Introduction

1 Introduction

In this course, the first mathematical objects we will consider are known as *lattices*. What is a lattice? It is a set of points in n-dimensional space with a periodic structure, such as the one illustrated in Figure 1. Three dimensional lattices occur naturally in crystals, as well as in stacks of oranges. Historically, lattices were investigated since the late 18th century by mathematicians such as Lagrange, Gauss, and later Minkowski.

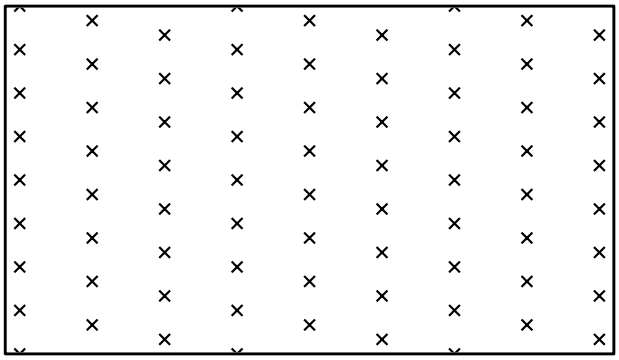


Figure 1: A lattice in $\mathbb{R}^2$

More recently, lattices have become a topic of active research in computer science. Algorithmic problems based on lattices (e.g. Shortest and Closest Vector Problems, ...) has found a wide variety of applications; they have used within optimization algorithms, in the design of wireless communication protocols, and perhaps the most active research area, in the development of secure cryptographic primitives (cryptography) and in establishing the insecurity of certain cryptographic schemes (cryptanalysis). We will cover topics in the all the above areas during this course.

The next mathematical objects we will study are *convex bodies*. A convex set is a subset of n-dimensional space for which the straight line between any two points in the set is fully contained within the set. For the simplest example, the interval $[0, 1]$ in the real line is a bounded convex set. In many important situations, convex sets represent the set of solutions to a mathematical optimization problem (e.g. linear / semidefinite programming), and the study of their geometry has lead to fundamental techniques for solving optimization problems (e.g. simplex method / ellipsoid method). From a slightly different perspective, the study of the metric and volumetric structure of convex bodies has lead to the discovery of many important high dimensional phenomena and inequalities (Brunn-Minkowski inequality, the isoperimetric inequality in $\mathbb{R}^n$, Dvorestky's theorem on spherical sections of convex bodies). We will cover many aspects of this study here.

In his seminal work on the Geometry of Numbers, Minkowski discovered many powerful relationships between the study of convex bodies and the study of lattices that continue to have many applications to this day (e.g. Minkowski's first theorem). Some focal points of this work have been on understanding questions such as "when does a convex set contain a lattice point?"

or "how many lattice points does a convex set contain?". A main area of focus throughout this course will be on understanding and exploring these relationships.

2 Definitions

The following concepts and notation will be useful in the segway. For $n \in \mathbb{N}$, we denote $[n] = \{1, \dots, n\}$. For sets $A, B \subseteq \mathbb{R}^n$, $s, t \in \mathbb{R}$, we define their Minkowski sum $sA + tB = \{s\mathbf{a} + t\mathbf{b} : \mathbf{a} \in A, \mathbf{b} \in B\}$. A set $K \subseteq \mathbb{R}^n$ is convex, if for all $\mathbf{x}, \mathbf{y} \in K$ and $\alpha \in [0,1]$ we have that $\alpha\mathbf{x} + (1-\alpha)\mathbf{y} \in K$. K is a convex body if K is a convex, compact (closed and bounded) set with non-empty interior. K is symmetric (or $\mathbf{0}$-symmetric) if $K = -K$. Lastly, we say that K is $\mathbf{0}$-centered if $\mathbf{0} \in \int(K)$. We write $\mathrm{bd}(K)$ and $\mathrm{int}(K)$ for the boundary and interior of K respectively.

For a linear transformation $T : \mathbb{R}^n \to \mathbb{R}^n$, let $\det(T)$ denote its determinant, and let $\ker(T) = \{\mathbf{x} \in R^n : T\mathbf{x} = \mathbf{0}\}$ denote the kernel of T. For any measurable set $A \subseteq \mathbb{R}^n$, a fundamental fact is that $\mathrm{vol}_n(T(A)) = |\det(T)|\mathrm{vol}_n(A)$, where $\mathrm{vol}_n(\cdot)$ denotes the standard n-dimensional volume (Lebesgue measure). In particular, $\mathrm{vol}_n(sA) = |s|^n\mathrm{vol}_n(A)$ for any $s \in \mathbb{R}$. An affine subspace H is the translation of a linear subspace, i.e. $A = W + \mathbf{t}$ for some vector $\mathbf{t} \in \mathbb{R}^n$ and linear subspace $W \subseteq \mathbb{R}^n$. In this setting, we let $\dim(A)$, the dimension of H, be $\dim(W)$, the dimension of its underlying linear space. For a set $S \subseteq \mathbb{R}^n$, let $\mathrm{span}(S)$ denotes is linear span, the smallest linear subspace containing S, and let $\mathrm{aff}(S)$ denotes is affine hull, the smallest affine subspace containing S, and let $\dim(S) = \dim(\mathrm{aff}(S))$ denote the dimension of its affine hull.

We define the euclidean (or ℓ_2) norm in the usual way by $\|\mathbf{y}\|_2 = \sqrt{\sum_{i=1}^n \mathbf{y}_i^2} = \sqrt{\langle \mathbf{y}, \mathbf{y} \rangle}$, where $\langle \cdot, \cdot \rangle$ is the standard inner product in $\mathbb{R}^n$. We write $B_2^n = \{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x}\|_2 \leq 1\}$ for the unit euclidean ball in $\mathbb{R}^n$, and S^{n-1} denote the unit sphere in $\mathbb{R}^n$. Note that $\{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x}\|_2 \leq t\} = tB_2^n$ for any $t \geq 0$. Using the triangle inequality and homogeneity of the ℓ_2 norm, one can easily check that B_2^n is a symmetric convex body.

Definition 1 (Lattice) *$\mathcal{L} \subseteq \mathbb{R}^n$ is discrete if for any $\mathbf{x} \in \mathcal{L}$ there exists an $\varepsilon > 0$ such that forall $\mathbf{y} \in \mathcal{L}$, $\mathbf{y} \neq \mathbf{x}$, $\|\mathbf{x} - \mathbf{y}\|_2 \geq \varepsilon$. $\mathcal{L}$ is an additive subgroup of $\mathbb{R}^n$ if $\mathbf{0} \in \mathcal{L}$, and for any $\mathbf{x}, \mathbf{y} \in \mathcal{L}$, $\mathbf{x} \pm \mathbf{y} \in \mathcal{L}$. $\mathcal{L}$ is a lattice if it is a discrete additive subgroup of $\mathbb{R}^n$.*

A lattice $\mathcal{L} \subseteq \mathbb{R}^n$ has rank k, $0 \leq k \leq n$, if $\dim(\mathrm{span}(\mathcal{L})) = k$, i.e. if $\mathcal{L}$ spans a k-dimensional linear subspace of $\mathbb{R}^n$. $\mathcal{L}$ has full rank if $k = n$.

