

Geometry of Numbers TCC Module 2024

Simon L Rydin Myerson

June 11, 2024

Contents

1	Introduction	2
2	Basic Notions and Definitions	4
2.1	The fundamentals	4
2.2	First results on bases	5
2.3	Gram Determinant and Determinant of a Lattice	5
2.4	What is lattice reduction?	6
2.5	Aside: Quadratic Forms	6
2.6	Sublattices	7
2.7	Orthogonal Projection of Lattices	7
2.8	Dual Lattices	8
2.9	Minkowski Minima	8
2.9.1	Bases achieving the minima	9
2.10	Gram-Schmidt Orthogonalization (GSO)	10
3	Lattice reduction	11
3.1	First applications of GSO	11
3.2	Size-Reduction and HKZ Reduction	12
3.3	How to think of a reduced basis?	13
3.4	Digression: The LLL Algorithm	14
3.5	The dual lattice again	15
3.6	Minima with respect to general norms	15
3.7	The classical perspective	16
4	Orthogonal lattices and an application	18
4.1	The orthogonal lattice	18
4.2	A short introduction to Manin's conjecture	19
4.2.1	Rational points	19

Chapter 1

Introduction

The geometry of numbers in its canonical form comes from work of Minkowski in the late 19th century. The general idea is to give estimates for the cardinality $N(B, M) = |B \cap M\mathbb{Z}^n|$ when M varies over some set of real matrices, B is a fixed domain containing the origin, and we want our bounds to have explicit dependence on both quantities. For example, if $M \in \text{SL}_n$, we may ask what conditions on B always guarantee a nonzero point in $B \cap M\mathbb{Z}^n$. Alternatively, if $M = mI_n$ is a dilation, we may ask for simple-looking upper and lower bounds with the same order of magnitude for $N(B, M)$. To answer the latter question, we need a 'reduced basis' for \mathbb{Z}^n .

Classical applications of geometry of numbers include Minkowski's results on discriminants and class numbers of number fields, the study of lattice packings, and a simple proof of the four-squares theorem. Historically, there was a lot of interest in the case $B = \{x : \prod x_i \leq 1\}$, in applying the 'reduced basis' idea to study 'reduction of quadratic forms', and in getting very good constants in bounds on discriminants of number fields, culminating perhaps in the work of Rogers and Mulholland.

More recent developments in the field include the LLL algorithm and algorithmic geometry of numbers, counting number fields with bounded discriminant and specified Galois group, and the reduction theory of quadratic forms as a key component of Bhargava's work ('Bhargavology'). The 'parametric geometry of numbers' goes beyond the cases when M varies over dilations, to a two-parameter family, which in the 2D case boils down to continued fractions, providing a way to extend the concept of continued fractions to vectors. Some reading on these ideas is available, but they aren't the focus of this module.

Instead we will focus on another direction, anticipated by Davenport: the geometry of numbers is a powerful tool for estimating the number of solutions to systems of linear and multilinear Diophantine equations and inequalities. This leads to many applications in modern number theory, where such quantities may control error terms even when they are not the main quantity of interest.

For example:

Theorem 1 (Maynard, 2019). *There exist infinitely many prime numbers whose decimal representation does not contain the digit 7.*

- Methods combine sieve theory, harmonic analysis, and combinatorial geometry.
- Maynard's contributions continue to inspire research in the field.

Another example:

Theorem 2. (Schmidt, 1966) *Given a fixed k , for $1 \leq k \leq n-1$, the number of primitive lattices \mathcal{L}^k of determinant $\leq H$ is asymptotically given by $P(n, k, H) = a(n, k)H^n + O(H^{n-\delta})$.*

Topics in this course include Minkowski minima, reduced bases, LLL algorithm, orthogonal and dual lattices, and point-counting results. The course will cover the proof of Manin's conjecture for $x_1y_1 + \dots + x_ny_n = 0$, which is a special case of a result of J L Thunder, and then examine a more advanced application of geometry of numbers, possibly in J Maynard's work on primes with missing digits or small fractional parts of polynomials.

Chapter 2

Basic Notions and Definitions

This chapter, and most of the next, are very closely based on Nguyen's Chapter 2 in Nguyen and Vallée (2010).

2.1 The fundamentals

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

Definition 2 (Lattice). A *lattice* of \mathbb{R}^n is a discrete subgroup of $(\mathbb{R}^n, +)$.

Examples of lattices include the zero lattice, the lattice of integers \mathbb{Z}^n , and its subgroups.

Definition 3 (Lattice generated by a set). If $b_1, \dots, b_d \in \mathbb{R}^n$, define

$$L(b_1, \dots, b_d) = \{n_1 b_1 + \dots + n_d b_d : n_i \in \mathbb{Z}\}$$

to be the group they generate.

Considering $\{x_1 b_1 + \dots + x_d b_d : x_i \in (-\frac{1}{2}, \frac{1}{2})\}$, we can show:

Theorem 3 (Nguyen, Theorem 1). *If the b_i are linearly independent, then $L(b_1, \dots, b_d)$ is a lattice.*

In this case we say that b_1, \dots, b_d is a *basis* of $L = L(b_1, \dots, b_d)$.

Definition 4. The *dimension* or *rank* of a lattice L in \mathbb{R}^n , denoted by $\dim(L)$, is the dimension of its linear span denoted by $\text{span}(L)$. The lattice is said to be *full-rank* when $d = n$; I will try to denote the dimension by n when the lattice is full-rank, and by d otherwise.

2.2 First results on bases

Theorem 4 (Nguyen, Theorem 2). *Let L be a d -dimensional lattice of \mathbb{R}^n . If c_1, \dots, c_d are linearly independent vectors in L , then there exists a lower triangular matrix $(u_{i,j})$ such that the vectors b_1, \dots, b_d defined by $b_i = \sum_{j=1}^i u_{i,j}c_j$ are linearly independent and $L = L(b_1, \dots, b_d)$.*

Proof idea: Induction on d , assuming the result for $L' = L \cap \text{span}(c_1, \dots, c_{d-1})$ and choosing $u_{i,d}$ such that $b_d \in L$ and $u_{d,d}$ is minimal.

