Lattices and LLL Algorithm

Seminar Report

Kushagra Shandilya
21Q050010
kushagras@cse.iitb.ac.in

Under the guidance of
Prof. Mrinal Kumar

Computer Science and Engineering
Indian Institute of Technology, Bombay
India
May 4, 2022

Abstract

Lattices and its related concepts have been highly influential in Mathematics and Computer Science. They have been used in areas such as Number Theory, factorizing algorithms, integer programming, sphere packing and many more.

This report surveys the basics of Lattices, its properties and bounds on the shortest lattice point. In addition, Gram-Schmidt procedure is discussed which orthonormalizes a set of vectors. The main work of this report focuses on one of the computational problems based on the shortest lattice point and the approximation algorithm which outputs the solution in polynomial time of input.

Acknowledgements

I thank Prof Mrinal for his guidance, advice and the valuable discussions we had regarding lattices and everything else. I thank Prof Sundar, Prof Rohit, Prof Abhiram for their teachings, for their advice whenever I felt stuck, and for helping me grow as a person of academia. Words are not enough to express my gratitude. I am very happy to be a part of the Theory lab at IIT Bombay.

Contents

1. Preliminaries

References: [Wik22a], [Wik22b], [Cam06]

1.1. Group-like structures.

Algebraic structures that contain a set of elements G and one binary operation $\circ$. Represented as $(G, \circ)$ or $(G, +)$.

Some properties

(G0) (Closure law) $\forall a, b \in G$, $a \circ b \in G$

(G1) (Associative law) $a \circ (b \circ c) = (a \circ b) \circ c \ \forall a, b, c \in G$

(G2) (Identity law) $\exists e \in G$ such that $a \circ e = e \circ a = a \ \forall a \in G$
where e is the identity element of a

(G3) (Inverse law) $\forall a \in G, \exists b \in G$ such that $a \circ b = b \circ a = e$
where e and b are the identity and inverse elements of a respectively

(G4) (Commutative law) $a \circ b = b \circ a \ \forall a, b \in G$

Some structures based on above laws

(1) $(G, \circ)$ is called a **Groupoid**.

(2) A Groupoid which follows Closure and Associative law is called a **Semigroup**.

(3) A Semigroup which follows Identity law is called a **Monoid**.

(4) A Monoid which follows Inverse law is called a **Group**.

(5) A Group which follows Commutative law is called an **Abelian Group** or a **Commutative group**..

1.2. Subgroup.

Given a group $(G, \circ)$. A subset H of G is called a subgroup of G if H also forms a group under the operation $\circ$.

This is represented as $H \leq G$.

1.3. Ring.

A **Ring** is a set of numbers R with two binary operations **addition**$(+)$ and **multiplication**$(.)$.

Properties on Addition

(A0) (Closure law) $\forall a, b \in R$, $a + b \in R$

(A1) (Associative law) $a + (b + c) = (a + b) + c \ \forall a, b, c \in R$

(A2) (Identity law) $\exists e \in R$ such that $a + e = e + a = a \ \forall a \in R$

(A3) (Inverse law) $\forall a \in R, \exists b \in R$ such that $a + b = b + a = e$

(A4) (Commutative law) $a + b = b + a \ \forall a, b \in R$

Properties on Multiplication

(M0) (Closure law) $\forall a, b \in R$, $a.b \in R$

(M1) (Associative law) $a.(b.c) = (a.b).c \ \forall a, b, c \in R$

Properties on Addition and Multiplication

(D) (Distributive law) $(a+b).c = a.c + b.c$ and $a.(b+c) = a.b + b.c \ \forall a, b, c \in R$

Alternate definition of Ring

A **Ring**$(R, +, .)$ is a

(1) set R
(2) two binary operations addition $(+)$ and multiplication $(.)$ defined on R such that following axioms are satisfied
(3) $(R, +)$ is an Abelian group
(4) $(R, .)$ is a semigroup
(5) (Distributive law) $(a+b).c = a.c + b.c$ and $a.(b+c) = a.b + b.c \ \forall a, b, c \in R$

Note: We can remove the dot as the multiplicative operator and can just append the two operands.

1.4. **Field.**

Some more properties on Multiplication

(M2) (Identity law) $\exists e \in R$ such that $ae = ea = a \ \forall a \in R$
(M3) (Inverse law) $\forall a \in R, \exists b \in R$ such that $ab = ba = e$
(M4) (Commutative law) $ab = ba \ \forall a, b \in R$

A **Field** is a Ring R which also satisfies M2, M3, M4. A Field is also called a **Commutative Division Ring**.

Alternate definition of Field

A **Field**$(F, +, .)$ is a

(1) set F
(2) two binary operations addition $(+)$ and multiplication $(.)$ defined on F such that following axioms are satisfied
(3) $(R, +)$ is an Abelian group
(4) $(R, .)$ is a Abelian group except for zero elements
(5) Multiplication distributes over Addition (Distributive law)

1.5. **Cauchy-Schwarz inequality.**

An inner product space is a real vector space or a complex vector space with an operation called an inner product. Inner product of two vectors a, b is a scalar product denoted by $\langle a, b \rangle$.

Cauchy-Schwarz inequality states that for all vectors a, b

$$|\langle a, b \rangle|^2 \leq \langle a, a \rangle \cdot \langle b, b \rangle \text{ or}$$
$$|\langle a, b \rangle| \leq ||a|| \cdot ||b||$$

where $||a||$ is the euclidean length/ norm of the vector a defined as $||a|| = \sqrt{\langle a, a \rangle}$.

2. Lattice

References: [Vai11], [Pei15], [Reg09], [Mic10]

2.1. Introduction.

Definition 2.1. *(Lattice) Given n linearly independent vectors $b_1, b_2, ...b_n \in \mathbb{R}^m$, the lattice generated by them is defined as*

$$\mathcal{L}(b_1, ..., b_n) = \left\{ \sum_{i=1}^{n} x_i b_i \;\mid\; x_i \in \mathbb{Z} \right\}$$

We will use $\mathcal{L}$ or Λ to denote a lattice. We can write the above definition in the matrix form as

$$\mathcal{L}(\mathbf{B}) = \{\mathbf{B}x \mid x \in \mathbb{Z}^n\}$$

where $\mathbf{B} = \begin{bmatrix} | & & | \\ b_1 & ... & b_n \\ | & & | \end{bmatrix}$ is called the basis matrix.

Properties of Lattices

- $b_1, ..., b_n$ are called the basis vectors of the lattice $\mathcal{L}$.
- $b_1, ..., b_n$ are linearly independent over $\mathbb{R}$.
- n is the rank of lattice
- m is the dimension of lattice
- In general, $n \leq m$. When $n = m$, lattice is called a full rank lattice. In this report, we will work with full rank lattices only.
- Lattice differentiates from Span as following

  $$\text{Span}(b_1, ..., b_n) = \left\{ \sum_{i=1}^{n} x_i b_i \;\mid\; x_i \in \mathbb{R} \right\}$$

  This further signifies the fact that lattice is a discrete set.

  $$\text{Span}(\mathbf{B}) \supset \mathcal{L}(\mathbf{B})$$

- Lattices are closed under addition and subtraction operations.