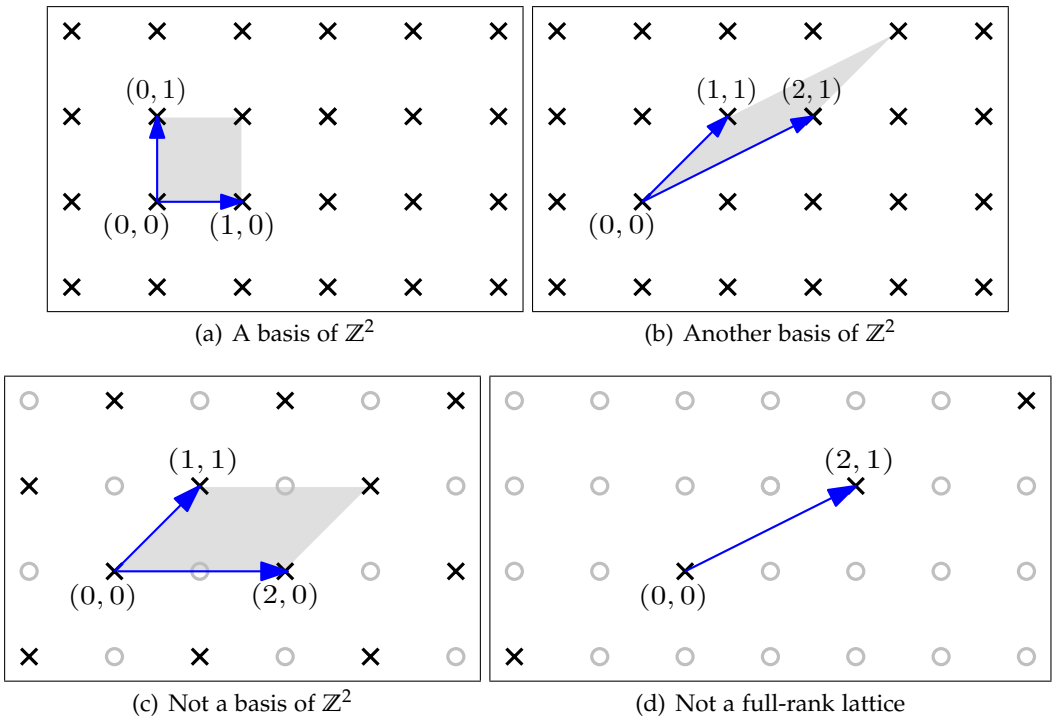
The above definition tell us the essential properties of a lattice, however it does not tell us how to build them. One of the most convenient ways to describe lattices is by giving an explicit set of generators. Another approach, we will become important later in the course, is describe lattices by a set of defining relations. We give some examples below.

To build a rank k lattice in $\mathbb{R}^n$, we may take any set $\mathbf{b}_1, \dots, \mathbf{b}_k \in \mathbb{R}^n$ of *linearly independent* vectors and define

$$\mathcal{L}(\mathbf{b}_1, \dots, \mathbf{b}_k) = \{\sum_{i=1}^k z_i\mathbf{b}_i : z_1, \dots, z_k \in \mathbb{Z}\}.$$

Here we say that $\mathbf{b}_1, \dots, \mathbf{b}_k$ is a basis for the lattice $\mathcal{L} = \mathcal{L}(\mathbf{b}_1, \dots, \mathbf{b}_k)$. We denote the basis matrix for $\mathbf{b}_1, \dots, \mathbf{b}_k$ as $B = (\mathbf{b}_1, \dots, \mathbf{b}_k)$, i.e. the matrix whose columns are $\mathbf{b}_1, \dots, \mathbf{b}_k$. For convenience, we will use the notation $\mathcal{L}(B) = B(\mathbb{Z}^k)$, where we note that $\mathcal{L}(B) = \mathcal{L}$. Furthermore, we will often interchangeably refer to B and $\mathbf{b}_1, \dots, \mathbf{b}_k$ as a basis for $\mathcal{L}$.

For the most basic example, we can take $\mathbb{Z}^2 = \{(x, y) : x, y \in \mathbb{Z}\}$, i.e. the standard integer lattice in 2 dimensions. Here it is easy to see that $\mathbb{Z}^2 = \mathcal{L}(\mathbf{b}_1, \mathbf{b}_2)$ where $\mathbf{b}_1 = (0, 1)$ and $\mathbf{b}_2 = (1, 0)$. Note that $\mathbb{Z}^2$ admits more than one basis, in particular the basis $(0, 1)$, $(1, 1)$ still generates the same lattice. In fact, for any lattice $\mathcal{L}$ of rank $n > 1$ admits *infinitely* many distinct bases.



(a) A basis of $\mathbb{Z}^2$

(b) Another basis of $\mathbb{Z}^2$

(c) Not a basis of $\mathbb{Z}^2$

(d) Not a full-rank lattice

Figure 2: Some lattice bases

We note that it is important that to use a set of linearly independent generators to build the lattice. Indeed, without this restriction one can easily build additive subgroups of $\mathbb{R}^n$ which are not lattices (see Exercise 1). Another useful way to define a lattice without using generators is via a set of defining relations. For example, we can define the lattice

$$\{\mathbf{x} \in \mathbb{Z}^n : \sum_{i=1}^{n} a_i \mathbf{x}_i \equiv 0 \bmod m\}$$

where $a_1, \ldots, a_n \in \mathbb{Z}_m$. Here we take all integers solutions to a single modular equation. For a specific example, we can look at $\{(x, y) \in \mathbb{Z}^n : x + y \equiv 0 (\mathrm{mod} 2)\}$, i.e. all integer points in $\mathbb{Z}^2$ whose coordinate sum is even. A little thought, will reveal that this lattice can in fact be generated by the basis $(1, 1)$ and $(0, 2)$.

3 Basic Properties of Lattices

In this section, we will derive some fundamental properties of lattices, and in particular, show that every lattice admits a basis of generators.

3.1 The Shortest Non-Zero Vector

We begin by defining the following crucial lattice parameter. For a lattice $\mathcal{L} \subseteq \mathbb{R}^n$, let the minimum distance of $\mathcal{L}$ under the euclidean norm be

$$\lambda_1(\mathcal{L}) = \inf_{\mathbf{y} \in \mathcal{L} \setminus \{\mathbf{0}\}} \|\mathbf{y}\|_2.$$

The following lemma establishes some simple equivalences for lattices.

