Lattices and Modular Forms

Dmytro Lyubka
Supervisor: Graham Ellis

August 2022

Contents

Chapter 1

Lattices

1.1 The Sphere Packing Problem

The sphere packing problem is stated as follows: *What is the maximal density of a sphere packing in a Euclidean space $\mathbb{R}^n$ of given dimension n?* Modern mathematics defines a "sphere" to be an $(n-1)$-dimensional boundary of its associated n-ball, embedded in n dimensions. In formal notation, a sphere is defined by

$$S^n(r) := \{x \in \mathbb{R}^{n+1} : |x| = r\},$$

while an n-ball is defined by

$$B^n(r) := \{x \in \mathbb{R}^n : |x| \leq r\}.$$

In the context of two dimensions, the sphere packing problem is in fact a problem of packing 2-balls (disks) in two dimensions rather than packing 1-spheres (circles) in two dimensions. As such, one ought to technically speak of "ball packing" as opposed to "sphere packing", for when we seek to identify coordinates that lie in the space $S \cup \mathbb{R}^n$ we are determining whether or not they are in the space $\{x \in \mathbb{R}^n : |x - c| \leq r\}$ for some sphere of radius r centred at $c \in \mathbb{R}^n$; this space is the region covered by a 2-ball (disk). With that being said, the pedanticism that comes along with modern terminology is unnecessary here, and one usually writes of S as a family of "spheres" rather than "balls" in this context. Accordingly, in this report a "sphere packing" is a collection S of balls of the same radius in a manner wherein their interiors are disjoint and do not overlap. The "density" of a sphere packing is the proportion of the volume of $\mathbb{R}^n$ covered by the collection of balls S.

Sphere packing for dimensions $n \leq 3$ provides a good model for many familiar real life problems and phenomena; layout of ommatadia in the eyes of arthropods, atom distribution in the crystal structure of single elements, packing fruit in a crate, etc. For $n = 1$, the problem becomes trivial once we establish what is meant by a sphere in a one-dimensional space (using the mathematics defined previously).

$$B^1(r) = \{x \in \mathbb{R} : |x| \leq r\}$$

This is the set of all points along a one-dimensional line segment that are within a distance r from some point $x \in \mathbb{R}$, i.e. an *interval* from $(x - r)$ to $(x + r)$. This packing achieves a density $\Delta(S) = 1$ and is attained, for instance, by $S = \{[2k - 1, 2k + 1] \mid k \in \mathbb{Z}\}$, see Figure 1.1.

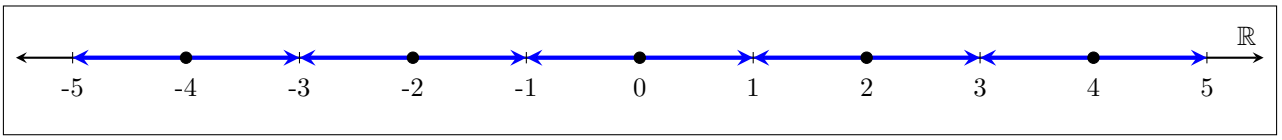


Figure 1.1: **The densest packing in one dimension. The collection S consists of all intervals $[2k - 1, 2k + 1]$, $k \in \mathbb{Z}$.**

Triviality ends once we look at the case $n = 2$. The densest packing for spheres in $\mathbb{R}^2$ has been known since antiquity, obtained by tiling the two-dimensional plane with regular hexagons of sides $\frac{2r}{\sqrt{3}}$. The collection S then consists of circles of radii r inscribed within these hexagons, yielding a packing density $\Delta(S) = \frac{\pi}{2\sqrt{3}} \approx 0.9069$, see Figure 1.2.

