1 A Brief History of Lattices in Cryptography

Lattices have been used in mathematics going back at least to the 18th century. However, computational aspects of lattices were not investigated much until the early 1980s, when they were successfully employed for breaking several proposed cryptosystems (among many other applications). In the mid-1990s, lattices were first used in the *design* of cryptographic schemes, and the last roughly 15 years has seen an explosion of new systems and applications. Here is a timeline of some of the more significant developments in the use of lattices in cryptography.

- 18th century through early 20th century: mathematicians such as Gauss and Lagrange use lattices in number theory (e.g., to give proofs of the quadratic reciprocity and four-square theorems); Minkowski greatly advances the study of lattices in his "geometry of numbers."
- Early 1980s: Lenstra, Lenstra and Lovász discover their famous "LLL" basis-reduction algorithm, whose applications include factoring integer polynomials and breaking several cryptosystems.
- Mid-1990s: Ajtai shows a remarkable "worst-case to average-case reduction" for lattice problems, yielding a cryptographic one-way function based on *worst-case* hardness conjectures; his follow-up work with Dwork gives a public-key encryption scheme supported by similar security theorems. However, due to their inefficiency and complexity, at the time these schemes are mainly of theoretical interest.

  Concurrently, Hoffstein, Pipher and Silverman introduce the NTRU public-key encryption scheme (and somewhat later, a related digital signature scheme), which is practically quite efficient, but lacks any theoretical evidence of security. After extensive cryptanalysis, some of the most efficient parameter sets are broken (and various iterations of the signature schemes are completely broken), though NTRU encryption appears to remain secure in an asymptotic sense.
- Early 2000s: researchers such as Regev and Micciancio dramatically simplify and improve the early theoretical works, obtaining much stronger security theorems and greatly improved efficiency.
- 2007–present: several researchers (e.g., Gentry, Brakerski, Vaikuntanathan, Lyubashevsky, your instructor) build a surprisingly rich toolbox of lattice-based cryptographic constructions, including powerful objects like trapdoor functions, signature schemes, identity- and attribute-based encryption, fully homomorphic encryption, and much more.
- 2016–present: organizations like the US National Institute of Standard and Technology (NIST), Google, Cloudflare, etc. solicit, test, and ultimately select lattice-based cryptosystems for the standardization and deployment of 'post-quantum' cryptography.

In this course we will cover a great deal of this history, especially the more modern developments.

2 Mathematical Background

We start with the definition of a lattice.

Definition 2.1. An n-dimensional *lattice* $\mathcal{L}$ is a discrete additive subgroup of $\mathbb{R}^n$.

Let's unpack this definition. First, a lattice $\mathcal{L}$ is a *subgroup* of $\mathbb{R}^n$, under standard component-wise addition. Recall that $\mathcal{L}$ is a subgroup if it contains the identity element $\mathbf{0} \in \mathbb{R}^n$ (the all-zeros vector), and if for any $\mathbf{x}, \mathbf{y} \in \mathcal{L}$, we have $-\mathbf{x} \in \mathcal{L}$ and $\mathbf{x} + \mathbf{y} \in \mathcal{L}$.

Second, a lattice is *discrete*: this means that every $\mathbf{x} \in \mathcal{L}$ has some "neighborhood" in which $\mathbf{x}$ is the only lattice point. Formally, for every $\mathbf{x} \in \mathcal{L}$ there exists some $\epsilon > 0$ such that $(\mathbf{x} + \epsilon\mathcal{B}) \cap \mathcal{L} = \{\mathbf{x}\}$, where $(\mathbf{x} + \epsilon\mathcal{B})$ denotes the (open) ball of radius ϵ centered at $\mathbf{x}$. Note that when $\mathcal{L}$ is a group, this condition is equivalent to one in which the quantifiers are reversed, i.e., there exists a single $\epsilon > 0$ that works for all $\mathbf{x} \in \mathcal{L}$. Moreover, this is equivalent to the existence of some $\epsilon > 0$ that works for the origin $\mathbf{0} \in \mathcal{L}$.

Example 2.2. Let us consider several examples of lattices and non-lattices.