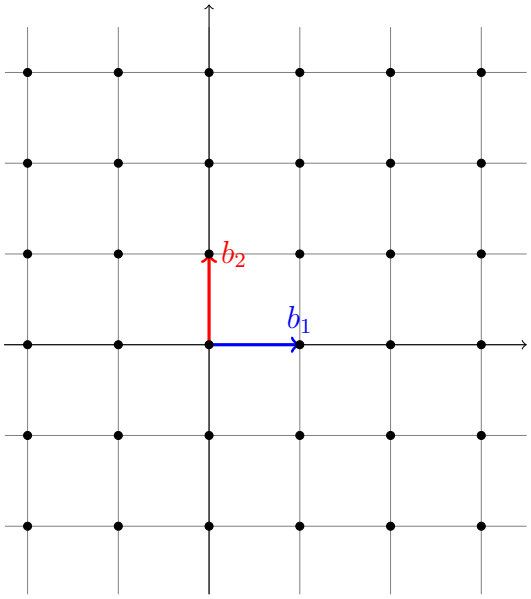
See the figures on the next page for some examples of lattices.

Lattice can also be defined as the following

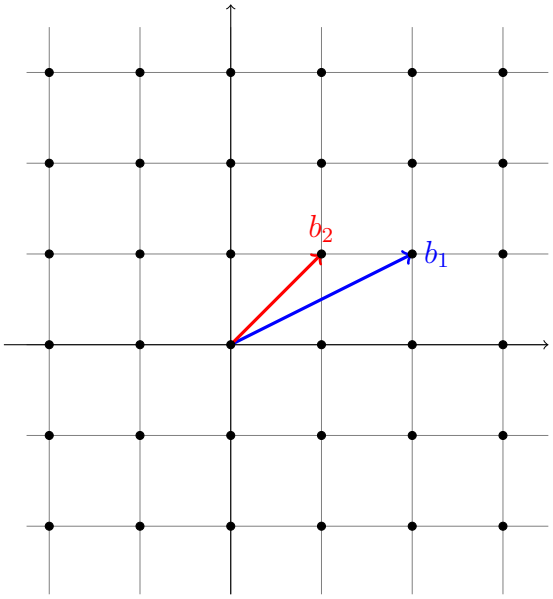
Definition 2.2. *(Lattice) A lattice $\mathcal{L}$ is a discrete additive subgroup of $\mathbb{R}^n$.*

In the above definition, by discrete we mean for any $x \in \mathcal{L}$ there is a space in which x is the only lattice point. Formally, for some $\epsilon > 0$, $||x - y|| \geq \epsilon$ where $x, y \in \mathcal{L}$ and $x \neq y$.

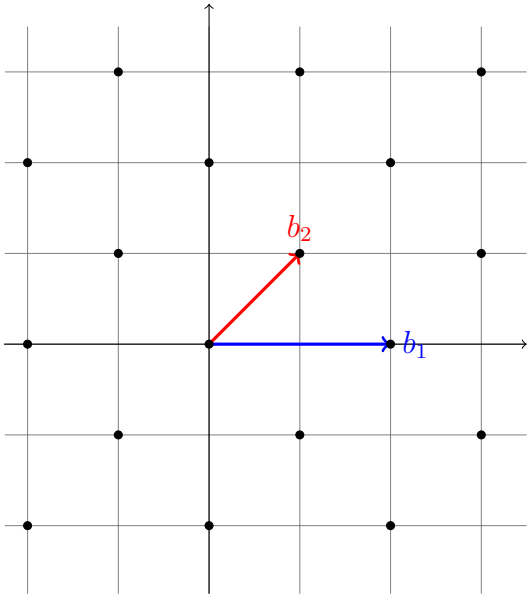
And by additive subgroup we mean $\mathcal{L}$ is a subgroup of $\mathbb{R}^n$ which is closed under addition, i.e. for any $x, y \in \mathcal{L}, x + y \in \mathcal{L}$.



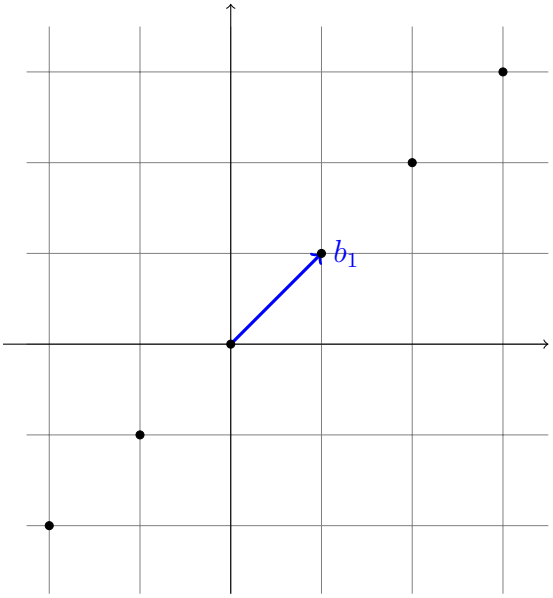
(A) The lattice $\mathbb{Z}^2$ with basis vectors $(0, 1)$ and $(1, 0)$.



(B) The lattice $\mathbb{Z}^2$ with a different basis consisting of vectors $(2, 1)$ and $(1, 1)$.



(C) A full-rank lattice generated by the basis vectors $(2, 0)$ and $(1, 1)$. This is a sub-lattice of $\mathbb{Z}^2$.



(D) A *non full-rank* lattice with basis vector $(1, 1)$

FIGURE 1. Various lattices and their bases [Vai11]

2.2. **Bases and Lattices.**

From the examples on the previous page, it is evident that different bases can generate the same lattice. In general, a lattice can have infinitely many bases.

- Some examples of basis of $\mathbb{Z}^2$

$$\mathbf{B}_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \text{ and } \mathbf{B}_2 = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \text{ and } \mathbf{B}_3 = \begin{pmatrix} 7 & 6 \\ 8 & 7 \end{pmatrix}$$

- While $\mathbf{B}_4$ produces a sublattice of $\mathbb{Z}^2$.

$$\mathbf{B}_4 = \begin{pmatrix} 2 & 1 \\ 0 & 1 \end{pmatrix}$$

But how can we verify whether two bases produce the same lattice or not?

