \leftarrow \deg u$.
19:         $n^* \leftarrow \deg F$.
20:         for $j = d + 1$ to $n^*$ do
21:             Using LLL algorithm with $\delta = 3/4$ compute a short vector $g$ in the lattice $L \subseteq \mathbb{Z}^j$ generated by the coefficient vectors of $\left\{uX^i : 0 \le i < j - d\right\} \cup \left\{p^k X^i : 0 \le i < d\right\}$, and denote the corresponding polynomial also by $g$.
22:             Determine by trial division the set $S \subseteq T$ of indices $i$ for which $h_i$ divides $g$ modulo $p$.
23:             Compute $h \in \mathbb{Z}[X]$ such that $h \equiv b \prod_{i \in T \setminus S} g_i \bmod p^k$ and $\|h\|_\infty < p^k/2$.
24:             if $\|g/c(g)\|_1 \|h/c(h)\|_1 \le B$ then
25:                 $T \leftarrow T \setminus S$.
26:                 $I \leftarrow I \cup \{g/c(g)\}$
27:                 $F \leftarrow h/c(h)$
28:                 $b \leftarrow \mathrm{lc}(F)$
29:                 break the loop 20 and goto Step 16.
30:             end if
31:         end for
32:         $T \leftarrow \emptyset$.
33:         $I \leftarrow I \cup \{F\}$.
34:     end while
35: end if
36: return $I$.
```

Theorem 15. *The LLL factoring algorithm is correct.*

Proof. It is enough to prove that the following are invariants:

- $F \equiv b \prod_{i \in T} g_i \bmod p^k$,
- $b = \mathrm{lc}(F)$,

- $f = \pm F \prod_{g \in I} g$,
- each polynomial in I is irreducible.

They hold initially, and we may assume they hold before Step 17. By Theorem 13, F is squarefree modulo p. Let $\tilde{g} \in \mathbb{Z}[X]$ be the irreducible factor of F divisible by u modulo p. By the following Lemma, and using invariants, we have that u divides $\tilde{g}$ modulo p^k.

Lemma 24. *Let R be an Euclidean domain and $p \in R$ a prime. Let $f, g \in R[X]$ such that $p \nmid lc(f)$, f modulo p is squarefree, and g divides f in $R[X]$. Let $u \in R[X]$ be monic and suppose that u divides f modulo p^k, for some $k \geq 1$, and g modulo p. Then u divides g modulo p^k.*

Proof. See [17]. □

First, an observation. Suppose that $v \in \mathbb{Z}[X]$ divides F in $\mathbb{Z}[X]$ and it is divisible by u modulo p. Since F is squarefree modulo p, then u does not divide F/v modulo p. Therefore, $\tilde{g}$ does not divide F/v in $\mathbb{Z}[X]$, and so it must divide v.

We claim that, if $j \leq \deg \tilde{g}$, then the condition in Step 24 is never fulfilled. Suppose that the condition is fulfilled. Then, by Lemma 19,

$$\frac{g}{c(g)} \frac{h}{c(h)} = \pm F,$$

and so g divides F. On the other hand g is divisible by u modulo p^k. Then, by the observation above, $\tilde{g}$ divides g in $\mathbb{Z}[X]$. But $\deg g < j \leq \deg \tilde{g}$, a contradiction.