1. The singleton set $\{\mathbf{0}\} \subset \mathbb{R}^n$ is a lattice (for any positive integer n).
2. The integers $\mathbb{Z} \subset \mathbb{R}$ form a 1-dimensional lattice, and the integer grid $\mathbb{Z}^n \subset \mathbb{R}^n$ is an n-dimensional lattice. The set $\mathbb{Z} \times \{0\} = \{(z, 0) : z \in \mathbb{Z}\} \subseteq \mathbb{R}^2$ is a two-dimensional lattice, though its rank is only one (see Definition 2.3 below).
3. For any lattice $\mathcal{L}$, its scaling $c\mathcal{L} = \{c\mathbf{x} : \mathbf{x} \in \mathcal{L}\}$ by any real c is also a lattice, e.g., the even integers $2\mathbb{Z}$. More generally, any linear transformation applied to a lattice is also a lattice.
4. The set $\{\mathbf{x} \in \mathbb{Z}^n : \sum_{i=1}^n x_i \in 2\mathbb{Z}\}$ is a lattice; it is often called the "checkerboard" or "chessboard" lattice, especially in two dimensions. (See Figure 1.)
5. The rationals $\mathbb{Q} \subset \mathbb{R}$ do *not* form a lattice, because although they form a subgroup, it is not discrete: there exist rational numbers that are arbitrarily close to zero.
6. The odd integers $2\mathbb{Z} + 1$ do *not* form a lattice, because although they are discrete, they do not form a subgroup of $\mathbb{R}$. (However, they do comprise a *coset* of the lattice $2\mathbb{Z}$; more on this point later.)
7. The group $G = \mathbb{Z} + \sqrt{2}\mathbb{Z}$ is *not* a lattice, because it is not discrete: since $\sqrt{2}$ admits arbitrarily good rational approximations a/b, there are values $a - b\sqrt{2} \in G$ that are arbitrarily close to zero.

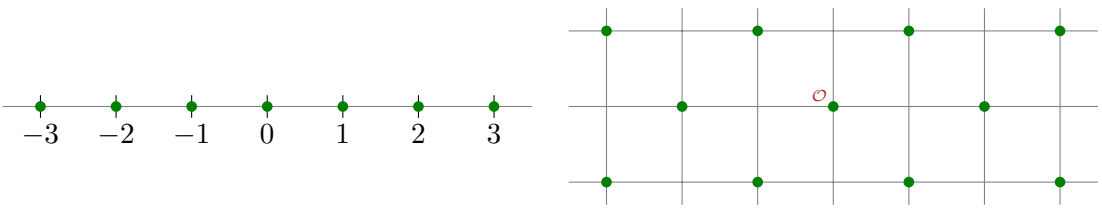

Figure 1: The integer and checkerboard lattices.

Definition 2.3. The *rank* k of a lattice $\mathcal{L} \subset \mathbb{R}^n$ is the dimension of its linear span, i.e., $k = \dim(\mathrm{span}(\mathcal{L}))$. When $k = n$, the lattice is said to be *full rank*.

In this course, all lattices are implicitly full rank unless stated otherwise. This assumption is usually (but not always!) without loss of generality, since we can restrict our attention to the subspace $\mathrm{span}(\mathcal{L})$, which can be mapped bijectively to $\mathbb{R}^k$ under a linear transformation that preserves inner products (and thus Euclidean norms, volumes, etc.).

Of particular interest are *rational* and *integer* lattices, which are subsets of $\mathbb{Q}^n$ and $\mathbb{Z}^n$, respectively. (Note that not all lattices are rational, e.g., $\sqrt{2}\mathbb{Z}$.) Up to scale, rational and integer lattices are equivalent: the discreteness of a rational lattice $\mathcal{L}$ guarantees a finite "common denominator" d such that the scaled lattice $d\mathcal{L}$ is integral. Rationality is important for computational purposes, because we need to be able to represent the lattice in an algorithm. However, most of the mathematical results in this course will apply to arbitrary (potentially non-rational) lattices, so we will explicitly identify when we are restricting to integral lattices.

2.1 Lattice Bases

Except for the degenerate case $\{\mathbf{0}\}$, a lattice is always an infinite set. However, a basic fact (which we will not prove) is that a lattice can always be finitely represented by a *basis*.

