2021

On properties of positive semigroups in lattices and totally real number fields

Siki Wang

Claremont McKenna College

On Properties of Positive Semigroups in Lattices and Totally Real Number Fields

Submitted to
Lenny Fukshansky

by
Siki Wang

for
B.A. Senior Thesis
03 May 2021

Abstract

In this thesis, we give estimates on the successive minima of positive semigroups in lattices and ideals in totally real number fields. In Chapter 1 we give a brief overview of the thesis, while Chapters 2 – 4 provide expository material on some fundamental theorems about lattices, number fields and height functions, hence setting the necessary background for the original results presented in Chapter 5. The results in Chapter 5 can be summarized as follows. For a full-rank lattice $L \subset \mathbb{R}^d$, we are concerned with the semigroup $L^+ \subseteq L$, which denotes the set of all vectors with nonnegative coordinates in L. Taking a basis $X \subseteq L^+$ for L and generating its $\mathbb{Z}_{\geq 0}$-span, we obtain a conical sub-semigroup $S(X)$ in L^+. We call the points in $L^+ \setminus S(X)$ the gaps of $S(X)$. We proceed to describe basic properties of these gaps, but the focus of this thesis is on the restrictive successive minima of L^+ and $L^+ \setminus S(X)$, for which we produce bounds in the spirit of Minkowski's successive minima theorem and its recent generalizations. Further, we apply these results to obtain analogous bounds for sub-semigroups of ideals in totally real number fields, whose image under the Minkowski embedding corresponds to L^+ for an appropriate lattice L. These bounds are obtained with respect to the Weil height of elements in the number field.

Acknowledgements

I would like to thank Professor Lenny Fukshansky for all his guidance and support throughout the year, as well as for the time he put into reading and editing this thesis.

Contents

1 Introduction

The Geometry of Numbers, initiated by Hermann Minkowski (1910), studies the interplay between compact convex 0-symmetric sets and lattices in a Euclidean space $\mathbb{R}^n$. Such lattices can be very general, but in the context of number theory, any ring of algebraic integers O_K in a number field K of degree n can be viewed as a lattice in $\mathbb{R}^n$ under the Minkowski embedding. This embedding maps an element of K to a vector of its n algebraic conjugates. It turns out that such a lattice construction can provide fundamental information about the ring itself. For example, the finiteness of the class number, a measure of the quotient group of fractional ideals that implies the existence of non-unique factorizations in O_K, was first proved by Minkowski with the use of his Geometry of Numbers. This result demonstrated the power of Minkowski's tool.

In this thesis, we are concerned with number fields that are totally real. In the spirit of Minkowski's successive minima theorems, we aim to obtain bounds for sets of successive minima that are restricted to subsemigroups of L, namely L^+ and $L \setminus S(X)$, where X represents a basis for L contained in L^+. The former set consists of lattice points with only non-negative entries, and the latter is generated by the difference between the non-negative $\mathbb{Z}$-span of X with L^+ itself, which we denote by $G(X)$. We will start off by describing some properties of the semigroup L^+ and the primitive gaps in $G(X)$, and prove the existence of a point in L^+ whose sup norm is bounded in between 1 and $2\mu(L) + 1$, where $\mu(L)$ stands for the inhomogeneous minimum of the standard unit sphere with respect to L. The estimates on the successive minima of L^+ and $G(X)$ are obtained with the help of that point, as well as the generic successive minima of L.

Next, we translate these results to a totally real number field K, because the image of an ideal I contained in its ring of integers O_K under the Minkowski's embedding will necessarily give us a lattice in $\mathbb{R}^d$. Analogous to L^+ and $S(X)$, we will look at I^+ and its the subsemigroup $S(\boldsymbol{\beta})$ generated by non-negative $\mathbb{Z}$-combinations of $\boldsymbol{\beta}$, a positive basis for I contained in I^+. Moreover, to have a comparable measure of the sizes of elements in K, we will look at the Weil height of $\mathbb{Q}$-linearly independent algebraic numbers in both I^+ and $I^+ \setminus S(\boldsymbol{\beta})$. Estimates on the product of their Weil heights will be obtained in relation to the discriminant of the number field Δ_K and the norm of the ideal $\mathbb{N}_K(I)$.

Chapters 2 – 4 are expository in nature and will survey some fundamental concepts and theorems related to lattices, number fields and height functions. These chapters are dedicated to equipping the reader with a working knowledge of the algebraic setup, in order to understand the original results presented in Chapter 5, which is based on [3]. Note that all the proofs of the theorems presented in Chapters 2 – 4 are omitted for simplicity: these are standard results that can be found in [1], [6], [4], and [2].

2 Lattices

A *lattice* is a discrete set of points in $\mathbb{R}^n$. It is generated by a collection of vectors, closed under vector addition and scalar multiplication by elements of $\mathbb{Z}$ (which is also induced by addition and subtraction). A lattice is fundamentally different from a vector space, since (real) vector spaces are closed under scalar multiplication by elements of $\mathbb{R}$, and hence are not discrete. In comparison, the vector addition operation in $\mathbb{R}^n$ gives the lattice a discreteness property: the lattice points are "spaced out" enough, so that around each point, we can find an open neighborhood that contains no other point of the lattice.

Having this property is nice, because it allows us to have the notion of distance between lattice points, to study how far and close they are from the origin, to compute areas, volumes, and to look at ways that geometric objects can be inscribed into a lattice and study their properties. To begin this journey, we first give a formal definition of a lattice.

Definition 2.1.

1. A *lattice* Λ is the set of all integer linear combinations of a fixed collection of linearly independent vectors $\boldsymbol{a}_1, ..., \boldsymbol{a}_r \in \mathbb{R}^n$. In other words, for all $1 \leq i \leq r$,

$$\Lambda = \text{span}_{\mathbb{Z}}\{\boldsymbol{a}_1, ..., \boldsymbol{a}_r\} := \Big\{ \sum_{i=1}^{r} z_i \boldsymbol{a}_i \mid z_i \in \mathbb{Z} \Big\}$$

2. Then, the lattice Λ has *rank r*, and the vectors $\boldsymbol{a}_1, ..., \boldsymbol{a}_r$ form a *basis* for Λ.
3. A lattice in $\mathbb{R}^n$ with rank n is of *full rank.*

The most natural way of organizing the set of basis vectors is to put them into a matrix. So for every lattice of full rank in $\mathbb{R}^n$, there will be a corresponding matrix M such that every column is a basis vector. As one may guess, such a matrix is not unique. We present an interesting fact about basis matrices for a lattice.

Lemma 2.1.

Let Λ be a lattice of full rank in $\mathbb{R}^n$, and let A be a basis matrix for Λ. Then B is another basis matrix for Λ if and only if there exists an $n \times n$ integral matrix U with determinant ± 1 such that $B = AU$.

This fact is saying that there exists some change-of-basis matrix $U \in \text{GL}_n(\mathbb{Z})$ that can relate one basis matrix of Λ to another. Because $|\text{det}(U)| = 1$, we can define an invariant of the lattice as the absolute value of the determinant of its basis matrix. We call such value the *determinant* of the lattice Λ and we denote it by det(Λ).

Now we know what a lattice in $\mathbb{R}^n$ looks like, something that we can explore is the interplay of a lattice with a set also contained $\mathbb{R}^n$. We can ask: does this set contain any points of the lattice? Is there any way to determine this only by looking at, say, the volume of this set?

