CSC 2414 Lattices in Computer Science September 20, 2011

Lecture 2

Lecturer: Vinod Vaikuntanathan *Scribe: Sergey Gorbunov*

1 Lecture Outline

- Alternative Definition of Lattices
- Upper bound on the length of a shortest lattice vector ($\lambda_1(\mathcal{L})$)
- Successive Minima
- Applications of Minkowski's Theorem

2 Alternative Definition of Lattices

Last lecture we saw the following definition of a lattice:

Definition 1 (Lattices)**.** *Given n* linearly independent *vectors $\mathbf{b}_1, \ldots, \mathbf{b}_n \in \mathbb{R}^m$, the lattice generated by them is defined as*

$$\mathcal{L}(\mathbf{b}_1, \ldots, \mathbf{b}_n) \stackrel{\text{def}}{=} \left\{ \sum_{i=1}^{n} x_i \mathbf{b}_i \mid x_i \in \mathbb{Z} \right\}$$

Now, consider the following set S:

Example 1.

$$S = \{(\mathbf{x}_1, \ldots, \mathbf{x}_n) \in \mathbb{Z}^n : \sum_{i=1}^{n} \mathbf{x}_i \text{ is even}\}$$

A typical question may ask whether S is a lattice. From the Definition 1, we know that if we are able to find a set of linearly independent vectors $\mathbf{b}_1, \ldots, \mathbf{b}_n \in \mathbb{R}^m$ such that for all $\mathbf{s} \in S$, $\mathbf{s} = a_1\mathbf{b}_1 + \ldots + a_n\mathbf{b}_n$ for some integer coefficients $a_1, \ldots, a_n$, then S is indeed a lattice. We will now give a new definition of a lattice that gives us a simpler way of verifying that S is a lattice.

Definition 2. *A lattice $\mathcal{L}$ is a discrete additive subgroup of $\mathbb{R}^n$.*

In the above definition, by discrete we mean that:

$$\exists \epsilon > 0, \text{ s.t. } \forall \mathbf{x} \neq \mathbf{y} \in \mathcal{L}, ||\mathbf{x} - \mathbf{y}|| \geq \epsilon.$$

And by additive subgroup we mean that:

$$\forall \mathbf{x}, \mathbf{y} \in \mathcal{L}, \mathbf{x} - \mathbf{y} \in \mathcal{L}.$$

Claim 1. *Definition 1 is equivalent to Definition 2.*

Proof. We will first show that Definition 1 $\implies$ Definition 2.

Assume $\mathcal{L}$ is a lattice defined as the set of all integer combinations of vectors $\mathbf{b}_1, \ldots, \mathbf{b}_n \in \mathbb{R}^m$ which are linearly independent (Definition 1). Then, clearly $\mathcal{L}$ is an additive subgroup of $\mathbb{R}^n$. In addition, $\forall \mathbf{x}, \mathbf{y} \in \mathcal{L}$, $\mathbf{x} - \mathbf{y} \in \mathcal{L}$. Therefore, from the lower bound on a shortest lattice vector,

$$||\mathbf{x} - \mathbf{y}|| \geq \lambda_1(\mathcal{L}) \geq \min_{i=1,\ldots,n} ||\tilde{\mathbf{b}_i}||.$$

In other words, the length of any lattice vector must be greater than the length of a shortest lattice vector. Therefore, we can let $\epsilon = \lambda_1$. So, both properties of Definition 2 are satisfied ($\mathcal{L}$ is a discrete additive subgroup of $\mathbb{R}^n$).

We show that Definition 2 $\implies$ Definition 1. Given a discrete additive subgroup $\mathcal{L}$ of $\mathbb{R}^n$, we construct a set of basis using the algorithm below.

We will use the following definition of a closed parallelepiped.