Definition 2.4. A *basis* $\mathbf{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\} \subset \mathbb{R}^n$ of a lattice $\mathcal{L}$ is a set of linearly independent vectors whose integer linear combinations generate the lattice:

$$\mathcal{L} = \mathcal{L}(\mathbf{B}) := \left\{ \sum_{i=1}^{n} z_i \mathbf{b}_i : z_i \in \mathbb{Z} \right\}. \tag{2.1}$$

Equivalently, if we interpret $\mathbf{B} \in \mathbb{R}^{n \times n}$ as a nonsingular matrix whose (ordered) columns are $\mathbf{b}_1, \ldots, \mathbf{b}_n$, then $\mathcal{L} = \mathbf{B} \cdot \mathbb{Z}^n = \{\mathbf{Bz} : \mathbf{z} \in \mathbb{Z}^n\}$.

Another way of reading the above definition is that any lattice can be obtained by applying some nonsingular linear transformation to the integer lattice $\mathbb{Z}^n$.

While a lattice always has a basis, such a basis is not unique. Essentially, the existence of multiple lattice bases is what makes lattice-based cryptography possible! (More on this later.) Recall that an integer matrix is said to be *unimodular* if it determinant is ± 1, i.e., if $\mathbf{U}^{-1}$ exists and is also an integer matrix.

Lemma 2.5. *Bases $\mathbf{B}_1, \mathbf{B}_2$ generate the same lattice $\mathcal{L}$ if and only if there exists a unimodular $\mathbf{U} \in \mathbb{Z}^{n \times n}$ such that $\mathbf{B}_1 = \mathbf{B}_2\mathbf{U}$.*

Proof. Suppose $\mathcal{L}(\mathbf{B}_1) = \mathcal{L}(\mathbf{B}_2)$, or equivalently, $\mathbf{B}_1 \cdot \mathbb{Z}^n = \mathbf{B}_2 \cdot \mathbb{Z}^n$. Then each column of $\mathbf{B}_1$ is an integer combination of the columns of $\mathbf{B}_2$, and vice-versa. That is, there exist $\mathbf{U}, \mathbf{V} \in \mathbb{Z}^{n \times n}$ such that $\mathbf{B}_1 = \mathbf{B}_2\mathbf{U}$ and $\mathbf{B}_2 = \mathbf{B}_1\mathbf{V} = \mathbf{B}_2\mathbf{UV}$. Because $\mathbf{B}_2$ is invertible, we have $\mathbf{UV} = \mathbf{I}$ and therefore $\det(\mathbf{U}) \det(\mathbf{V}) = 1$, so $\det(\mathbf{U}) = \det(\mathbf{V}) = \pm 1$ because both matrices are integral.

For the other direction, suppose $\mathbf{B}_1 = \mathbf{B}_2\mathbf{U}$ for some unimodular $\mathbf{U}$. Then we have

$$\mathcal{L}(\mathbf{B}_1) = \mathbf{B}_1 \cdot \mathbb{Z}^n = \mathbf{B}_2 \cdot (\mathbf{U} \cdot \mathbb{Z}^n) = \mathbf{B}_2 \cdot \mathbb{Z}^n = \mathcal{L}(\mathbf{B}_2),$$

where the equality $\mathbf{U} \cdot \mathbb{Z}^n = \mathbb{Z}^n$ follows from the fact that $\mathbf{Ux} = \mathbf{y} \iff \mathbf{x} = \mathbf{U}^{-1}\mathbf{y}$, and that $\mathbf{U}, \mathbf{U}^{-1}$ are both integral. □

Corollary 2.6. *The bases of $\mathbb{Z}^n$ are exactly the unimodular matrices $\mathbf{U} \in \mathbb{Z}^{n \times n}$.*

Corollary 2.7. *We can efficiently test whether two given matrices $\mathbf{B}_1, \mathbf{B}_2 \in \mathbb{Z}^{n \times n}$ generate the same lattice, by checking whether $\mathbf{B}_1^{-1} \cdot \mathbf{B}_2$ is unimodular.*

2.2 Fundamental Regions

Since a lattice is an infinite periodic "grid" in $\mathbb{R}^n$, it is often useful to consider a corresponding periodic "tiling" of $\mathbb{R}^n$ by copies of some body. For example, consider the typical tiling of bricks in a wall (which corresponds to the checkerboard lattice), or a tiling of the plane by regular hexagons (with lattice points at their centers). The notion of a fundamental region formalizes this concept.