To make these questions meaningful, you might have noticed that these sets can't just be any kind of sets. A rectangular strip that happens to be lying in the "gap" of the lattice will never touch any lattice points, no matter how much we stretch it to increase its volume. For our purposes, we want these sets to be compact, convex, and $\boldsymbol{0}$-symmetric in $\mathbb{R}^n$. Informally, it means that the set is closed, bounded, has a positive volume, and is symmetric with respect to the origin. A common example in $\mathbb{R}^n$ is the closed unit ball $B_n(r)$ of radius r centered at the origin.

Now that we got better sets, we define some useful values of a lattice $\Lambda \in \mathbb{R}^n$ in relation to an arbitrary set *M*, which is compact, convex, and $\boldsymbol{0}$-symmetric, as described above.

Definition 2.2.

1. The *first successive minimum* of *M* with respect to Λ, denoted λ_1, is the smallest λ such that $\lambda_1 M$ contains a nonzero point of the lattice Λ. In other words,

$$\lambda_1 := \inf\{\lambda \in \mathbb{R}_{>0} \mid \lambda M \cap \Lambda \setminus \{\boldsymbol{0}\} \neq \emptyset\}.$$

2. For $1 \leq i \leq n$, the *i-th successive minimum* of *M* is defined to be the smallest λ such that $\lambda_i M$ contains at least i linearly independent points of Λ. In other words,

$$\lambda_i := \inf\{\lambda \in \mathbb{R}_{>0} \mid \dim(\operatorname{span}_{\mathbb{R}}\{\lambda M \cap \Lambda \setminus \{\boldsymbol{0}\} \geq i\})\}.$$

Since a set *M* in $\mathbb{R}^n$ is nothing but a collection of n-dimensional vectors, we can multiply each of their entries by a real number λ to get a *homogeneously expanded* set λM. The successive minima is a measure of how much the set needs to be enlarged or shrunk so that it contains the desired number of lattice points. Some celebrated theorems of Minkowski gave explicit bounds on the values of the λ_i's. For example, the Minkowski Convex Body Theorem (page 18 of [2]) stated that

$$0 \leq \lambda_1 \leq 2\left(\frac{\det(\Lambda)}{\mathrm{Vol}(M)}\right)^{1/n}. \tag{1}$$

Minkowski's Successive Minima Theorem (page 18 of [2]) also gave bounds on the product of all the λ_i's,

$$\frac{2^n \det(\Lambda)}{n!\mathrm{Vol}(M)} \leq \lambda_1 \ldots \lambda_n \leq \frac{2^n \det(\Lambda)}{\mathrm{Vol}(M)}. \tag{2}$$

In Chapter 5 of this paper, we will give an estimate on the successive minima restricted to the lattice points lying in the positive orthant of $\mathbb{R}^n$, namely, the bounds on λ_i's such that $\lambda_i B_n(1)$ contains i-linearly independent lattice points with coordinates that are all non-negative. The bounds will be given in relation to the *inhomogeneous minimum* of a lattice Λ. First, we define what it is.

Definition 2.3.

The *inhomogeneous minimum* of the set M with respect to the lattice Λ is defined as

$$\mu := \inf \{\gamma \in \mathbb{R}_{>0} \mid \gamma M + \Lambda = \mathbb{R}^n\}.$$

In other words, it is the smallest positive real number γ such that translates of γM by all points in the lattice Λ covers the entire $\mathbb{R}^n$. Another name for μ is the *covering radius*. Historically, many upper and lower bounds on μ are given in relation to the successive minima, such as this one given by Jarnik (see, for instance [4], Section 13.1, Theorem 1):

$$\mu \leq \frac{1}{2} \sum_{i=1}^{n} \lambda_i. \tag{3}$$

In the next chapter, we will give some background on number fields, and introduce a way to embed ideals of the ring of integers of a number field into Euclidean spaces so that they could be viewed as lattices there, as well as some invariants of a number field such as its discriminant. All of these machinery will be useful for interpreting the results on the restricted successive minima of the positive semigroups of the lattices in $\mathbb{R}^n$, and for comparing the results on the heights of the corresponding algebraic numbers back in the number field.

3 Number Fields

Recall that a number field *K* is a subfield of $\mathbb{C}$ which is a finite algebraic extension over $\mathbb{Q}$. By the primitive element theorem, *K* can always be written as $K = \mathbb{Q}(\beta)$ for some $\beta \in K$. Since β is algebraic, let us denote the degree of *K* over $\mathbb{Q}$ to be d. Then, we will necessarily have a set of d linearly independent elements $\{1, \beta_1, \beta_2, ..., \beta_{d-1}\}$ that spans *K* as a finite-dimensional vector space over $\mathbb{Q}$.

To motivate the study of lattices in the setting of number fields, we narrow our attention to a particularly nice subring of $\mathbb{C}$: the algebraic integers, denoted by $\mathbb{I}$. An algebraic integer $\alpha \in \mathbb{I}$ is a number that is a root of a monic polynomial with coefficients in $\mathbb{Z}$. For example, $\frac{1}{3}$ is not an algebraic integer, because its minimal polynomial over $\mathbb{Z}$ is not monic: $3x - 1$. It follows that the algebraic integers of $\mathbb{Q}$ is just the ring of integers $\mathbb{Z}$, as one may expect.

Given an arbitrary number field *K*, we define its ring of integers O_K to be all the numbers in *K* which are algebraic integers, i.e., $O_K = K \cap \mathbb{I}$. In a way, they are the integer-equivalents of the "fractions" in *K*, just like the case with $\mathbb{Z}$ in $\mathbb{Q}$. Now, we state without proof an important fact about O_K.

Lemma 3.1.

> The ring O_K is a free $\mathbb{Z}$-module of rank $d = [K : \mathbb{Q}]$, i.e. there exist elements $\beta_1, \ldots, \beta_d \in O_K$ such that
>
> $$O_K = \{c_1\beta_1 + \cdots + c_d\beta_d : c_1, \ldots, c_d \in \mathbb{Z}\}.$$
>
> Further, K is the fraction field of O_K, and for any $\beta \in K$ there exists some $c \in \mathbb{Z}$ such that $c\beta \in O_K$.

Moreover, an ideal *I* in the ring O_K forms a *sublattice* of the same rank, because its index in O_K is necessarily finite, i.e., $[O_K : I] < \infty$. This description of lattices is different from what we gave in Section 2, because those lattices are viewed in an ambient Euclidean space equipped with a norm, which we don't yet have for O_K. Luckily, there is a natural way to view the lattice of algebraic integers as a lattice in the Euclidean space through the *Minkowski embedding*, which is a tool that helps us transform an algebraic integer into a vector in $\mathbb{R}^d$.

Before we formally introduce the Minkowski embedding, we define what an *embedding* is in the general sense, and how an embedding of a number field into $\mathbb{C}$ looks like. To do this, we will need the notion of an algebraic conjugate. Recall that for an algebraic number α, its algebraic conjugates, $\alpha_1, ..., \alpha_d$ are the roots of its minimal polynomial $m_\alpha(x)$, which are all distinct. Now, we define what an embedding is.

Definition 3.1.