LEMMA 2 *Let $\mathcal{L}$ be a non-trivial additive subgroup of $\mathbb{R}^n$. Then the following are equivalent:*

1. *$\mathcal{L}$ is a lattice.*
2. *$\lambda_1(\mathcal{L}) > 0$.*
3. *$\forall r > 0$, $|\mathcal{L} \cap rB_2^n| < \infty$.*
4. *$\mathcal{L}$ contains a shortest non-zero vector.*

PROOF:

$(1 \Rightarrow 2)$. Since $\mathcal{L}$ is discrete and $\mathbf{0} \in \mathcal{L}$, we know that there exists $\varepsilon > 0$ such that forall $\mathbf{x} \in \mathcal{L}$, $\mathbf{x} \neq \mathbf{0}$, $\|\mathbf{x} - \mathbf{0}\|_2 = \|\mathbf{x}\|_2 \geq \varepsilon$. Furthermore, since $\mathcal{L}$ is non-trivial there exists a non-zero vector. Therefore $\lambda_1(\mathcal{L}) \geq \varepsilon > 0$ as needed.

$(2 \Rightarrow 3)$. Let $\lambda = \lambda_1(\mathcal{L}) > 0$, and let $A = \mathcal{L} \cap rB_2^n$. For any distinct $\mathbf{x}, \mathbf{y} \in A$ note that $\mathbf{x} - \mathbf{y} \in \mathcal{L} \setminus \{\mathbf{0}\}$ and hence $\|\mathbf{x} - \mathbf{y}\|_2 \geq \lambda$. From here, we must have that open balls of radius $\lambda/2$ around $\mathbf{x}$ and $\mathbf{y}$ must be interior disjoint, i.e. $\mathbf{x} + \lambda/2 \int (B_2^n) \cap \mathbf{x} + \lambda/2 \int (B_2^n) = \emptyset$. Furthermore, $\mathbf{x} + \lambda/2 B_2^n \subseteq rB_2^n + \lambda/2 B_2^n = (r + \lambda/2) B_2^n$. If $|\mathcal{L} \cap rB_2^n| = \infty$, then

$$\mathrm{vol}_n((r + \lambda/2) B_2^n)) \geq \sum_{\mathbf{x} \in \mathcal{L} \cap rB_2^n} \mathrm{vol}_n(\lambda/2 B_2^n) = \infty,$$

a clear contradiction.

$(3 \Rightarrow 4)$. Since $\mathcal{L}$ is non-trivial, we can pick $\mathbf{y} \in \mathcal{L} \setminus \{\mathbf{0}\}$. Let $r = \|\mathbf{y}_2\|$, and examine $A = (\mathcal{L} \cap rB_2^n) \setminus \{\mathbf{0}\}$. Notice that non-zero vector shorter than $\mathbf{y}$ must be contained in A. Since $\mathcal{L}$ is discrete we have that $|A| < \infty$. Since A is finite, we can pick $\mathbf{y} \in A$ of minimum ℓ_2 norm. By construction, $\mathbf{y}$ is a shortest-nonzero vector of $\mathcal{L}$, as needed.

$(4 \Rightarrow 1)$. Follows from the proof of $2 \Rightarrow 3$. $\square$

The following exercise helps illustrate certain canonical situations where additive groups are or are not lattices.

EXERCISE 1

1. Let $A = \{x + \alpha y : x, y \in \mathbb{Z}\} \subseteq \mathbb{R}$, where $\alpha > 0$ is irrational. Show that A is not a lattice.
2. Let $\mathbf{v}_1, \ldots, \mathbf{v}_m \in \mathbb{Q}^n$. Show that $\mathcal{L}(\mathbf{v}_1, \ldots, \mathbf{v}_m)$ is a lattice (note that $\mathbf{v}_i$'s need not be linearly independent, in particular m maybe greater than n).

3.2 Gram Schmidt Orthogonalization and Bounds on the Shortest Vector

For a linear subspace $W \subseteq \mathbb{R}^n$, we denote the orthogonal complement of W as $W^\perp = \{\mathbf{x} \in \mathbb{R}^n : \langle \mathbf{x}, \mathbf{y} \rangle = 0 \ \ \forall \mathbf{y} \in W\}$. From standard linear algebra, we know that $W \cap W^\perp = \{\mathbf{0}\}$ and that $W + W^\perp = \mathbb{R}^n$. Let $\pi_W : \mathbb{R}^n \to W$ denote the orthogonal projection on W. More precisely, for $\mathbf{x} \in \mathbb{R}^n$, letting $\mathbf{x} = \mathbf{w} + \mathbf{w}^\perp$ denote the unique decomposition of $\mathbf{x}$ such that $\mathbf{w} \in W$ and $\mathbf{w}^\perp \in W^\perp$, π_W is the map that sends $\mathbf{x}$ to $\mathbf{w}$. Put differently, π_W is the unique linear map having as kernel $W^\perp$ that is the identity on W.

The following exercise establishes some elementary property of orthogonal projections.

Exercise 2 The following holds:

1. For vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, and a linear subspace $W \in \mathbb{R}^n$, $\langle \pi_W(\mathbf{x}), \mathbf{y} \rangle = \langle \mathbf{x}, \pi_W(\mathbf{y}) \rangle$.
2. For a linear subspace $W \subseteq \mathbb{R}^n$ and vector $\mathbf{x} \in \mathbb{R}^n$, $\pi_W(\mathbf{x}) = \arg\min_{\mathbf{z} \in \mathbf{x} - W^\perp} \|\mathbf{z}\|_2 = \arg\min_{\mathbf{w} \in W} \|\mathbf{x} - \mathbf{w}\|_2$.
3. $(W_1 \cap W_2)^\perp = W_1^\perp + W_2^\perp$.
4. For linear subspaces $W_1, W_2 \subseteq \mathbb{R}^n$, such that $W_1^\perp$ and $W_2^\perp$ are orthogonal, then $\pi_{W_1} \circ \pi_{W_2} = \pi_{W_2} \circ \pi_{W_1} = \pi_{W_1 \cap W_2}$.

For a set of linearly independent vectors $\mathbf{b}_1, \ldots, \mathbf{b}_k \in \mathbb{R}^n$, we let $\tilde{\mathbf{b}}_1, \ldots, \tilde{\mathbf{b}}_k$ denote the Gram Schmidt Orthogonalization of $\mathbf{b}_1, \ldots, \mathbf{b}_k$, which we define by the following recursive relation: $\tilde{\mathbf{b}}_1 = \mathbf{b}_1$, and for $2 \leq i \leq k$:

$$\tilde{\mathbf{b}}_i = \mathbf{b}_i - \sum_{j=1}^{i-1} \mu_{ji} \tilde{\mathbf{b}}_j \quad \text{where } \mu_{ji} = \frac{\langle \tilde{\mathbf{b}}_j, \mathbf{b}_i \rangle}{\langle \tilde{\mathbf{b}}_j, \tilde{\mathbf{b}}_j \rangle}. \tag{1}$$

Gram-Schmidt orthogonalization is a basic procedure in linear algebra that takes any set of n linearly independent vectors, and creates a set of n *orthogonal* vectors. It works by projecting each vector on the space orthogonal to the span of the previous vectors. See Figure 3 for an illustration.