Definition 2.8. A set $\mathcal{F} \subseteq \mathbb{R}^n$ is a *fundamental region* of a lattice $\mathcal{L}$ if its translates $\mathbf{x} + \mathcal{F} := \{\mathbf{x} + \mathbf{y} : \mathbf{y} \in \mathcal{F}\}$, taken over all $\mathbf{x} \in \mathcal{L}$, form a partition of $\mathbb{R}^n$.

For example, the half-open unit intervals $[0,1)$ and $[-\frac{1}{2},\frac{1}{2})$ are fundamental regions of the integer lattice $\mathbb{Z}$: any $x \in \mathbb{R}$ is in the unique translate $\lfloor x \rfloor + [0,1)$ or $\lfloor x \rceil + [-\frac{1}{2},\frac{1}{2})$, respectively (where $\lfloor x \rceil := \lfloor x + \frac{1}{2} \rfloor$ denotes rounding to the nearest integer, with ties broken upward). Similarly, the half-open cubes $[0,1)^n$ and $[-\frac{1}{2},\frac{1}{2})^n$ are fundamental regions of $\mathbb{Z}^n$. Note that in general, a fundamental region need not be convex or even consist of a single connected component (though all the fundamental regions we consider in this course will satisfy both of those properties).

A lattice basis naturally yields a fundamental region:

Definition 2.9. The *fundamental parallelepiped* of a lattice basis $\mathbf{B}$ is defined as

$$\mathcal{P}(\mathbf{B}) := \mathbf{B} \cdot [-\tfrac{1}{2},\tfrac{1}{2})^n = \left\{ \sum_{i=1}^{n} c_i \mathbf{b}_i : c_i \in [-\tfrac{1}{2},\tfrac{1}{2}) \right\}. \tag{2.2}$$

Alternatively, the fundamental parallelepiped is sometimes defined as $\mathcal{P}(\mathbf{B}) = \mathbf{B} \cdot [0,1)^n$. However, in this course we will typically use the formulation from Equation (2.2), since it has the convenient property of being essentially symmetric about the origin (ignoring its boundary).

Lemma 2.10. *Let $\mathcal{F}$ be a fundamental region of $\mathbb{Z}^n$ and $\mathbf{B}$ be a lattice basis. Then $\mathbf{B} \cdot \mathcal{F} = \{\mathbf{B}\mathbf{x} : \mathbf{x} \in \mathcal{F}\}$ is a fundamental region of $\mathcal{L} = \mathcal{L}(\mathbf{B})$. In particular, $\mathcal{P}(\mathbf{B})$ is a fundamental region of $\mathcal{L}$.*

Proof. We need to show that each $\mathbf{x} \in \mathbb{R}^n$ is in exactly one translate $\mathbf{v} + \mathbf{B} \cdot \mathcal{F}$, where $\mathbf{v} = \mathbf{B}\mathbf{z} \in \mathcal{L}$ for some $\mathbf{z} \in \mathbb{Z}^n$. Because $\mathbf{B}$ is nonsingular, we have $\mathbf{x} \in \mathbf{B}\mathbf{z} + \mathbf{B} \cdot \mathcal{F}$ if and only if $\mathbf{B}^{-1}\mathbf{x} \in \mathbf{z} + \mathcal{F}$. Since $\mathbf{B}^{-1}\mathbf{x} \in \mathbb{R}^n$ and $\mathcal{F}$ is a fundamental region of $\mathbb{Z}^n$, there is exactly one $\mathbf{z} \in \mathbb{Z}^n$ for which the latter inclusion holds, which proves the claim. □

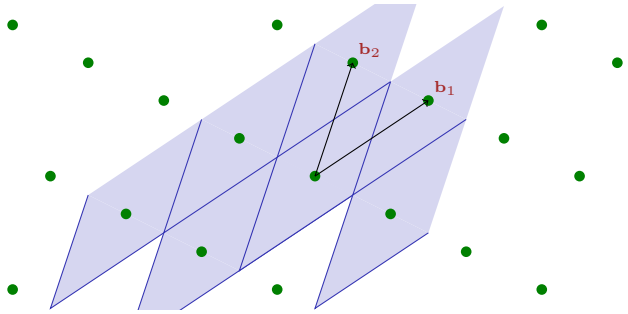

Figure 2: A lattice basis with a (partial) tiling by its fundamental parallelepiped.