1. Let *E*, *F* be fields. An *embedding* of *E* into *F* is an injective ring homomorphism $\sigma : E \to F$.
2. For $1 \leq i \leq d$, an *embedding* from *K* into $\mathbb{C}$ is a map $\sigma_i : K \to \mathbb{C}$ given by

$$\sigma_i \left(\sum_{m=0}^{d-1} c_m \alpha^m \right) = \sum_{m=0}^{d-1} c_m \cdot \sigma_i(\alpha)^m = \sum_{m=0}^{d-1} c_m \alpha_i^m.$$

This definition tell us that for every algebraic conjugate α_i ($1 \leq i \leq d$) of a number, there is an embedding σ_i defined by it. For an element in $\mathbb{Q}$, the map σ_i does nothing, because the algebraic conjugates of a rational is just itself. Hence, the map σ_i fixes the subfield $\mathbb{Q}$ of *K* and maps $\alpha \in K$ to its *i*-th algebraic conjugate, α_i. It can be further verified that σ_i is an injective homomorphism of *K* into $\mathbb{C}$, and that $K \cong \sigma_n(K)$. In fact, $\sigma_1, ..., \sigma_d$ are the only possible embeddings of *K* into $\mathbb{C}$.

Next, let us classify some basic properties of the σ_i maps. Since the algebraic conjugates of a number can be either real or complex, for the σ_i maps, we will make an effort to distinguish those two: σ_i is said to be real if the the field $\sigma_i(K)$ is contained in $\mathbb{R}$, and it is complex otherwise, which we denote by $\overline{\sigma}_i$. We will order the embeddings by

$$\sigma_1, ..., \sigma_r, \sigma_{r+1}, \overline{\sigma}_{r+1}, ..., \sigma_{r+s}, \overline{\sigma}_{r+s},$$

where $\sigma_1, ..., \sigma_r$ are real and $\sigma_{r+1}, \overline{\sigma}_{r+1}, ..., \sigma_{r+s}, \overline{\sigma}_{r+s}$ are complex, and hence always come in conjugate pairs. We are now ready to give a definition of the Minkowski embedding.

Definition 3.2.

The Minkowski embedding $\Sigma : K \to \mathbb{R}^r \times \mathbb{C}^s$ is defined as

$$\Sigma(\alpha) := (\sigma_1(\alpha), ..., \sigma_r(\alpha), \sigma_{r+1}(\alpha), ..., \sigma_{r+s}(\alpha)),$$

for some $\alpha \in K$.

Notice that each complex embedding σ_{r+i} which maps into $\mathbb{C}$ can also be viewed as an embedding into $\mathbb{R}^2$ by splitting $\sigma_{r+i}(\alpha)$ into a real and an imaginary part. More can be said about these splittings, but in this paper, we will be only concerned with algebraic number fields that are totally real, i.e., for every $\alpha \in K$, its conjugate $\sigma_i(\alpha)$ is a real number, for all $1 \leq i \leq d$. An example will be the pair $\sqrt{2}$ and $-\sqrt{2}$, which are roots of the polynomial $x^2 - 2 = 0$. A totally real field *K* allows us to view any of its ring of integers O_K as lattices in the d-dimensional real Eulidean space, because under the Minkowski embedding, any algebraic number $\alpha \in K$ will become a vector with coordinates that are all real numbers. This allows us to borrow geometric tools on lattices to study properties about this number fields under the Minkowski embedding.

Finally, we use the embeddings to define as useful invariant on K. A $\mathbb{Z}$-basis for O_K is also a basis for K as a $\mathbb{Q}$-vector space: such a basis is called an *integral basis* for K.

Definition 3.3.

Let $\{\alpha_1, ..., \alpha_d\}$ be an integral basis for K, the *discriminant* of K is defined to be

$$\Delta_K := (\det(\sigma_n(\alpha_k))_{1\leq n,k\leq d})^2$$

Since O_K is a lattice, multiplication by a change of basis matrix $A \in \mathrm{GL}_d(\mathbb{Z})$ will preserve the determinant of the lattice, and hence preserving the discriminant of K.

Using the discriminant, we will cite without proof a property related to the index of an ideal $I \subseteq O_K$. We call this index the *norm* of the ideal $\mathbb{N}_K(I) = |O_K/I|$. It is a fact that $\mathbb{N}_K(I)$ is always finite, and that $\mathbb{N}_K(I) = |\Delta(\beta_1, ..., \beta_d)/\Delta_K|^{1/2}$, where $\beta_1, ..., \beta_d$ is an integral basis for I.

In Chapter 5, we will use the Minkowski embedding, the norm of an ideal in O_K and the discriminant of K to give bounds on linearly independent elements in a number field. This will heavily rely on the notion of the height of an algebraic number, which we will introduce in the next section.

4 Absolute Values and Height Functions

A height function is a tool that quantifies the complexity of mathematical objects. There are height functions for polynomials, for algebraic varieties, etc. Here, we are only concerned with a specific kind of height function, the Weil height $h(\alpha)$ of an algebraic number $\alpha \in K$.

To understand the components of $h(\alpha)$, we first describe what height functions look like in general. To properly define one, we have to patiently look back to the definition of an *absolute value* of an algebraic number, a concept necessarily for defining its height.

Definition 4.1.

Let K be a field. An *absolute value* on K is a function $|\cdot| : K \to \mathbb{R}_{\geq 0}$ such that for all $x, y \in K$, we have:

1. $|x| \geq 0$, with equality if and only if $x = 0$.
2. $|xy| = |x||y|$.
3. *Triangle inequality*: $|x + y| \leq |x| + |y|$.
4. * *Ultrametric inequality*: $|x + y| \leq \max\{|x|, |y|\}$.

The ultrametric inequality only applies to some absolute values. So, if an absolute value $|\cdot|$ satisfies (1), (2), (3) but fails (4), we call it an *archimedean* absolute value. An example of an archimedean absolute value would be the L_2 norm defined on a Eulidean vector space, which we generally refer to as the distance function. If an absolute value $|\cdot|$ satisfies (1), (2), (3), and also (4), we call it a *non-archimedean* absolute value. An example would be the discrete metric defined on a metric space.

Next, we define equivalence classes of absolute values and introduce the *places* of K.

Definition 4.2.

1. Two absolute values $|\cdot|_1$ and $|\cdot|_2$ on K are *equivalent*, denoted $|\cdot|_1 \sim |\cdot|_2$, if there exists $r \in \mathbb{R}_{>0}$ such that $|x|_1 = |x|_2^r$, for all $x \in K$.
2. Equivalence classes of nontrivial absolute values are called *places* of K.
3. The set of all places of K is denoted by $M(K)$.

Based on these definitions, an archimedean absolute value cannot be equivalent to a non-arhimedean one. Also, it can be verified that this relation $\sim$ is an actual equivalence relation by checking reflexivity, symmetry and transitivity.

We now introduce two types of absolute values that are standard on $\mathbb{Q}$. Later, we will extend them to absolute values on *K*.

First, the usual absolute value we are familiar with, which is denoted by $|\cdot|_\infty$, where $|x|_\infty = x$ if $x \geq 0$ and $|x|_\infty = -x$ if $x < 0$, is an archimedean absolute value on $\mathbb{Q}$. The completion of $\mathbb{Q}$ with respect to $|\cdot|_\infty$ is $\mathbb{R}$.

