**CSE 206A: Lattice Algorithms and Applications** Fall 2017

# Point Lattices

Instructor: *Daniele Micciancio* UCSD CSE

Lattices are regular arrangements of points in Euclidean space. The simplest example of lattice in $n$-dimensional space is $\mathbb{Z}^n$, the set of all $n$-dimensional vectors with integer entries. (See Figure 1, left.) More generally, a lattice is the result of applying an *injective*[1] linear transformation $\mathbf{B}\colon \mathbb{R}^n \to \mathbb{R}^d$ to the integer lattice $\mathbb{Z}^n$, to obtain the set $\mathcal{L}(\mathbf{B}) = \mathbf{B}(\mathbb{Z}^n) = \{\mathbf{Bx}\colon \mathbf{x} \in \mathbb{Z}^n\}$. (See Figure 1, right.) Equivalently, the lattice $\mathcal{L}(\mathbf{B})$ can be described as the set of all integer linear combinations $\sum_i \mathbf{b}_i x_i$ (with $x_i \in \mathbb{Z}$) of the columns of $\mathbf{B} = [\mathbf{b}_1, \ldots, \mathbf{b}_n]$.

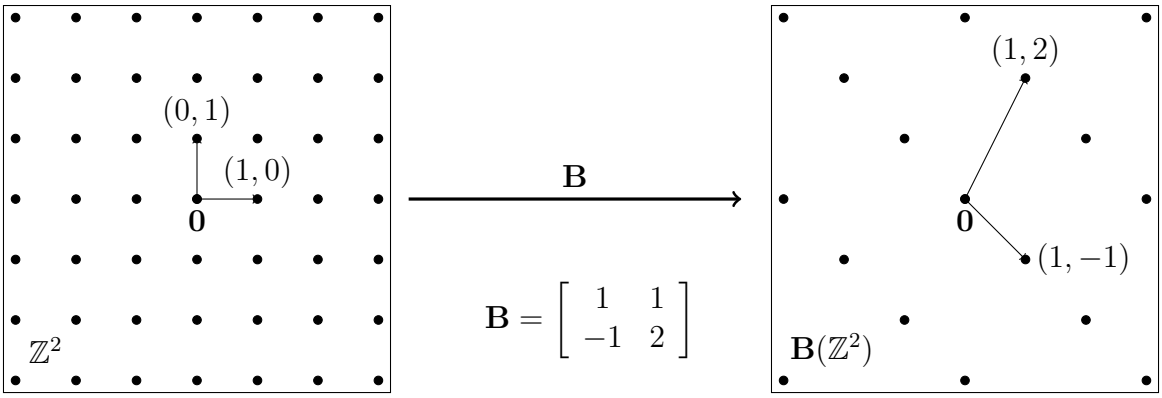



Figure 1: The integer lattice $\mathbb{Z}^2$, and the two dimensional lattice $\mathbf{B}(\mathbb{Z}^2)$.

Despite the simplicity of their definition, lattices are powerful mathematical objects that allow to apply geometric techniques to the solution of hard combinatorial problems. Lattices naturally occur in many settings, like crystallography, communication theory, (algebraic) number theory, etc. They have many applications in computer science and mathematics, including the solution of integer programming problems, diophantine approximation, cryptanalysis, the design of error correcting codes for multi antenna systems, and many more. Recently, lattices have also attracted much attention as a source of computational hardness for the design of secure cryptographic functions, and they are a powerful tool for the construction of advanced cryptographic primitives, like fully homomorphic encryption. This course offers an introduction to lattices, and their relation to other areas of computer science, like algorithms, computational complexity and cryptography. We begin with the definition of lattices and their most important mathematical properties.

[1] We recall that if the linear transformation is represented as a matrix $\mathbf{B} \in \mathbb{R}^{d \times n}$, the injectivity of $\mathbf{B}$ is equivalent to requiring the columns of $\mathbf{B}$ to be linearly independent.

# 1 Point Lattices

**Definition 1** *Let* $\mathbf{B} = [\mathbf{b}_1, \ldots, \mathbf{b}_n] \in \mathbb{R}^{d \times n}$ *be linearly independent vectors in* $\mathbb{R}^d$*. The lattice generated by* $\mathbf{B}$ *is the set*

$$\mathcal{L}(\mathbf{B}) = \{\mathbf{B}\mathbf{x} \colon \mathbf{x} \in \mathbb{Z}^n\} = \left\{ \sum_{i=1}^{n} x_i \cdot \mathbf{b}_i \colon \forall i. x_i \in \mathbb{Z} \right\}$$

*of all the* integer *linear combinations of the columns of* $\mathbf{B}$*. The matrix* $\mathbf{B}$ *is called a* basis *for the lattice* $\mathcal{L}(\mathbf{B})$*. The integer* $n$ *is called the* rank *or* dimension[2] *of the lattice. If* $n = d$ *then* $\mathcal{L}(\mathbf{B})$ *is called a* full rank *or* full dimensional *lattice in* $\mathbb{R}^d$*.*

Definition 1 gives a simple way to represent a lattice (which is an infinite set of points) by a finite object: lattices can be represented by a basis matrix $\mathbf{B}$. For example, the 2-dimensional integer lattice $\mathbb{Z}^2$ (Figure 1, left) is represented by the basis $\mathbf{I} = [\mathbf{e}_1, \mathbf{e}_2]$ where $\mathbf{e}_1 = (1, 0), \mathbf{e}_2 = (0, 1)$ are the standard unit vectors, while the lattice $\mathbf{B}(\mathbb{Z}^2) = \mathcal{L}(\mathbf{B})$ (Figure 1, right) is represented by the basis $\mathbf{B} = \begin{bmatrix} 1 & 1 \\ -1 & 2 \end{bmatrix}$. In computer science applications, the basis matrix typically has integer or rational entries, and can be easily represented as an array of integers.

Notice the similarity between the definition of a lattice $\mathcal{L}(\mathbf{B}) = \{\mathbf{B}\mathbf{x} \colon \mathbf{x} \in \mathbb{Z}^n\}$ and the definition of vector space generated by $\mathbf{B}$:

$$\text{span}(\mathbf{B}) = \{\mathbf{B}\mathbf{x} \colon \mathbf{x} \in \mathbb{R}^n\}.$$

The difference is that in a vector space you can combine the columns of $\mathbf{B}$ with arbitrary real coefficients, while in a lattice only integer coefficients are allowed, resulting in a discrete set of points. Notice that, since vectors $\mathbf{b}_1, \ldots, \mathbf{b}_n$ are linearly independent, any point $\mathbf{y} \in \text{span}(\mathbf{B})$ can be written as a linear combination $\mathbf{y} = \mathbf{B}\mathbf{x}$ with $\mathbf{x} \in \mathbb{R}^n$ in a unique way. Therefore $\mathbf{B}\mathbf{x} \in \mathcal{L}(\mathbf{B})$ if and only if $\mathbf{x} \in \mathbb{Z}^n$.