Now let $j = 1 + \deg \tilde{g}$. Therefore, $d + 1 \leq j = \deg \tilde{g} + 1 \leq n^*$, and Step 20 is executed for such a j. We claim that for such a j the condition in Step 24 is fulfilled, and the invariants hold at the next pass through Step 16. By Lemma 24, the coefficient vector of $\tilde{g}$ is in L. By Corollary 3 and Mignotte's bound, we have

$$\|g\|_2 \leq 2^{(j-1)/2} \|\tilde{g}\|_2 \leq 2^{\deg \tilde{g}/2} (n+1)^{1/2} 2^{\deg \tilde{g}} \|f\|_\infty = 2^{3(\deg \tilde{g})/2 - n} B < 2^{n/2} B.$$

By the choice of k, we have that $\|g\|_2^{\deg \tilde{g}} \|\tilde{g}\|_2^{\deg g} < (2^{n/2}B)^n B^n \leq p^k$. Therefore, by Lemma 23, $\gcd(g, \tilde{g}) \in \mathbb{Z}[X]$ is nonconstant. Since $\tilde{g}$ is irreducible and $\deg g \leq j - 1 = \deg \tilde{g}$, we have $\tilde{g} = \pm g/c(g)$. Now let $\tilde{h} = F/\tilde{g}$ and $S \subseteq T$ be as in Step 22. As in the proof of Theorem 14, the uniqueness in Hensel's Lemma implies that $\mathrm{lc}(\tilde{g})\tilde{h} \equiv h \bmod p^k$ in Step 23. On the other hand

$$\left\|\mathrm{lc}(\tilde{g})\tilde{h}\right\|_\infty = |\mathrm{lc}(\tilde{g})| \left\|\tilde{h}\right\|_\infty \leq |\mathrm{lc}(F)| (n^* + 1)^{1/2} 2^{n^*} \|F\|_\infty \leq bB < \frac{p^k}{2},$$

from which $\mathrm{lc}(\tilde{g})\tilde{h} = h$, $\tilde{h} = h/c(h)$, and $F = \pm(g/c(g))(h/c(h))$. Therefore, the condition in Step 24 is fulfilled, and the execution of Steps $25 - 28$ implies that the invariants hold at the next pass through Step 16.

Finally, Steps 32 and 33 guarantee that the invariants hold at the end of the algorithm if $|T| = 1$ and $\tilde{g} = F$ is irreducible. □

Theorem 16. *The LLL factoring algorithm is polynomial time.*

Proof. This is clear since we have replaced the exhaustive search in Zassenhaus's algorithm by polynomial time calls to the LLL algorithm. □

Remark 20. Even though the LLL factoring algorithm was fast in theory, it was not the algorithm most often used in practice. Indeed, Zassenhaus' algorithm was preferred. On the other hand, Belabas, van Hoeij, Klüners and Steel [9] found an algorithm in 2004, built on previous ideas by van Hoeij [16], which is both good in theory and in practice. It is based on the LLL reduction, but it uses a different type of lattices compared to the ones in the LLL factoring algorithm. It was subsequently improved in 2008 by Novocin [15].

4. Factoring over number fields

Knowing how to factor in $\mathbb{Q}[X]$ allows us to factor over any number field. If needed, for a nice introduction to number fields see [12]. Let K be a number field of degree n and denote by $\mathcal{O}_K$ its ring of integers. Recall that K is the field of fractions of $\mathcal{O}_K$ and that $K = \mathbb{Q}(\theta)$, for some $\theta \in K$. Let σ_j, for $1 \leq j \leq n$, be the n embeddings of K into $\mathbb{C}$.

We would like to factor a polynomial $P \in K[X]$. As remarked earlier, we can reduce to the case P monic and squarefree. On the contrary, we cannot assume the polynomial to have coefficients in $\mathcal{O}_K$, because $\mathcal{O}_K$ is not necessarily a UFD. Let us extend the σ_j to $K[X]$ by acting on the coefficients and define the norm of $Q \in K[X]$ by

$$\mathcal{N}(Q) = \prod_{j=1}^{n} \sigma_j(Q).$$

It is clear that if Q is monic then $\mathcal{N}(Q)$ is also monic.

Lemma 25. $\mathcal{N}(Q) \in \mathbb{Q}[X]$.