Second, for any prime number p, the *p-adic absolute* value of a rational number $\frac{a}{b}$ is $\frac{|a|_p}{|b|_p}$. More specifically, for an integer z, its p-adic absolute value is defined by $|z|_p = \frac{1}{p^{\mathrm{ord}_p(z)}}$, where $\mathrm{ord}_p(z)$ denotes the highest power of p that factors into z. For example, $\mathrm{ord}_2(12) = 2$ and and $\mathrm{ord}_3(12) = 1$, since $12 = 2^2 \cdot 3$. Since every integer has only finitely many prime factors, it follows that $|z|_p = \frac{1}{p^{\mathrm{ord}_p(z)}} = \frac{1}{p^0} = 1$ for all but finitely many primes $p \nmid z$, and $|z|_p \leq 1$ for all $z \in \mathbb{Z}$. It can be verified that $|\cdot|_p$ is a non-archimedean absolute value, and $|\cdot|_p$ and $|\cdot|_q$ are not equivalent for distinct primes p and q. Moreover, the completion of $\mathbb{Q}$ with respect to $|\cdot|_p$ is $\mathbb{Q}_p$, which is called the field of p-adic numbers.

There could be many more absolute values on $\mathbb{Q}$ besides the two that are given, but the Ostrowski's Theorem states that any non-trivial absolute value on $\mathbb{Q}$ is equivalent to either $|\cdot|_\infty$ (if it is archimedean) or $|\cdot|_p$ for some prime p (if it is non-arhimedean). This is a great result because it nicely classifies all absolute values on $\mathbb{Q}$ and significantly reduced our work in search for more. Therefore, we can write the places of $\mathbb{Q}$ as $M(\mathbb{Q}) = \{\infty\} \cup P$, where the equivalence class of archimedean places is indexed by ∞, and all the equivalence classes of non-archimedean places are indexed by p's for each $p \in P$, the set of prime numbers in $\mathbb{Z}$.

With all the machinery introduced so far, we are ready to define the absolute values on a number field *K*. To begin with, for any place $v \in M(K)$, it is *extended* from a place $u \in M(\mathbb{Q})$, which we know is either archimedean or non-archimedean. As a result, we say that *v lies over* ∞ or p, which are denoted by $v \mid \infty$ or $v \mid p$ respectively. We first introduce all the archimedean places of *K*.

Definition 4.3.

1. Let $a \in K$ and let $\sigma_1, ..., \sigma_{r+2s}$ be the embeddings on *K*. The *archimedean absolute values* on *K* is defined by
$$|a|_{\sigma_i} = |\sigma_i(a)|_\infty$$
for $1 \leq i \leq r + 2s$.
2. The completion of *K* at v is denoted by K_v. If $v \mid u$ for some $u \in M(\mathbb{Q})$, then K_v/Q_u is a field extension. We will define the *local degree* of *K* at v to be the degree of this extension, and denote it by
$$d_v = [K_v : \mathbb{Q}_u]$$

3. For each $v \mid \infty$,

$$d_v = [K_v : \mathbb{Q}_\infty] = [K : \mathbb{R}] = \begin{cases} 1 & \text{if } v \text{ is real} \\ 2 & \text{if } v \text{ is complex} \end{cases}$$

Therefore

$$\sum_{v|\infty} d_v = r + 2s = d.$$

We now proceed to describe the non-archimedean places of K which are extended from the prime numbers of $\mathbb{Z}$ and the local degrees of the respective field extensions. Recall that O_K is the ring of integers of the number field K. Even though a prime number $p \in \mathbb{Z}$ may no longer be prime in O_K, we can factorize ideal $\langle p \rangle \subseteq O_K$ into its unique factorization of prime ideals, i.e., $\langle p \rangle = P_1^{e_1}...P_k^{e_k}$, where P_i denotes the prime ideals of O_K and e_i is its respective power, known as the *ramification degree* of P_i over p. Note that a unique factorization exists because O_K is a Dedekind domain. In a Dedekind domain, every nonzero prime ideal is maximal, hence P_i is a maximal ideal, which means that O_K/P_i is a field; in fact, it is a finite field of characteristic p, i.e., $|O_K/P_i| = p^{f_i}$ for some $f_i \in \mathbb{Z}_{>0}$. Each f_i is called the *inertia degree* of P_i over p.

Definition 4.4.

1. Let $p \in \mathbb{Z}$ be prime and let $\langle p \rangle = P_1^{e_1}...P_k^{e_k}$ be the unique factorization of the ideal $\langle p \rangle \subseteq O_K$, where e_i is the ramification degree of P_i over p. The *non-archimedean absolute values* on K is defined by

$$\left|\frac{a}{b}\right|_{P_i} = \frac{|a|_{P_i}}{|b|_{P_i}}$$

Specifically, for $a \in O_K$,

$$|a|_{P_i} = p^{-\frac{\operatorname{ord}_{P_i} a}{e_i}}$$

for all P_i lying over p, and where $\operatorname{ord}_{P_i} a$ is defined by

$$\operatorname{ord}_{P_i} a = \max\{j \in \mathbb{Z} \mid a \in P_i^j\}$$

2. Any non-archimedean place $v \in M(K)$ lying over a rational prime p corresponds to some prime ideal P_i as above, and its local degree is

$$d_v = [K_v : \mathbb{Q}_p] = e_i f_i$$

where each e_i, f_i are respectively the ramification degree and the inertia degree of P_i over p.

Now that we have a general ideal of the archimedean and non-archimedean absolute values of an algebraic number, we state without proof a generalization of the Artin-Whaples product formula over a number field *K*. The proof relies on the understanding that any algebraic integer will only have finitely many prime factors.

Lemma 4.1.

If $0 \neq a \in K$, then

$$\prod_{v \in M(K)} |a|_v^{d_v} = 1.$$

In the next chapter, we will primarily use the Weil height to quantify the complexity of an algebraic number $\alpha \in K$, and we are now ready to formally introduce it.

Definition 4.5.

Let $\alpha \in K$. The *Weil height* of α, denoted $h(\alpha)$, is defined as

$$h(\alpha) = \prod_{v \in M(K)} \max\{1, |\alpha|_v\}^{d_v/d},$$

where $d_v = [K_v : \mathbb{Q}_v]$ is the local degree of *K* at the place $v \in$ *M(K)*. Notice that for each $v \mid \infty$, $d_v = 1$ if *K* is a totally real field.

5 Results on Semigroups in Lattices and Totally Real Number Fields

5.1 Restricted successive minima

Let $d \geq 2$ and L be a lattice of full rank in $\mathbb{R}^d$, and let us write $\mathbb{R}^d_{\geq 0}$ for the positive orthant of the Euclidean space $\mathbb{R}^d$. Define $L^+ = L \cap \mathbb{R}^d_{\geq 0}$, which are all points of the lattice L with non-negative coordinates. Then, L^+ is an additive semigroup.

A *positive basis* of a lattice consists of basis vectors that are all non-negative. In other words, if $\{\boldsymbol{x}_1, ..., \boldsymbol{x}_d\}$ is such a basis, then $\boldsymbol{x}_i \in \mathbb{R}^d_{\geq 0}$ for all i. We first prove a lemma about the existence of a positive basis for a lattice.

Lemma 5.1.

Every lattice has a positive basis; in fact, there exist infinitely many positive bases for any lattice.