Definition 3 (Closed Fundamental Parallelepiped)**.** *Given* n linearly independent *vectors* $\mathbf{b}_1, \ldots, \mathbf{b}_n \in \mathbb{R}^m$, *their* closed fundamental parallelepiped *is defined as*

$$\bar{\mathcal{P}}(\mathbf{b}_1, \ldots, \mathbf{b}_n) \stackrel{\text{def}}{=} \left\{ \sum_{i=1}^{n} x_i \mathbf{b}_i \mid x_i \in \mathbb{R}, 0 \leq x_i \leq 1 \right\}$$

Pick $\mathbf{y} \in \mathcal{L}$ such that there is no lattice vector between the zero vector and $\mathbf{y}$. Let $\mathbf{b}_1 = \mathbf{y}$.
Iterate for all i, $1 \leq i < n$: Assume we have already chosen $\mathbf{b}_1, \ldots, \mathbf{b}_i$. Choose $\mathbf{y}$ not in the $\mathsf{Span}(\mathbf{b}_1, \ldots, \mathbf{b}_i)$.
Consider a $\bar{\mathcal{P}}(\mathbf{b}_1, \ldots, \mathbf{b}_i, \mathbf{y})$ (See Figure-1 for an example).
Now, $\bar{\mathcal{P}}$ contains at least one lattice point (namely $\mathbf{y}$) and it contains finitely many lattice points. Now, choose a vector $\mathbf{z} \in \bar{\mathcal{P}}(\mathbf{b}_1, \ldots, \mathbf{b}_i, \mathbf{y}) \setminus \mathsf{Span}(\mathbf{b}_1, \ldots, \mathbf{b}_i)$ such that

$$dist(\mathbf{z}, \mathsf{Span}(\mathbf{b}_1, \ldots, \mathbf{b}_i)) \text{ is the smallest.}$$

Note, that we can do this since we have only finitely many points to choose from. Let $\mathbf{b}_{i+1} = \mathbf{z}$.

We will now show that the above algorithm returns a basis $\mathbf{b}_1, \ldots \mathbf{b}_n \in \mathbb{R}^m$ for the lattice. Clearly all $\mathbf{b}_i s \in \mathbb{R}^m$ and they are linearly independent by the algorithm that we used. We are left to show that

$$\mathcal{L} \subseteq \{\Sigma x_i \mathbf{b}_i : x_i \in \mathbb{Z}\}.$$

Let $\mathbf{z} = \Sigma z_i \mathbf{b}_i$ be an arbitrary lattice vector (where $z_i \in \mathbb{R}$). Let $\mathbf{z}' = \Sigma \lfloor z_i \rfloor \mathbf{b}_i \in \mathcal{L}$. Then, $\mathbf{z} - \mathbf{z}' = \Sigma (z_i - \lfloor z_i \rfloor) \mathbf{b}_i \in \mathcal{L}$. We will show that all coefficients z_i must be integers. Express

$$\mathbf{z} - \mathbf{z}' = (z_n - \lfloor z_n \rfloor)\mathbf{b}_n + \mathsf{Span}(\mathbf{b}_1, \ldots, \mathbf{b}_{n-1}) = (z_n - \lfloor z_n \rfloor)\tilde{\mathbf{b}_n} + \mathsf{Span}(\mathbf{b}_1, \ldots, \mathbf{b}_{n-1})$$

In other words, vector $\mathbf{z} - \mathbf{z}'$ is in the $\mathsf{Span}(\mathbf{b}_1, \ldots, \mathbf{b}_{n-1})$ plus a multiple of $\tilde{\mathbf{b}_n}$ with coefficients $0 \leq \lfloor z_n \rfloor < 1$.
Now,

$$dist(\mathbf{z} - \mathbf{z}', \mathsf{Span}(\mathbf{b}_1, \ldots, \mathbf{b}_{n-1})) = (z_n - \lfloor z_n \rfloor)||\tilde{\mathbf{b}_n}||$$

This follows because the distance is defined as the orthogonal component of $\mathbf{z} - \mathbf{z}'$ to the span $\mathsf{Span}(\mathbf{b}_1, \ldots, \mathbf{b}_{n-1})$, which is precisely $(z_n - \lfloor z_n \rfloor)||\tilde{\mathbf{b}_n}||$. Similarly,

$$dist(\mathbf{b}_n, \mathsf{Span}(\mathbf{b}_1, \ldots, \mathbf{b}_{n-1})) = ||\tilde{\mathbf{b}_n}||$$

In addition, since $0 \leq (z_n - \lfloor z_n \rfloor) < 1$,

$$dist(\mathbf{z} - \mathbf{z}', \mathsf{Span}(\mathbf{b}_1, \ldots, \mathbf{b}_{n-1})) < dist(\mathbf{b}_n, \mathsf{Span}(\mathbf{b}_1, \ldots, \mathbf{b}_{n-1}))$$