Another useful fundamental region is given by the *Voronoi cell* of a lattice, which is the set of all points in $\mathbb{R}^n$ that are closer to the origin than to any other lattice point:

$$\mathcal{V}(\mathcal{L}) := \{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x}\| < \|\mathbf{x} - \mathbf{v}\| \ \forall\, \mathbf{v} \in \mathcal{L} \setminus \{\mathbf{0}\}\}. \tag{2.3}$$

Actually, the open set $\mathcal{V}(\mathcal{L})$ itself is not quite a fundamental region: its translates by lattice points are pairwise disjoint and cover all of $\mathbb{R}^n$, *except* for those points on the boundaries of the translates (which form a set of measure zero). The Voronoi cell can be made into a true fundamental region by making it "half-open" in some appropriate sense, but this involves some technical subtleties.

A basic fact is that all fundamental regions of a lattice have the *same* volume; this volume is an essential lattice invariant.

Definition 2.11. The *volume* (or *determinant*) of a lattice $\mathcal{L}$, denoted $\det(\mathcal{L})$, is defined as $\mathrm{vol}(\mathcal{F})$ where $\mathcal{F}$ is any fundamental region of $\mathcal{L}$.

Claim 2.12. *The volume of a lattice $\mathcal{L}$ is $\det(\mathcal{L}) = |\det(\mathbf{B})|$, where $\mathbf{B}$ is any basis of $\mathcal{L}$.*

Proof. By Lemma 2.10, $\mathcal{P}(\mathbf{B})$ is a fundamental region of $\mathcal{L}$, and by basic linear algebra, $\mathrm{vol}(\mathcal{P}(\mathbf{B})) = |\det(\mathbf{B})|$. □

Note that $|\det(\mathbf{B})|$ is invariant under choice of lattice basis $\mathbf{B}$, because by Lemma 2.5, any other basis is of the form $\mathbf{B}' = \mathbf{B}\mathbf{U}$ for some unimodular $\mathbf{U}$, and $|\det(\mathbf{B}')| = |\det(\mathbf{B})| \cdot |\det(\mathbf{U})| = |\det(\mathbf{B})|$.

2.3 Minimum Distance

Because a lattice $\mathcal{L}$ is discrete, it has a nonzero element $\mathbf{v} \in \mathcal{L}$ of minimum length under some specified norm, e.g., the Euclidean norm. This element is not unique, because also $-\mathbf{v} \in \mathcal{L}$.

Definition 2.13. The *minimum distance* of a lattice $\mathcal{L}$ is defined as

$$\lambda_1(\mathcal{L}) := \min_{\mathbf{v} \in \mathcal{L} \setminus \{\mathbf{0}\}} \|\mathbf{v}\| = \min_{\text{distinct } \mathbf{x},\mathbf{y} \in \mathcal{L}} \|\mathbf{x} - \mathbf{y}\|. \tag{2.4}$$

(The second equality follows from the fact that $\mathbf{x} - \mathbf{y}$ is itself a nonzero lattice vector.)

A central result of Minkowski relates the minimum distance of a lattice to its determinant. Recall that a *centrally symmetric* body S is one for which $\mathbf{x} \in S \iff -\mathbf{x} \in S$, and a *convex* body S is one for which $\mathbf{x}, \mathbf{y} \in S \Rightarrow \alpha\mathbf{x} + (1-\alpha)\mathbf{y} \in S$ for any $\alpha \in [0,1]$, i.e., the line segment connecting $\mathbf{x}$ and $\mathbf{y}$ is entirely contained in S.