Proof. Let $\{\boldsymbol{y}_1, ..., \boldsymbol{y}_d\}$ be a basis for the lattice L and let Y denote the corresponding basis matrix. Then, $L = Y\mathbb{Z}^d$. Pick $\boldsymbol{x}_1 \in L \cap \mathbb{R}^d_{\geq 0}$ such that $\boldsymbol{x}_1 = \sum_{i=1}^d a_i \boldsymbol{y}_i$, where all $\boldsymbol{x}_i$ are relatively prime, and not all $x_i = 0$. Let $\mathbf{a} = (a_1 ... a_d)$ be a row vector, and by Lemma 2, p. 15, [1], $\mathbf{a}$ is extendable to a matrix A with $\det(A) = \pm 1$. Hence, by Lemma 3.1, YA is another basis matrix for L with $\boldsymbol{x}_1$ as the first column. Denote the rest of the columns of YA by $\boldsymbol{x}_2, ..., \boldsymbol{x}_d$. For each $\boldsymbol{x}_i$ and some $z_i \in \mathbb{Z}$, let $\boldsymbol{x}_i + z_i \boldsymbol{x}_1$ be the new column vector with all positive coordinates. The resulting vectors $\{\boldsymbol{x}_1, \boldsymbol{x}_2 + z_2 \boldsymbol{x}_1, ..., \boldsymbol{x}_d + z_d \boldsymbol{x}_1\}$ is a totally positive basis for L. Infinitely many such basis for L can be obtained by adding positive integer multiples of any basis vector to the rest.

□

With a specific choice of a positive basis X of L, we can look at all lattice points spanned by X with non-negative integer coefficients. In other words, we can define

$$S(X) = \left\{ \sum_{i=1}^{n} a_i \boldsymbol{x}_i \mid a_i \in \mathbb{Z}_{\geq 0} \right\} = X\mathbb{Z}^d_{\geq 0}.$$

It can be easily checked that $S(X)$ is a subsemigroup of L^+.

In addition to $S(X)$, we define two more structures on $\mathbb{R}^d$. First, let

$$C(X) = \left\{ \sum_{i=0}^{n} r_i \boldsymbol{x}_i \mid r_i \in \mathbb{R}_{\geq 0} \right\} = X\mathbb{R}^d_{\geq 0},$$

which we call the *positive cone spanned by* X. We also define the set of *gaps* of $S(X)$ in L^+ to be $G(X) = L^+ \setminus S(X)$. In other words, $G(X)$ consists of non-negative lattices points where some of its integer coefficients are necessarily negative. A simple fact about the relationship between $S(X)$ and the positive cone $C(X)$ is that $S(X) = L^+ \cap C(X)$. This means that no gaps of $S(X)$ can be contained in the positive cone $C(X)$. We now prove this fact, which relies on the basic linear algebra property that for an element $y \in L$ with $y = \sum_{i=1}^{d} z_i \boldsymbol{x}_i$, this representation is necessarily unique.

Lemma 5.2.

Let $X = \{\boldsymbol{x}_1, ..., \boldsymbol{x}_d\}$ be a positive basis for L, then $S(X) = L^+ \cap C(X)$.

Proof. $S(X) \subseteq L^+ \cap C(X)$ is obvious. We now show that $L^+ \cap C(X) \subseteq S(X)$. Suppose not, then there exists some $\boldsymbol{y} \in L^+ \cap C(X)$ where $y \notin S(X)$. So $\boldsymbol{y} = \sum_{i=1}^{d} z_i \boldsymbol{x}_i$ for all $z_i \in \mathbb{Z}$ where some $z_i \neq 0$. On the other hand, since $\boldsymbol{y} \in C(X)$, we also have $\boldsymbol{y} = \sum_{i=1}^{d} r_i \boldsymbol{x}_i$ for some $r_i \in \mathbb{R}_{\geq 0}$. So $\boldsymbol{y} = \sum_{i=1}^{d} z_i \boldsymbol{x}_i = \sum_{i=1}^{d} r_i \boldsymbol{x}_i$, which means that $\sum_{i=1}^{d} (z_i - r_i) \boldsymbol{x}_i = 0$ for some $z_i - c_i \neq 0$. Since the representation of $\boldsymbol{y}$ in L is unique, this contradicts the fact that the basis vectors $\boldsymbol{x}_1, ..., \boldsymbol{x}_d$ are linearly independent in $\mathbb{R}^d$. Hence, $S(X) = L^+ \cap C(X)$.

□

The implication for Lemma 5.2 is that all the gaps of $S(X)$ in L^+ are outside of the cone $\mathbb{C}(X)$. It can be observed that there are infinitely many such gaps, but most of them are integer multiples of each other. Hence we projectively define a subset $G_{\text{pr}}(X)$ to be the *primitive gaps* of $G(X)$, i.e., $G_{\text{pr}}(X) = \{\boldsymbol{y} \in G(X) : \boldsymbol{y} \neq z\boldsymbol{y}' \text{ for any } \boldsymbol{y}' \in L \text{ and integer } z > 1\}$.

Lemma 5.3.

The set $G_{\text{pr}}(X)$ is finite if and only if the positive basis X is orthogonal.

Proof. Suppose the positive basis X is orthogonal. Since $X \subseteq R^d_{\geq 0}$ and the lattice is of full rank, X aligns with the coordinate axes of $R^d_{\geq 0}$, which means that the number of gaps is 0, and hence $G_{\text{pr}}(X)$ is finite. Now suppose the positive basis is not orthogonal. This means that $C(X)$ is an acute cone, and so the set $R^d_{\geq 0} \setminus C(X)$ is unbounded. Thus its interior contains Euclidean balls of arbitrarily large radius, hence the union of all such balls must contain infinitely many primitive points of the lattice L.

□

Lemma 5.4.

All gaps of $S(X)$ are positive integer multiples of some primitive gaps, i.e., $G(X) = \{z\boldsymbol{y} \mid \boldsymbol{y} \in G_{\text{pr}}(X),\ z \in \mathbb{Z}^+\}$.

Proof

Let $X = \{\boldsymbol{x}_1, ..., \boldsymbol{x}_d\}$ be a positive basis for L. Let $\boldsymbol{v}, \boldsymbol{y} \in L^+$ such that $\boldsymbol{y} = z\boldsymbol{v}$ for some $z \in \mathbb{Z}^+$. We just need to show that either $\boldsymbol{v}, \boldsymbol{y}$ are both in $S(X)$, or neither is in $S(X)$. By proof of contradiction, suppose $\boldsymbol{v} \in S(X)$ and $\boldsymbol{y} \notin S(X)$. Let $\boldsymbol{v} = \sum_{i=1}^{d} a_i\boldsymbol{x}_i$ and $\boldsymbol{y} = \sum_{i=1}^{d} b_i\boldsymbol{x}_i$ for some $a_i \in \mathbb{Z}^+$ and $b_i \in \mathbb{Z}$ and not all b_i's are positive. Then, we have that $\boldsymbol{y} = z\boldsymbol{v} = \sum_{i=1}^{d} b_i\boldsymbol{x}_i = z\sum_{i=1}^{d} a_i\boldsymbol{x}_i$, i.e., $\sum_{i=1}^{d}(za_i - b_i)\boldsymbol{x}_i = 0$ for some $za_i - b_i \neq 0$. This contradicts the linear independence of $\boldsymbol{x}_1, ..., \boldsymbol{x}_d$. Hence, either $\boldsymbol{v}, \boldsymbol{y} \in S(X)$ or $\boldsymbol{v}, \boldsymbol{y} \notin S(X)$.