2.3. **An Algebraic Characterization using Unimodular Matrices.**

Definition 2.3. *(Unimodular matrix) A matrix $\boldsymbol{U} \in \mathbb{Z}^{n\times n}$ is Unimodular if $|det(\boldsymbol{U})| = 1$ where $|.|$ represents the absolute value.*

Lemma 2.4. *If $\boldsymbol{U}$ is Unimodular, so is $\boldsymbol{U}^{-1}$.*

Proof. $\boldsymbol{U}^{-1} = adj(\boldsymbol{U})/det(\boldsymbol{U})$. Since every element of $\boldsymbol{U}$ is an integer, then the determinant of each minor matrix will also be an integer. Furthermore, $|det(\boldsymbol{U})| = 1$. Thus, each element of $\boldsymbol{U}^{-1}$ will also be an integer. Also,
$det(\boldsymbol{U}^{-1}) = 1/det(\boldsymbol{U}) = 1$ □

Theorem 2.5. *Given two full rank bases $B, B' \in \mathbb{R}^{n\times n}$, $\mathcal{L}(B) = \mathcal{L}(B')$ if and only if there exists an Unimodular matrix U such that $B' = BU$.*

Proof. ($\Rightarrow$) Suppose $\mathcal{L}(B) = \mathcal{L}(B')$. Then for each of the b_i columns of B', $b_i \in \mathcal{L}(B)$. This implies there exists an integer matrix $U \in \mathbb{Z}^{n\times n}$ such that $B' = BU$. Similarly, there exists an integer matrix $V \in \mathbb{Z}^{n\times n}$ such that $B = B'V$.

$$B = B'V = BUV$$
$$UV = I_n$$
$$det(U)det(V) = 1$$

Since U and V are integer matrices their determinants will also be of integer value. Therefore, $|det(U)| = |det(V)| = 1$.

($\Leftarrow$) Suppose $\exists$ an Unimodular matrix U such that $B' = BU$. Therefore, each column of B' is contained in $\mathcal{L}(B)$, we get $\mathcal{L}(B') \subseteq \mathcal{L}(B)$.

Similarly, $B = B'U^{-1}$ for some Unimodular matrix U^{-1}(2.4). By following the above argument, we get $\mathcal{L}(B) \subseteq \mathcal{L}(B')$.

Therefore, $\mathcal{L}(B) = \mathcal{L}(B')$. □

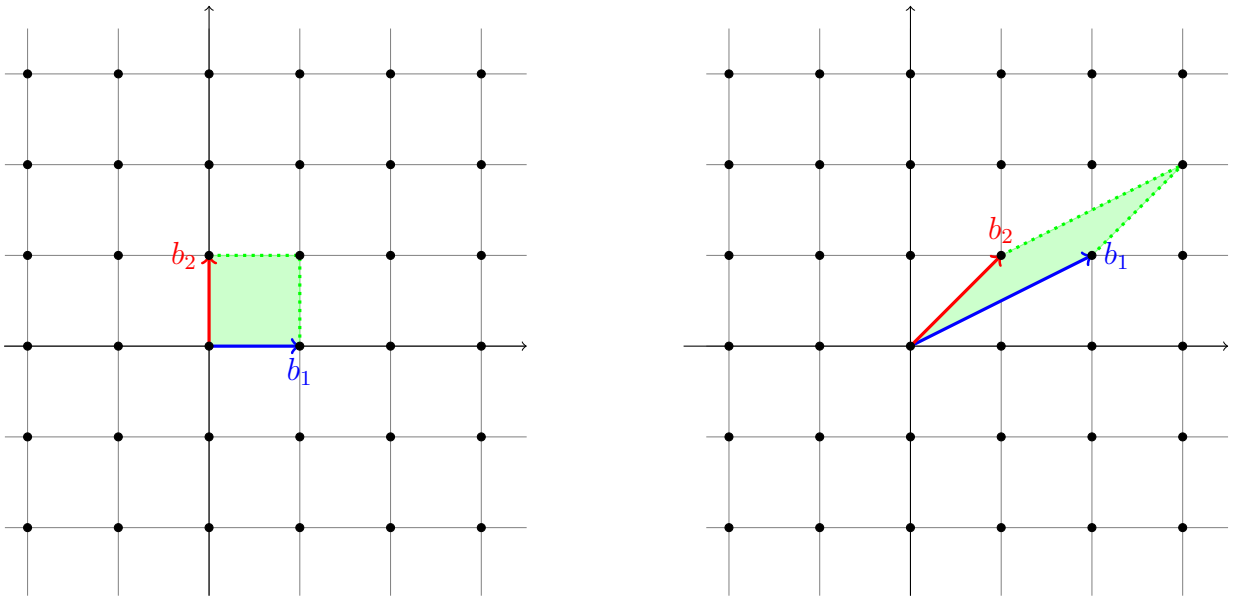
2.4. **A Geometric Characterization using Fundamental Parallelepiped.**

Definition 2.6. *(Fundamental Parallelepiped) Given n linearly independent vectors* $b_1, b_2, ...b_n \in \mathbb{R}^m$*, their fundamental parallelepiped is defined as*

$$\mathcal{P}(b_1, ..., b_n) = \left\{ \sum_{i=1}^{n} x_i b_i \;\mid\; 0 \leq x_i < 1 \right\}$$

Properties of Fundamental Parallelepiped

- Fundamental Parallelepiped is the region enclosed by the basis vectors.
- Different bases of the same lattice generate different fundamental parallelepipeds.



(A) The lattice $\mathbb{Z}^2$ with basis vectors $(0, 1)$ and $(1, 0)$ and the associated fundamental parallelepiped.

(B) The lattice $\mathbb{Z}^2$ with a different basis consisting of vectors $(2, 1)$ and $(1, 1)$, and the associated fundamental parallelepiped.

FIGURE 2. Parallelepipeds for various bases of the lattice $\mathbb{Z}^2$. Notice that the parallelepipeds in either case do not contain any non-zero lattice point [Vai11].

Theorem 2.7. *Let Λ be a full rank n-dimensional lattice, and let $b_1, ..., b_n \in \Lambda$ be n linearly independent vectors. Then, $b_1, ..., b_n$ forms a basis of Λ if and only if $\mathcal{P}(b_1, ..., b_n) \cap \Lambda = \{0\}$.*

Proof. ($\Rightarrow$) Suppose $b_1, ...b_n$ forms a basis of Λ. Let $a = \mathcal{P}(b_1, ..., b_n) \cap \Lambda$. Then,

$$\Lambda = \left\{ \sum_{i=1}^{n} x_i b_i \ \mid \ x_i \in \mathbb{Z} \right\} \text{ and}$$

$$\mathcal{P}(b_1, ..., b_n) = \left\{ \sum_{i=1}^{n} x_i b_i \ \mid \ 0 \leq x_i < 1 \right\}$$

This implies, in the intersection of Λ and $\mathcal{P}(b_1, ..., b_n)$, all $x_i = 0$. Thus, $a = \{0\}$.