But since $\mathbf{b}_n$ was chosen as the closest vector to $\mathsf{Span}(\mathbf{b}_1, \ldots, \mathbf{b}_{n-1})$, $\mathbf{z} - \mathbf{z}'$ must be linearly dependent of $\mathbf{b}_1, \ldots, \mathbf{b}_{n-1}$. Therefore, $z_n - \lfloor z_n \rfloor = 0$ and so $z_n \in \mathbb{Z}$. By recursively repeating the above argument for $\mathbf{z} = \mathbf{z} - z_i \mathbf{b}_i \in \mathsf{Span}(\mathbf{b}_1, \ldots, \mathbf{b}_{i-1})$ for all $1 < i \leq n$ we obtain that all coefficients z_j for $1 \leq j \leq n$ must be integers. □

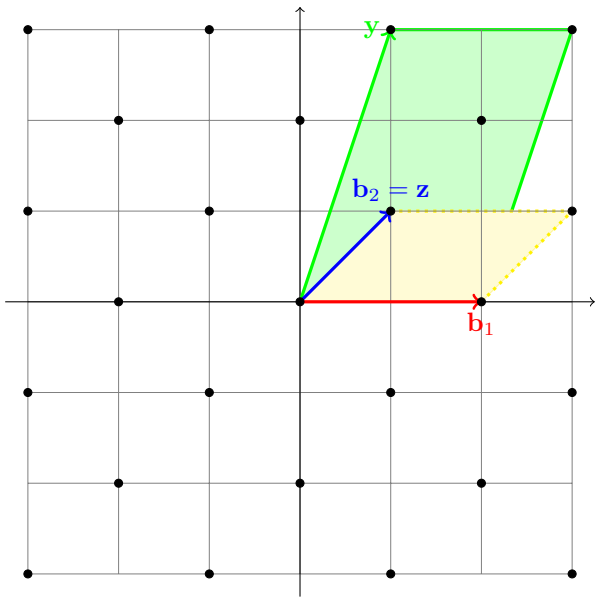

Figure 1: Constructing lattice basis from a discrete additive subgroup of $\mathbb{Z}^2$. After the first iteration if we choose $\mathbf{y} = (1, 3)$, then in the $\bar{\mathcal{P}}(\mathbf{b}_1, \mathbf{y})$ we choose $\mathbf{z} = \mathbf{b}_2$ which is at the minimum distance from $\mathsf{Span}(\mathbf{b}_1)$. We can see that there are no non-zero lattice vectors in $\mathcal{P}(\mathbf{b}_1, \mathbf{b}_2)$. Therefore, $\{\mathbf{b}_1, \mathbf{b}_2\}$ forms a basis for this lattice.

3 Upper bound on a shortest lattice vector

Last time we saw how to compute a lower bound on the length of a shortest lattice vector. In particular, we saw that

$$\lambda_1(\mathcal{L}) \geq \min_{i=1,\ldots,n} ||\tilde{\mathbf{b}}||$$

Now, we will compute an upper bound on $\lambda_1(\mathcal{L})$.

Before we state the theorem, we give some intuition on what a shortest lattice vector might depend on. Recall, that the determinant of a lattice is an inverse of its density. So larger determinant implies less dense lattice, while smaller determinant implies denser lattice. Therefore, we should be able to express an upper bound for $\lambda_1(\mathcal{L})$ in terms of the determinant of a lattice. But we know that $\lambda_1(\mathcal{L}) \leq det(\mathcal{L})$. We will now prove a stronger bound. In addition, our expression should scale well. In particular, if $\mathcal{L}$ is an arbitrary lattice with a shortest vector of length $\lambda_1(\mathcal{L})$, then lattice defined by scaling every vector in $\mathcal{L}$ by k should have a shortest vector of length $k\lambda_1(\mathcal{L})$. Similarly, $det(c\mathcal{L}) = c^n\mathcal{L}$ should satisfy. One can check that Minkowski's first theorem holds well on these properties.

Theorem 1 (Minkowski's First Theorem)**.** *For every full rank lattice $\mathcal{L}$:*

$$\lambda_1(\mathcal{L}) \leq \sqrt{n} * det(\mathcal{L})^{1/n}.$$

In order to prove the theorem, we will need to use the following 2 theorems.

Theorem 2 (Blichfeld)**.** *For all full rank lattice $\mathcal{L}$ and measurable set $S \subseteq \mathbb{R}^n$ s.t. $vol(S) > det(\mathcal{L})$,*

$$\exists \mathbf{x}, \mathbf{y} \in S, s.t.\ \mathbf{x} - \mathbf{y} \in \mathcal{L}.$$

See Figure-2 for an example.