We are now ready to introduce the main results of this paper. Recall that the i-th successive minima λ_i of a set M gives us a homogeneously expanded set $\lambda_i M$, whose intersection with Λ captures i linearly independent lattice points. We now define the *restricted successive minima* with respect to L^+ using the regular successive minima, which are defined in relation to the unit cube instead of the unit sphere.

Here is the construction. We first let the convex, compact, 0-symmetric set be a d-dimensional cube with side length $2t$ centered at the origin, i.e.,

$$\mathrm{C}_d(t) = \{\boldsymbol{x} \in \mathbb{R}^d \mid |x| \leq t\}.$$

Then, let $0 < \lambda_1 \leq ... \leq \lambda_d$ be the usual successive minima of L with respect to $\mathrm{C}_d(1)$, i.e.

$$\lambda_i := \inf\{t \in \mathbb{R}_{>0} \mid \dim(\mathrm{span}_{\mathbb{R}}\{\mathrm{C}_d(t) \cap \mathrm{L} \setminus \{\boldsymbol{0}\} \geq i\})\}.$$

Next, we restrict our attention to L^+ and let $0 < \lambda_1(L^+) \leq ... \leq \lambda_d(L^+)$ be the successive minima of L^+ with respect to $\mathrm{C}_d(1)$, i.e.

$$\lambda_i(L^+) := \inf\{t \in \mathbb{R}_{>0} \mid \dim(\mathrm{span}_{\mathbb{R}}\{\mathrm{C}_d(t) \cap (L^+) \setminus \{\boldsymbol{0}\} \geq i\})\}.$$

Finally, for a positive basis X of L, we define the restricted successive minima of $G(X) = L^+ \setminus S(X)$ in the spirit of Henk and Thiel's definition in [5]. In this paper, we define them to be

$$\lambda_i(L^+, X) := \inf\{t \in \mathbb{R}_{>0} \mid \dim(\mathrm{span}_{\mathbb{R}}\{\mathrm{C}_d(t) \cap G(X) \setminus \{\boldsymbol{0}\} \geq i\})\}.$$

In other words, $\lambda_i(L^+)$ (respectively, $\lambda_i(L^+, X)$) is the minimal non-negative real number t such that there exist i linearly independent vectors in L^+ (respectively, gaps of $S(X)$ in L^+) with supnorm no bigger than t. As a result, we necessarily have that

$$0 < \lambda_1(L^+) \leq ... \leq \lambda_d(L^+),\ 0 < \lambda_1(L^+, X) \leq ... \leq \lambda_d(L^+, X).$$

The following theorems are bounds obtained on these two special kinds of successive minima in terms to the usual inhomogenenous minimum of L, i.e.,

$$\mu(L) := \min\{t \in \mathbb{R}_{>0} \mid \mathrm{B}_d(t) + L = \mathbb{R}^d\}.$$

Theorem 5.5.

Let $L \subset \mathbb{R}^d$ be a lattice of full rank. Then

$$\lambda_1(L^+) \leq 2\mu(L) + 1, \ \lambda_i(L^+) \leq 2\lambda_i(\mu(L) + 1),$$

for all $2 \leq i \leq d$.

Proof. To prove $\lambda_1(L^+) \leq 2\mu(L) + 1$, we first show that there exists a lattice point $\boldsymbol{y} \in L$ with $1 \leq |\boldsymbol{y}| \leq 2\mu(L) + 1$ where $|\boldsymbol{y}|$ denotes the sup-norm of $\boldsymbol{y}$, i.e., the Euclidean distance of its largest coordinate. Then, we will use this $\boldsymbol{y}$ to construct the bound on $\lambda_1(L^+)$.

Let $r = \mu(L)$ and let $\boldsymbol{z}_r = (r+1)(1, ..., 1)^T \in \mathbb{R}^d$. Since $\mathbb{R}^d = \bigcup_{\boldsymbol{x} \in L}(B_d(r) + \boldsymbol{x})$ by definition, the ball $B_d(r) + \boldsymbol{z}_r$ must be covered by some translates of $B_d(r)$ by points of the lattice L. Specifically, at least one of them will have its center in $B_d(r) + \boldsymbol{z}_r$ because they are balls of the same radius. Hence, there must exist $\boldsymbol{y} \in L \cap (B_d(r) + \boldsymbol{z}_r) \subset C_d(r) + \boldsymbol{z}_r$, where $C_d(r) + \boldsymbol{z}_r = \{\boldsymbol{x} \in \mathbb{R}^d \mid 1 \leq x_i \leq 2r + 1\}$ is the corresponding cube in the positive orthant of $\mathbb{R}^d$. As a result, for the given $\boldsymbol{y} \in L$, we have $1 \leq |\boldsymbol{y}| \leq 2\mu(L) + 1$.

Let $\boldsymbol{x}_1, ..., \boldsymbol{x}_d$ be the vectors corresponding to the successive minima $\lambda_1, ..., \lambda_d$. Then, for at least one $1 \leq j \leq d$, $\boldsymbol{y} \notin \operatorname{span}_{\mathbb{R}}\{\boldsymbol{x}_i, i \neq j\}$. For this j, let $I_j = \{1 \leq i \leq d : i \neq j\}$. Then the collection of d vectors $\{\boldsymbol{y}\} \cup \{\lambda_i \boldsymbol{y} + \boldsymbol{x}_i : i \in I_j\}$ is linearly independent, and for each $i \in I_j$, $|\lambda_i \boldsymbol{y} + \boldsymbol{x}_i| \leq \lambda_i |\boldsymbol{y}| + |\boldsymbol{x}_i| \leq 2\lambda_i(\mu(L) + 1)$. Further, since $|\boldsymbol{y}| \geq 1$, for each $1 \leq k \leq d$, the k-th coordinate of each of such vector $\boldsymbol{y} + \boldsymbol{x}_i$ is greater than or equal to $\lambda_i + x_{ik} \geq 0$, so all of these vectors are in L^+.

□

Now that we have estimated the successive minima for L^+, we would like to proceed by giving estimates to the restricted successive minima of $G(X)$ by manipulating the coefficients of the linear combinations.

Theorem 5.6.

Let $X = \{\boldsymbol{x}_1, ..., \boldsymbol{x}_d\}$ be a positive basis for L. There exist linearly independent vectors $\boldsymbol{z}_1, ..., \boldsymbol{z}_d \in G_{\text{pr}}(X)$ with

$$|\boldsymbol{z}_i| = \max_{1 \leq m \leq d} \left\{ \left(\max_{1 \leq k \leq d} \left[\frac{x_{ik}}{\sum_{j=1, j \neq i}^{d} x_{jk}} \right] + 1 \right) \sum_{j=1, j \neq i}^{d} x_{jm} - x_{im} \right\}$$

Proof. It is enough to prove that there exist such vectors $|\boldsymbol{z}_i| \in G(X)$ because $\boldsymbol{z}_i = m\boldsymbol{z}'_i$ for some $m \in \mathbb{Z}^+$ and $\boldsymbol{z}'_i \in G_{\text{pr}}(X)$, so $|\boldsymbol{z}'_i| \leq |\boldsymbol{z}_i|$. Suppose that a vector

$\boldsymbol{z} = \sum_{i=1}^{d} a_i \boldsymbol{x}_i \in L^+$ with at least one of the integer coefficients $a_i = -1$. Then $z \in G(X)$.