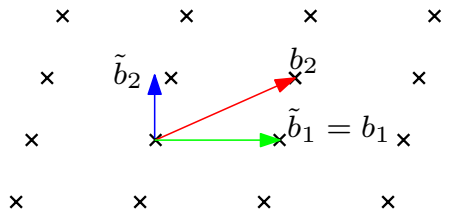


Figure 3: Gram-Schmidt orthogonalization

We note that the order of the vectors is important in the definition of the Gram Schmidt vectors. Changing the order of the vectors of $\mathbf{b}_1, \ldots, \mathbf{b}_k$ will generally change the set of outputed vectors. Furthermore if $\mathbf{b}_1, \ldots, \mathbf{b}_k$ form a basis of a lattice $\mathcal{L}$, the vectors $\tilde{\mathbf{b}}_1, \ldots, \tilde{\mathbf{b}}_k$ may not be lattice vectors (in fact, this is generally the case). Even so, as we will see in Lemma 4, the Gram-Schmidt vectors of a lattice basis provide useful geometric information.

To begin we first establish the fundamental properties of the Gram-Schmidt vectors.

LEMMA 3 *Let $\mathbf{b}_1, \ldots, \mathbf{b}_k \in \mathbb{R}^n$ be linearly independent vectors. For $i \in [k+1]$, let $W_i = \text{span}(\mathbf{b}_1, \ldots, \mathbf{b}_{i-1})$ (where $W_1 = \{\mathbf{0}\}$) and let $\pi_i = \pi_{W_i^\perp}$. Then the following holds:*

1. *For $i \in [k+1]$, $W_{i+1} = \text{span}(\tilde{\mathbf{b}}_1, \ldots, \tilde{\mathbf{b}}_i)$. Furthermore, $\tilde{\mathbf{b}}_1, \ldots, \tilde{\mathbf{b}}_k$ are linearly independent.*
2. *$\langle \tilde{\mathbf{b}}_i, \tilde{\mathbf{b}}_j \rangle = 0$ for all $1 \leq i < j \leq k$.*
3. *For $i \in [k]$, the map $\mathbf{x} \to \sum_{j=1}^{i} \frac{\langle \mathbf{x}, \tilde{\mathbf{b}}_j \rangle}{\langle \tilde{\mathbf{b}}_j, \tilde{\mathbf{b}}_j \rangle} \tilde{\mathbf{b}}_j$ is equal to $\pi_{W_{i+1}}$. Furthermore, the vectors $\frac{\tilde{\mathbf{b}}_1}{\|\tilde{\mathbf{b}}_1\|}, \ldots, \frac{\tilde{\mathbf{b}}_i}{\|\tilde{\mathbf{b}}_i\|_2}$ form an orthornomal basis for W_{i+1}.*
4. *For $i \in [k+1]$, the map $\mathbf{x} \to \mathbf{x} - \sum_{j=1}^{i-1} \frac{\langle \mathbf{x}, \tilde{\mathbf{b}}_j \rangle}{\langle \tilde{\mathbf{b}}_j, \tilde{\mathbf{b}}_j \rangle} \tilde{\mathbf{b}}_j$ is equal to $\pi_{W_i^\perp} = \pi_i$. Furthermore for $i \in [k]$, $\pi_i(\mathbf{b}_i) = \tilde{\mathbf{b}}_i$, and $\pi_i(\mathbf{b}_i), \ldots, \pi_i(\mathbf{b}_k)$ are linearly independent.*
5. *The gram schmidt orthogonalization of $\pi_i(\mathbf{b}_i), \ldots, \pi_i(\mathbf{b}_k)$ is $\tilde{\mathbf{b}}_i, \ldots, \tilde{\mathbf{b}}_k$ for $i \in [k]$.*
6. *The matrix $C_{i,j} = \frac{\langle \mathbf{b}_j, \tilde{\mathbf{b}}_i \rangle}{\|\tilde{\mathbf{b}}_i\|_2}$, $i, j \in [k]$, yields the coordinates of $\mathbf{b}_1, \ldots, \mathbf{b}_k$ under the orthonormal basis $\frac{\tilde{\mathbf{b}}_1}{\|\tilde{\mathbf{b}}_1\|_2}, \ldots, \frac{\tilde{\mathbf{b}}_k}{\|\tilde{\mathbf{b}}_k\|_2}$ of W_{k+1}. Furthermore, C is a non-singular upper triangular matrix satisfying $C_{i,i} = \|\tilde{\mathbf{b}}_i\|_2$ for $i \in [k]$.*

PROOF: To prove 1, note that the each $\tilde{\mathbf{b}}_i$ is a linear combination of $\mathbf{b}_1, \ldots, \mathbf{b}_i$, hence we clearly get that $W_{i+1} = \text{span}(\tilde{\mathbf{b}}_1, \ldots, \tilde{\mathbf{b}}_i) \subseteq \text{span}(\mathbf{b}_1, \ldots, \mathbf{b}_i)$. To show that the spans are equal, we need only show that each $\mathbf{b}_i$ is a linear combination of $\tilde{\mathbf{b}}_i$, but this is obvious from equation 1. The linear independence follows directly from the equality $\text{span}(\tilde{\mathbf{b}}_1, \ldots, \tilde{\mathbf{b}}_k) = \text{span}(\mathbf{b}_1, \ldots, \mathbf{b}_k)$ and the linear independence of the $\mathbf{b}_i$'s.

To prove 2, we prove by induction on l that $\langle \tilde{\mathbf{b}}_i, \tilde{\mathbf{b}}_j \rangle = 0$ for $1 \leq j < i \leq l$. The base case $l = 1$ is trivial, so assume $2 \leq l \leq k$. Given the induction hypothesis, we need only show that $\langle \tilde{\mathbf{b}}_i, \tilde{\mathbf{b}}_j \rangle = 0$ for $1 \leq i < l$. To see this note that

$$\langle \tilde{\mathbf{b}}_i, \tilde{\mathbf{b}}_l \rangle = \langle \tilde{\mathbf{b}}_i, \mathbf{b}_l \rangle - \sum_{j=1}^{l-1} \frac{\langle \tilde{\mathbf{b}}_j, \mathbf{b}_l \rangle}{\langle \tilde{\mathbf{b}}_j, \tilde{\mathbf{b}}_j \rangle} \langle \tilde{\mathbf{b}}_i, \tilde{\mathbf{b}}_j \rangle = \langle \tilde{\mathbf{b}}_i, \mathbf{b}_l \rangle - \frac{\langle \tilde{\mathbf{b}}_i, \mathbf{b}_l \rangle}{\langle \tilde{\mathbf{b}}_i, \tilde{\mathbf{b}}_i \rangle} \langle \tilde{\mathbf{b}}_i, \tilde{\mathbf{b}}_i \rangle = 0$$

as needed.