If $\mathbf{B}$ is a basis for the lattice $\mathcal{L}(\mathbf{B})$, then it is also a basis for the vector space $\text{span}(\mathbf{B})$ spanned by the lattice. However, not every set of lattice vectors $\mathbf{C} \subseteq \mathcal{L}(\mathbf{B})$ which is basis for the vector space $\text{span}(\mathbf{B})$ is a lattice basis for $\mathcal{L}(\mathbf{B})$. For example $2\mathbf{B} = [2\mathbf{b}_1, \ldots, 2\mathbf{b}_n]$ is a basis for $\text{span}(\mathbf{B})$ as a vector space, but it is not a lattice basis for $\Lambda = \mathcal{L}(\mathbf{B})$ because it only generates the sublattice $2\Lambda = \{2\mathbf{v} \colon \mathbf{v} \in \Lambda\} \subset \Lambda$. Another example can be seen in Figure 1 (right), where the matrix $\mathbf{B} \in \mathbb{Z}^{2\times 2}$ is a basis for $\mathbb{R}^2$ (the vector space spanned by the integer lattice $\mathbb{Z}^2$), but it only generates a sublattice of $\mathbb{Z}^2$ because the unit vectors $\mathbf{e}_1 = (1, 0)$ and $\mathbf{e}_2 = (0, 1)$ do not belong to $\mathcal{L}(\mathbf{B}) = \mathbf{B}(\mathbb{Z}^2)$.

For any lattice $\Lambda$, a *sublattice* of $\Lambda$ is a subset $\Lambda' \subseteq \Lambda$ which is itself a lattice, i.e., $\Lambda' = \mathcal{L}(\mathbf{B})$ for some matrix $\mathbf{B}$ with linearly indepedent columns. For example, the lattice $\mathbf{B}(\mathbb{Z}^2)$ in Figure 1 is a sublattice of $\mathbb{Z}^2$ because all of its (basis) vectors have integer coordinates. Sublattices are easily characterized in terms of their bases.

[2]Sometime, the term "dimension" is also used to refer to the number of coordinates of the lattice vectors, i.e., the dimension $d$ of the Euclidean space $\mathbb{R}^d$ containing the lattice.

**Exercise 1** *Show that for any two lattice bases* $\mathbf{B} \in \mathbb{R}^{d\times k}$ *and* $\mathbf{C} \in \mathbb{R}^{d\times n}$*, the first generates a sublattice of the second (*$\mathcal{L}(\mathbf{B}) \subseteq \mathcal{L}(\mathbf{C})$*) if and only if there is an integer matrix* $\mathbf{U} \in \mathbb{Z}^{n\times k}$ *such that* $\mathbf{B} = \mathbf{CU}$*.*

Two important special cases are *full rank sublattices*, i.e., sublattices $\Lambda' \subseteq \Lambda$ that have the same rank as $\Lambda$, and *full sublattices*, i.e., sublattices $\Lambda' = \Lambda \cap \mathrm{span}(\Lambda')$ obtained as the intersection of the lattice with a linear subspace. Clearly, the only sublattice of $\Lambda$ which is both full and full rank, is $\Lambda$ itself.

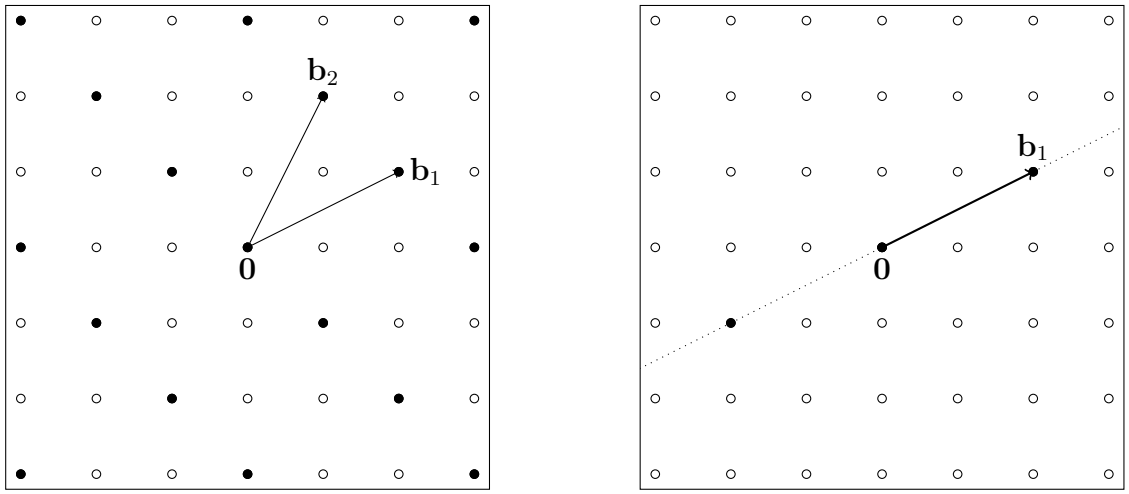


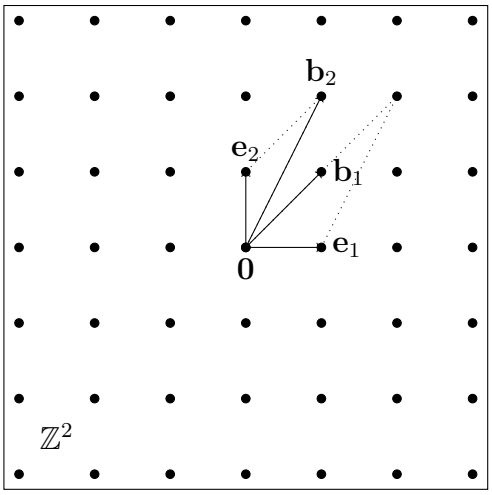

Figure 2: Two sublattices of $\mathbb{Z}^2$. $\mathcal{L}(\mathbf{B})$, on the left, is a *full rank* sublattice because it is also 2-dimensional. $\mathcal{L}(\mathbf{b}_1)$ on the right is a *full* sublattice of $\mathbb{Z}^2$ because it contains all lattice points in its linear span. On the other hand $\mathcal{L}(\mathbf{B})$ is not a full sublattice of $\mathbb{Z}^2$.

# 2 Bases

The same lattice can be represented by several different bases. For example, the integer lattice $\mathbb{Z}^2$ is also generated by the basis $\mathbf{B} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}$. (See Figure 3.)

To a large extent, most algorithmic problems on lattices reduce to the problem of transforming an arbitrary basis for a lattice into another basis for the same lattice with some special properties. So, in order to study lattice problems, we need first to get a good understanding of how lattice bases can be modified without changing the lattice they generate. Different bases of the same lattice are related by invertible integer transformations.