Now, let $\boldsymbol{y} = \boldsymbol{x}_1 + ... + \boldsymbol{x}_d$. Since all coordinates of the $\boldsymbol{x}_i$'s are nonnegative and these vectors form a basis for $\mathbb{R}^d$, the sum of their k-th coordinates has to be positive for each $1 \leq k \leq d$. For each $1 \leq i \leq d$, define

$$\boldsymbol{z}_i = a_i \boldsymbol{y} - (a_i + 1)\boldsymbol{x}_i = a_i \sum_{j=1, j \neq i}^{d} \boldsymbol{x}_j - \boldsymbol{x}_i,$$

where a_i is a positive integer to be specified. In order for such $\boldsymbol{z}_i$ to be in $G(X)$, we only need it to be in L^+, meaning that for each $1 \leq k \leq d$, we must have $a_i x_{jk} > x_{ik}$. Hence, let

$$a_i = \max_{1 \leq k \leq d} \left\lceil \frac{x_{ik}}{\sum_{j=1, j \neq i}^{d} x_{jk}} \right\rceil + 1.$$

With this choice of a_i, we will have linearly independent $\boldsymbol{z}_i \in G(X)$ where

$$|\boldsymbol{z}_i| = \max_{1 \leq m \leq d} \left\{ \left(\max_{1 \leq k \leq d} \left\lceil \frac{x_{ik}}{\sum_{j=1, j \neq i}^{d} x_{jk}} \right\rceil + 1 \right) \sum_{j=1, j \neq i}^{d} x_{jm} - x_{im}. \right\}$$

This concludes the proof.

□

5.2 Positive semigroups in number fields

In this section, we will translate the bounds on the successive minima of L^+ into their equivalents in a number field using Weil height. It makes sense to do so because a lattice in a totally real number field corresponds to the L^+ lying in the positive orthant of Euclidean space, under the Minkowski embedding. To make these statements precise, we first introduce the general setup.

Let *K* be a totally real number field of degree d over $\mathbb{Q}$ and $\sigma_1, ..., \sigma_d : K \to \mathbb{R}$ be the embeddings of *K*, which are all real because *K* is a totally real number field. Let *I* be an ideal in the ring of integers O_K. Let I^+ be the additive semigroup of totally positive elements in *I*, i.e., $I^+ = \{\alpha \in I \mid \sigma_i(\alpha) \geq 0\}$ for all $1 \leq i \leq d$.

Earlier in the section, we have defined a positive basis $X = \{\boldsymbol{x}_1, ..., \boldsymbol{x}_d\}$ for a general lattice *L* to contain only non-negative basis vector, i.e., $\boldsymbol{x}_i \in \mathbb{R}^d_{\geq 0}$ for all i. Also recall that an ideal $I \subseteq O_K$ is a lattice of full rank in the Euclidean space $\mathbb{R}^d$ under the Minkowski embedding $\Sigma = (\sigma_1, ..., \sigma_d) : K \to \mathbb{R}^d$, which we can write as $L_I = \Sigma(I)$. Then, it follows naturally that a positive basis for L_I as a lattice in $\mathbb{R}^d$ corresponds to a positive basis for the ideal *I* as a lattice in the number field *K*. In other words, a basis $\boldsymbol{\beta} = \{\beta_1, ..., \beta_d\}$ of *I* is contained in I^+ if and only if its image $\Sigma(\boldsymbol{\beta}) = \{\Sigma(\beta_1), ..., \Sigma(\beta_d)\}$ is a basis for L_I contained in L_I^+.

With this choice of basis $\boldsymbol{\beta}$, define $S(\boldsymbol{\beta}) = \{a_i\beta_i \mid a_i \in \mathbb{Z}_{\geq 0}\}$ to be the subsemigroup of I^+ spanned by $\boldsymbol{\beta}$ with non-negative integer coefficients. It follows from lemma XX that I have infinitely many positive basis. Also, define the set of gaps of $S(\boldsymbol{\beta})$ in I^+ to be $G(\boldsymbol{\beta}) = I^+ \setminus S(\boldsymbol{\beta})$. It follows from definition that α is a gap of $S(\boldsymbol{\beta})$ if and only if $\Sigma(\alpha)$ is a gap of $S(\Sigma(\boldsymbol{\beta}))$.

It is legitimate to say that L^+ is a positive semigroup, since the basis $\boldsymbol{\beta}$ is never orthogonal. For $\boldsymbol{\beta}$ to be orthogonal the vectors $X = \{, \boldsymbol{x}_1, ..., \boldsymbol{x}_d\}$ must be lying along the coordinate axes in $\mathbb{R}^d$, meaning that these vectors have some entries that are zero. This is not possible, since the entries of any nonzero vector $y = \Sigma(\alpha) \in L_I$ are conjugates of a nonzero element $\alpha \in I$, hence cannot be zero. Therefore, we get the analogous result that the set of gaps $G(\boldsymbol{\beta}) := I^+ \setminus S(\boldsymbol{\beta})$ is infinite.

Let us write $L_I = \Sigma(I)$ to denote this lattice in $\mathbb{R}^d$. Then, because $\mathbb{N}_K(I) = \frac{\det(L_I)}{\det(O_K)}$, we have

$$\det(L_I) = \mathbb{N}_K(I) \cdot |\Delta_K|^{1/2}. \tag{4}$$

From (2), (3) and our earlier choice of $M = C_d(1)$, we have the following inequalities for the successive minima and the inhomogeneous minimum of L:

$$\prod_{i=1}^{d} \lambda_i \leq \det(L),\ \mu(L) \leq \frac{\sqrt{d}}{2} \sum_{i=1}^{d} \lambda_i. \tag{5}$$

One last thing we want to introduce is a measure of the relative "sizes" of the algebraic numbers in K. In a general lattice L, lattice points can be viewed as d-dimensional vectors, so it makes sense to measure them by sup norm or the usual Euclidean distance. But in a number field K, we will use the Weil height as a measure of sizes for the algebraic numbers, which conveys information about their arithmetic complexity.

Let $\alpha \in K$. The Weil height of α is $h(\alpha) = \prod_{v \in M(K)} \max\{1, |\alpha|_v\}^{d_v/d}$, where $d_v = [K_v : \mathbb{Q}_v]$ is the local degree of K at the place $v \in M(K)$. Since K is totally real, $d_v = 1$ for each $v \mid \infty$.We first give a lemma that nicely compares Weil height of algebraic numbers with its sup norm as a lattice point.

Lemma 5.7.

1. For every nonzero $\alpha \in O_K$, $1 \leq h(\alpha) \leq |\Sigma(\alpha)|$.
2. For every nonzero $\beta \in K$, $|\Sigma(\beta)| \leq h(\beta)^d$.