We prove 3. For $i \in [k+1]$, we let $\pi'(\mathbf{x}) = \sum_{j=1}^{i} \frac{\langle \mathbf{x}, \tilde{\mathbf{b}}_j \rangle}{\langle \tilde{\mathbf{b}}_j, \tilde{\mathbf{b}}_j \rangle} \tilde{\mathbf{b}}_j$. To show that $\pi' = \pi_{W_{i+1}}$ it suffices to show that π' is the identity on W_{i+1} and that $W_{i+1}^\perp \subseteq \ker(\pi')$ (note that this uniquely determines π'). For $\mathbf{x} \in W_i^\perp$, since $\tilde{\mathbf{b}}_j \in W_{i+1}$, for $j \in [i]$, we clearly have $\langle \mathbf{x}, \tilde{\mathbf{b}}_j \rangle = 0$ and hence $\pi'(\mathbf{x}) = \mathbf{0}$. Therefore $W_{i+1}^\perp \subseteq \ker(\pi')$. Take $\mathbf{x} \in W_{i+1}$. Since $\tilde{\mathbf{b}}_1, \ldots, \tilde{\mathbf{b}}_i \in W_{i+1}$, we clearly have that $\mathbf{x} - \pi'(\mathbf{x}) \in W_{i+1}$. Using the orthogonality of $\tilde{\mathbf{b}}_1, \ldots, \tilde{\mathbf{b}}_i$, we see that for $l \in [i]$

$$\begin{aligned} \langle \mathbf{x} - \pi'(\mathbf{x}), \tilde{\mathbf{b}}_l \rangle &= \langle \mathbf{x} - \sum_{j=1}^{i} \frac{\langle \mathbf{x}, \tilde{\mathbf{b}}_j \rangle}{\langle \tilde{\mathbf{b}}_j, \tilde{\mathbf{b}}_j \rangle} \tilde{\mathbf{b}}_j, \tilde{\mathbf{b}}_l \rangle = \langle \mathbf{x}, \tilde{\mathbf{b}}_l \rangle - \frac{\langle \mathbf{x}, \tilde{\mathbf{b}}_l \rangle}{\langle \tilde{\mathbf{b}}_l, \tilde{\mathbf{b}}_l \rangle} \langle \tilde{\mathbf{b}}_l, \tilde{\mathbf{b}}_l \rangle \\ &= \langle \mathbf{x}, \tilde{\mathbf{b}}_l \rangle - \langle \mathbf{x}, \tilde{\mathbf{b}}_l \rangle = 0. \end{aligned}$$

Since $\text{span}(\tilde{\mathbf{b}}_1, \ldots, \tilde{\mathbf{b}}_i) = W_{i+1}$, we see that $\mathbf{x} - \pi'(\mathbf{x}) \in W_{i+1}^\perp$. Since $\mathbf{x} - \pi'(\mathbf{x}) \in W_{i+1} \cap W_{i+1}^\perp = \{\mathbf{0}\}$, we have that $\mathbf{x} = \pi'(\mathbf{x})$, and hence π' is the identity on W_{i+1}, as needed. Let $\bar{\mathbf{b}}_j = \frac{\tilde{\mathbf{b}}_j}{\|\tilde{\mathbf{b}}_j\|}$ for $j \in [i]$. Note that since $\tilde{\mathbf{b}}_1, \ldots, \tilde{\mathbf{b}}_i$ are linearly independent, we have that $\|\tilde{\mathbf{b}}_j\| > 0$, $j \in [i]$, and

hence $\tilde{\mathbf{b}}_1, \ldots, \tilde{\mathbf{b}}_i$ are well defined. By construction, the $\tilde{\mathbf{b}}_j$'s are unit vectors, and since they are non-zero scalings of the $\tilde{\mathbf{b}}_j$'s, they remain orthogonal and have the same span. Hence they form an orthonormal basis of W_{i+1} as needed.

We prove 4. For $i \in [k+1]$, let $\pi'(\mathbf{x}) = \mathbf{x} - \sum_{j=1}^{i-1} \frac{\langle \mathbf{x}, \tilde{\mathbf{b}}_j \rangle}{\langle \tilde{\mathbf{b}}_j, \tilde{\mathbf{b}}_j \rangle}$. By part 3, we know that $\pi'(\mathbf{x}) = \mathbf{x} - \pi_{W_i}(\mathbf{x})$. To show that $\pi' = \pi_i$, it suffices to show that π' is the identity of $W_i^\perp$ and that $W_i \subseteq \ker(\pi')$. Take $\mathbf{x} \in W_i$, then $\pi'(\mathbf{x}) = \mathbf{x} - \pi_{W_i}(\mathbf{x}) = \mathbf{x} - \mathbf{x} = \mathbf{0}$, as needed. Take $\mathbf{x} \in W_i^\perp$. Then $\pi'(\mathbf{x}) = \mathbf{x} - \pi_{W_i}(\mathbf{x}) = \mathbf{x} - \mathbf{0} = \mathbf{x}$, as needed. For the furthermore, since $\pi' = \pi_i$, it follows directly that $\pi_i(\mathbf{b}_i) = \tilde{\mathbf{b}}_i$. We show that $\pi_i(\mathbf{b}_i), \ldots, \pi_i(\mathbf{b}_k)$ are linearly independent. If $\sum_{j=i}^n a_j \pi_i(\mathbf{b}_j) = \mathbf{0}$, then by linearity we have that $\sum_{j=i}^n a_j \mathbf{b}_j \in \ker(\pi_i) = W_{i-1} = \text{span}(\mathbf{b}_1, \ldots, \mathbf{b}_{i-1})$. Since $\mathbf{b}_1, \ldots, \mathbf{b}_k$ are linearly independent, we have that $W_i \cap \text{span}(\mathbf{b}_i, \ldots, \mathbf{b}_k) = \{\mathbf{0}\}$, and hence $\sum_{j=1}^n a_j \mathbf{b}_j = \mathbf{0} \Rightarrow a_i, \ldots, a_k = 0$ as needed.

We prove 5. Let $Z_j = \text{span}(\pi_i(\mathbf{b}_i), \ldots, \pi_i(\mathbf{b}_{i+j-1}))$ for $0 \leq j \leq n-i$ and let τ_j denote the orthogonal projection onto $Z_j^\perp$ (noting that $Z_0 = \{\mathbf{0}\}$). By part 4, we have that gram schmidt orthogonalization of $\pi_i(\mathbf{b}_i), \ldots, \pi_i(\mathbf{b}_k)$ is equal to $\tau_0(\pi_i(\mathbf{b}_i)), \ldots, \tau_{n-i}(\pi_i(\mathbf{b}_k))$. It now suffices to prove $\tau_j \circ \pi_i = \pi_{i+j}$, since then $\tau_j(\pi_i(\mathbf{b}_{i+j})) = \pi_{i+j}(\mathbf{b}_{i+j}) = \tilde{\mathbf{b}}_{i+j}$ as needed. By exercise 2, since Z_j and W_i are orthogonal we have that $\tau_j \circ \pi_i = \pi_{Z_j^\perp \cap W_i^\perp}$. Again by exercise 2, we see that $Z_j^\perp \cap W_i^\perp = (Z_j + W_i)^\perp$. Now note that $W_i + Z_j = \text{span}(\mathbf{b}_1, \ldots, \mathbf{b}_{i-1}, \pi_i(\mathbf{b}_i), \ldots, \pi_i(\mathbf{b}_{i+j-1})) = \text{span}(\mathbf{b}_1, \ldots, \mathbf{b}_{i+j-1}) = W_{i+j}$, as needed.