**Definition 2** *A integer square matrix* $\mathbf{U} \in \mathbb{Z}^{n\times n}$ *is* invertible *if there is some other matrix* $\mathbf{V} \in \mathbb{Z}^{n\times n}$ *(the inverse of* $\mathbf{U}$*) such that* $\mathbf{VU} = \mathbf{UV} = \mathbf{I}$*.*




$$[\mathbf{e}_1, \mathbf{e}_2] = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \qquad [\mathbf{b}_1, \mathbf{b}_2] = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}$$

$$\mathbf{e}_1 = 2\mathbf{b}_1 - \mathbf{b}_2$$

$$\mathbf{e}_2 = \mathbf{b}_2 - \mathbf{b}_1$$

Figure 3: $[\mathbf{e}_1, \mathbf{e}_2]$ and $[\mathbf{b}_1, \mathbf{b}_2]$ are two different bases for the integer lattice $\mathbb{Z}^2$. Clearly, $\mathcal{L}([\mathbf{b}_1, \mathbf{b}_2]) \subseteq \mathbb{Z}^2$ because $\mathbf{b}_1, \mathbf{b}_2$ are integer vectors. We also have $\mathbb{Z}^2 \subseteq \mathcal{L}([\mathbf{b}_1, \mathbf{b}_2])$ becaue both unit vectore $\mathbf{e}_1, \mathbf{e}_2$ can be expressed as integer linear combinations of $\mathbf{b}_1, \mathbf{b}_2$.

The inverse matrix is necessarily unique, and it is denoted $\mathbf{U}^{-1}$. The set of invertible $n \times n$ integer matrices is denoted $\mathrm{GL}(n, \mathbb{Z})$, which stands for "general linear group". It is easy to check that $\mathrm{GL}(n, \mathbb{Z})$ is indeed a group under matrix multiplication.

**Exercise 2** *Prove that $GL(n, \mathbb{Z})$ is a group with respect to matrix multiplication with the identity matrix $\mathbf{I}$ as its neutral element, i.e., for all $\mathbf{U}$,$\mathbf{V} \in GL(n, \mathbb{Z})$, the matrices $\mathbf{U} \cdot \mathbf{V}$ and $\mathbf{U}^{-1}$ are also in $GL(n, \mathbb{Z})$.*

**Theorem 3** *Let $\mathbf{B}$ and $\mathbf{C}$ be two lattice bases. Then $\mathcal{L}(\mathbf{B}) = \mathcal{L}(\mathbf{C})$ if and only if there exists an invertible matrix $\mathbf{U} \in GL(n, \mathbb{Z})$ such that $\mathbf{B} = \mathbf{CU}$.*

*Proof.* First assume $\mathbf{B} = \mathbf{CU}$ for some integer matrix $\mathbf{U} \in \mathrm{GL}(n, \mathbb{Z})$. By assumption, there is also an integer matrix $\mathbf{U}^{-1}$ such that $\mathbf{C} = \mathbf{BU}^{-1}$. It follows from Exercise 1 that

$$\mathcal{L}(\mathbf{B}) = \mathcal{L}(\mathbf{CU}) \subseteq \mathcal{L}(\mathbf{C}) = \mathcal{L}(\mathbf{BU}^{-1}) \subseteq \mathcal{L}(\mathbf{B}).$$

This proves that $\mathcal{L}(\mathbf{B}) = \mathcal{L}(\mathbf{C})$, i.e., the two bases generate the same lattice.

Now assume $\mathbf{B}$ and $\mathbf{C}$ are two bases for the same lattice. Then, again by Exercise 1, there exist integer matrices $\mathbf{V}$ and $\mathbf{U}$ such that $\mathbf{B} = \mathbf{CU}$ and $\mathbf{C} = \mathbf{BV}$. Combining these two equations we get $\mathbf{B} = \mathbf{CU} = \mathbf{BVU}$, or equivalently, $\mathbf{B}(\mathbf{I} - \mathbf{VU}) = \mathbf{O}$. Since $\mathbf{B}$ is injective, it must be $\mathbf{I} - \mathbf{VU} = \mathbf{O}$, i.e., $\mathbf{VU} = \mathbf{I}$. The proof that $\mathbf{UV} = \mathbf{I}$ is similar. This proves that $\mathbf{U} \in \mathrm{GL}(n, \mathbb{Z})$ is an invertible matrix, with inverse $\mathbf{U}^{-1} = \mathbf{V}$. □

Theorem 3 shows that invertible integer matrices $\mathbf{U}$ can be used to transform a lattice basis $\mathbf{B}$ into any other basis for the same lattice. In practice, it is often easier or more convenient to transform $\mathbf{B}$ into a different basis through a sequence of simpler "local" operations.

**Definition 4** *An elementary (integer) column operation on a matrix* $\mathbf{B} \in \mathbb{R}^{d\times n}$ *is one of the following:*

SWAP*(i,j):* *Exchange two basis vectors* $(\mathbf{b}_i, \mathbf{b}_j) \leftarrow (\mathbf{b}_j, \mathbf{b}_i)$ *for any* $i \neq j$.

INVERT*(i):* *Change the sign of a basis vector* $\mathbf{b}_i \leftarrow (-\mathbf{b}_i)$ *for any* $i$.

ADD*(i,c,j):* *Add an integer multiple of a basis vector to another* $\mathbf{b}_i \leftarrow (\mathbf{b}_i + c \cdot \mathbf{b}_j)$ *for any* $i \neq j$ *and* $c \in \mathbb{Z}$.

Notice that elementary column operations $\sigma$ act on the right of matrices, so that for any matrices $\mathbf{B}$ and $\mathbf{A}$, $\sigma(\mathbf{B} \cdot \mathbf{A}) = \mathbf{B} \cdot \sigma(\mathbf{A})$. In particular, any elementary column operation $\sigma$ corresponds to right multiplication by an integer matrix $\sigma(\mathbf{I}) \in \mathbb{Z}^{n\times n}$ because

$$\sigma(\mathbf{B}) = \sigma(\mathbf{B} \cdot \mathbf{I}) = \mathbf{B} \cdot \sigma(\mathbf{I}).$$

It is easy to see that elementary column operations do not change the lattice generated by the basis because they are invertible.

**Exercise 3** *Show that each elementary column operation* $\sigma$ *(*SWAP*(i,j),* INVERT*(i) or* ADD*(i,c,j) for* $c \in \mathbb{Z}$ *and* $i, j \in \{1, ..., n\}$, $i \neq j$*) is invertible, and its inverse is also an elementary column operation. Conclude that for any elementary column operation* $\sigma$*, the corresponding matrix is invertible, i.e.,* $\sigma(\mathbf{I}) \in GL(n, \mathbb{Z})$.

It follows from Exercises 2 and 3 that any *sequence* of elementary integer column operations $\sigma = [\sigma_1, \ldots, \sigma_k]$ can be expressed as right multiplication by a invertible matrix $\sigma(\mathbf{I}) = \sigma_1(\mathbf{I}) \cdot \sigma_2(\mathbf{I}) \cdots \sigma_k(\mathbf{I}) \in \mathrm{GL}(n, \mathbb{Z})$. In particular, by Theorem 3, any sequence of elementary column operations $\sigma$ turns a lattice basis $\mathbf{B}$ into an equivalent basis $\sigma(\mathbf{B})$ for the same lattice $\mathcal{L}(\mathbf{B}) = \mathcal{L}(\sigma(\mathbf{B}))$. At this point, it is natural to ask: can any invertible matrix $\mathbf{U} \in \mathrm{GL}(n, \mathbb{Z})$ be expressed as a sequence of elementary column operations? Equivalently, can any basis of a lattice $\mathcal{L}(\mathbf{B})$ be obtained by applying a sequence of elementary column operations to $\mathbf{B}$? As we will see, the answer is yes. As a first step, we show that invertible matrices must have unit determinant.