If $b \in L$ then by subtracting a multiple of b_d we can get a vector that is in L' , and conclude.

Theorem 5 (Nguyen, Theorem 3). *Let (b_1, \dots, b_d) be a basis of a lattice L in \mathbb{R}^n . Let c_1, \dots, c_d be vectors of L . Then there exists a unique $d \times d$ integral matrix $U = (u_{i,j})$ such that $c_i = \sum_{j=1}^d u_{i,j}b_j$ for all $1 \leq i \leq d$. And (c_1, \dots, c_d) is a basis of L if and only if the matrix U has determinant ± 1 .*

2.3 Gram Determinant and Determinant of a Lattice

Definition 5. If $b_1, \dots, b_d \in \mathbb{R}^n$, define the *Gram matrix* $(b_i \cdot b_j)_{1 \leq i, j \leq d}$.

The *Gram determinant* is the determinant of the Gram matrix and is written $\Delta(b_1, \dots, b_d)$.

Two bases (b_1, \dots, b_d) and (c_1, \dots, c_d) of a lattice L in \mathbb{R}^n are related by a U of determinant ± 1 , and hence

$$\Delta(b_1, \dots, b_d) = \Delta(c_1, \dots, c_d).$$

We can interpret $\Delta(b_1, \dots, b_d)$ as the square of the volume of $\{x_1b_1 + \dots + x_db_d : x_i \in (0, 1)\}$, hence it is positive. We will prove this for $d = n$, I leave it to you to deduce the general case.

Definition 6. Thus, the *determinant* $\det(L)$ is defined as $\Delta(b_1, \dots, b_d)^{1/2}$ and is independent of the choice of the basis. (Also called the volume or covolume of L .)

Lemma 1 (Nguyen, Lemma 2). *Let L be a full-rank lattice in \mathbb{R}^n :*

1. *For any basis (b_1, \dots, b_n) of L , $\det(L) = |\det(b_1, \dots, b_n)|$.*
2. *For any $r > 0$, the number of points of L in a ball of radius r , $s_L(r)$, satisfies*

$$\lim_{r \rightarrow \infty} \frac{s_L(r)}{r^n \text{Vol}(B(0, 1))} = \frac{1}{\det(L)}.$$

Proof idea: for any $y \in \mathbb{R}^n$ there is exactly one point of L in the region $\{y + x_1b_1 + \dots + x_nb_n : x_i \in [0, 1)\}$, which has volume $\det(L)$.

Definition 7 (Fundamental region). A region like $\{x_1b_1 + \dots + x_db_d : x_i \in (0, 1)\}$, or $\{y + x_1b_1 + \dots + x_nb_n : x_i \in [0, 1)\}$, or $\{x_1b_1 + \dots + x_db_d : x_i \in (-\frac{1}{2}, \frac{1}{2})\}$ is called a fundamental region or fundamental domain for $L(b_1, \dots, b_d)$. These are often useful in proofs of all kinds, but especially when counting lattice points is concerned.

Definition 8 (Gaussian Heuristic). For a full-rank lattice L in \mathbb{R}^n , and C a measurable subset of \mathbb{R}^n , the *Gaussian Heuristic* predicts that the number of points of $L \cap C$ is roughly $\text{vol}(C)/\text{vol}(L)$.

2.4 What is lattice reduction?

- Lattice reduction aims to find a basis with shorter and nearly orthogonal vectors.
- Corresponds to: make $\|\sum_{i=1}^d x_i b_i\|^2$ approximately diagonal with small coefficients.
- To measure “shorter” we need successive minima.
- To measure “nearly orthogonal” we need orthogonalisation.
- Before those, we need a few more basic definitions.

2.5 Aside: Quadratic Forms

Definition 9 (Quadratic Form). A function $q : \mathbb{R}^d \rightarrow \mathbb{R}$ defined by $q(x_1, \dots, x_d) = \|\sum_{i=1}^d x_i b_i\|^2$, where (b_1, \dots, b_d) is a basis of a lattice L , is called a *positive definite quadratic form* over \mathbb{R}^d .

This is equivalent to the usual definition, in which $q(x) = x^T Q x$ for some symmetric positive definite matrix Q .

Indeed $\|\sum_{i=1}^d x_i b_i\|^2 = x^T Q x$ where Q is the Gram matrix of (b_1, \dots, b_d) .

And we can write any positive definite Q as the Gram matrix of some d linearly independent vectors by orthogonally diagonalising it:

$$Q = O^T \text{diag}(d_1, \dots, d_d) O, \quad B = \text{diag}(\sqrt{d_1}, \dots, \sqrt{d_d}) O \\ B = (b_1 \mid \dots \mid b_d) \in \text{Mat}_{n \times d}(\mathbb{R})$$

The values at integer points of $q(x)$ are the same as the values of $q(Ux)$ for any fixed $U \in \text{GL}_d(\mathbb{Z})$. This in turn is the same as changing basis of the lattice $L(b_1, \dots, b_d)$. So we can try to choose a basis to make $q(Ux)$ as ‘close to diagonal’ as possible. This is reduction of quadratic forms.

2.6 Sublattices

Definition 10 (Sublattice). A *sublattice* is a subgroup of a lattice; a *full-rank sublattice* of L is a sublattice with the same dimension as L .

Lemma 2 (Nguyen, Lemma 3). *The sublattice M is a full-rank sublattice of L if and only if the group index $[L : M]$ is finite, in which case $\det(M) = \det(L) \times [L : M]$.*

Definition 11 (Primitive Sublattice). A sublattice M of L is said to be *primitive* if there exists a subspace E of \mathbb{R}^n such that $M = L \cap E$.

Note that for a one-dimensional sublattice of \mathbb{Z}^n , this is the usual notion of primitive i.e. if $v \in \mathbb{Z}^n$ then $L(v)$ is primitive iff $\gcd(v) = 1$.