We prove 6. That C represents the coordinates of $\mathbf{b}_1, \ldots, \mathbf{b}_k$ under $\frac{\tilde{\mathbf{b}}_1}{\|\tilde{\mathbf{b}}_1\|_2}, \ldots, \frac{\tilde{\mathbf{b}}_k}{\|\tilde{\mathbf{b}}_k\|_2}$ follows directly from the fact that the latter is an orthonormal basis of W_{k+1} (part 3). To see that see C is upper triangular, we note that for $i > j$, $\tilde{\mathbf{b}}_i \in W_i^\perp$ and $\mathbf{b}_j \in W_i$, and hence $C_{i,j} = \frac{\langle \mathbf{b}_j, \tilde{\mathbf{b}}_i \rangle}{\|\tilde{\mathbf{b}}_i\|_2} = 0$ as needed. Next, since $\tilde{\mathbf{b}}_i = \pi_i(\mathbf{b}_i)$, and since π_i is an orthogonal projection, we have that

$$C_{i,i} = \frac{\langle \mathbf{b}_i, \tilde{\mathbf{b}}_i \rangle}{\|\tilde{\mathbf{b}}_i\|_2} = \frac{\langle \mathbf{b}_i, \pi_i(\tilde{\mathbf{b}}_i) \rangle}{\|\tilde{\mathbf{b}}_i\|_2} = \frac{\langle \pi_i(\mathbf{b}_i), \tilde{\mathbf{b}}_i \rangle}{\|\tilde{\mathbf{b}}_i\|_2} = \frac{\langle \tilde{\mathbf{b}}_i, \tilde{\mathbf{b}}_i \rangle}{\|\tilde{\mathbf{b}}_i\|_2} = \|\tilde{\mathbf{b}}_i\|_2$$

as needed. That C is non-singular follows from the fact that it is upper triangular and has strictly positive diagonal. □

We now show that if $\mathcal{L}$ is generated by linearly independent vectors, then one can use the basis to get a useful lower bound on the length of the shortest non-zero vector.

LEMMA 4 *Let $\mathcal{L} = \mathcal{L}(\mathbf{b}_1, \ldots, \mathbf{b}_k)$, where $\mathbf{b}_1, \ldots, \mathbf{b}_k \in \mathbb{R}^n$ are linearly independent. Then*

$$\lambda_1(\mathcal{L}) \geq \min_{1 \leq i \leq k} \|\tilde{\mathbf{b}}_i\|_2$$

In particular, $\mathcal{L}$ is a lattice.

PROOF: Take $\mathbf{y} \in \mathcal{L} \setminus \{\mathbf{0}\}$. We express $\mathbf{y}$ as $\sum_{i=1}^k a_i \mathbf{b}_i$ for some $a_1, \ldots, a_k \in \mathbb{Z}$. Since $\mathbf{y} \neq \mathbf{0}$, we have that not all the $a_i's$ are 0. Let $W_i = \text{span}(\mathbf{b}_1, \ldots, \mathbf{b}_{i-1})^\perp$, and let $j \in [k]$ denote the index of the last non-zero a_i. Note that since $a_j \in \mathbb{Z} \setminus \{0\}$, we have that $|a_j| \geq 1$. By exercise 2 and Lemma 3, we have that

$$\|\mathbf{y}\|_2 = \|\sum_{i=1}^j a_i \mathbf{b}_i\|_2 \geq \|\pi_{W_j}(\sum_{i=1}^j a_i \mathbf{b}_i)\|_2 = \|a_i \tilde{\mathbf{b}}_j\|_2 = |a_j| \|\tilde{\mathbf{b}}_j\|_2 \geq \|\tilde{\mathbf{b}}_j\|_2, \geq \min_{1 \leq i \leq k} \|\tilde{\mathbf{b}}_i\|_2,$$

as needed. That $\mathcal{L}$ is a lattice now follows from Lemma 2. □

3.3 Building a Lattice Basis.

THEOREM 5 *Let $\mathcal{L} \subseteq \mathbb{R}^n$ be a $k \geq 1$ dimensional lattice. For any set $\mathbf{y}_1, \ldots, \mathbf{y}_i \in \mathcal{L}$ of linearly independent lattice vectors, there exists linearly independent vectors $\mathbf{b}_1, \ldots, \mathbf{b}_k$ of $\mathcal{L}$, such that for all $j \in [i]$, $\mathcal{L}(\mathbf{b}_1, \ldots, \mathbf{b}_j) = \mathrm{span}(\mathbf{y}_1, \ldots, \mathbf{b}_j) \cap \mathcal{L}$, and $\mathcal{L}(\mathbf{b}_1, \ldots, \mathbf{b}_k) = \mathcal{L}$.*

PROOF: Since $\mathcal{L}$ is k-dimensional, any set of $i < k$ linearly independent vectors $\mathbf{y}_1, \ldots, \mathbf{y}_i \in \mathcal{L}$, can be extended to a set of k linearly independent lattice vectors. Therefore it suffices to prove the statement for $i = k$.

We shall prove the statement by induction on k. Let $\mathcal{L}_1 = \mathrm{span}(\mathbf{y}_1) \cap \mathcal{L}$. Clearly $\mathcal{L}_1$ is a 1 dimensional sublattice of $\mathcal{L}$. By Lemma 2, we can pick a $\mathbf{b}_1 \in \mathcal{L}_1 \setminus \{\mathbf{0}\}$ to be a shortest non-zero vector of $\mathcal{L}_1$.

CLAIM 6 $\mathcal{L}^1 = \mathcal{L}(\mathbf{b}_1)$.

PROOF: Assume not, then there exists $\mathbf{w} \in \mathcal{L}^1$ such that $\mathbf{w} \notin \mathbb{Z}\mathbf{b}_1$. Since $\mathcal{L}^1$ is 1-dimensional and $\mathbf{b}_1$ is non-zero, we have that $\mathbf{w} \in \mathrm{span}(\mathbf{b}_1)$ and hence $\mathbf{w} = \alpha\mathbf{b}_1$ for some $\alpha \in \mathbb{R} \setminus \mathbb{Z}$. Examine $\mathbf{w}' = (\alpha - \lfloor \alpha \rfloor)\mathbf{b}_1$. Since $\lfloor \alpha \rfloor \in \mathbb{Z}$ and $\mathbf{b}_1 \in \mathcal{L}^1$, we must have that $\mathbf{w}' = \mathbf{w} - \lfloor \alpha \rfloor \mathbf{b}_1 \in \mathcal{L}^1$. Since $\alpha \in \mathbb{R} \setminus \mathbb{Z}$, we have that $\alpha - \lfloor \alpha \rfloor \in (0, 1)$ and hence

$$\|\mathbf{b}_1\|_2 > (\alpha - \lfloor \alpha \rfloor)\|\mathbf{b}_1\|_2 = \|\mathbf{w}'\|_2 > 0,$$

a clear contradiction to $\mathbf{b}_1$ being a shortest non-zero vector of $\mathcal{L}^1$. Therefore $\mathcal{L}^1 = \mathbb{Z}\mathbf{b}_1 = \mathcal{L}(\mathbf{b}_1)$ as needed. □

Note that if $k = 1$, the above claim proves the assertion. Hence we may assume that $k \geq 2$.