**Definition 5** *A matrix* $\mathbf{B} \in \mathbb{Z}^{n\times n}$ *is* unimodular *if* $|\det(\mathbf{B})| = 1$.

**Lemma 6** *Any invertible matrix* $\mathbf{U} \in GL(n, \mathbb{Z})$ *is unimodular.*

*Proof.* Recall that $\det(\mathbf{AB}) = \det(\mathbf{A}) \det(\mathbf{B})$ for any square matrices $\mathbf{A}, \mathbf{B}$. Therefore we have $\det(\mathbf{U}) \cdot \det(\mathbf{U}^{-1}) = \det(\mathbf{U}\mathbf{U}^{-1}) = \det(\mathbf{I}) = 1$. Since $\mathbf{U} \in \mathrm{GL}(n, \mathbb{Z})$, $\mathbf{U}^{-1}$ is also an integer matrix, and $\det(\mathbf{U}), \det(\mathbf{U}^{-1})$ are both integers. But for their product to be 1, it must be either $\det(\mathbf{U}) = \det(\mathbf{U}^{-1}) = 1$ or $\det(\mathbf{U}) = \det(\mathbf{U}^{-1}) = -1$. □

At this point we have shown that elementary column operations define invertible matrices, and that invertible matrices are unimodular. In order to close the circle, and prove that all

these properties are equivalent, we need to show that any unimodular matrix can be expressed as a sequence of elementary column operations. To this end, it is useful to define the Hermite normal form (HNF) of an integer matrix, which will also find more applications later in the course. Here we define the HNF only for square nonsingular integer matrices, as this is all we need for now. In pictures, an HNF matrix is an upper triangular matrix

$$\mathbf{H} = \begin{bmatrix} h_{1,1} & \cdots & \cdots & \cdots & \cdots & \cdots \\ & \ddots & & & & \\ & & h_{i,i} & \cdots & h_{i,j} & \cdots \\ & & & \ddots & & \\ & & & & \ddots & \\ & & & & & h_{n,n} \end{bmatrix}$$

with positive diagonal elements $h_{i,i} > 0$, and all other nonzero entries $0 \leq h_{i,j} < h_{i,i}$ reduced modulo the corresponding diagonal element $h_{i,i}$ on the same row.

**Definition 7** *A nonsingular square matrix $\mathbf{H} \in \mathbb{R}^{n \times n}$ is in Hermite normal form (HNF) iff*

1. *$h_{i,j} = 0$ for all $i > j$, i.e., $\mathbf{H}$ is upper triangular*
2. *$h_{i,i} > 0$ for all $i$, i.e., the diagonal elements are positive*
3. *$0 \leq h_{i,j} < h_{i,i}$ for all $i < j$, i.e., the remaining elements are reduced modulo the diagonal elements on the same row.*

We will show that every nonsingular square integer matrix can be put in HNF using elementary column operations. The core of the method is the following generalization of Euclid's algorithm to compute the greatest common divisor of two integers. Recall the (centered) Euclidean algorithm for gcd computation:

$$\text{GCD}(a, b) = \begin{cases} |a| & \text{if } b = 0 \\ \text{GCD}(b, a) & \text{if } |a| < |b| \\ \text{GCD}(a - cb, b) & \text{if } 0 < |b| \leq |a|,\ c = \lfloor a/b \rceil \end{cases}$$

The algorithm performs elementary operations SWAP(1,2) and ADD(1,-c,2) on a pair of integers $(a, b)$, possibly followed by an INVERT(1) operation. The sequence of elementary column operations transforms the initial input $(a, b)$ into $(g, 0)$ where $g = \gcd(a, b)$.

**Exercise 4** *Show that for any nonzero integer vector $\mathbf{v} = [v_1, \ldots, v_n]$, there is a sequence of elementary column operations $\sigma$ such that $\sigma(\mathbf{v}) = [0, \ldots, 0, g]$ where $g = \gcd(v_1, \ldots, v_n) > 0$.* [Hint: Use Euclid's gcd algorithm for $n = 2$, and proceed by induction.]

**Theorem 8** *For any nonsingular square integer matrix $\mathbf{B} \in \mathbb{Z}^{n \times n}$ there is a sequence of elementary integer column operations $\sigma$ such that $\sigma(\mathbf{B})$ is in HNF.*

*Proof.* Let $\sigma$ be a sequence of elementary operations from Exercise 4 which, when applied to the last row of $\mathbf{B}$ gives a vector of the form $[0, \ldots, 0, d_n]$ with $d_n \geq 0$. Apply $\sigma$ to the whole matrix to yield

$$\sigma(\mathbf{B}) = \left[\begin{array}{c|c} \mathbf{B}' & \mathbf{b} \\ \mathbf{0}^\top & d_n \end{array}\right].$$

Notice that since $\mathbf{B}$ is nonsingular, $\det(\mathbf{B}) = \pm \det(\mathbf{B}') \cdot d_n \neq 0$, and it must be $d_n > 0$ and $\det(\mathbf{B}') \neq 0$. By induction, there is a sequence of elementary column operations $\sigma'$ such that $\sigma'(\mathbf{B}') = \mathbf{H}'$ is in HNF. Applying this sequence of operations to $\sigma(\mathbf{B})$ one gets

$$\sigma'(\sigma(\mathbf{B})) = \left[\begin{array}{cc} \mathbf{H}' & \mathbf{b} \\ \mathbf{0}^\top & d_n \end{array}\right].$$

At this point, we have a matrix that satisfies all properties of the HNF definition, except that the entries of $\mathbf{b}$ may not be reduced modulo the corresponding diagonal elements. Let $d_1, \ldots, d_{n-1} > 0$ be the diagonal elements of $\sigma'(\mathbf{B})$, and $\mathbf{b} = (b_1, \ldots, b_{n-1})$. In order to satisfy this last property, we start by reducing $b_{n-1}$ modulo $d_{n-1}$ by adding an appropriate multiple of the last column of $\mathbf{H}'$, and proceed similarly for the other entries $b_i$. More specifically, for $i = n-1, n-2, \ldots, 1$, we apply the operations $\textsc{add}(n, -\lfloor b_i/d_i \rfloor, i)$ sequentially to the matrix. Each operation reduces $b_i$ modulo $d_i$, without modifying $b_j$ for $j > i$ because $\mathbf{H}'$ is upper triangular. So, after all $n-1$ operations, all $b_i$ satisfy $0 \leq b_i < d_i$. □

As an example, the HNF of the matrix $\mathbf{B}$ from Figure 1 is $\mathbf{H} = \begin{bmatrix} 3 & 2 \\ 0 & 1 \end{bmatrix}$ and can be obtained from $\mathbf{B}$ by applying the sequence of operations $\textsc{add}(2,2,1)$, $\textsc{invert}(1)$, $\textsc{swap}(1,2)$, $\textsc{add}(2,1,1)$.