($\Leftarrow$) Suppose $\mathcal{P}(b_1, ..., b_n) \cap \Lambda = \{0\}$. We want to show $(b_1, ..., b_n)$ is a basis of Λ.

Λ is a lattice of rank n and $b_1, ..., b_n$ are linearly independent vectors of Λ. Thus, $\mathcal{L}(b_1, ..., b_n) \subseteq \Lambda$.

Λ is a lattice of rank n. For any $a \in \Lambda$

$$a = \sum_{i=1}^{n} x_i b_i \text{ where } x_i \in \mathbb{R}$$

Consider the vectors

$$a' = \sum_{i=1}^{n} \lfloor x_i \rfloor b_i \text{ and}$$

$$a'' = a - a' = \sum_{i=1}^{n} x_i - \lfloor x_i \rfloor b_i$$

a' is the integral combination of the vectors $(b_1, ..., b_n)$, thus $a' \in \mathcal{L}(b_1, ..., b_n)$. From our previous statement, $\mathcal{L}(b_1, ..., b_n) \subseteq \Lambda$. We can then imply $a' \in \Lambda$. We know that lattices are closed under addition and subtraction. Therefore, $a'' \in \Lambda$.

Since, $0 \leq x_i - \lfloor x_i \rfloor < 1$ for all x_i. Thus, $a'' \in \mathcal{P}(b_1, ..., b_n)$ which further implies $a'' \in \mathcal{P}(b_1, ..., b_n) \cap \Lambda$.

By our assumption, $a'' = \{0\}$. Vectors $b_1, ..., b_n$ are linearly independent. This implies $x_i - \lfloor x_i \rfloor = 0$ for all i. Therefore, $\Lambda \subseteq \mathcal{L}(b_1, ..., b_n)$.

Together, we can say $\Lambda = \mathcal{L}(b_1, ..., b_n)$.

Our proof concludes. □

Lemma 2.8. *Two bases are equivalent if and only if one can be obtained from the other by the following operations on columns*

- $b_i \leftarrow b_i + k b_j$ *for some* $k \in \mathbb{Z}$
- $b_i \leftrightarrow b_j$
- $b_i \leftarrow -b_i$

Rather than a proof, we present the intuition behind the lemma.

Each of the above operation can be converted to an integer matrix by performing some column/ row operations on the identity matrix. These matrices are called elementary matrices. Let some elementary matrix be E.

Suppose we want to perform swapping of some basis vectors. This can be done by calculating BE where E is an elementary matrix corresponding to the swapping of some columns.

Since E is formed from the identity matrix by performing column/ row operations. Then, $|det(E)| = |det(I)|$. We can say E is an Unimodular matrix.

Therefore performing the above operations on basis vectors is equivalent to right multiplying the basis matrix by a series of Unimodular matrices.

2.5. **Determinant of a Lattice.**

Definition 2.9. *(Determinant of a Lattice) Let $\mathcal{L}(B)$ be a lattice. Then the determinant of $\mathcal{L}(B)$ is defined as the volume of the fundamental parallelepiped formed by its basis vectors.*

Determinant of a lattice, denoted $det(\mathcal{L}(B))$, is computed as the absolute value of the determinant of its basis matrix.

Properties of the Determinant of a lattice

- Fundamental parallelepipeds corresponding to different basis of the same lattice produces the same volume. 2.5 Thus, determinant is a lattice invariant.
- Intuitively, if the volume of some fundamental parallelepiped is large, it means the lattice points are far away from the origin and from each other. Thus, greater the determinant, the sparser the lattice.

3. **Gram-Schmidt Orthogonalization**

References: [Vai11], [Pei15], [Reg09], [Mic10]

Definition 3.1. *(Gram-Schmidt Orthogonalization) For a sequence of n linearly independent vectors $b_1, ..., b_n \in \mathbb{R}^m$, Gram-Schmidt Orthogonalization is the procedure to convert $b_1, ..., b_n$ into a sequence of orthogonal vectors $\tilde{b_1}, ..., \tilde{b_n}$. It is computed as the following*

$$\tilde{b_i} = b_i - \sum_{j=1}^{i-1} \mu_{i,j} \tilde{b_j} \qquad where\ \mu_{i,j} = \frac{\langle b_i, \tilde{b_j} \rangle}{\langle \tilde{b_j}, \tilde{b_j} \rangle}$$

A Gram-Schmidt vector $\tilde{b_i}$ is the component of b_i which is orthogonal to the span of $(\tilde{b_1}, \tilde{b_2}, ..., \tilde{b_{i-1}})$. Thus, the first Gram-Schmidt vector $\tilde{b_1}$ is b_1 itself.

In the formula, $\mu_{i,j}$ are called the Gram-Schmidt coefficients where μ is called the projection coefficient.

$$Proj_u(v) = \text{Projection of v on u} = \frac{\langle u, v\rangle}{\langle u, u\rangle} u$$

where $\langle u, v\rangle$ represents the inner/ dot product of u and v.

If we were to expand the above written formula, it will be as follows

$$\tilde{b_1} = b_1 \qquad e_1 = \frac{\tilde{b_1}}{||\tilde{b_1}||}$$

$$\tilde{b_2} = b_2 - \mu_{2,1}\tilde{b_1} \qquad e_2 = \frac{\tilde{b_2}}{||\tilde{b_2}||}$$

$$\tilde{b_3} = b_3 - \mu_{3,1}\tilde{b_1} - \mu_{3,2}\tilde{b_2} \qquad e_3 = \frac{\tilde{b_3}}{||\tilde{b_3}||}$$

and so on. Here, $(\tilde{b_1}, ..., \tilde{b_n})$ are the set of orthogonal vectors and $(e_1, ..., e_n)$ are the set of unit orthogonal vectors.

Let us see this procedure through an example. Let b_1 and b_2 be our set of basis vectors as the following

$$b_1 = \begin{bmatrix} 2 \\ 3 \end{bmatrix} \quad b_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

$$\tilde{b_1} = b_1 = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$$

$$\tilde{b_2} = b_2 - \mu_{2,1}\tilde{b_1}$$

$$\mu_{2,1} = \frac{\langle \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ 3 \end{bmatrix} \rangle}{\langle \begin{bmatrix} 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 2 \\ 3 \end{bmatrix} \rangle} = \frac{8}{13}$$

$$\tilde{b_2} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} - \frac{8}{13}\begin{bmatrix} 2 \\ 3 \end{bmatrix} = \begin{bmatrix} -3/13 \\ 2/13 \end{bmatrix}$$

It is easily verifiable that $\langle \tilde{b_1}, \tilde{b_2} \rangle$ is 0. Therefore, $(\tilde{b_1}, \tilde{b_2})$ are our required Gram-Schmidt vectors.