Let π_2 denote the orthogonal projection onto $W_2 = \mathrm{span}(\mathbf{b}_1)^\perp$, and let $\mathcal{L}_2 = \pi_2(\mathcal{L})$. By linearity of π_2, we have that $\mathcal{L}_2$ is an additive subgroup of $\mathbb{R}^n$. Since $\mathrm{span}(\mathbf{y}_1) = \mathrm{span}(\mathbf{b}_1)$, we clearly have that $\pi_2(\mathbf{y}_1) = \mathbf{0}$. Therefore, we have that $\mathrm{span}(\mathcal{L}_2) = \pi_2(\mathrm{span}(\mathcal{L})) = \pi_2(\mathrm{span}(\mathbf{y}_1, \ldots, \mathbf{y}_k)) = \mathrm{span}(\pi_2(\mathbf{y}_2), \ldots, \pi_2(\mathbf{y}_k))$. Furthermore, since $\pi_2(\mathbf{y}_2), \ldots, \pi_2(\mathbf{y}_k)$ are linearly independent (Lemma 3), we have that $\mathcal{L}_2$ is $k - 1$ dimensional.

CLAIM 7 *$\mathcal{L}_2$ is a lattice.*

PROOF: It suffices to show that $\mathcal{L}_2$ is discrete. In particular, by Lemma 2 it suffices to show that for $r > 0$, $|rB_2^n \cap \mathcal{L}_2| < \infty$. To do this, we will show that each point $\mathbf{x} \in \mathcal{L}_2$ can be lifted to a point $\mathbf{w}_\mathbf{x} \in \mathcal{L}$ of norm bounded by $\|\mathbf{x}\|_2 + \frac{1}{2}\|\mathbf{b}_1\|_2$, such that $\mathbf{w}_\mathbf{x} \neq \mathbf{w}_\mathbf{y}$ for any $\mathbf{y} \in \mathcal{L}_2$, $\mathbf{y} \neq \mathbf{x}$. Assuming this lifting exists, note that if $\mathbf{x} \in \mathcal{L}_2 \cap rB_2^n$, then by our assumption the lifting $\mathbf{w}_\mathbf{x} \in (r + \frac{1}{2}\|\mathbf{b}_1\|_2)B_2^n \cap \mathcal{L}$. Since the lifting is an injective map from $\mathcal{L}_2$ to $\mathcal{L}$ we get that

$$|\mathcal{L}_2 \cap rB_2^n| = |\{\mathbf{x} \in \mathcal{L}_2 : \|\mathbf{x}\|_2 \leq r\}| = |\{\mathbf{w}_\mathbf{x} : \mathbf{x} \in \mathcal{L}_2, \|\mathbf{x}\|_2 \leq r\}| \leq |\mathcal{L} \cap (r + \frac{1}{2}\|\mathbf{b}_1\|_2)B_2^n| < \infty,$$

since $\mathcal{L}$ is a lattice (using Lemma 2).

We now construct the lifting. Take $\mathbf{x} \in \mathcal{L}_2$. By definition of $\mathcal{L}_2$, there exists $\mathbf{y}_\mathbf{x} \in \mathcal{L}$ such that $\pi_2(\mathbf{y}_\mathbf{x}) = \mathbf{x}$. Since $\mathbf{y}_\mathbf{x} - \mathbf{x} \in \mathrm{span}(\mathbf{b}_1)$, we may write $\mathbf{y}_\mathbf{x} - \mathbf{x} = \alpha\mathbf{b}_1$ for some $\alpha \in \mathbb{R}$. Let $\mathbf{w}_\mathbf{x} = \mathbf{y}_\mathbf{x} - \lfloor \alpha \rceil \mathbf{b}_1$ (here $\lfloor \cdot \rceil$ denotes rounding to the nearest integer), and note that $\mathbf{w}_\mathbf{x} \in \mathcal{L}$. Furthermore, by the triangle inequality

$$\|\mathbf{w}_\mathbf{x}\|_2 = \|(\mathbf{w}_\mathbf{x} - \mathbf{x}) + \mathbf{x}\|_2 \leq \|\mathbf{w}_\mathbf{x} - \mathbf{x}\|_2 + \|\mathbf{x}\|_2 = \|(\alpha - \lfloor \alpha \rceil)\mathbf{b}_1\|_2 + \|\mathbf{x}\|_2 < \frac{1}{2}\|\mathbf{b}_1\|_2 + \|\mathbf{x}\|_2,$$

as needed. Lastly, note that the map $\mathbf{x} \to \mathbf{x_w}$ is one to one, since for distinct $\mathbf{x}_1, \mathbf{x}_2 \in \mathcal{L}_2$, we have by construction that $\pi_2(\mathbf{w}_{\mathbf{x}_1}) = \mathbf{x}_1 \neq \mathbf{x}_2 = \pi_2(\mathbf{w}_{\mathbf{x}_2})$. It now follows that $\mathcal{L}_2$ is a lattice, as claimed. □

Since $\mathcal{L}_2$ is a $k-1$ dimensional lattice, and $\pi_2(\mathbf{y}_2), \ldots, \pi_2(\mathbf{y}_k)$ are linearly independent vectors in $\mathcal{L}_2$, by the induction hypothesis there exists a basis $\tilde{\mathbf{b}}_2, \ldots, \tilde{\mathbf{b}}_k$ of $\mathcal{L}_2$ satisfying $\mathcal{L}(\tilde{\mathbf{b}}_2, \ldots, \tilde{\mathbf{b}}_j) = \text{span}(\pi_2(\mathbf{y}_2), \ldots, \pi_2(\mathbf{y}_j)) \cap \mathcal{L}_2$ for all $2 \leq j \leq k$. Let $\mathbf{b}_2, \ldots, \mathbf{b}_k \in \mathcal{L}$ be arbitrary liftings of $\tilde{\mathbf{b}}_2, \ldots, \tilde{\mathbf{b}}_k$ satisfying $\pi_2(\mathbf{b}_j) = \tilde{\mathbf{b}}_j$ for $2 \leq j \leq k$.