**Corollary 9** *For any unimodular matrix $\mathbf{U}$, there is a sequence of elementary integer column operations $\sigma$ such that $\sigma(\mathbf{U}) = \mathbf{I}$.*

*Proof.* Let $\sigma$ be the sequence of operations such that $\sigma(\mathbf{U}) = \mathbf{H}$ is in HNF. This matrix has determinant $\det(\mathbf{H}) = \det(\mathbf{U}) \det(\sigma(\mathbf{I})) = \pm \det(\mathbf{U}) = \pm 1$. But $\mathbf{H}$ is upper triangular. So, its determinant is the product of the diagonal elements. Since these diagonal elements are positive integers, they must all be equal to $h_{i,i} = 1$. Also, all other entries must be 0 because they are integers in the range $[0, h_{i,i}) = [0, 1)$. So, $\mathbf{H} = \mathbf{I}$ is the identity matrix. □

As an example, the matrix $\mathbf{B} = [\mathbf{b}_1, \mathbf{b}_2]$ from Figure 3 has determinant $\det(\mathbf{B}) = 2-1 = 1$, and it can be transformed into the identity matrix $\mathbf{I} = [\mathbf{e}_1, \mathbf{e}_2]$ (the standard basis for $\mathbb{Z}^2$) using the following sequence of elementary column operations: $\textsc{add}(2,-2,1)$, $\textsc{swap}(1,2)$, $\textsc{invert}(1)$, $\textsc{add}(2,-1,1)$.

**Corollary 10** *For any matrix $\mathbf{U} \in \mathbb{Z}^{n \times n}$, the following conditions are equivalent:*

1. *$\mathbf{U} = \sigma(\mathbf{I})$ for some sequence of elementary column operations $\sigma$.*
2. *$\mathbf{U}$ is invertible, i.e., $\mathbf{U} \in GL(n, \mathbb{Z})$.*

*3.* $\mathbf{U}$ *is unimodular, i.e.,* $\det(\mathbf{U}) = \pm 1$.

*Proof.* We have already seen that $1 \to 2$ (Exercises 2 and 3) and $2 \to 3$ (Lemma 6). It remains to prove $3 \to 1$. So, let $\mathbf{U}$ be a unimodular matrix. By Corollary 9 there is a sequence of elementary integer column operations $\sigma$ such that $\sigma(\mathbf{U}) = \mathbf{I}$ is the identity matrix. But by Exercise 3 elementary column operations are invertible, so there is an inverse sequence $\sigma'$ such that $\mathbf{U} = \sigma'(\sigma(\mathbf{U})) = \sigma'(\mathbf{I})$. □

Together with Theorem 3, this shows that any two bases of the same lattice can be related by a sequence of elementary integer column operations.

The HNF can be very useful to check algebraic properties of lattices, e.g., two bases generate the same lattice if and only if they have the same HNF. However, HNF typically gives a basis with very bad geometric properties, as shown in the next exercise.

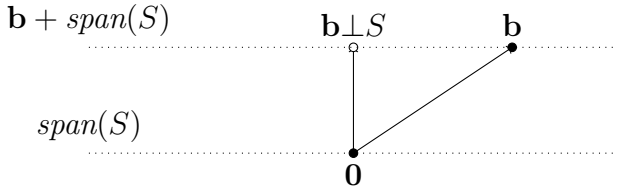
**Exercise 5** *Show that for any positive integer* $n$*, there is an* $n \times n$ *integer basis matrix* $\mathbf{B} = [\mathbf{b}_1, \ldots, \mathbf{b}_n]$ *with short vectors* $\mathbf{b}_i \in \{0, 1, -1\}^n$ *(in particular,* $\|\mathbf{b}_i\| \leq \sqrt{n}$*), while* all *vectors in the HNF basis* $\mathbf{H} = HNF(\mathbf{B})$ *have length* $\|\mathbf{h}_i\| \geq 2^{\Omega(n)}$ *that is exponentially bigger.*

# 3 Gram-Schmidt orthogonalization

Any basis $\mathbf{B}$ can be transformed into an orthogonal basis for the same vector space using the well-known Gram-Schmidt orthogonalization method. Suppose we have vectors $\mathbf{B} = [\mathbf{b}_1, \ldots, \mathbf{b}_n] \in \mathbb{R}^{d \times n}$ generating a vector space $V = \text{span}(\mathbf{B})$. These vectors are not necessarily orthogonal (or even linearly independent), but we can always find an orthogonal basis $\mathbf{B}^* = [\mathbf{b}_1^*, \ldots, \mathbf{b}_n^*]$ for $V$ where $\mathbf{b}_i^*$ is the component of $\mathbf{b}_i$ orthogonal to the linear span $\text{span}([\mathbf{b}_1, \ldots, \mathbf{b}_{i-1}])$ of the previous vectors. We recall the definition of orthogonal projection.

**Definition 11** *For any vector* $\mathbf{b} \in \mathbb{R}^d$ *and subset* $S \subseteq \mathbb{R}^d$*, let* $\mathbf{b}\perp S$ *be the component of* $\mathbf{b}$ *orthogonal to* $S$*, i.e., the (necessarily unique) vector defined by the following conditions:*

1. $(\mathbf{b}\perp S)$ *belongs to* $\mathbf{b} + span(S)$*,*
2. $(\mathbf{b}\perp S)$ *is orthogonal to all elements of* $S$*.*




Geometrically, $\mathbf{b}\perp S$ is the shortest vector in $\mathbf{b} + \text{span}(S)$, and can be thought of as the standard representative for the coset $\mathbf{b} + \text{span}(S)$, or the result of reducing $\mathbf{b}$ modulo $\text{span}(S)$.

**Definition 12** *The Gram-Schmidt orthogonalization of a sequence of vectors* $\mathbf{B} = [\mathbf{b}_1, \ldots, \mathbf{b}_n]$ *is the sequence* $\mathbf{B}^* = [\mathbf{b}_1^*, \ldots, \mathbf{b}_n^*]$ *obtained by taking the component* $\mathbf{b}_i^* = \mathbf{b}_i \perp [\mathbf{b}_1, \ldots, \mathbf{b}_{i-1}]$ *of each vector, orthogonal to the previous vectors in the sequence.*