Lemma 3 (Nguyen, Lemma 4). *A sublattice M of L is primitive if and only if every basis of M can be completed to a basis of L , that is, for any basis (b_1, \dots, b_r) of M , there exist additional vectors such that (b_1, \dots, b_d) is a basis of L .*

It suffices to prove the lemma for $\iota^{-1}(L)$, where $\iota : \mathbb{R}^d \rightarrow \mathbb{R}^n$ is an isometry onto $\text{Span}(L)$. This means it suffices to prove the lemma for $d = n$. This is a trick we will use constantly!

The lemma then follows from

Theorem 6 (Nguyen, Theorem 2). *Let L be a d -dimensional lattice of \mathbb{R}^n . If c_1, \dots, c_d are linearly independent vectors in L , then there exists a lower triangular matrix $(u_{i,j})$ such that the vectors b_1, \dots, b_d defined by $b_i = \sum_{j=1}^i u_{i,j} c_j$ are linearly independent and $L = L(b_1, \dots, b_d)$.*

Proof idea for the theorem: One chooses b_i in such a way that u_{ii} is positive and minimal.

Proof idea for the lemma: extend b_1, \dots, b_r to a linearly independent d -tuple $b_1, \dots, b_d \in L$. Make a basis $b'_i = \sum u_{ij} b_j$. As u_{ij} is lower triangular, $b'_1, \dots, b'_r \in \text{Span } M$ and since M is primitive they must be in M . Using this, we can replace b'_1, \dots, b'_r with b_1, \dots, b_r .

Definition 12 (Primitive Vectors and Primitive Sublattice). Let (b_1, \dots, b_r) be an r -tuple of vectors of a lattice L . We say (b_1, \dots, b_r) is primitive (in L) if $L(b_1, \dots, b_r)$ is a primitive sublattice of L .

2.7 Orthogonal Projection of Lattices

Lemma 4 (Nguyen, Lemma 5). *Let L be a d -rank lattice in \mathbb{R}^n , and let M be a r -rank primitive sublattice of L . Let π_M be the orthogonal projection of L along the span of M , that is $\pi_M : \mathbb{R}^n \rightarrow \text{span}(M)^\perp$.*

Then $\pi_M(L)$ is a lattice of rank $d - r$, denoted by $\pi_M(L)$, with determinant $\det(L)/\det(M)$.

Proof. The proof relies on extending a basis of M to a basis of L and showing that the additional basis vectors project to a basis of $\pi_M(L)$. Some column operations then prove the volume statement.

Definition 13. If (b_1, \dots, b_d) is a basis of a lattice L in \mathbb{R}^n , and we let π_i be the projection along the first $i - 1$ of these basis vectors, that is $\pi_1 = \text{id}$ and $\pi_i = \pi_{L(b_1, \dots, b_{i-1})}$.

Corollary 1 (Nguyen, Corollary 2). *The image $\pi_i(L)$ has rank $d - i + 1$ and determinant $\det(L)/\det(L(b_1, \dots, b_{i-1}))$.*

- We'll use this later in various inductive arguments.

2.8 Dual Lattices

Definition 14 (Dual Lattice L'). The *dual lattice* L' of a lattice L consists of all vectors \mathbf{y} in the span of L such that $\langle \mathbf{y}, \mathbf{x} \rangle \in \mathbb{Z}$ for all $\mathbf{x} \in L$.

If $d = n$, and B has columns that are a basis of L , then L^* has a basis given by the columns of B^{-T} . So the dual 'is' the inverse transpose.

Lemma 5 (Nguyen, Lemma 6). *The dual lattice L' has the same rank and a reciprocal volume to L , satisfying $\text{vol}(L) \cdot \text{vol}(L') = 1$.*

Proof idea: If $d = n$ and $A \in \text{GL}_n(\mathbb{R})$ then $(AL)' = A^{-T}L'$. Thus we can reduce to the case $L = \mathbb{Z}^n$. Now we check that $(\mathbb{Z}^n)' = \mathbb{Z}^n$, that is $(y \in \mathbb{R}^n, \sum x_i y_i \in \mathbb{Z} \forall x \in \mathbb{Z}^n) \iff y \in \mathbb{Z}^n$.

2.9 Minkowski Minima

Definition 15 (Successive Minima/Minkowski Minima). For a lattice L of dimension d , the i -th minimum $\lambda_i(L)$ is the smallest radius such that a closed ball of that radius centered at the origin contains i linearly independent lattice vectors.

That is the $\lambda_i(L)$ are minimal such that for each i , there exist linearly independent $x_1, \dots, x_i \in L$ with $\|x_i\| \leq \lambda_i(L)$.

A lattice is *balanced* if all the minima have similar sizes. If the largest and smallest minima are very different in size it is *unbalanced*.

- The minima are increasing: $\lambda_1(L) \leq \lambda_2(L) \leq \dots \leq \lambda_d(L)$.
- They provide information about the lattice's density and packing.
- They describe the "shape" of the lattice. We can think of the lattice as a grid of dots separated in one direction by λ_1 , in another by λ_2 and so on. The dots are evenly distributed if all the λ_i are around the same size (balanced).

- A basis consists of “small” vectors if the basis vectors have norms close to λ_i . For if b_1, \dots, b_d is a basis of L with $\|b_1\| \leq \dots \leq \|b_d\|$ then we must have $\|b_i\| \geq \lambda_i(L)$, by the definition above.

2.9.1 Bases achieving the minima

- The four-dimensional lattice L with basis vectors satisfying $\sum x_i = \text{even}$ has all minima equal to $\sqrt{2}$.
- It serves as a canonical example illustrating that not all linearly independent d -tuples of lattice vectors of minimal length form a basis.
- The columns of the following matrix are linearly independent, minimal, but not a basis for \mathbb{Z}^4 .

NB my vectors are always column vectors!

$$\begin{pmatrix} 1 & 1 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & -1 & 1 \end{pmatrix}$$

- For the lattice generated by these vectors is $\{x \in \mathbb{Z}^4 : 2 \mid x_1 + x_2, 2 \mid x_3 + x_4\}$ which is a proper sublattice of L .
- There is however a minimal basis given by the columns of

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ -1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

- In five dimensions, no basis reaches all minima simultaneously. From the work of Korkine and Zolotarev, shows that the shortest vector problem becomes more complex with higher dimensions.
- Let L be the lattice generated by the columns of

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

L contains five linearly independent vectors of norm 2, but every basis includes a vector with all five entries nonzero (hence of norm at least $\sqrt{5} \approx 2.24$).