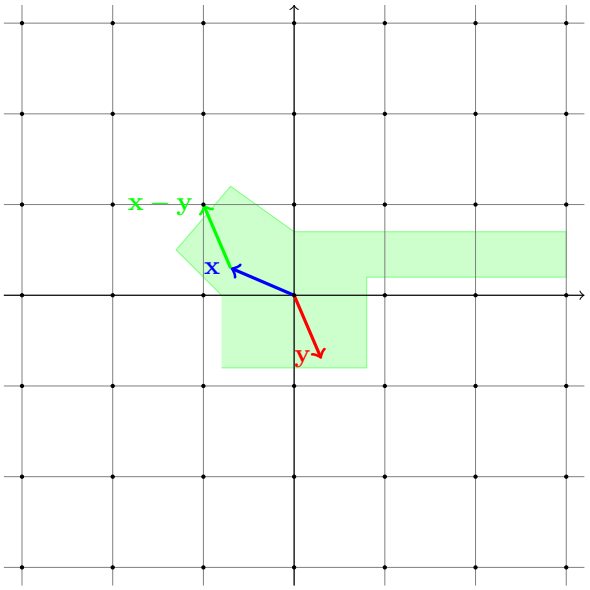

Figure 2: By the Blichfeld's theorem we can find $\mathbf{x}$ and $\mathbf{y}$ in this set such that $\mathbf{x} - \mathbf{y} \in \mathcal{L}$.

Proof. Let B be a basis for the lattice $\mathcal{L}$. Define $f : \mathbb{R}^n \to \mathcal{P}(B)$ as follows: $f(\Sigma x_i \mathbf{b}_i) = \Sigma(x_i - \lfloor x_i \rfloor)\mathbf{b}_i$. First, note that $\Sigma x_i \mathbf{b}_i - f(\Sigma x_i \mathbf{b}_i) = \Sigma \lfloor x_i \rfloor \mathbf{b}_i \in \mathcal{L}$. Now consider the following 2 cases:

Case 1: If $\exists \mathbf{x}, \mathbf{y} \in S$ s.t. $f(x) = f(y)$ (i.e. we have a collision from two vectors). Then, $\mathbf{x} - \mathbf{y} = (\mathbf{x} - f(\mathbf{x})) - (\mathbf{y} - f(\mathbf{y}))$. But as noted above, $\mathbf{x} - f(\mathbf{x}) \in \mathcal{L}$ and $\mathbf{y} - f(\mathbf{y}) \in \mathcal{L}$. Therefore, $\mathbf{x} - \mathbf{y} \in \mathcal{L}$.

Case 2: Assume there are no collisions. Let $S = \bigcup_{\mathbf{x} \in \mathcal{L}} S_x$. Define $\tilde{S}_\mathbf{x} = S_\mathbf{x} - \mathbf{x}$. By definition, $\tilde{S}_\mathbf{x} \subseteq \mathcal{P}(B)$. Also, $vol(S) = \Sigma vol(S_\mathbf{x})$ and $vol(\tilde{S}_\mathbf{x}) = vol(S_\mathbf{x})$. Therefore, $vol(S) = \Sigma vol(S_\mathbf{x}) = \Sigma vol(\tilde{S}_\mathbf{x})$. But since we assume that we do not have any collisions, then for all $\mathbf{x}, \mathbf{y}$, $\tilde{S}_\mathbf{x} \bigcap \tilde{S}_\mathbf{y} = \emptyset$. And so,

$$vol(S) = \Sigma vol(\tilde{S}_\mathbf{x}) = vol(\bigcup_{\mathbf{x} \in \mathcal{L}} \tilde{S}_\mathbf{x}) \leq vol(\mathcal{P}(B)) = det(\mathcal{L})$$

Therefore, $vol(S) \leq det(\mathcal{L})$ which contradicts with our assumption. □

Definition 4 (Convex Set)**.** *A set S is convex if:*

$$\forall x \neq y \in S, \forall \alpha \in [0, 1], \alpha x + (1 - \alpha) y \in S.$$

Informally, the above definition say that if we take any two points from the set, any point that lies on the straight line between the two points must also be in the set.