In 1773, Joseph-Louis Lagrange proved that the highest density achievable via a lattice packing in $\mathbb{R}^2$ is attained by a hexagonal lattice. This was not a claim that the highest achievable density of *all possible packings* is attained by a hexagonal lattice; it is of course possible to pack spheres in a manner such that their centres do not correspond to a lattice. And so, the quest of obtaining sphere packing optimality in two dimensions did not end with Lagrange. It wasn't until 1890 that a Norwegian mathematician Axel Thue published a proof which claimed that a hexagonal lattice packing achieves the highest density out of *all packings*, complementing Lagrange's work on lattices in an attempt to close the century-old problem of two-dimensional sphere packing. Unfortunately, Thue's proof was considered by many to be incomplete, and so the problem remained open. Mathematics had to wait over 50 years before the hexagonal packing in $\mathbb{R}^2$ was shown to be optimal when the mathematician László Fejes Tóth finally constructed a watertight proof in 1942.

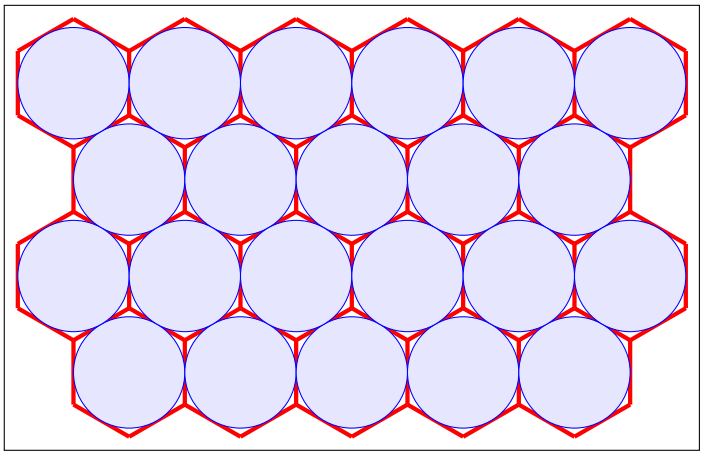


Figure 1.2: **Densest packing in two dimensions.**

In 1611, Johannes Kepler claimed that no arrangement of equally sized spheres filling $\mathbb{R}^3$ has a greater packing density than that of the cubic close packing (face-centered cubic) and hexagonal close packing arrangements. Both of these packings have a density $\Delta(S) \approx 0.7405$. Kepler did not have a proof for his Conjecture, and the next advancement in three-dimensional sphere packing was made by Carl Friedrich Gauss in 1831 who proved that the Kepler Conjecture is true if the spheres are arranged in a regular lattice. This meant that in order to disprove Kepler, one would need to explore irregular packings, a plethoric task making Kepler's Conjecture notoriously difficult to prove. So difficult, in fact, that the Conjecture's rigour was shrouded in arduousness for the next 100 years. 1953 marked a breakthrough discovery that would eventually lead to a comprehensive proof of Kepler's renowned Conjecture; the Hungarian mathematician László Fejes Tóth (who constructed the optimality proof for $\mathbb{R}^2$) showed that the problem of determining the maximum density of all arrangements, in conformity to lattices or not, could be reduced to a finite (but very large) number of calculations. This means that a proof by exhaustion was, in principle, possible and Fejes Tóth theorized that a fast enough computer could turn this theoretical result into a practical approach to solving the problem. Fejes Tóth was a scholar of the future for the rapid advancement in processor speed meant that iterating over all possible arrangements was possibly before even the turn of the century. Deciding to follow the approach outlined by Fejes Tóth, an American mathematician at the University of Michigan, Thomas Hales, determined that the maximum density of all arrangements could be found by minimizing a function of 150 variables. In 1992, assisted by his graduate student Samuel Ferguson, Hales took on the task of systematically applying linear programming methods to find a lower bound on the value of this function for each one of a set of over 5,000 different configurations of spheres. If a lower bound for the function value could be found for every one of these configurations that was greater than the value of the function for the conjectured optimal configuration (hexagonal close), then Kepler's Conjecture would be proved. Although a crude brute-forced method, it was rigorous enough for Hales to spend the next 6 years solving about 100,000 linear programming problems in aims of finding the lower bounds. In 1998, Hales announced that the proof was complete, offering 250 pages of notes and 3 gigabytes of computer programs, data and results. The mathematical community was at that point "99% certain" of the correctness of the proof, but not all of the computer calculations could be verified. In January 2003, Hales announced the inception of a collaborative project called *Flyspeck* - the F, P and K standing for *Formal Proof of Kepler*. In 2012, Hales published a "blueprint" for the formal proof of the Conjecture. The completion of the project was