2.10 Gram-Schmidt Orthogonalization (GSO)

Definition 16. Gram-Schmidt orthogonalization (GSO) is a process that takes a set of linearly independent vectors b_1, \dots, b_d and produces an **orthogonal** set of vectors b_1^*, \dots, b_d^* that spans the same subspace.

- Each b_i^* is obtained by subtracting from b_i its projection onto $\text{span}(b_1^*, \dots, b_{i-1}^*)$.
- $b_1^* = b_1$ and $b_j^* = b_j - \sum_{u=1}^{j-1} \mu_{i,j} b_u^*$ for $1 \leq i < j \leq d$, where $\mu_{i,j} = \frac{b_j \cdot b_i^*}{\|b_i^*\|^2}$.
- The coefficients $\mu_{i,j}$ represent the b_i^* -coefficient of b_j .
- If they are small, the b_j are close to orthogonal.
- GSO allows us to represent the original basis $B = (b_1 \mid \dots \mid b_d)$ as $B = B^* \mu$, where B^* is the matrix with columns b_1^*, \dots, b_d^* .
- μ is an upper triangular matrix with 1's on the diagonal and $\mu_{i,j}$'s as the off-diagonal entries:

$$\mu = \begin{pmatrix} 1 & \mu_{1,2} & \cdots & \mu_{1,d} \\ 0 & 1 & \cdots & \mu_{2,d} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix},$$

$$b_j^* = b_j - \sum_{i=1}^{j-1} \mu_{i,j} b_i^*, \quad \mu_{i,j} = \frac{b_j \cdot b_i^*}{\|b_i^*\|^2}.$$

- Hence $\text{span}(b_1^*, \dots, b_j^*) = \text{span}(b_1, \dots, b_j) \dots$
- so $b_j^* = b_j - \sum_{i=1}^{j-1} \frac{b_j \cdot b_i^*}{\|b_i^*\|^2} b_i^* = \pi_j(b_j)$ (project $L(b_1, \dots, b_d)$ along $L(b_1, \dots, b_{j-1})$)

Theorem 7. If (b_1, \dots, b_d) is a basis of a lattice L , then the volume of L is equal to the product of the norms of the GSO vectors:

$$\det(L) = \prod_{i=1}^d \|b_i^*\|$$

- Proof idea: WLOG $d = n$. Now $\text{diag}(\|b_1^*\|, \dots, \|b_n^*\|)^{-1} B^* \in O_n(\mathbb{R})$, and so $\det(B) = \det(B^* \mu) = \det(O \text{diag}(\|b_1^*\|, \dots, \|b_n^*\|) \mu) = \prod \|b_i^*\|$.

Chapter 3

Lattice reduction

We continue to follow Chapter 2 of Nguyen-Vallée.

3.1 First applications of GSO

Lemma 6 (Gram-Schmidt and Minima). *If $(\mathbf{b}_1, \dots, \mathbf{b}_d)$ is a basis of a lattice L , then its Gram-Schmidt Orthogonalization satisfies $\lambda_1(L) \geq \min \|b_i^*\|$ for all $1 \leq i \leq d$.*

Proof idea: Let $v \in L$, then

$$\|v\| \geq \|\pi_2(v)\| \geq \|\pi_3\pi_2(\pi_1(v))\| = \|\pi_2(v)\| \geq \dots \geq \|\pi_d(v)\|.$$

Now $\pi_d(L) = L(b_d^*)$. If $\pi_d(v) \neq 0$ we therefore have $\|\pi_d(v)\| \geq \|b_d^*\|$. If $\pi_d(v) = 0$ then $v \in L(b_1, \dots, b_{d-1})$. We proceed by induction and find that $\|v\| \geq \|b_i^*\|$ for some i .

Definition 17 (Size-Reduced Basis). We say $(\mathbf{b}_1, \dots, \mathbf{b}_d)$ is size-reduced if its Gram-Schmidt orthogonalization coefficients satisfy $|\mu_{i,j}| \leq \frac{1}{2}$ for all $1 \leq j < i \leq d$.

- This ensures that each vector is as short as possible in the direction of the preceding vectors
- Every L has a size-reduced basis (double induction! b_1, b_2, \dots but $b_j^*, b_{j-1}^* \dots$)

Proof of existence of a size-reduced basis. By induction on k we show that $L(b_1, \dots, b_k)$ has a size-reduced basis. If $k = 1$ this is easy. Suppose by induction that b_1, \dots, b_{k-1} is size-reduced. We find b'_k so that $b_1, \dots, b_{k-1}, b'_k$ is a basis and size-reduced.

Recall $b_j^* = b_j - \sum_{i < j} u_{ij} b_i^*$. We know $|\mu_{ij}| \leq 1/2$ for $1 \leq i < j < k$.

First we replace b_k by $b_j - n_{k-1} b_{k-1}$ for some integer n_{k-1} to make $|\mu_{k-1,k}| \leq 1/2$, then by $b_j - n_{k-1} b_{k-1} - n_{k-2} b_{k-2}$ for some integer n_{k-2} to make $|\mu_{k-2,k}| \leq 1/2$, and so on.

The key fact is that

$$b_i^* \cdot b_j = \begin{cases} \|b_j^*\|^2 & i = j \\ 0 & i > j. \end{cases}$$

So

$$b_{k-t}^* \cdot (v - n_{k-t} b_{k-t}) = b_{k-t}^* \cdot v - n_{k-t} \|b_{k-t}^*\|^2$$

and there is n_{k-t} for which the expression above is at most $\frac{1}{2} \|b_{k-t}^*\|^2$ in absolute value. □

Size-reduction is sensitive to re-ordering, and the elements of a size-reduced basis can be much bigger than the minima λ_i .

For example let $c > 0$ be small. Then $b_1 = \begin{pmatrix} c \\ 1 \end{pmatrix}, b_2 = \begin{pmatrix} c \\ 1/2 \end{pmatrix}$ is size-reduced but b_2, b_1 is not.