Proof. The coefficients of $\mathcal{N}(Q)$ are symmetric polynomials in the coefficients of $\sigma_j(Q)$, for $1 \leq j \leq n$. Therefore, they are symmetric polynomials in the conjugates of θ (i.e. the roots of the minimal polynomial of θ over $\mathbb{Q}$) with rational coefficients. On the other hand, we know that every symmetric polynomial with rational coefficients can be written as a rational polynomial in the elementary symmetric polynomials (for a proof see [2]). The elementary symmetric polynomials in the conjugates of θ are nothing but the coefficients of the minimal polynomial of θ over $\mathbb{Q}$, hence they are rational numbers. □

Lemma 26. *If $Q \in K[X]$ is monic and irreducible, then $\mathcal{N}(Q)$ is the power of an irreducible monic polynomial in $\mathbb{Q}[X]$.*

Proof. Let

$$\mathcal{N}(Q) = \prod_{i=1}^{r} T_i^{e_i}$$

be the factorization of $\mathcal{N}(Q)$ into monic irreducible factors in $\mathbb{Q}[X]$. Since Q divides $\mathcal{N}(Q)$ in $K[X]$ and it is irreducible in $K[X]$, then it divides T_i in $K[X]$ for some i. But $T_i \in \mathbb{Q}[X]$ implies that $\sigma_j(Q)$ divides T_i for every j. Therefore, $\mathcal{N}(Q)$ divides T_i^n in $K[X]$. But both polynomials have rational coefficients. Hence $\mathcal{N}(Q)$ divides T_i^n in $\mathbb{Q}[X]$. On the other hand, T_i is irreducible and monic. Then $\mathcal{N}(Q) = T_i^m$, for some $m \leq n$. □

Lemma 27. *Let $P \in K[X]$ be a monic and squarefree polynomial of degree d. Then for all but finitely many $k \in \mathbb{Q}$ the polynomial $\mathcal{N}(P(X - k\theta))$ is squarefree.*

Proof. For $1 \leq i \leq n$, $1 \leq j \leq d$, denote by $\alpha_{i,j}$ the roots of $\sigma_i(P)$ in $\mathbb{C}$. Let $k \in \mathbb{Q}$. The roots of $\sigma_i(P(X - k\theta))$ are given by $\alpha_{i,j} + k\sigma_i(\theta)$. Therefore, $\mathcal{N}(P(X - k\theta))$ is not squarefree if and only if two such roots coincide, i.e. there exist i_1, i_2, j_1, j_2 such that $\alpha_{i_1,j_1} + k\sigma_{i_1}(\theta) = \alpha_{i_2,j_2} + k\sigma_{i_2}(\theta)$, or equivalently

$$k = \frac{\alpha_{i_1,j_1} - \alpha_{i_2,j_2}}{\sigma_{i_2}(\theta) - \sigma_{i_1}(\theta)}.$$

But there are only a finite number of such k. □

The following Theorem gives us the desired factorization of P in $K[X]$.

Theorem 17. *Let $P \in K[X]$ be monic, squarefree, and assume that $\mathcal{N}(P)$ is squarefree. Let $\mathcal{N}(P) = \prod_{i=1}^{r} T_i$ be the factorization of $\mathcal{N}(P)$ into monic irreducible factors in $\mathbb{Q}[X]$. Then*

$$P = \prod_{i=1}^{r} \gcd(P, T_i)$$

is the factorization of P into monic irreducible factors in $K[X]$.

Proof. Let $P = \prod_{i=1}^{s} P_i$ be the factorization of P into monic irreducible factors in $K[X]$. Since $\mathcal{N}(P)$ is squarefree, then $\mathcal{N}(P_i)$ is also squarefree. Therefore, by the proof of Lemma 26, $\mathcal{N}(P_i) = T_{j(i)}$, for some $1 \leq j(i) \leq r$. On the other hand, for $i \neq j$, $\mathcal{N}(P_i)\mathcal{N}(P_j) = \mathcal{N}(P_iP_j)$ divides $\mathcal{N}(P)$, which is squarefree. Then $\mathcal{N}(P_i)$ and $\mathcal{N}(P_j)$ are coprime, and so, by eventually reordering the indices, we have $\mathcal{N}(P_i) = T_i$ and $r = s$. Finally, since for $j \neq i$, P_j is coprime to T_i, and since P_i is monic, we have that $P_i = \gcd(P, T_i)$. □

It is easy to see that a polynomial over a field of characteristic 0 is squarefree if and only if $\gcd(f, f') = 1$. On the other hand, recall that, thanks to Corollary 7, we have a practical way to check if such a polynomial is squarefree. Finally, we can state our algorithm.