Properties of Gram-Schmidt Orthogonalization

- Span of $(\tilde{b_1}, ..., \tilde{b_i})$ is the same as of $(b_1, ..., b_i) \quad 1 \leq i \leq n$
- Gram-Schmidt vectors do not necessarily form a lattice basis. They might not even be part of the lattice.

- Gram-Schmidt vectors depend on the order in which basis vectors are sequenced.

We can represent the procedure of Gram-Schmidt orthogonalization in the matrix form as $B = \tilde{B}.U$ where U is an Upper Unitriangular matrix

$$\begin{pmatrix} | & & | \\ b_1 & \dots & b_n \\ | & & | \end{pmatrix} = \begin{pmatrix} | & & | \\ \tilde{b_1} & \dots & \tilde{b_n} \\ | & & | \end{pmatrix} . \begin{pmatrix} 1 & \mu_{2,1} & \mu_{3,1} & \dots & \mu_{n,1} \\ 0 & 1 & \mu_{3,2} & \dots & \mu_{n,2} \\ 0 & 0 & 1 & \dots & \mu_{n,3} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} | & & | \\ \frac{\tilde{b_1}}{||\tilde{b_1}||} & \dots & \frac{\tilde{b_n}}{||\tilde{b_n}||} \\ | & & | \end{pmatrix} . \begin{pmatrix} ||\tilde{b_1}|| & \mu_{2,1}||\tilde{b_1}|| & \mu_{3,1}||\tilde{b_1}|| & \dots & \mu_{n,1}||\tilde{b_1}|| \\ 0 & ||\tilde{b_2}|| & \mu_{3,2}||\tilde{b_2}|| & \dots & \mu_{n,2}||\tilde{b_2}|| \\ 0 & 0 & ||\tilde{b_3}|| & \dots & \mu_{n,3}||\tilde{b_3}|| \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & ||\tilde{b_n}|| \end{pmatrix}$$

Determinant of matrix with column vectors $\frac{\tilde{b_1}}{||\tilde{b_1}||}$ is 1 because they are orthogonal unit vectors.

$$\text{Thus, } det(\mathcal{L}(B)) = \prod_{i=1}^{n} ||\tilde{b_i}||$$

4. **Successive Minima and Bounds**

References: [Vai11], [Pei15], [Reg09], [Mic10]

One basic parameter of the lattice is the length of the shortest non-zero vector in the lattice. This is called the first successive minima of the lattice $\mathcal{L}$. It is denoted by $\lambda_1(\mathcal{L})$. Here, by length we mean the euclidean norm of the vector. We can define other successive minimas as well.

Definition 4.1. *(Successive minima) Let $\mathcal{L}$ be a lattice and integer $k \leq rank(\mathcal{L})$. Then $\lambda_k(\mathcal{L})$ is the smallest $r > 0$ such that $\mathcal{L}$ contains at least k linearly independent vectors of length bounded by r .*

Theorem 4.2. *Let B be a rank-n lattice basis and $\tilde{B}$ be its Gram-Schmidt orthogonalization. Then,*

$$\lambda_1(\mathcal{L}(B)) \geq \min_{i=1,..,n} ||\tilde{b_i}||$$

Proof. Let $x \in \mathbb{Z}^n$ be any non-zero integer vector. We want to show any vector $Bx \in \mathcal{L}(B)$ has length at least $min_i||\tilde{b_i}||$. We calculate $|\langle Bx, \tilde{b_j}\rangle|$ in two different ways.

1. Let $j \in 1, ..., n$ be the largest index such that $x_j \neq 0$. Then,

$$|\langle Bx, \tilde{b_j}\rangle| = \left|\langle \sum_{i=1}^{n} x_i b_i, \tilde{b_j}\rangle\right| = \left|\sum_{i=1}^{n} \langle x_i b_i, \tilde{b_j}\rangle\right| = |x_j \langle b_j, \tilde{b_j}\rangle| = |x_j \langle \tilde{b_j}, \tilde{b_j}\rangle| = |x_j| \cdot ||\tilde{b_j}||^2$$

where the first equality follows by rewriting Bx as the integral combination of basis vectors and the second equality due to the linearity of inner product.

The third equality follows because for $j < i, \langle b_i, \tilde{b_j}\rangle = 0$ and for $j > i, x_j = 0$ by the definition of j.

The fourth equality follows again from the linearity of inner product because b_j has two components, one along $\tilde{b_j}$ and the other perpendicular to $\tilde{b_j}$.

2. By the Cauchy-Schwarz, inequality

$$|\langle Bx, \tilde{b_j}\rangle| \leq ||Bx|| \cdot ||\tilde{b_j}||$$

Through the above two equations, we get

$$||Bx|| \geq |x_j| \cdot ||\tilde{b_j}|| \geq ||\tilde{b_j}|| \geq \min_{i=1,..,n} ||\tilde{b_i}||$$

where the second inequality follows because x is a non-zero integer vector.

The above statement states that the length of any lattice vector is at least $min_i||\tilde{b_i}||$, therefore we can conclude

$$\lambda_1(\mathcal{L}) \geq \min_{i=1,..,n} ||\tilde{b_i}||$$

□

Corollary 4.3. *4.2 leads to the definition of 2.2.*

Let $\mathcal{L}$ be a lattice. Then there exists some $\epsilon > 0$ such that $||x - y|| \geq \epsilon$ for any two non-equal lattice points $x, y \in \mathcal{L}$.

Proof. From 4.2, $||x - y|| \geq \lambda_1(\mathcal{L})$. Thus, we can set $\epsilon = \lambda_1(\mathcal{L})$ to obtain the statement of corollary. □

After computing the lower bound on the length of the shortest vector, we move to its upper bound. We present Minkowski's theorems without their proofs.

Theorem 4.4. *(Minkowski's First Theorem) For any full rank lattice $\mathcal{L}$ of rank n,*

$$\lambda_1(\mathcal{L}) \leq \sqrt{n} * det(\mathcal{L})^{1/n}$$

Theorem 4.5. *(Minkowski's Second Theorem) For any full rank lattice $\mathcal{L}$ of rank n,*

$$\left(\prod_{i=1}^{n} \lambda_i(\mathcal{L})\right)^{1/n} \leq \sqrt{n} * det(\mathcal{L})^{1/n}$$

5. Lattice Computational Problems

References: [Reg09], [Mic10]

From the previous theorems, we have a bound on the length of the shortest vector in a lattice but we do not have any algorithm to find such a lattice point.

We define some computational problems to address these issues.