Also $\lambda_1(b_1, b_2) = 2c$ which could be much smaller than $\|b_1\|, \|b_2\|$.

Our goal is to make a small, approximately orthogonal basis: so size-reduction is not enough!

3.2 Size-Reduction and HKZ Reduction

Definition 18 (HKZ-Reduced Basis). A basis is Hermite-Korkine-Zolotarev (HKZ)-reduced if it is size-reduced and each basis vector \mathbf{b}_i^* is the shortest vector in $\pi_i(L)$, the projection of the lattice onto the orthogonal complement of $\mathbf{b}_1, \dots, \mathbf{b}_{i-1}$.

Note that we need not have $\|b_1\| \leq \|b_2\| \leq \dots$

1. We prove by induction on j : there is a j tuple b_1, \dots, b_j which is size-reduced as a basis of $L(b_1, \dots, b_j)$, primitive, and such that $\mathbf{b}_i^* = \pi_i(b_i)$ is the shortest vector in $\pi_i(L)$ for all $i \leq j$.
2. for $j = 1$ let b_1 be the shortest vector of $L = \pi_1(L)$.
3. Assume the result for $j - 1$.
4. Let c^* be a shortest vector of $\pi_j(L)$. Let $c \in L$ with $\pi_j(c) = c^*$.
5. I claim that b_1, \dots, b_{j-1}, c is primitive.

Proof idea: If $d \in \text{Span}(b_1, \dots, b_{j-1}, c)$ then $\pi_j(d) \in \text{Span}(c^*)$. If $d \in L$, there is $n \in \mathbb{Z}$ such that $\pi_j(d - nc) \in \{xc^* : x \in [0, 1)\}$. We must have $x = 0$, so $\pi_j(d - nc) = 0$. By induction b_1, \dots, b_{j-1} are primitive so $d - nc \in L(b_1, \dots, b_{j-1})$.

6. Now size-reduce b_1, \dots, b_{j-1}, c : replace c with $b_j = c - \sum_{i=1}^{j-1} n_i b_i$ for suitable $n_i \in \mathbb{Z}$.

We show that HKZ-reduced bases are “really” reduced, their elements are about as short as possible.

Notice: $\|b_{j+k}^*\| \geq \|\pi_j(b_{j+k})\| \geq \|b_j^*\|$

Lemma 7. *For an HKZ-reduced basis $\lambda_j(L) \geq \lambda_1(\pi_j(L)) = \|b_j^*\|$.*

Proof. Indeed let $v_1, \dots, v_j \in L$ be linearly independent with $\|v_i\| \leq \lambda_j(L)$. At least one of the $\pi_j(v_i) \neq 0$. Then $\|v_i\| \geq \|\pi_j(v_i)\| \geq \lambda_1(\pi_j(L))$. \square

Theorem 8 (Ngyuen, Theorem 6 - Mahler/Korkine-Zolotarev). *Let (b_1, \dots, b_d) be an HKZ-reduced basis of a lattice L . Then for all indices j such that $1 \leq j \leq d$, the following inequality holds:*

$$\frac{4}{j+3} \leq \left(\frac{\|b_j\|}{\lambda_j(L)} \right)^2 \leq \frac{j+3}{4}$$

Proof sketch. • Last upper bound: the b_i^* are orthogonal, $\|b_i^*\| \leq \lambda_i(L)$, and $b_j = b_j^* + \sum_{i=1}^{j-1} \mu_{i,j} b_i^*$ with $|\mu_{i,j}| \leq \frac{1}{2}$.

- If we had $\|b_1\| \leq \dots \leq \|b_d\|$, then immediately $\|b_i\| \geq \lambda_i(L)$.
- First upper bound: $\|b_1\|, \dots, \|b_j\| \leq \sqrt{\frac{j+3}{4}} \|b_j\|$ as $b_i = b_i^* + \sum_{k=1}^{i-1} \mu_{k,i} b_k^*$ and $\|b_i^*\| \leq \|b_i\|$.

\square

In fact, once we have proved that an HKZ reduced basis exists and has the properties we want, all we ever need in this course is a size-reduced basis with $\|b_i\| \asymp_d \lambda_i(L)$, that is $c\lambda_i(L) \leq \|b_i\| \leq C\lambda_i(L)$ for some $0 < c < C$ depending only on d . One can relate such a basis to an HKZ-reduced basis to see that it has all properties we will need, with worse implicit constants.

3.3 How to think of a reduced basis?

Definition 19 (Big-O notation). We use $A \ll_{a,\dots,z} X$ or $A = O_{a,\dots,z}(X)$ to mean $|A| < CX$ for some constant C depending at most on a, \dots, z , we use $A \asymp X$ to denote $X \ll A \ll X$.

Lemma 8. *Let (b_1, \dots, b_d) be an HKZ-reduced basis of a lattice L . Then $\|b_j^*\| \geq \frac{1}{2} \|b_{j-1}^*\|$.*

Proof. $\|b_j^*\| = \|\pi_{j-1}(b_j^*)\| = \|\pi_{j-1}(b_j - \sum_{i<j} \mu_{i,j} b_i^*)\| = \|\pi_{j-1}(b_j) - \mu_{(j-1)j} b_{j-1}^*\| \geq \|\pi_{j-1}(b_j)\| - \frac{1}{2} \|b_{j-1}^*\| \geq \frac{1}{2} \|b_{j-1}^*\|$ as $\|b_{j-1}^*\|$ is minimal in $\pi_{j-1}(L) \setminus \{0\}$. \square

Lemma 9. *Let (b_1, \dots, b_d) be an HKZ-reduced basis of a lattice L . Then $\|x_1 b_1 + \dots + x_d b_d\| \asymp_n \|x_1 b_1^* + \dots + x_d b_d^*\|$ for all $x \in \mathbb{R}^d$.*

Proof. WLOG $d = n$, by using an isometry $\iota : \mathbb{R}^d \rightarrow \text{Span } L$. As b_i is size reduced and $\|b_{j+1}\| \geq \frac{1}{2} \|b_j\|$, we can show: the linear maps $b_i^* \leftrightarrow b_i$ are $B^* \mu^{\pm 1} B^{*-1} = OD \mu^{\pm 1} D^{-1} O^T$ with entries $\ll_n 1$. \square

Corollary 2. $\|b_j^*\| \asymp_n \|b_j\| \asymp_n \lambda_i(L)$

Corollary 3 (My Favorite Corollary). $\|x_1b_1 + \dots + x_db_d\| \asymp_n \max_i |x_i|\lambda_i(L)$ for all $x \in \mathbb{R}^d$.