Algorithm 7 Factoring in $K[X]$

Input: a number field $K = \mathbb{Q}[X]$ and a monic squarefree polynomial $P \in K[X]$.
Output: the monic irreducible factors of P in $K[X]$.

```
 1: $k \leftarrow 0$.
 2: $F \leftarrow \emptyset$.                                              ▷ monic irreducible factors
 3: if $\mathrm{res}(\mathcal{N}(P(X - k\theta)), \mathcal{N}(P(X - k\theta))') = 0$ then    ▷ $\mathcal{N}(P(X - k\theta))$ is not squarefree
 4:     $k \leftarrow k + 1$.
 5:     goto Step 2.
 6: end if
 7: $T \leftarrow \mathcal{N}(P(X - k\theta))$.
 8: Using Zassenhaus's algorithm or LLL factoring algorithm, factorize $T$ into monic irreducible factors: $T = \prod_{i=1}^{r} T_i$.
 9: for $i = 1$ to $r$ do
10:     Calculate $P_i := \gcd(P, T_i(X + k\theta))$ in $K[X]$.
11:     $F \leftarrow F \cup \{P_i\}$.
12: end for
13: return $F$.
```

Theorem 18. *The algorithm above is correct.*

Proof. Clear from Theorem 17 and Lemma 27. □

References

1. J. Abbott, *Bounds on Factors in* $\mathbb{Z}[x]$, arXiv:0904.3057.
2. M. Artin, *Algebra*, Prentice Hall, 1991.
3. E. Bombieri and W. Gubler, *Heights in Diophantine geometry*, Cambridge University Press, 2007.
4. M. R. Bremner, *Lattice Basis Reduction: An Introduction to the LLL Algorithm and Its Applications*, CRC Press, 2011.
5. J. W. S. Cassels, *An Introduction to the Geometry of Numbers*, Springer-Verlag, Berlin, 1997.
6. A. M. Cohen, H. Cuypers, and H. Sterk (Eds.), *Some Tapas of Computer Algebra*, Springer-Verlag, Berlin, 1999.
7. H. Cohen, *A Course in Computational Algebraic Number Theory*, third corrected printing, Springer, 1996.
8. R. Crandall and C. Pomerance, *Prime Numbers: A Computational Perspective*, second edition, Springer, 2005.
9. K. Belabas, van Hoeij, Klüners and Steel, *Factoring polynomials over global fields*, Journal de théorie des nombres de Bordeaux **21** (2009), no. 1, 15-39.
10. G. Everest and T. Ward, *Heights of polynomials and entropy in algebraic dynamics*, Springer-Verlag, London, 1999.
11. N. Koblitz, *p-adic Numbers, p-adic Analysis, and Zeta-Functions*, Springer-Verlag, New York, 1984.
12. D. A. Marcus, *Number Fields*, Springer-Verlag, New York, 1977.
13. J. Neukirch, *Algebraic Number Theory*, Springer-Verlag, Berlin, 1999.
14. P. Q. Nguyen and B. E. Vallée (Eds.), *The LLL Algorithm. Survey and Applications*, Springer, Berlin, 2010.
15. A. Novocin, *Factoring Univariate Polynomials over the Rationals*, PhD thesis, http://andy.novocin.com/pro/dissertation.pdf.
16. M. van Hoeij, *Factoring polynomials and the knapsack problem*, J. Number Theory **95** (2002), 167-189.
17. J. von zur Gathen and J. Gerhard, *Modern Computer Algebra*, Cambridge University Press, 2003.