We claim that $\mathbf{b}_1, \ldots, \mathbf{b}_k$ satisfy the conditions of the lemma. For j, $1 \leq j \leq k$, pick any $\mathbf{x} \in \text{span}(\mathbf{y}_1, \ldots, \mathbf{y}_j) \cap \mathcal{L}$. Note that $\pi_2(\mathbf{x}) \in \text{span}(\pi_2(\mathbf{y}_1), \ldots, \pi_2(\mathbf{y}_j)) \cap \mathcal{L}_2$, and hence $\pi_2(\mathbf{x}) = \sum_{i=2}^{j} z_i \tilde{\mathbf{b}}_i$ for some $z_2, \ldots, z_j \in \mathbb{Z}$. Let $\hat{\mathbf{x}} = \sum_{i=2}^{j} z_i \mathbf{b}_i$, noting that $\hat{\mathbf{x}} \in \mathcal{L}$. Since by construction, $\pi_2(\hat{\mathbf{x}}) = \sum_{i=2}^{j} z_i \pi_2(\mathbf{b}_i) = \sum_{i=2}^{j} z_i \tilde{\mathbf{b}}_i = \pi_2(\mathbf{x})$, we must have that $\pi_2(\mathbf{x} - \hat{\mathbf{x}}) = \mathbf{0} \Rightarrow \mathbf{x} - \hat{\mathbf{x}} \in \ker(\pi_2) = \text{span}(\mathbf{b}_1)$. Therefore $\mathbf{x} - \hat{\mathbf{x}} \in \mathcal{L} \cap \text{span}(\mathbf{b}_1) = \mathcal{L}^1$, and by Claim 6 we can write $\mathbf{x} - \hat{\mathbf{x}} = z_1 \mathbf{b}_1$ for some $z_1 \in \mathbb{Z}$. Therefore, $\mathbf{x} = z_1 \mathbf{b}_1 + \hat{\mathbf{x}} = \sum_{i=1}^{j} z_i \mathbf{b}_i$ as needed. Lastly, since $\text{span}(\mathbf{b}_1) = \text{span}(\mathbf{y}_1)$ we have that

$$\begin{aligned} \text{span}(\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_j) &= \text{span}(\mathbf{y}_1, \mathbf{b}_2, \ldots, \mathbf{b}_j) = \text{span}(\mathbf{y}_1, \tilde{\mathbf{b}}_2, \ldots, \tilde{\mathbf{b}}_j) \\ &= \text{span}(\mathbf{y}_1, \pi_2(\mathbf{y}_2), \ldots, \pi_2(\mathbf{y}_j)) = \text{span}(\mathbf{y}_1, \ldots, \mathbf{y}_j) \end{aligned}$$

Therefore $\mathcal{L}(\mathbf{b}_1, \ldots, \mathbf{b}_j) \subseteq \text{span}(\mathbf{y}_1, \ldots, \mathbf{y}_j) \cap \mathcal{L}$ as needed. □

Equivalence of Lattice Definitions. As a direct corollary of Lemmas 4 and 5 we get:

COROLLARY 8 *Let $\mathcal{L} \subseteq \mathbb{R}^n$ be a k dimensional additive subgroup of $\mathbb{R}^n$. Then the following are equivalent:*

1. *$\mathcal{L}$ is discrete.*
2. *$\mathcal{L} = \mathcal{L}(\mathbf{b}_1, \ldots, \mathbf{b}_k)$ for some linearly independent $\mathbf{b}_1, \ldots, \mathbf{b}_k \in \mathbb{R}^n$.*

Equivalent lattice bases. From Lemma 5, we see that in fact a lattice can have many equivalent bases. A first question we ask is given two lattices bases B_1, B_2, when is it that $\mathcal{L}(B_1) = \mathcal{L}(B_2)$? In this section, we derive the basic relationship between equivalent lattices bases.

DEFINITION 9 (UNIMODULAR MATRIX) *A matrix $U \in \mathbb{Z}^{n \times n}$ is unimodular if $\det(U) = \pm 1$.*

For example, the matrix

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

is unimodular. The following lemma tells us that the inverse of a unimodular matrix is also unimodular (so it follows that the set of unimodular matrices forms a group under matrix multiplication).

LEMMA 10 *$U \in \mathbb{Z}^{n \times n}$ is unimodular iff $U^{-1} \in \mathbb{Z}^{n \times n}$.*

PROOF: Assume U is unimodular. Let $M_{ij} \in \mathbb{Z}^{(n-1) \times (n-1)}$ denote the principal minor of U obtained by deleting the i^{th} row and j^{th} column. Then by Cramer's rule, we know that $U^{-1}_{ij} = -1^{i+j} \det(M_{ij}) / \det(U)$. Since the determinant of an integer matrix is an integer, and since $\det(U) = \pm 1$, we have that $U^{-1} \in \mathbb{Z}^{n \times n}$ as needed.

Assume $U^{-1} \in \mathbb{Z}^{n\times n}$. Then note that $1 = \det(I) = \det(UU^{-1}) = \det(U)\det(U^{-1})$. Since both U and U^{-1} are integer matrices, we know that both $\det(U), \det(U^{-1}) \in \mathbb{Z}$. Since the only integers dividing 1 are ± 1, we must have that $\det(U) = \pm 1$ as needed. □

The next lemma tells us the two lattices bases generate the same lattice if and only if they are related by a unimodular transformation.

LEMMA 11 *For non-singular matrices $B_1, B_2 \in \mathbb{R}^{n\times k}$, $\mathcal{L}(B_1) = \mathcal{L}(B_2)$ iff and only if $B_1 = B_2U$ for some unimodular matrix $U \in \mathbb{Z}^{k\times k}$.*

PROOF: Assume $\mathcal{L}(B_1) = \mathcal{L}(B_2)$. Since each column of B_1 is in $\mathcal{L}(B_2)$, we can write $B_1 = B_2U$ for some $U \in \mathbb{Z}^{k\times k}$. Similarly, we get that $B_1 = B_2U'$, for some $U' \in \mathbb{Z}^{k\times k}$. Hence we get that $B_1 = B_2U = B_1U'U$. Since B_1 is non-singular, $B_1U'U = B_1 \Leftrightarrow U'U = I_k$, where I_k is the $k \times k$ identity. Hence U is unimodular as needed.

Assume $B_1 = B_2U$ for some unimodular matrix $U \in \mathbb{Z}^{k\times k}$. Clearly $\mathcal{L}(B_1) \subseteq \mathcal{L}(B_2)$. Since U is unimodular, $B_2 = B_1U^{-1}$, where $U^{-1} \in \mathbb{Z}^{k\times k}$, and hence $\mathcal{L}(B_2) \subseteq \mathcal{L}(B_1)$, as needed. □

As an immediate corollary, we obtain that B is a basis of $\mathbb{Z}^n$ if and only if it is unimodular (verify this with the examples in Figure 2).

The following lemma, which we give as an exercise, provides a different way to check whether two lattice bases are equivalent.

EXERCISE 3 Two bases are equivalent if and only if one can be obtained from the other by the following operations on columns:

1. $\mathbf{b}_i \leftarrow \mathbf{b}_i + k\mathbf{b}_j$ for some $k \in \mathbb{Z}$,
2. $\mathbf{b}_i \leftrightarrow \mathbf{b}_j$,
3. $\mathbf{b}_i \leftarrow -\mathbf{b}_i$.