Definition 5 (Centrally Symmetric Set)**.** *A set S is centrally symmetric if:*

$$\forall x \in S, -x \in S.$$

Theorem 3 (Minkowski's Convex Body Theorem)**.** *For all full-rank lattice $\mathcal{L}$, and a convex centrally symmetric set S with $vol(S) > 2^n det(\mathcal{L})$, S contains a non-zero lattice point.*

Proof. Let $\tilde{S} = \{\mathbf{x}/2 : \mathbf{x} \in S\}$. Then,

$$vol(\tilde{S}) = 2^{-n} vol(S) > det(\mathcal{L})$$

Therefore, by the Blichfeld's theorem $\exists \mathbf{x}, \mathbf{y} \in \tilde{S}$ s.t. $\mathbf{x} - \mathbf{y} \in \mathcal{L}$. We will show that $\mathbf{x} - \mathbf{y} \in S$. Now, $2\mathbf{x} \in S$ and $2\mathbf{y} \in S$ by the construction of $\tilde{S}$. Therefore, $-2\mathbf{y} \in S$. And $\mathbf{x} - \mathbf{y} = \frac{2\mathbf{x} - 2\mathbf{y}}{2} \in S$.

□

We are now ready to prove Minkowski's first theorem (Theorem 1).

Proof. (Minkowski's first theorem) Let $S = \mathcal{B}(0, \lambda_1(\mathcal{L}))$, where $\mathcal{B}(\mathbf{x}, r)$ is an n-dimensional open ball of radius r centred at $\mathbf{x}$.
Note that using l_2 norm this n-dimensional ball contains an n-dimensional cube of length $\frac{2r}{\sqrt{n}}$. Therefore,

$$vol(\mathcal{B}(0, r)) \geq (\frac{2r}{\sqrt{n}})^n.$$

Therefore, we get $vol(\mathcal{B}(0, \lambda_1(\mathcal{L}))) \geq (\frac{2\lambda_1(\mathcal{L})}{\sqrt{n}})^n$. But from Minkowski's convex body theorem and the fact that the ball is open and hence contains no non-zero lattice points, we get

$$(\frac{2\lambda_1(\mathcal{L})}{\sqrt{n}})^n \leq vol(\mathcal{B}(0, \lambda_1(\mathcal{L})) \leq 2^n det(\mathcal{L}).$$

Rearranging,

$$\lambda_1(\mathcal{L}) \leq \sqrt{n} * det(\mathcal{L})^{1/n}.$$

□

4 Successive Minima

Given a lattice, it is natural to ask questions such as: What is a shortest vector in the lattice? What is a second shortest vector in the lattice? In general, what is the length of the ith shortest vector in the lattice? As we saw from the last lecture, the length of a shortest lattice vector is defined as $\lambda_1(\mathcal{L})$. We now extend this definition:

Definition 6 (Successive Minima)**.** *Let $\mathcal{L}$ be an arbitrary lattice of rank n. Then $\forall i, 1 \leq i \leq n$:*

$$\lambda_i(\mathcal{L}) \stackrel{\text{def}}{=} inf\{r : \mathcal{B}(0, r) \text{ contains } \geq i \text{ linearly independent lattice vectors}\}$$

Following the above the definition and the lattice described in Figure-3, we can see that $\lambda_1(\mathcal{L}) = ||\mathbf{x}_1||$ and $\lambda_2(\mathcal{L}) = ||\mathbf{x}_3||$, since neither $\mathbf{x}_2$ nor $2\mathbf{x}_1$ are linearly independent of $\mathbf{x}_1$.

We saw that Minkowski's first theorem gives us an upper bound on $\lambda_1(\mathcal{L})$. In fact, his second theorem strengthens the results by considering a geometric mean of $\lambda_1(\mathcal{L}), \lambda_2(\mathcal{L}), \ldots, \lambda_n(\mathcal{L})$. We leave the proof of the theorem to the reader.