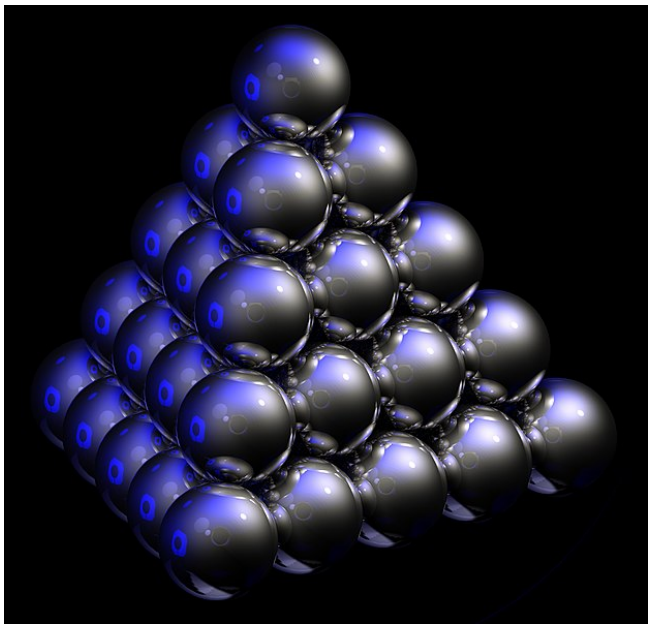


Figure 1.3: **Sphere packing in $\mathbb{R}^n$ which optimizes packing density: $\Delta(S) \approx 74\%$**

announced in August 2014. In January 2015, Hales and his fellow collaborators submitted a paper titled "A formal proof of the Kepler conjecture". In 2017, the formal proof was accepted into the Forum of Mathematics journal, marking the end of a strenuous 406 year long journey seeking rigour in simplicity. As we progress to generality, we lose access to the tangible dimensions $n \leq 3$ which support comprehensible visualizations, but sphere packing in higher dimensions is still of great mathematical interest. While the problem of sphere packing in $n = 4$ is more recent, it is still at least a century old and is part of problem 18 in the famous list of 23 problems presented by David Hilbert in 1900. At present, the sphere packing problem has not been solved for dimensions $n \geq 4$, with the exception of $n = 8$ and $n = 24$.

1.2 Lattice Theory Fundamentals

To explore sphere packing through the lens of lattices, one would first need to establish what is meant by a lattice. Closely related is the *hyperplane*, defined below.

Definition 1.2.1. A hyperplane is a subspace whose dimension is one less than that of its ambient space. In the case of $\mathbb{R}^2$, a hyperplane is a straight line.

Definition 1.2.2. A set of points Λ is a lattice if and only if Λ is a discrete additive subgroup of $\mathbb{R}^2$ not contained in any hyperplane.

It's difficult to work with this definition however, we need to translate it to a numerical system that allows for fundamental operations so that we can manipulate lattices in a succinct manner. Consider two linearly independent vectors $v_1, v_2 \in \mathbb{R}^2$. A lattice Λ is then composed of all integer combinations of v_1 and v_2.

$$\Lambda := \{\alpha \cdot v_1 + \beta \cdot v_2 \mid \alpha, \beta \in \mathbb{Z}\}$$

We can combine the two lattice vectors into a single matrix $A = \begin{pmatrix} | & | \\ v_1 & v_2 \\ | & | \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. This matrix is then called the *basis of lattice* Λ. A is invertible unless the two vectors are linearly dependent, in which case the vectors form a hyperplane in $\mathbb{R}^2$ and so do not correspond to a lattice. One can then represent the collection of integer combinations by multiplying the aforementioned matrix by all vectors $\gamma = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} : \alpha, \beta \in \mathbb{Z} \implies \gamma \in \mathbb{Z}^2$, i.e. two-dimensional column vectors with integer elements.