Corollary 4.

- $\det(L) = \prod \|b_i^*\| \asymp_n \prod \lambda_i(L)$.
- If $i \neq j$ the angle between b_i and b_j is $\gg_n 1$.
- For any $v \in L$, $v = \sum n_i b_i$ for some $n_i \in \mathbb{Z}$, and then $\|v\| \asymp_n \max_i |n_i|\lambda_i(L)$.
- For suitable ϵ, C depending on n , we have

$$\left\{ \sum n_i b_i : |n_i| \leq \epsilon X / \lambda_i(L) \right\} \subseteq \{v \in L : \|v\| \leq X\} \subseteq \left\{ \sum n_i b_i : |n_i| \leq CX / \lambda_i(L) \right\}.$$
- $\#\{v \in L : \|v\| \leq X\} \asymp_n \prod_{i=1}^d (1 + X/\lambda_i(L)) \asymp_n \max\{1, \frac{X}{\lambda_1(L)}, \dots, \frac{X^d}{\lambda_1(L)\dots\lambda_d(L)}\}$.

We also have, see e.g. Barroero and Widmer (2018) or Davenport (1951):

Theorem 9. $\#\{v \in L : \|v\| \leq X\} = X^d \frac{\text{Vol}(B(0;1))}{\det L} (1 + O_n(\frac{\lambda_d(L)}{X}))$ if $X > \lambda_d(L)$.

Proof idea: By tiling with translated copies of $\{x_1b_1 + \dots + x_db_d : x_i \in (-\frac{1}{2}, \frac{1}{2})\}$, show that

$$\begin{aligned} \frac{\text{Vol}\{v \in \mathbb{R}^n : \|v\| \leq X - C\lambda_d(L)\}}{\det(L)} &\leq \#\{v \in L : \|v\| \leq X\} \\ &\leq \frac{\text{Vol}\{v \in \mathbb{R}^n : \|v\| \leq X + C\lambda_d(L)\}}{\det(L)}. \end{aligned}$$

3.4 Digression: The LLL Algorithm

Definition 20 (Lovász Condition). The Lovász condition for LLL reduction requires that for basis vectors \mathbf{b}_i and \mathbf{b}_{i+1} :

$$\|\mathbf{b}_{i+1}^* + \mu_{i,i+1}\mathbf{b}_i^*\|^2 \geq \frac{3}{4}\|\mathbf{b}_i^*\|^2.$$

Equivalently,

$$\|\mathbf{b}_{i+1}^*\|^2 \geq (3/4 - \mu_{i,i+1}^2)\|\mathbf{b}_i^*\|^2.$$

- This condition prevents \mathbf{b}_{i+1} from becoming too short after size reduction.
1. The LLL algorithm is named after (Arjen) Lenstra, (Hendrik) Lenstra, and Lovász, who introduced it in 1982.

2. Starts with an initial lattice basis.
3. Size reduction is applied.
4. The Lovász condition is checked to potentially swap adjacent vectors.
5. These steps are repeated until the basis is LLL-reduced.

Theorem 10 (Nguyen’s Corollary 4). *A LLL-reduced basis (b_1, \dots, b_d) satisfies:*

$$\begin{aligned} \|b_1\| &\leq 2^{(d-1)/2} \cdot \det(L)^{1/d}, \\ \|b_i\| &\leq 2^{d-1} \cdot \lambda_i(L), \\ \prod_{i=1}^d \|b_i\| &\leq 2^{d(d-1)/2} \cdot \det(L). \end{aligned}$$

- In a similar way to HKZ-reduced bases above, we can interpret this as making the basis approximately orthogonal in a nice way.
- The LLL algorithm has polynomial-time complexity, making it practical for many applications.
- (That is, it takes a number of operations that is bounded by a polynomial function of the number of bits used to store the b_i , in the case when $L \subseteq \mathbb{Z}^n$.)

3.5 The dual lattice again

Definition 21 (Dual basis). L' has a dual basis defined by $b'_i \cdot b_j = \delta_{ij} = \mathbf{1}_{i=j}$.

If $n = d$, $L = L(b_1, \dots, b_n)$ and B has columns b_i , then b'_i are the columns of B^{-T} .

Lemma 10. *HKZ reduced $\implies \|x_1 b'_1 + \dots + x_d b'_d\| \asymp_n \|x_1 \frac{b_1^*}{\|b_1^*\|^2} + \dots + x_d \frac{b_d^*}{\|b_d^*\|^2}\|$ for all $x \in \mathbb{R}^d$.*

Proof. The dual basis of b_i^* is $b_i^* / \|b_i^*\|^2$. The maps $b_i^{*'} \leftrightarrow b_i$ are the inverse transpose of $b_i^* \leftrightarrow b_i$, with entries $\ll_n 1$. □

Corollary 5. $\lambda_i(L') \asymp_n \|b_{n+1-i}^*\|^{-1} \asymp_n \lambda_{n+1-i}(L)^{-1}$.

3.6 Minima with respect to general norms

Lemma 11. *Let $F : \mathbb{R}^n \rightarrow [0, \infty)$ be a norm. There is $M \in \text{GL}_n(\mathbb{R})$ such that*

$$F(x) \asymp_n \|Mx\|.$$

- Very important to note: the constant only depends on n , not F .

- Another way to say it: every F is $\asymp_n \sqrt{Q}$ for a quadratic form $Q = M^T M$.
- Proof idea: Take m_1 such that $F(m_1)/\|m_1\|$ is maximal, restrict to m_1^\perp , use induction. Show that $F(\lambda m_1 + v) \asymp_n \lambda F(m) + F(v)$ for $v \cdot m_1 = 0$.