Theorem 4 (Minkowski's Second Theorem)**.** *For all full rank lattices $\mathcal{L}$,*

$$(\prod_{i=1}^{n} \lambda_i(\mathcal{L}))^{1/n} \leq \sqrt{n} * (det(\mathcal{L}))^{1/n}.$$

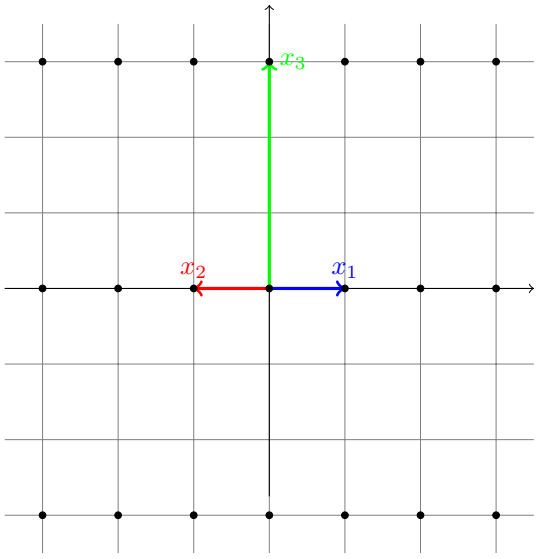

Figure 3: The first and second successive minima in the lattice generated by $(1, 0)$ and $(0, 3)$. Knowing that $\lambda_1(\mathcal{L}) = ||\mathbf{x}_1||$, we can ask whether $\lambda_2(\mathcal{L}) = ||\mathbf{x}_2||, \lambda_2(\mathcal{L}) = 2\lambda_1(\mathcal{L})$ or $\lambda_2(\mathcal{L}) = ||\mathbf{x}_3||$. By the definition, $\lambda_2(\mathcal{L}) = ||\mathbf{x}_3||$.

5 Applications of Minkowski's Theorem

Minkowski's Theorem is widely used in computer and mathematical sciences. For example, we now can prove the following theorems:

- Dirichlets theorem on Diophantine approximation. This theorem allows us to approximate real numbers with rationals (See Figure-4 for an example).
- Lagranges four-square theorem. Intuitively, the theorem states that we can express every positive integer as the sum of four squares of integers. We leave the proof of this theorem to the reader.

Theorem 5 (Dirichlet's Theorem on Diophantine Approximation)**.** *For all $\lambda \in \mathbb{R}$ and all $Q \in \mathbb{N}$, there exists p, q such that*

$$q < Q \text{ and } |\lambda - \frac{p}{q}| \leq \frac{1}{qQ}.$$

Proof. Consider the lattice $\mathcal{L} = \mathbb{Z}^2$. Let $S = \{(x, y) : -Q \leq x \leq Q, -\frac{1}{Q} \leq \lambda x - y \leq \frac{1}{Q}\}$.
Now, $vol(S) =$ base*height$= \frac{2}{Q} * 2Q = 4$. Therefore, $vol(S) \geq 2^2 det(\mathbb{Z}^2)$.
Applying Minkowski's first theorem, we know that there exists $(q, p) \in \mathbb{Z}^2$ such that,

$$-Q \leq q \leq Q \text{ and } -\frac{1}{Q} \leq \lambda q - p \leq \frac{1}{Q}.$$

Therefore, $|\lambda q - p| \leq \frac{1}{Q}$, which implies $|\lambda - \frac{p}{q}| \leq \frac{1}{qQ}$. □

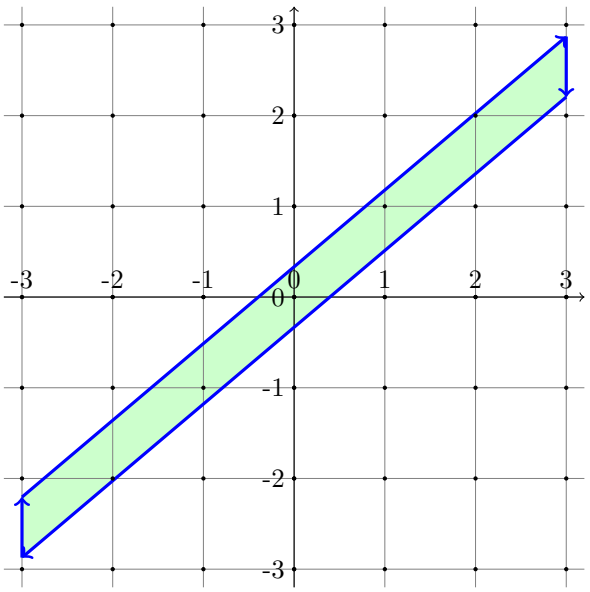

Figure 4: Applying Dirichlet's theorem to approximate 0.846153846 for $Q = 3$, we can see that there exists (p, q) satisfying the requirements. For example, $(p, q) = (1, 1)$, since $1 \leq 3$ and $|0.846153846 - 1| = 0.153846154 \leq \frac{1}{1*3}$.