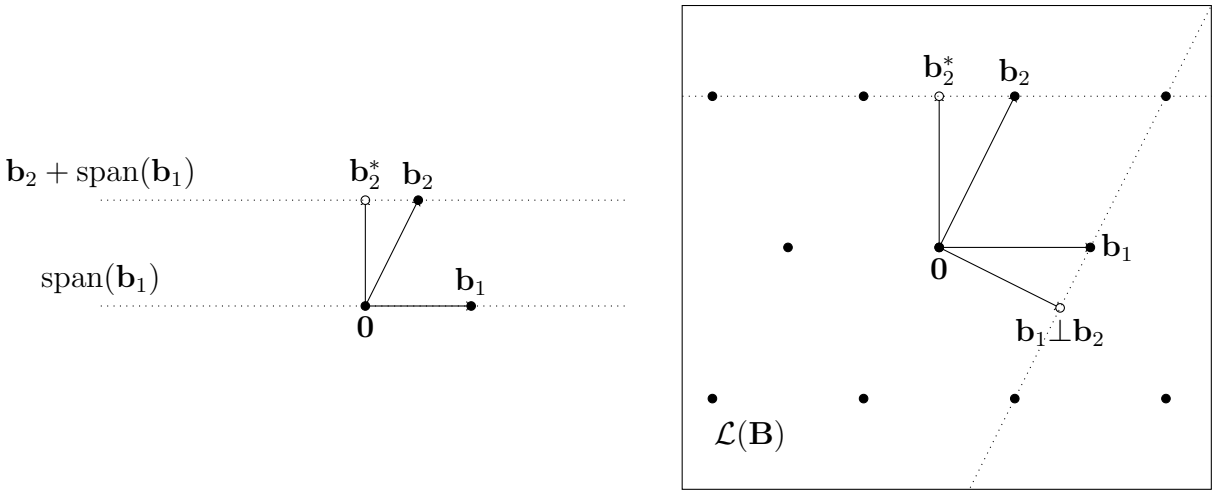



Figure 4: A basis $\mathbf{b}_1, \mathbf{b}_2$ and its Gram-Schmidt orthogonalization $\mathbf{b}_1^* = \mathbf{b}_1$,$\mathbf{b}_2^*$. The orthogonalized basis depends on the order of the basis vectors: the basis $\mathbf{b}_2, \mathbf{b}_1$ has orthogonalization $\mathbf{b}_2^* = \mathbf{b}_2$, $\mathbf{b}_1^* = \mathbf{b}_1 \perp \mathbf{b}_2$

When a basis $\mathbf{B}$ is clear from the context, the orthogonal projection modulo the first $i-1$ basis vectors mapping $\mathbf{x}$ to $\mathbf{x} \perp \mathrm{span}(\mathbf{b}_1, \ldots, \mathbf{b}_{i-1})$ is usually denoted $\pi_i$, so that $\mathbf{b}_i^* = \pi_i(\mathbf{b}_i)$.

It immediately follows from the definition that if $\mathbf{B}^* = [\mathbf{b}_1^*, \ldots, \mathbf{b}_n^*]$ is the orthogonalization of $\mathbf{B} = [\mathbf{b}_1, \ldots, \mathbf{b}_n]$, then $\mathrm{span}([\mathbf{b}_1^*, \ldots, \mathbf{b}_i^*]) = \mathrm{span}([\mathbf{b}_1, \ldots, \mathbf{b}_i])$ for every $i = 1, \ldots, n$. Also, the vectors $\mathbf{B}$ are linearly independent if and only if $\mathbf{B}^*$ are linearly independent. In particular, if $\mathbf{B}$ is a lattice basis, then $\mathbf{B}^*$ is a basis for the vector space $\mathrm{span}(\mathbf{B})$ spanned by the lattice $\mathcal{L}(\mathbf{B})$. However, generally, $\mathbf{B}^*$ is not a lattice basis for $\mathcal{L}(\mathbf{B})$ because the orthogonalized vectors $\mathbf{B}^*$ may not belong to the lattice.

As an example, consider the vectors $\mathbf{b}_1 = (2, 0)$ and $\mathbf{b}_2 = (1, 2)$. Then, the Gram-Schmidt orthogonalization of $[\mathbf{b}_1, \mathbf{b}_2]$ is given by $\mathbf{b}_1^* = \mathbf{b}_1$ and $\mathbf{b}_2^* = \mathbf{b}_2 \perp \mathbf{b}_1 = (0, 2)$. (See Figure 4.) Notice that the orthogonalized vector $\mathbf{b}_2^*$ does not belong to the lattice $\mathcal{L}(\mathbf{B})$. So, $\mathbf{B}^*$ is not a lattice basis for $\mathcal{L}(\mathbf{B})$. It is also important to notice that the Gram-Schmidt orthogonalization of a basis depends on the order of the basis vectors. E.g., if we invert the order of $\mathbf{b}_1, \mathbf{b}_2$ in the above example, we get $\mathbf{b}_2^* = \mathbf{b}_2 = (1, 2)$ and $\mathbf{b}_1^* = \mathbf{b}_1 \perp \mathbf{b}_2 = (8/5, -4/5)$.

Definition 11 can be naturally formulated as a recursive definition, setting the Gram-Schmidt orthogonalization of $[\mathbf{B}, \mathbf{b}]$ to the matrix $[\mathbf{B}^*, \mathbf{b}^*]$ where $\mathbf{b}^* = \mathbf{b} \perp \mathbf{B}$ and $\mathbf{B}^*$ is the Gram-Schmidt orthogonalization of $\mathbf{B}$. This recursive formulation admits a very natural geometric description/interpretation as follows: Any lattice with basis $[\mathbf{B}, \mathbf{b}]$ can be decomposed into layers $\mathcal{L}([\mathbf{B}, \mathbf{b}]) = \bigcup_{c \in \mathbb{Z}} (c\mathbf{b} + \mathcal{L}(\mathbf{B}))$, where each layer $c\mathbf{b} + \mathcal{L}(\mathbf{B}) \subset c\mathbf{b} + \mathrm{span}(\mathbf{B})$ is a shifted copy of a lower dimensional lattice $\mathcal{L}(\mathbf{B}) \subset \mathrm{span}(\mathbf{B})$. (See Figure 5.) Then, the Gram-Schmit vector $\mathbf{b}^*$ is a vector orthogonal to these layers, and its length is precisely the distance between any two consecutive layers $c\mathbf{b} + \mathrm{span}(\mathbf{B})$ and $(c+1)\mathbf{b} + \mathrm{span}(\mathbf{B})$.

The following lemma gives some useful formulas for the computation of the Gram-Schmidt orthogonalization.

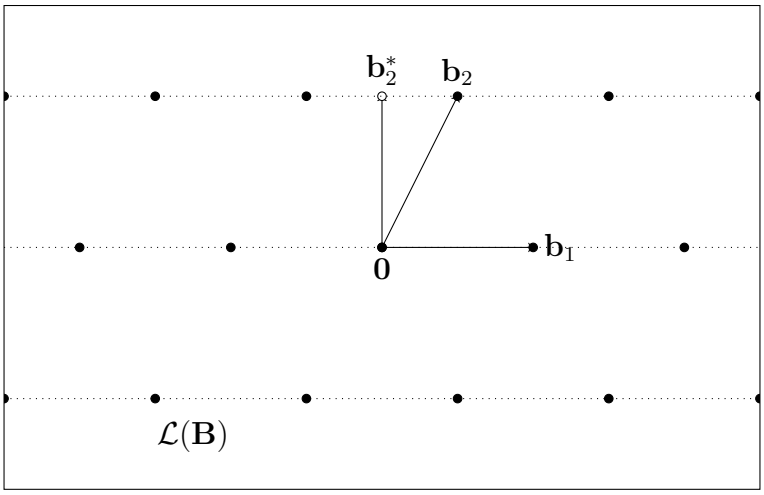



Figure 5: The lattice $\mathcal{L}(\mathbf{B})$ can be decomposed into layers orthogonal to $\mathbf{b}_n^*$, at distance $\|\mathbf{b}_n^*\|$ from each other. Each layer is a shifted copy of lower dimensional lattice $\mathcal{L}([\mathbf{b}_1, \ldots, \mathbf{b}_{n-1}])$.