Definition 1.2.3. For invertible generator matrix $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, a lattice Λ is defined by

$$\Lambda = \{A\gamma \mid \gamma \in \mathbb{Z}^2\}.$$

For any given lattice Λ, there are infinitely many choices for the basis matrix as if (w_1, w_2) is a basis for a lattice then so too is $(aw_1 + bw_2, cw_1 + dw_2)$ for infinitely many $a, b, c, d \in \mathbb{Z}$. This idea of switching a basis to another one that is contained within the span of $\{A\gamma \mid \gamma \in \mathbb{Z}^2\}$ can be condensed into *elementary column operations*, i.e. switching columns, adding columns, and multiplying columns by some non-zero integer scalar. Elementary column (and row) operations can be represented by the group action of *unimodular matrices*, i.e. matrices with determinant ± 1. This then gives rise to the following lemma.

Lemma 1.2.1. Let A and B be two bases. Let $\Lambda(A)$ and $\Lambda(B)$ denote the lattices generated by A and B respectively. Then, $\Lambda(B) = \Lambda(A)$ if and only if there exists a unimodular integer matrix $M \in GL_2(\mathbb{Z})$ such that $B = AM$.

Proof. The lemma makes a claim of bijection, meaning that we need to prove the relation in both directions. First, assume that $B = AM$ for some unimodular matrix $M \in GL_2(\mathbb{Z})$. Then,

$$\Lambda(B) \subseteq \Lambda(A). \tag{1.1}$$

This is because all the column vectors of B can be written as linear combinations of the columns of A, and so the set of all integer combinations of the columns of B (i.e. the lattice $\Lambda(B)$) is contained within the set of all integer combinations of the columns of A. Note that if M is unimodular, so too is M^{-1} as $\det(M^{-1}) = \frac{1}{\det(M)} = \frac{1}{\pm 1} = \pm 1$. Thus, $A = BM^{-1}$, and so

$$\Lambda(A) \subseteq \Lambda(B). \tag{1.2}$$

(1.1) and (1.2) are only possible if the two generated lattices are equal. Thus:

$$\Lambda(A) = \Lambda(B)$$

Now assume that $\Lambda(B) = \Lambda(A)$. Then there exist integer matrices M and M' such that:

$$B = AM \qquad \text{and} \qquad A = BM'$$

It follows that:

$$\begin{aligned} & B = AM = (BM')M = B(M'M) \\ & \Rightarrow B = B(M'M) \\ & \Rightarrow M'M = I_2 \\ & \Rightarrow \det(M'M) = \det(I_2) \\ & \Rightarrow \det(M') \cdot \det(M) = 1. \end{aligned}$$

This is only possible in two cases:

- $\det(M) = \det(M') = 1$, or
- $\det(M) = \det(M') = -1$.

In both cases, $|\det(M)| = |\det(M')| = 1$, and thus M and M' are unimodular lattices, proving the bijection. □

We took care to note that generator matrix A for lattice Λ ought to be invertible, i.e. it must have a non-zero determinant. This has a natural geometric interpretation when we recall that two vectors v_1, v_2 construct a parallelogram in the Euclidean plane $\mathbb{R}^2$ if v_1 and v_2 are linearly independent, see Figure 1.4. A non-invertible matrix would have determinant zero, corresponding to a tile area of zero, leading to mathematical degeneracy. This occurs when v_1 and v_2 are linearly dependent. Thus, the assumption of matrix invertibility is not as restrictive as may initially seem, for it simply allows a lattice to exist in the first place.