Definition 5.1. *(Search SVP) Given a lattice basis $B \in \mathbb{Z}^{m\times n}$, find a lattice vector v such that $||v|| = \lambda_1(\mathcal{L}(B))$.*

Definition 5.2. *(Optimization SVP) Given a lattice basis $B \in \mathbb{Z}^{m\times n}$, compute $\lambda_1(\mathcal{L}(B))$.*

Definition 5.3. *(Decisional SVP) Given a lattice basis $B \in \mathbb{Z}^{m\times n}$ and a number $r \in \mathbb{Q}$, determine whether $\lambda_1(\mathcal{L}(B)) \leq r$ or not.*

Note that we have restricted ourselves to integer basis matrix to efficiently represent those numbers in bit form.

These computational problems are hard to solve. Therefore, we relax the search SVP problem to define a γ-approximation problem.

Definition 5.4. *(Search SVP_γ) Given a lattice basis $B \in \mathbb{Z}^{m\times n}$, find a non-zero lattice vector v such that $||v|| \leq \gamma \cdot \lambda_1(\mathcal{L}(B))$.*

We will now look at the LLL algorithm to tackle the Search SVP_γ computational problem.

6. LLL Algorithm

References: [Reg09], [Pei15], [Mic10], [LLL82]

The LLL algorithm gives a polynomial time solution to the *Search SVP_γ* problem. The best approximation factor we can get using LLL is $\gamma = (2/\sqrt{3})^n$ where n is the dimension of the lattice. Though this bound is exponential, it is non-trivial because it only depends on the dimension of the lattice. Furthermore, in many applications of LLL, the number of dimensions remains constant.

The goal of the LLL algorithm is to output a set of "nearly" orthogonalized short vectors. The LLL algorithm converts an input set of basis vectors into a set of "reduced" vectors which generates the same lattice as the original basis.

We will first define the δ-LLL-reduced basis, then present an algorithm to find such a reduced basis and finally analyse its running time complexity.

6.1. δ-LLL-reduced basis.

Definition 6.1. *(δ-LLL-reduced basis) A basis $B \in \mathbb{R}^{n\times n}$ is δ-LLL reduced if it satisfies the following properties*

- $|\mu_{i,j}| \leq \frac{1}{2} \quad \forall\, 1 \leq i \leq n, j < i$
- $\delta||\tilde{b_i}||^2 \leq ||\mu_{i+1,i} \cdot \tilde{b_i} + \tilde{b}_{i+1}||^2 \quad \forall\, 1 \leq i < n$

where $\tilde{b_i}$ are the orthogonal vectors we get after performing the Gram-Schmidt orthogonalization.

Hereon, we will refer to these properties as the first and second property of the δ-LLL-reduced basis.

The first property corresponds to the size reduced basis and the second property is the Lovász condition. Let us look at the second property in some detail. It states

$$\delta||\tilde{b_i}||^2 \leq ||\mu_{i+1,i} \cdot \tilde{b_i} + \tilde{b}_{i+1}||^2 = \mu_{i+1,i}^2||\tilde{b_i}||^2 + ||\tilde{b}_{i+1}||^2$$

Here, the second equality follows because $\tilde{b_i}$ and $\tilde{b}_{i+1}$ are orthogonal vectors. Rearranging the above equation, we get

$$||\tilde{b}_{i+1}||^2 \geq \big(\delta - \mu_{i+1,i}^2\big)||\tilde{b_i}||^2$$

If our basis vectors follows the first property, i.e. the reduction property, then $\mu_{i+1,i} \leq 1/2$. We get

$$||\tilde{b}_{i+1}||^2 \geq \left(\delta - \frac{1}{4}\right)||\tilde{b_i}||^2$$

This informally states that $\tilde{b}_{i+1}$ is not much shorter than $\tilde{b_i}$.

LLL algorithm works for $\frac{1}{4} < \delta < 1$. It will be helpful to keep the case of $\delta = \frac{3}{4}$ in mind as we continue.

After applying the first property, the matrix U in Gram-Schmidt orthogonalization will look as follows

$$\begin{pmatrix} ||\tilde{b_1}|| & \leq \frac{1}{2}||\tilde{b_1}|| & \leq \frac{1}{2}||\tilde{b_1}|| & \dots & \leq \frac{1}{2}||\tilde{b_1}|| \\ 0 & ||\tilde{b_2}|| & \leq \frac{1}{2}||\tilde{b_2}|| & \dots & \leq \frac{1}{2}||\tilde{b_2}|| \\ 0 & 0 & ||\tilde{b_3}|| & \dots & \leq \frac{1}{2}||\tilde{b_3}|| \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & ||\tilde{b_n}|| \end{pmatrix}$$

Lemma 6.2. *Let $b_1, ..., b_n \in \mathbb{R}^n$ be a reduced basis. Then*

$$||b_1|| \leq \left(\delta - \frac{1}{4}\right)^{-(n-1)/2} \cdot \lambda_1(\mathcal{L})$$

Proof. Rearranging the second property, we got

$$||\tilde{b_n}||^2 \geq \left(\delta - \frac{1}{4}\right)||\tilde{b}_{n-1}||^2 \geq \cdots \geq \left(\delta - \frac{1}{4}\right)^{n-1}||\tilde{b_1}||^2 \geq \left(\delta - \frac{1}{4}\right)^{n-1}||b_1||^2$$

where the last equality follows because according to the definition of Gram-Schmidt orthogonalization $\tilde{b_1} = b_1$. Thus, after rearranging the above equation we get

$$||b_1|| \leq \left(\delta - \frac{1}{4}\right)^{-(i-1)/2} ||\tilde{b_i}|| \leq \left(\delta - \frac{1}{4}\right)^{-(n-1)/2} ||\tilde{b_i}||$$

where the last inequality follows because $\frac{1}{4} < \delta < 1$.

The above inequality should be satisfiable for any $\tilde{b_i}$ and we know for any $i, ||\tilde{b_i}|| \geq \min_j ||\tilde{b_j}||$, we get

$$||b_1|| \leq \left(\delta - \frac{1}{4}\right)^{-(n-1)/2} \min_j ||\tilde{b_j}||$$

From 4.2 $\lambda_1(\mathcal{L}) \geq \min_j ||\tilde{b_j}||$. Plugging this into the above we get

$$||b_1|| \leq \left(\delta - \frac{1}{4}\right)^{-(n-1)/2} \cdot \lambda_1(\mathcal{L}).$$

Thus, $||b_1|| \leq \gamma \cdot \lambda_1(\mathcal{L})$ where $\gamma = \left(\delta - \frac{1}{4}\right)^{-(n-1)/2}$ □

The above theorem gives us an approximate solution to the *Search SVP*$_\gamma$ problem. Assuming we can compute a δ-LLL-reduced basis, we can just output b_1 as the solution to the *Search SVP*$_\gamma$ problem. For $\delta = 3/4$ we get a $\gamma = 2^{(n-1)/2}$ approximation factor.