Proof. Let $\alpha \in O_K$ be nonzero. First, notice that for each v over a non-archimedean place, $|\alpha|_v$ is at most 1. So $\max\{1, |\alpha|_v\} = 1$ for all $v \nmid \infty$. This leads to

$$h(\alpha) = \prod_{v|\infty} \max\{1, |\alpha|_v\}^{1/d} = \left(\prod_{i=1}^{d} \max\{1, |\sigma_i(\alpha)|_\infty\} \right)^{1/d}.$$

Next, by the Artin-Whaples product formula combined with the arithmetic-geometric mean inequality, we obtain the following bound:

$$1 = \prod_{v \in M(K)} |\alpha|_v^{d_v/d} \le \prod_{i=1}^{d} |\sigma_i(\alpha)|^{1/d} \le \frac{1}{d} \sum_{i=1}^{d} |\sigma_i(\alpha)| \le \max_{1\le i\le d} |\sigma_i(\alpha)| = |\Sigma(\alpha)|.$$

This implies that

$$h(\alpha) \le \left(\prod_{i=1}^{d} \max_{1\le j\le d} \{|\sigma_j(\alpha)|\} \right)^{1/d} \le |\Sigma(\alpha)|.$$

On the other hand, for a nonzero $\beta \in K$,

$$|\Sigma(\beta)| = \max_{1\le i\le d} |\sigma_i(\beta)| \le \prod_{i=1}^{d} \max\{1, |\sigma_i(\beta)|\} \le \prod_{v \in M(K)} \max\{1, |\beta|_v^{d_v}\} = h(\beta)^d.$$

□

With the help of this lemma, we are ready to state the bounds on the successive minima of the ideal *I* with respect to the Weil height.

Lemma 5.8.

1. There exist $\mathbb{Q}$-linearly independent elements $s_1, \dots, s_d \in I$ such that

$$\prod_{i=1}^{d} h(s_i) \le \prod_{i=1}^{d} |\Sigma(s_i)| \le \mathbb{N}_K(I)|\Delta_K|^{1/2}.$$

2. The inhomogeneous minimum of the lattice L_I satisfies the inequality

$$\mu(L_I) \le \frac{d^{3/2}}{2} \mathbb{N}_K(I)|\Delta_K|^{1/2}.$$

Proof. Notice that the successive minima λ_i of the lattice L_I, as we defined them earlier in this chapter, are with respect to the set $C_d(t)$, which includes all lattice points $\boldsymbol{x}$ such that $|\boldsymbol{x}| \le t$. Therefore, the i-th successive minima λ_i is the smallest sup-norms of the i-th linearly independent vector in L_I. Hence, let $s_1, \dots, s_d \in I$ be a collection of elements so that $\Sigma(s_1), \dots, \Sigma(s_d) \in L_I$ the linearly independent vectors corresponding to these successive minima. Using the fact that $\prod_{i=1}^{d} \lambda_i \le$

det(L), we have

$$\prod_{i=1}^{d} h(s_i) \leq \prod_{i=1}^{d} |\Sigma(s_i)| = \prod_{i=1}^{d} \lambda_i \leq \det(L_I) \leq \mathbb{N}_K(I) \cdot |\Delta_K|^{1/2}.$$

Combining this bound with $\mu(L) \leq \frac{\sqrt{d}}{2} \sum_{i=1}^{d} \lambda_i$ we obtain

$$\mu(L_I) \leq \frac{\sqrt{d}}{2} \sum_{i=1}^{d} \lambda_i \leq \frac{\sqrt{d}}{2} d \cdot \prod_{i=1}^{d} \lambda_i \leq \frac{d^{3/2}}{2} \mathbb{N}_K(I) \cdot |\Delta_K|^{1/2}.$$

□

We proceed to give bounds on the product of linearly independent elements in the semigroup L^+ of L.

Lemma 5.9.

There exist $\mathbb{Q}$-linearly independent elements $\alpha_1, \ldots, \alpha_d \in I^+$ such that

$$\prod_{i=1}^{d} h(\alpha_i) \leq \left(2d\sqrt{d}\right)^d \left(\mathbb{N}_K(I)|\Delta_K|^{1/2}\right)^{d+1}.$$

Proof.

As explained earlier, the set of elements $\alpha_1, \ldots, \alpha_d$ is $\mathbb{Q}$-linearly independent in I^+ if and only if $\Sigma(\alpha_1), \ldots, \Sigma(\alpha_d)$ are $\mathbb{Q}$-linearly independent in L_I^+. Let $s_1, \ldots, s_d$ be a set of linearly independent elements in L. Combining Lemma 5.7 with (4), (5), there exist $\mathbb{Q}$-linearly independent in I^+ such that

$$\begin{aligned}
\prod_{i=1}^{d} h(\alpha_i) &\leq \prod_{i=1}^{d} |\Sigma(\alpha_i)| \leq (2\mu(L_I)+1) \prod_{i=2}^{d} 2|\Sigma(s_i)|(\mu(L_I)+1) \\
&\leq 4^d \mu(L_I)^d \prod_{i=2}^{d} |\Sigma(s_i)| \leq 4^d \mu(L_I)^d \prod_{i=1}^{d} |\Sigma(s_i)| \\
&\leq \left(2d^{3/2}\right)^d (\mathbb{N}_K(I)|\Delta_K|)^{d+1},
\end{aligned}$$

where the last inequality follows by Lemma 5.8.

□

Lemma 5.10.

Let $\boldsymbol{\beta} = \{\beta_1, \ldots, \beta_d\}$ be a positive basis for the ideal I and $G(\boldsymbol{\beta}) = I^+ \setminus S(\boldsymbol{\beta})$ be the corresponding set of gaps. For each $1 \leq i \leq d$, let $\beta_i' = \sum_{j=1, j\neq i}^{d} \beta_j$. Then there exist $\mathbb{Q}$-linearly independent gaps $\alpha_1, ..., \alpha_d \in G(\boldsymbol{\beta})$ such that

$$h(\alpha_i) \leq \Big(h(\beta_i/\beta_i')^d + 1\Big)\, h(\beta_i')^d.$$

Proof.

Notice that $G(\Sigma(\boldsymbol{\beta})) = \Sigma(G(\boldsymbol{\beta}))$. Then there exist $\mathbb{Q}$-linearly independent elements $\alpha_1, \ldots, \alpha_d \in G(\boldsymbol{\beta})$ such that

$$\begin{aligned} |\Sigma(\alpha_i)| &= \max_{1\leq m\leq d}\left\{\left(\max_{1\leq k\leq d}\left[\frac{\sigma_k(\beta_i)}{\sum_{j=1, j\neq i}^{d}\sigma_k(\beta_j)}\right]+1\right)\sum_{j=1,j\neq i}^{d}\sigma_m(\beta_j)-\sigma_m(\beta_i)\right\} \\ &\leq \max_{1\leq m\leq d}\left\{\left(\max_{1\leq k\leq d}\sigma_k\left(\frac{\beta_i}{\beta_i'}\right)+1\right)\sigma_m\left(\beta_i'\right)-\sigma_m(\beta_i)\right\} \\ &\leq \big(|\Sigma(\beta_i/\beta_i')|+1\big)\,|\Sigma(\beta_i')| \leq \Big(h(\beta_i/\beta_i')^d+1\Big)\,h(\beta_i')^d, \end{aligned}$$

by Lemma 5.7. This completes the proof.

□

References

[1] J. W. S. Cassels. *An Introduction to the Geometry of Numbers.* Springer, 1997.

[2] L. Fukshansky. Geometric Number Theory. `https://www1.cmc.edu/pages/faculty/lenny/papers/GNT_lecture_notes.pdf`.

[3] L. Fukshansky and S. Wang. Positive semigroups in lattices and totally real number fields. *in preparation*, 2021.

[4] Peter M. Gruber and C. G. Lekkerkerker. *Geometry of Numbers.* North-Holland, 1987.

[5] M. Henk and C. Thiel. Restricted successive minima. *Pacific J. Math.*, 269(2):341–354, 2014.

[6] Serge Lang. *Algebraic Number Theory.* Springer Science Business Media, 2013.