Theorem 2.14 (Minkowski). *Let $\mathcal{L} \subset \mathbb{R}^n$ be a lattice. Any convex, centrally symmetric body $S \subset \mathbb{R}^n$ of volume $\mathrm{vol}(S) > 2^n \cdot \det(\mathcal{L})$ contains a nonzero lattice point, i.e., $S \cap \mathcal{L} \neq \{\mathbf{0}\}$.*

Proof. Let $S' = S/2$, so $\mathrm{vol}(S') > \det(\mathcal{L})$. (Recall that in $\mathbb{R}^n$, volumes scale by the nth power of the scaling factor.) By a "volumetric pigeonhole argument," we claim that there exist distinct $\mathbf{x}, \mathbf{y} \in S'$ such that $\mathbf{x} - \mathbf{y} \in \mathcal{L}$. To see this, consider any fundamental region $\mathcal{F}$ of $\mathcal{L}$, and partition S' into sets $S'_{\mathbf{v}} := S' \cap (\mathbf{v} + \mathcal{F})$ for each $\mathbf{v} \in \mathcal{L}$. Then the translates $S'_{\mathbf{v}} - \mathbf{v} \subseteq \mathcal{F}$ have total volume $\mathrm{vol}(S') > \mathrm{vol}(\mathcal{F})$, so they must overlap somewhere, i.e., there exists some $\mathbf{z} \in (S'_{\mathbf{u}} - \mathbf{u}) \cap (S'_{\mathbf{v}} - \mathbf{v})$ for distinct lattice points $\mathbf{u}, \mathbf{v} \in \mathcal{L}$. It follows that $\mathbf{x} = \mathbf{z} + \mathbf{u}, \mathbf{y} = \mathbf{z} + \mathbf{v} \in S'$ are two distinct points in S' whose difference $\mathbf{x} - \mathbf{y} = \mathbf{u} - \mathbf{v} \in \mathcal{L}$ is a lattice point.

Finally, we have $2\mathbf{x}, -2\mathbf{y} \in S$ by definition of S' and central symmetry of S, and their midpoint $(2\mathbf{x} - 2\mathbf{y})/2 = \mathbf{x} - \mathbf{y} \in S$ by convexity. □

Corollary 2.15 (Minkowski's First Theorem). *For any lattice $\mathcal{L}$, we have $\lambda_1(\mathcal{L}) \leq \sqrt{n} \cdot \det(\mathcal{L})^{1/n}$.*

Notice that this bound "scales properly": if we scale the lattice (and hence its minimum distance) by a c factor, the determinant scales by a c^n factor, hence the bound also scales by a c factor.

Proof. Without loss of generality, assume that $\det(\mathcal{L}) = 1$ by scaling the lattice by a $\det(\mathcal{L})^{-1/n}$ factor, which scales λ_1 by the same factor. Now take $S = \sqrt{n} \cdot \bar{\mathcal{B}}$ to be the (closed) Euclidean ball of radius $\sqrt{n}$. The cube $[-1,1]^n$, which has side length 2 and therefore volume 2^n, is strictly contained in S (when $n > 1$; for $n = 1$ the corollary holds trivially), hence $\mathrm{vol}(S) > 2^n$. Therefore, S contains a nonzero lattice point by Theorem 2.14, and the bound on λ_1 immediately follows. □

Note that the bound we used on the volume of a ball was somewhat loose. Using a more precise formula for the volume of an n-dimensional ball, we can obtain a slightly tighter bound $\lambda_1(\mathcal{L}) \leq \sqrt{n/(2\pi e)} \cdot \det(\mathcal{L})^{1/n}$, which is better than the one given in Corollary 2.15 by only a constant factor.

Also notice that Minkowski's bound can be arbitrarily loose: for example, consider the lattice $\mathcal{L} \subset \mathbb{R}^2$ of unit determinant generated by the basis vectors $(2^{100}, 0)$, $(0, 2^{-100})$. Then $\lambda_1(\mathcal{L}) = 2^{-100} \ll \sqrt{2}$. However, it is known that in general, Minkowski's bound cannot be improved beyond small constant factors: there exist infinite families of n-dimensional, unit-determinant lattices $\mathcal{L}$ for which $\lambda_1(\mathcal{L}) \geq C\sqrt{n}$, where $C \approx \sqrt{1/(\pi e)}$ is some universal positive constant.