6.2. **LLL algorithm.**

Given integral basis vectors $b_1, ..., b_n$ as input, do the following:

1 Compute $\tilde{b_1}, ..., \tilde{b_n}$ using Gram-Schmidt Orthogonalization

2 Reduction step:

for $i = 2$ *to* n **do**

for $j = i - 1$ *to* 1 **do**

$b_i \leftarrow b_i - c_{i,j} \cdot b_j$ where $c_{i,j} = \left\lceil \frac{\langle b_i, \tilde{b_j} \rangle}{\langle \tilde{b_j}, \tilde{b_j} \rangle} \right\rfloor$

end

end

3 Swap step:

if *for any* $i, \delta||\tilde{b_i}||^2 > ||\mu_{i+1,i} \cdot \tilde{b_i} + \tilde{b}_{i+1}||^2$ **then**

swap b_i with b_{i+1}

goto step 1

end

4 Output $b_1, ..., b_n$

where $\lceil . \rfloor$ is the rounding operator which rounds to the nearest integer.

We make some observations on the above procedure. First, we perform column operations of type $b_i \leftarrow b_i + ab_j, a \in \mathbb{Z}$. These operations do not change the lattice structure 2.8.

The second property is enforced by the swap step. If the algorithm ever terminates, the output must satisfy the second property.

Since, we perform column operations of type $b_i \leftarrow b_i + ab_j$, for $i > j, a \in \mathbb{Z}$, these operations do not change the Gram-Schmidt vectors. To see this, remember the matrix form of Gram-Schmidt orthogonalization which is $B = \tilde{B}U$, where U is an upper-triangular matrix. We mentioned in the lemma 2.8 that these column operations are elementary matrices formed from the identity matrix. Since we restrict the operation $b_i \leftarrow b_i + ab_j$ for $i > j$, our elementary matrix is an uppertriangular matrix with one non-zero off-diagonal entry. Let this matrix be E. Thus,

$$B = \tilde{B}U$$
$$BE = \tilde{B}UE$$
$$B' = \tilde{B}U'$$

where the last equality follows because multiplication of two upper-triangular matrices results in an upper-triangular matrix. This implies that the original basis vectors and the Gram-Schmidt coefficients (μ) change during the reduction step but the $\tilde{b_i}$ vectors remain the same.

In the i^{th} iteration of the outer loop, the reduction step makes sure that the projection of b_i on $\tilde{b_j}$ for any $j < i$ is at most $\frac{1}{2}||\tilde{b_j}||$. This is ensured by subtracting the appropriate integer multiple of column b_j from column b_i such that the absolute value of the j^{th} coordinate, in the column b_i, becomes at most $\frac{1}{2}||\tilde{b_j}||$.

We now formally look at the correctness of the LLL algorithm.

Theorem 6.3. *(Correctness) If the LLL procedure ever terminates, then the output vectors produce the same lattice as the original basis vectors. Furthermore, the output vectors follows the properties of δ-LLL-reduced basis.*

Proof. The output of the algorithm generates the same lattice and the swap step enforces the second property of the δ-LLL-reduced basis both of which we have already explained in the above observations.

We just need to show now that after the reduction step vectors $b_1, ..., b_n$ satisfy $\mu_{i,j} \leq 1/2$ for $i > j$.

We have already discussed that the Gram-Schmidt vectors do not change during the reduction step. So, consider the outer loop running in some i^{th} iteration and the inner loop is in some j^{th} iteration and clearly $j < i$. Then $\mu_{i,j}$ can be written

as,

$$\begin{aligned}
|\mu_{i,j}| &= \left| \frac{\langle b_i - c_{i,j} \cdot b_j, \tilde{b_j} \rangle}{\langle \tilde{b_j}, \tilde{b_j} \rangle} \right| \\
&= \left| \frac{\langle b_i, \tilde{b_j} \rangle}{\langle \tilde{b_j}, \tilde{b_j} \rangle} - \left\lceil \frac{\langle b_i, \tilde{b_j} \rangle}{\langle \tilde{b_j}, \tilde{b_j} \rangle} \right\rfloor \frac{\langle b_j, \tilde{b_j} \rangle}{\langle \tilde{b_j}, \tilde{b_j} \rangle} \right| \\
&= \left| \frac{\langle b_i, \tilde{b_j} \rangle}{\langle \tilde{b_j}, \tilde{b_j} \rangle} - \left\lceil \frac{\langle b_i, \tilde{b_j} \rangle}{\langle \tilde{b_j}, \tilde{b_j} \rangle} \right\rfloor \right| \\
&\leq \frac{1}{2}
\end{aligned}$$

where the first equality follows from the reduction step, the second equality by plugging the value of $c_{i,j}$ and linearity of inner product. The third equality follows because $\langle b_j, \tilde{b_j} \rangle = \langle \tilde{b_j}, \tilde{b_j} \rangle$ and finally the first inequality comes from the definition of rounding operator. □

6.3. **Runtime Analysis.**

The analysis of LLL consists of two steps. We want to bound the number of iterations and then bound the time an iteration can take. We want to show that the running time of the algorithm is polynomial in the input size. We were given n vectors in the input and a vector of norm r requires at least $\log r$ bits to represent. So we define $M = \{n, \log(\max_i ||b_i||\}$ and associate a "potential value" for a basis as follows.

Definition 6.4. *(Potential of a lattice basis) Given a lattice basis $B = (b_1, ..., b_n)$, the potential of B is defined as*

$$\begin{aligned}
\mathcal{D}_b &= \prod_{i=1}^{n} ||\tilde{b_i}||^{n-i+1} \\
&= \prod_{i=1}^{n} ||\tilde{b_1}||\ ||\tilde{b_2}|| \cdots ||\tilde{b_i}|| \\
&= \prod_{i=1}^{n} \mathcal{D}_{b,i}
\end{aligned}$$

where $\mathcal{D}_{b,i} = det(\Lambda_i)$ and Λ_i is the lattice generated by vectors $b_1, ..., b_i$

Our objective is to show that the initial value of $\mathcal{D}_b$ is not too large and with each iteration it decays quick enough. We know that $||\tilde{b_i}|| \leq ||b_i||$, thus we can upper bound the value of $\mathcal{D}_b$ by $(\max_i ||b_i||^{(n(n+1)/2)})$. Thus, the logarithm of $\mathcal{D}_b$ is polynomial in M.

Theorem 6.5. *The number of iterations is polynomial in M.*

Proof. We explained earlier that during the reduction step the Gram-Schmidt vectors do not change, thus, $\mathcal{D}_b$ also remains the same during the reduction step. But, in the swap step we change the sequence of vectors which does affect the Gram-Schmidt vectors. We will show that after the swapping, $\mathcal{D}_b$ decreases and by using this fact we will bound the number of iterations. Suppose during some iteration b_i is swapped with b_{i+1}.

This swapping does not affect the lattice Λ_k for $k < i$ because b_i and b_{i+1} are not used as basis vectors for Λ_k.