**Lemma 13** *The Gram-Schmidt orthogonalization of* $\mathbf{B} = [\mathbf{b}_1, \ldots, \mathbf{b}_n]$ *is computed by the following formulas*

$$\mathbf{b}_i^* = \mathbf{b}_i - \sum_{j<i} \mu_{i,j} \mathbf{b}_j^* \text{ where } \mu_{i,j} = \frac{\langle \mathbf{b}_i, \mathbf{b}_j^* \rangle}{\langle \mathbf{b}_j^*, \mathbf{b}_j^* \rangle}.$$

**Exercise 6** *Verify that the vectors computed in Lemma 13 are indeed the Gram-Schmidt orthogonalization of* $\mathbf{B}$*, i.e., they satisfy the condition* $\mathbf{b}_i^* = \mathbf{b}_i \perp [\mathbf{b}_1, \ldots, \mathbf{b}_{i-1}]$.

In matrix notation, the basis $\mathbf{B}$ and its orthogonalization $\mathbf{B}^*$ satisfy

$$\mathbf{B} = \mathbf{B}^*\mathbf{T} \qquad \text{where} \qquad \mathbf{T} = \begin{bmatrix} 1 & \mu_{2,1} & \cdots & \mu_{n,1} \\ & \ddots & & \vdots \\ & & 1 & \mu_{n,n-1} \\ & & & 1 \end{bmatrix}$$

i.e., $\mathbf{T}$ is the upper triangular matrix with 1 along the diagonal and $t_{j,i} = \mu_{i,j}$ for all $j < i$.

# 4 The determinant

**Definition 14** *Given a basis* $\mathbf{B} = [\mathbf{b}_1, ..., \mathbf{b}_n] \in \mathbb{R}^{d \times n}$*, the* fundamental parallelepiped *associated to* $\mathbf{B}$ *is the set of points*

$$\mathcal{P}(\mathbf{B}) = \mathbf{B}\mathbb{T}^n = \{\Sigma_{i=1}^n x_i \cdot \mathbf{b}_i : \forall i. 0 \leq x_i < 1\}$$

*where* $\mathbb{T} = [0, 1)$ *is the half-open unit interval.*

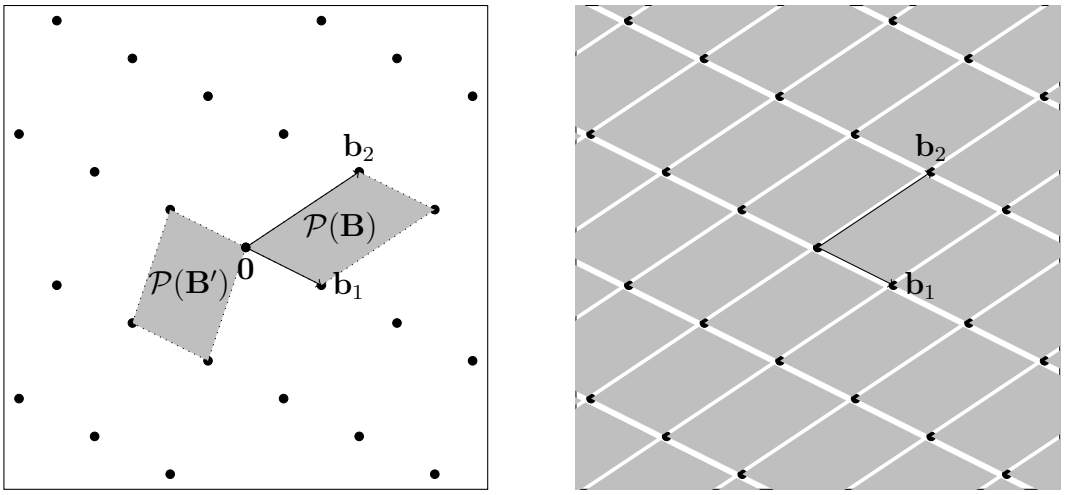



Figure 6: The fundamental parallelepiped $\mathcal{P}(\mathbf{B})$ defined by a basis. Different bases define different parallelepipeds, but they all have the same volume (area). The whole vector space spanned by the lattice $\mathcal{L}(\mathbf{B})$ can be tiled with shifted copies of $\mathcal{P}(\mathbf{B})$, one per lattice vector.

Note that $\mathcal{P}(\mathbf{B})$ is half-open, so that the translates $\mathcal{P}(\mathbf{B}) + \mathbf{v}$ (for $\mathbf{v} \in \mathcal{L}(\mathbf{B})$) form a partition of the whole space span($\mathbf{B}$). More precisely, for any $\mathbf{x} \in \text{span}(\mathbf{B})$ in the linear span of the lattice, there exists a unique lattice point $\mathbf{v} \in \mathcal{L}(\mathbf{B})$, such that $\mathbf{x} \in (\mathbf{v} + \mathcal{P}(\mathbf{B}))$. (See Figure 6.) A region of space $S \subset \text{span}(\Lambda)$ such that $\{\mathbf{x} + S : \mathbf{x} \in \Lambda\}$ is a partition of span($\Lambda$) is called a "fundamental region" of the lattice $\Lambda$. Another important example of fundamental region is the orthogonalized parallelepiped $\mathcal{P}(\mathbf{B}^*)$. (See Figure 7.)

**Exercise 7** *Prove that for any basis* $\mathbf{B}$*, the orthogonalized parallelepiped* $\mathcal{P}(\mathbf{B}^*)$ *is a fundamental region of the lattice* $\mathcal{L}(\mathbf{B})$*, i.e., for any point in space* $\mathbf{t} \in$ *span*$(\mathbf{B})$*, there exists a unique lattice point* $\mathbf{v} \in \mathcal{L}(\mathbf{B}^*)$ *such that* $\mathbf{t} \in \mathbf{v} + \mathcal{P}(\mathbf{B}^*)$*.* [Hint: use induction on the dimension of the lattice.]

**Exercise 8** *Prove that for any lattice* $\Lambda$ *and vector* $\mathbf{x} \in$ *span*$(\Lambda)$*, if* $S$ *is a fundamental region of* $\Lambda$*, then also* $\mathbf{x} + S$ *is a fundamental region.*

It immediately follows from the previous exercise that other examples of fundamental regions are given by the centered parallelepiped

$$\mathcal{C}(\mathbf{B}) = \mathbf{B}\left[-\frac{1}{2}, \frac{1}{2}\right)^n = -\frac{1}{2}\sum_i \mathbf{b}_i + \mathcal{P}(\mathbf{B})$$

or the centered orthogonalized parallelepiped $\mathcal{C}(\mathbf{B}^*) = -\frac{1}{2}\sum_i \mathbf{b}_i^* + \mathcal{P}(\mathbf{B}^*)$.

We now define a fundamental quantity associated to any lattice, the determinant.