Definition 22 (Minima with respect to F). Let $\lambda_j(L, F)$ be minimal such that there are j linearly independent vectors of L with $F(v_1), \dots, F(v_j) \leq \lambda_j(L, F)$.

If $F(x) \asymp_n \|Mx\|$ then $\lambda_j(L, F) \asymp \lambda_j(ML)$, $\det(ML) \asymp \frac{\det(L)}{\text{Vol}_{B_{L,F}(0;1)}}$ where $B_{L,F}(0;1) = \{x \in \text{Span}(L) : F(x) \leq 1\}$.

If $F(x) \asymp_n \|Mx\|$ then $\lambda_j(L, F) \asymp_n \lambda_j(ML)$, $\det(ML) \asymp \frac{\det(L)}{\text{Vol}_{B_{L,F}(0;1)}}$ so we already know:

- $\prod \lambda_i(L, F) \asymp_n \frac{\det(L)}{\text{Vol}_{B_{L,F}(0;1)}}$,
- $\lambda_1(L, F)^n \ll_n \frac{\det(L)}{\text{Vol}_{B_{L,F}(0;1)}}$,
- $\#\{v \in L : F(v) \leq X\} \asymp_n \prod_{i=1}^d (1 + \frac{X}{\lambda_i(L, F)}) \asymp_n \max\{1, \frac{X}{\lambda_1(L, F)}, \dots, \frac{X^d}{\lambda_1(L, F) \dots \lambda_d(L, F)}\}$,

The way to think of the order of magnitude result is that $\log \#\{v \in L : F(v) \leq X\}$ is a piecewise linear, continuous, convex function (“convex up”, i.e. with increasing derivative), with derivative d for all sufficiently large n .

The asymptotic is more delicate. We have $\#\{v \in L : F(v) \leq X\} = X^d \frac{\text{Vol}(B_{L,F}(0;1))}{\det L} (1 + O_M(\frac{\lambda_d(L, F)}{X}))$ if $X > \lambda_d(L, F)$ by using the minima of ML as above. But in practise this isn’t much use, since we often have some M which is varying in a family. Barroero and Widmer (2018) or Davenport (1951) give much more useful versions.

Concerning reduced bases relative to a norm, considering ML shows that there is a basis b_i of L such that

- $F(b_i) \asymp_n \lambda_i(L, F)$,
- $F(x_1 b_1 + \dots + x_d b_d) \asymp_n \max_i |x_i| \lambda_i(L, F)$,
- for any $v \in L$, $v = \sum n_i b_i$ for some $n_i \in \mathbb{Z}$, and then $F(v) \asymp_n \max_i |n_i| \lambda_i(L, F)$,
- $\{\sum n_i b_i : n_i \leq \epsilon \frac{X}{\lambda_i(L, F)}\} \subseteq \{v \in L : F(v) \leq X\} \subseteq \{\sum n_i b_i : n_i \leq C \frac{X}{\lambda_i(L, F)}\}$.

3.7 The classical perspective

The geometry of numbers has its origins as a distinct field in attempts to estimate the minima, and related quantities, as optimally as possible.

Theorem 11 (Minkowski’s First Theorem). $\lambda_1(L, F) \leq 2 \frac{\det(L)}{\text{Vol}(B_{L,F}(0;1))}$

Theorem 12 (Minkowski's Second Theorem).

$$\frac{2^d}{d!} \frac{\det(L)}{\text{Vol}(B_{L,F}(0;1))} \leq \lambda_1(L, F) \cdots \lambda_d(L, F) \leq 2^d \frac{\det(L)}{\text{Vol}(B_{L,F}(0;1))}.$$

The latter result has a generalisation to general non-negative symmetric (i.e. $F(-x) = F(x)$) 1-homogeneous functions F , not necessarily norms. For example $F(x) = \prod_{i=1}^n |L_i(x)|^{1/n}$ for some linear forms L_i .

What about the results for dual bases?

Recall $b'_i \cdot b_i = \delta_{ij}$.

Theorem 13 (Banaszczyk (1993), Theorem 2.1). $1 \leq \lambda_i(L) \lambda_{d+1-i}(L') \leq d$.

Actually there is a version for a general norm F , but one must define the dual norm $F'(u) = \sup\{u \cdot v / F(v) : v \in \text{Span } L\}$.

Note $\|\cdot\|$ is self-dual. One can show the dual norm of $\sqrt{x^T A x}$ is $\sqrt{x^T A^{-1} x}$.

Theorem 14 (Cassels, chapter VIII Theorem VI). $1 \leq \lambda_i(L, F) \lambda_{d+1-i}(L', F') \leq d!$.

The classical idea of a 'reduced' basis involves the dual basis:

Theorem 15 (Riesz/Mahler, Cassels chapter VIII Theorem VII). L has a basis such that

$$\begin{aligned} \|b_1\| &= \lambda_1(L), \\ \|b_i\| &\leq \frac{i}{2} && (i > 1), \\ \|b_i\| \cdot \|b'_i\| &\leq \left(\frac{1}{2}\right)^{d-1} n!^2 && (i \geq 1). \end{aligned}$$

NB: $v = \sum n_i b_i \implies n_i = b'_i \cdot v$.

So $\|b_i\| \cdot \|b'_i\| \ll 1$ is essentially the same as $\|x_1 b_1 + \cdots + x_d b_d\| \asymp_n \max_i |x_i| \cdot \|b_i\|$.

This theorem is actually stated for minima relative to a norm, I just took $F = \|\cdot\|$.

Chapter 4

Orthogonal lattices and an application

We are going to give an asymptotic for

$$Z(N) = \#\{x, y \in \mathbb{Z}^n \setminus \{0\} : x \cdot y = 0, \|x\| \cdot \|y\| \leq N\},$$

for each fixed $n \geq 3$, as $N \rightarrow \infty$. We will set this in the context of Manin's conjecture.

4.1 The orthogonal lattice

Definition 23 (Orthogonal lattice). Given L a primitive, rank $d < n$ sublattice of \mathbb{Z}^n , let $L^\perp = \{v \in \mathbb{Z}^n : v \cdot u = 0 \forall u \in L\}$.