This swapping also does not affect the lattice Λ_k for $k > i$. This follows because the initial sequence of basis vectors was $(b_1, ..., b_i, b_{i+1}, ..., b_k)$ and the new sequence is $(b_1, ..., b_{i+1}, b_i, ..., b_k)$. We know from 2.8 that these operations do not change the lattice.

We know that the determinant is a lattice invariant. Thus, we can infer for $k \neq i, \Lambda_k$ and $\mathcal{D}_{B,k}$ remains the same. The only change occurs is in $\mathcal{D}_{b,i}$. Let Λ'_i and $\mathcal{D}'_{b,i}$ be the new values of Λ_i and $\mathcal{D}_{b,i}$ respectively. Then,

$$\begin{aligned}\frac{\mathcal{D}'_{b,i}}{\mathcal{D}_{b,i}} &= \frac{det\Lambda'_i}{det\Lambda_i} \\ &= \frac{det\Lambda(b_1, ..., b_{i-1}, b_{i+1})}{det\Lambda(b_1, ..., b_{i-1}, b_i)} \\ &= \frac{\left(\prod_{j=1}^{i-1} ||\tilde{b_j}||\right)||\tilde{b'}_{i+1}||}{\prod_{j=1}^{i} ||\tilde{b_j}||}\end{aligned}$$

Before swapping, the old value of Gram-Schmidt vector corresponding to b_{i+1} is $\tilde{b}_{i+1} = b_{i+1} - \sum_{j=1}^{i} \mu_{i+1,j}\tilde{b_j}$ and after swapping new value is $\tilde{b}'_{i+1} = b_{i+1} - \sum_{j=1}^{i-1} \mu_{i+1,j}\tilde{b_j} = \tilde{b}_{i+1} + \mu_{i+1,i} \cdot \tilde{b_i}$. After plugging this value in the above equation, we get

$$\begin{aligned}\frac{\mathcal{D}'_{b,i}}{\mathcal{D}_{b,i}} &= \frac{||\tilde{b}_{i+1} + \mu_{i+1,i} \cdot \tilde{b_i}||}{||\tilde{b_i}||} \\ &< \frac{\sqrt{\delta} \cdot ||\tilde{b_i}||}{||\tilde{b_i}||} \\ &= \sqrt{\delta}\end{aligned}$$

where the first inequality follows from $||\tilde{b}_{i+1} + \mu_{i+1,i} \cdot \tilde{b_i}||^2 < \delta \cdot ||\tilde{b_i}||^2$ which holds true because swapping occurred due to the violation of second property.

Through the above statement we can say that $\mathcal{D}_b$ decreases by $\sqrt{\delta}$ multiplicative factor.

Let $\mathcal{D}_{B,0}$ be the initial value of $\mathcal{D}_b$. Since $\mathcal{D}_b$ is a positive integer, therefore, we can bound the number of iterations to

$$\log_{\frac{1}{\sqrt{\delta}}} \mathcal{D}_{B,0} = \frac{\log \mathcal{D}_{B,0}}{\log \frac{1}{\sqrt{\delta}}} \leq \frac{1}{\log \frac{1}{\sqrt{\delta}}} \cdot \frac{n(n+1)}{2} \cdot \log(\max_i ||b_i||)$$

Since $1/4 < \delta < 1$, the above quantity is polynomial in M. □

We have bound the number of iterations to polynomial in input size. Now, we need to show that each iteration also takes polynomial time in input. In the algorithm, we perform only basic operations, i.e. addition, multiplication, for a polynomial number of times. Thus, we need to show that the bit size of the numbers which occur during the algorithm is polynomial in M. In addition, we also need to show that Gram-Schmidt orthogonalization can be computed in polynomial time. We present the statement of the theorem without a proof at the moment.

Theorem 6.6. *The running time of each iteration is polynomial in M.*

Remark 6.7. *The best approximation factor LLL can offer is* $\left(\frac{2}{\sqrt{3}}\right)^n$ *using* $\delta = \frac{1}{4} + \frac{3}{4}^{n-1}$.

7. Conclusions and Future work

We introduced a mathematical object named Lattice. We studied some basic properties of Lattices such as Equivalent bases, Fundamental Parallelepiped, Determinant etc. We looked at the Gram-Schmidt Orthogonalization which orthonormalizes a set of arbitrary vectors. We defined another interesting property of lattices which were successive minima and gave bounds on it. We then defined some computational problems on lattices and studied the LLL algorithm which solves one of the problems in polynomial time.

Lattices and its related concepts have interesting applications. Lattices and specifically Minkowski's theorems are used to solve some flavours of diophantine equations. The LLL algorithm has an application in polynomial factorization. Lattices have provided the fastest algorithm to solve Integer Programming in fixed variables. Furthermore, they are also being used heavily in modern cryptography.

My interests align more towards the Number theoretic side of lattices and the algorithms such as polynomial factorization, integer programming etc. My future goal would be to read up more on the above and to understand the bit time complexity of the LLL algorithm which I did not prove earlier. I also did not prove Minkowski's theorems because they use some advanced geometry concepts which I was not familiar with. I would like to add this to my future goal list too.

References

[Cam06] Peter J. Cameron. *Notes on Algebraic Structures*. `https://webspace.maths.qmul.ac.uk/p.j.cameron/notes/algstr.pdf`. 2006.

[LLL82] Arjen K Lenstra, Hendrik Willem Lenstra, and László Lovász. "Factoring polynomials with rational coefficients". In: *Mathematische annalen* 261.ARTICLE (1982), pp. 515–534.

[Mic10] Daniele Micciancio. *CSE206A: Lattices Algorithms and Applications*. https://cseweb.ucsd.edu//classes/wi10/cse206a/. 2010.

[Pei15] Chris Peikert. *EECS 598: Lattices in Cryptography*. https://web.eecs.umich.edu/~cpeikert/lic15/index.html. 2015.

[Reg09] Oded Regev. *Lattices in Computer Science*. https://cims.nyu.edu/~regev/teaching/lattices_fall_2009/. 2009.

[Vai11] Vinod Vaikuntanathan. *CSC2414: Topics in Applied Discrete Mathematics: Lattices in Computer Science*. https://people.csail.mit.edu/vinodv/COURSES/CSC2414-F11/index.html. 2011.

[Wik22a] Wikipedia contributors. *Algebraic structure — Wikipedia, The Free Encyclopedia*. [Online; accessed 6-April-2022]. 2022. URL: https://en.wikipedia.org/w/index.php?title=Algebraic_structure&oldid=1081186117.

[Wik22b] Wikipedia contributors. *Cauchy–Schwarz inequality — Wikipedia, The Free Encyclopedia*. [Online; accessed 6-April-2022]. 2022. URL: https://en.wikipedia.org/w/index.php?title=Cauchy%E2%80%93Schwarz_inequality&oldid=1072458704.