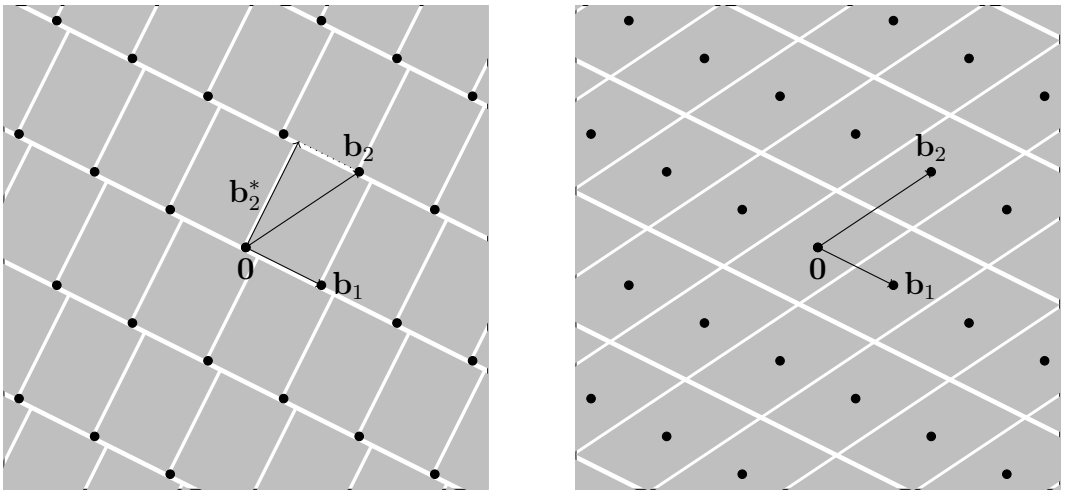



Figure 7: Other fundamental regions of a lattice are given by the orthogonalized parallelepiped $\mathcal{P}(\mathbf{B}^*)$ and the centered parallelepiped $\mathcal{C}(\mathbf{B})$. The orthogonalized parallelepipeds are not face-to-face, but they still tile the plane.

**Definition 15** *Let $\mathbf{B} \in \mathbb{R}^{d \times n}$ be a basis with Gram-Schmidt orthogonalization $\mathbf{B}^* = [\mathbf{b}_1^*, \ldots, \mathbf{b}_n^*]$. The determinant of a lattice $\Lambda = \mathcal{L}(\mathbf{B})$ is defined as the $n$-dimensional volume of the fundamental parallelepiped associated to the basis $\mathbf{B}$*

$$\det(\Lambda) = vol(\mathcal{P}(\mathbf{B})) = \prod_i \|\mathbf{b}_i^*\|$$

*and does not depend (as we will prove) on the choice of the basis $\mathbf{B}$.*

The expression $\prod_i \|\mathbf{b}_i^*\|$ for the determinant of a lattice is a generalization of the well known formula for the area of a parallelepiped. Geometrically, the determinant represents the inverse of the density of lattice points in space (e.g., the number of lattice points in a large and sufficiently regular region of space $A$ is approximately equal to the volume of $A$ divided by the determinant.) The next simple upper bound on the determinant immediately follows from the fact that $\|\mathbf{b}_i^*\| \leq \|\mathbf{b}_i\|$.

**Theorem 16 (Hadamard Inequality)** *For any lattice $\Lambda = \mathcal{L}(\mathbf{B})$, $\det(\Lambda) \leq \prod_i \|\mathbf{b}_i\|$.*

In the next lecture we will prove that the Gram-Schmidt orthogonalization of a basis can be computed in polynomial time. So, the determinant of a lattice can be computed in polynomial time by first computing the orthogonalized vectors $\mathbf{B}^*$, and then taking the product of their lengths. But there are simpler ways to express the determinant of a lattice that do not involve the Gram-Schmidt orthogonalized basis. The following proposition shows that the determinant of a lattice can be obtained from a simple matrix determinant

computation.[3] The matrix $\mathbf{B}^\top\mathbf{B}$ used in the next proposition is called the *Gram* matrix of the basis $\mathbf{B}$.

**Lemma 17** *For any lattice basis* $\mathbf{B} \in \mathbb{R}^{d\times n}$

$$\det(\mathcal{L}(\mathbf{B})) = \sqrt{\det(\mathbf{B}^\top\mathbf{B})}.$$

*In particular, if* $\mathbf{B} \in \mathbb{R}^{n\times n}$ *is a (nonsingular) square matrix then* $\det(\mathcal{L}(\mathbf{B})) = |\det(\mathbf{B})|$.

*Proof.* Remember the Gram-Schmidt orthogonalization procedure. In matrix notation, it shows that the orhogonalized vectors $\mathbf{B}^*$ satisfy $\mathbf{B} = \mathbf{B}^*\mathbf{T}$, where $\mathbf{T}$ is an upper triangular matrix with ones on the diagonal. So, our formula for the determinant of a lattice can be written as

$$\sqrt{\det(\mathbf{B}^\top\mathbf{B})} = \sqrt{\det(\mathbf{T}^\top\mathbf{B}^{*\top}\mathbf{B}^*\mathbf{T})} = \sqrt{\det(\mathbf{T}^\top)\det(\mathbf{B}^{*\top}\mathbf{B}^*)\det(\mathbf{T})}.$$

The matrices $\mathbf{T}^\top, \mathbf{T}$ are triangular, and their determinant can be easily computed as the product of the diagonal elements, which is 1. Now consider $\mathbf{B}^{*\top}\mathbf{B}^*$. This matrix is diagonal because the columns of $\mathbf{B}^*$ are orthogonal. So, its determinant can also be computed as the product of the diagonal elements which is

$$\det(\mathbf{B}^{*\top}\mathbf{B}^*) = \prod_i \langle \mathbf{b}_i^*, \mathbf{b}_i^*\rangle = (\prod_i \|\mathbf{b}_i^*\|)^2 = \det(\mathcal{L}(\mathbf{B}))^2.$$

Taking the square root we get $\sqrt{\det(\mathbf{T}^\top)\det(\mathbf{B}^{*\top}\mathbf{B}^*)\det(\mathbf{T})} = \det(\mathcal{L}(\mathbf{B}))$. □

Now it is easy to show that the determinant does not depend on the particular choice of the basis, i.e., if two bases generate the same lattice then their lattice determinants have the same value.

**Theorem 18** *Suppose* $\mathbf{B}$, $\mathbf{C}$ *are bases of the same lattice* $\mathcal{L}(\mathbf{B}) = \mathcal{L}(\mathbf{C})$. *Then, vol*$(\mathcal{P}(\mathbf{B})) =$ *vol*$(\mathcal{P}(\mathbf{C}))$.

*Proof.* Suppose $\mathbf{B}, \mathbf{C}$ are two bases of the same lattice. Then $\mathbf{B} = \mathbf{C}\cdot\mathbf{U}$ for some unimodular matrix $\mathbf{U} \in \mathbb{Z}^{n\times n}$. It follows that

$$\det(\mathbf{B}^\top\mathbf{B}) = \det((\mathbf{C}\mathbf{U})^\top(\mathbf{C}\mathbf{U})) = \det(\mathbf{U}^\top)\det(\mathbf{C}^\top\mathbf{C})\det(\mathbf{U}) = \det(\mathbf{C}^\top\mathbf{C})$$

because $\det(\mathbf{U}) = \pm 1$. □

**Exercise 9** *Let* $\Lambda$ *be a full rank sublattice of* $\Lambda'$. *Prove that*

- $\det(\Lambda)$ *divides* $\det(\Lambda')$.
- $\Lambda = \Lambda'$ *if and only if* $\det(\Lambda) = \det(\Lambda')$.

[3] Recall that the determinant of an integer matrix can be computed in polynomial time by computing $\det(\mathbf{B})$ modulo many small primes, and combining the results using the Chinese reminder theorem.