For example if $n = 2$ and $L = L(b)$, then $L^\perp = \{(x, y) \in \mathbb{Z}^2 : b_1x = -b_2y\}$ and so $L^\perp = L\left(\frac{(b_2, -b_1)^T}{\gcd(b)}\right)$.

Lemma 12. Let $\pi_L : \mathbb{R}^n \rightarrow \text{Span}(L)^\perp$ be the orthogonal projection along L . Then $(L^\perp)' = \pi_L(\mathbb{Z}^n)$.

Since $\det(M') = \det(M)^{-1}$, we get:

Corollary 6. L^\perp has rank $n - d$ and $\det(L^\perp) = \det(L)$.

Proof of lemma. We start by observing that $(L^\perp)^\perp \subseteq L$. It suffices to find $n - d$ linearly independent vectors in L^\perp , which can be done for example by Gaussian elimination.

For convenience we put $d' = n - d$.

Let $L^\perp = L(b_1, \dots, b_{d'})$, extend to a basis b_1, \dots, b_n of \mathbb{Z}^n , let b'_i be dual basis of \mathbb{Z}^n .

Beware: $b'_1, \dots, b'_{d'}$ is not the dual basis of $b_1, \dots, b_{d'}$, it is unlikely that $b'_1, \dots, b'_{d'} \in \text{Span}(L^\perp)$.

Idea: Construct two maps to $L(b'_1, \dots, b'_{d'})$ and show they are inverses. It may help to think of $L(b'_1, \dots, b'_{d'})$ as the lattice of integral linear forms on L^\perp .

Observe that $b'_{d'+1}, \dots, b'_n \in (L^\perp)^\perp = L$.

So $\pi_L(\mathbb{Z}^n) = \pi_L(L(b'_1, \dots, b'_{d'}))$. One of our maps will be given by the restriction of π_L to a map $L(b'_1, \dots, b'_{d'}) \rightarrow \pi_L(\mathbb{Z}^n)$.

Define $\phi : (L^\perp)' \rightarrow L(b'_1, \dots, b'_{d'})$ by $v \mapsto \sum(v \cdot b_i) b'_i$.

Note ϕ is a bijection: in particular $v \in (L^\perp)'$ is determined uniquely by $b_i \cdot v$ ($i \leq d'$). The reason is that if M has basis c_i , then M' has basis c'_i with

Idea of application

- $Z(N) = \#\{x, y \in \mathbb{Z}^n \setminus \{0\} : x \cdot y = 0, \|x\| \cdot \|y\| \leq N\}$, let $L = L(x/\gcd(x))$.
- $\#\{y \in L^\perp : \|y\| \leq \frac{N}{\|x\|}\} = \left(\frac{N}{\|x\|}\right)^{n-1} \frac{\text{Vol}(B(0;1))}{\det L} (1 + O_n(\frac{\lambda_{n-1}(L^\perp)}{N/\|x\|}))$ if $\frac{N}{\|x\|} > \lambda_1(L^\perp)$.
- WLOG $\|x\| \leq \|y\|$ so $\|x\| \leq \sqrt{N}$. We will (eventually) prove that most lattices are ‘balanced’, in the sense that $\lambda_i(M) \asymp \det(M)^{1/\text{rank}(M)}$. One has to be careful: for given constants in \asymp , a positive proportion of lattices violate this.
- Taking $X = N/\|x\|$, and assuming the lattice $L(x)^\perp$ is balanced, we need $N/\|x\| > (\|x\|/\gcd(x))^{\frac{1}{n-1}}$, that is $N^{n-1} \gcd(x) > \|x\|^n$. For most $\|x\| \leq \sqrt{N}$ this is true.
- So for $n \geq 3$, we will prove $Z(N)$ behaves like

$$\sum_{x \in \mathbb{Z}^n \setminus \{0\}, \|x\| \leq N} (N/\|x\|)^{n-1} \frac{\gcd(x)}{\|x\|} \sim c_n N^{n-1} \log N$$

for some explicit $c_n > 0$.

- If we took $n = 2$ then we could get $c_2 N (\log N)^2$, but it wouldn’t really use the geometry of numbers in any meaningful way.
- Different proofs were given by Franke-Manin-Tschinkel (89) using L-functions, Thunder (93) using geometry of numbers, Robbiani (01) and Spencer (08) both using the geometry of numbers. Morally speaking we follow Thunder.

4.2 A short introduction to Manin’s conjecture

4.2.1 Rational points

We can understand $x, y \in \mathbb{Z}^n \setminus \{0\} : x \cdot y = 0$ as rational points on a projective variety.

\mathbb{P}^{n-1} has rational points $[x] = \{\lambda x : \lambda \in \mathbb{Q} \setminus \{0\}\}$, parametrised by $x \in \mathbb{Z}^n$, $(x_1, \dots, x_n) = 1$, with height $h(P) = \|x\|$.

Moreover $\mathbb{P}^{n-1} \times \mathbb{P}^{n-1}$ has rational points

$$\{([x], [y]) : x, y \in \mathbb{Z}^n, \gcd(x) = \gcd(y) = 1\},$$

with height $\|x\| \cdot \|y\|$.

The equation $x \cdot y = 0$ defines a hypersurface H in $\mathbb{P}^{n-1} \times \mathbb{P}^{n-1}$, with the number of points of height $\leq N$ given by

$$\begin{aligned} & \frac{1}{4} \#\{x, y \in \mathbb{Z}^n : x \cdot y = 0, \gcd(x) = \gcd(y) = 1, \|x\| \cdot \|y\| \leq N\} \\ &= \frac{1}{4} \sum_{d_1, d_2 \in \mathbb{N}} \mu(d_1) \mu(d_2) Z(N/d_1 d_2). \end{aligned}$$

Here $\mu(d) = \pm 1$ is the Möbius function, and the proof uses the identity $\sum_{d|m} \mu(d) = \mathbf{1}_{|m|=1}$ valid for $m \in \mathbb{Z} \setminus \{0\}$, creating sums over divisors $d_1 | x, d_2 | y$. The factor $1/4$ comes from the fact that for primitive x, x' we have $[x] = [x']$ iff $x = \pm x'$.