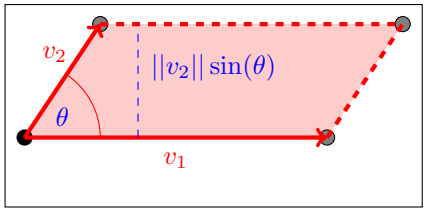


Figure 1.4: **Parallelogram tile formed by vectors $v_1, v_2 \in \mathbb{R}^2$.**

Lemma 1.2.2. The parallelogram described by an invertible 2×2 matrix A has area $|\det(A)|$.

Proof. Consider invertible matrix $A = \begin{pmatrix} | & | \\ v_1 & v_2 \\ | & | \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$.
The columns of A describe a parallelogram in $\mathbb{R}^2$, see Figure 1.4. Recall that the area of a parallelogram is given by base $\times$ height. Here, base$= ||v_1||$ and height$= ||v_2|| \sin(\theta)$. Thus:

$$\text{Area} = ||v_1|| \cdot ||v_2|| \sin(\theta)$$

which corresponds to the definition of the *cross product* $v_1 \times v_2$. Consider the parallelogram to be embedded in the three-dimensional space where $z = 0$. Then:

$$\begin{aligned} |v_1 \times v_2| &= \begin{vmatrix} +\hat{i} & -\hat{j} & +\hat{k} \\ v_{1x} & v_{1y} & 0 \\ v_{2x} & v_{2y} & 0 \end{vmatrix} \\ &= \begin{vmatrix} +\hat{i} & -\hat{j} & +\hat{k} \\ a & c & 0 \\ b & d & 0 \end{vmatrix} \\ &= |+\hat{i}(c \cdot 0 - 0 \cdot d) - \hat{j}(a \cdot 0 - 0 \cdot b) + \hat{k}(a \cdot d - c \cdot b)| \\ &= |\det(A) \cdot \hat{k}| \\ &= |\det(A)|. \end{aligned}$$

□

Closely related to the determinant of a lattice (i.e. the determinant of its generator matrix) is the *density of a lattice*, defined as follows:

Definition 1.2.4. For lattice Λ generated by invertible matrix A,

$$\text{Density of } \Lambda = \frac{1}{|\det(A)|}$$

This gives us some measure of how many lattice points are squeezed into a particular area in $\mathbb{R}^2$. Recall that the parallelogram tiles formed by vectors $v_1, v_1 \in \mathbb{R}^2$ and their integer combinations completely span the two-dimensional plane, see Figure 1.5. Consider the set of all lattice points in the lattice Λ,

$$C = \{A\gamma \mid \gamma \in \mathbb{Z}^2\},$$

and the set of all tiles spanning the two-dimensional plane,

$$P = \left\{(v_1, v_2)\begin{pmatrix}\epsilon_1 \\ \epsilon_2\end{pmatrix} + \lambda \mid 0 \leq \epsilon_i < 1,\ \forall \lambda \in \Lambda\right\}$$

These two sets have equal cardinality, easily seen when we translate the parallelogram tiling (which spans $\mathbb{R}^2$) such that each parallelogram centre P_i is a lattice point, see Figure 1.6. Each lattice point has a distinct parallelogram tile surrounding it.

$$\begin{aligned}\text{Density } &= 1 \text{ lattice point per parallelogram area} \\ &= \frac{1}{|\det(A)|}.\end{aligned}$$

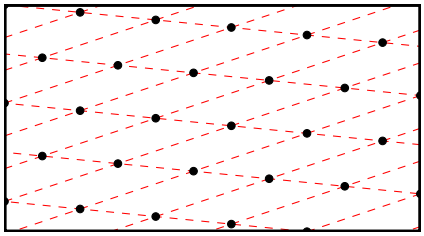


Figure 1.5: **Parallelogram tiles formed by lattice span the two-dimensional Euclidean plane $\mathbb{R}^2$.**

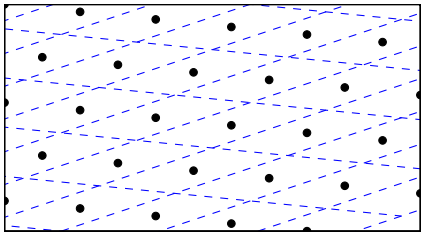
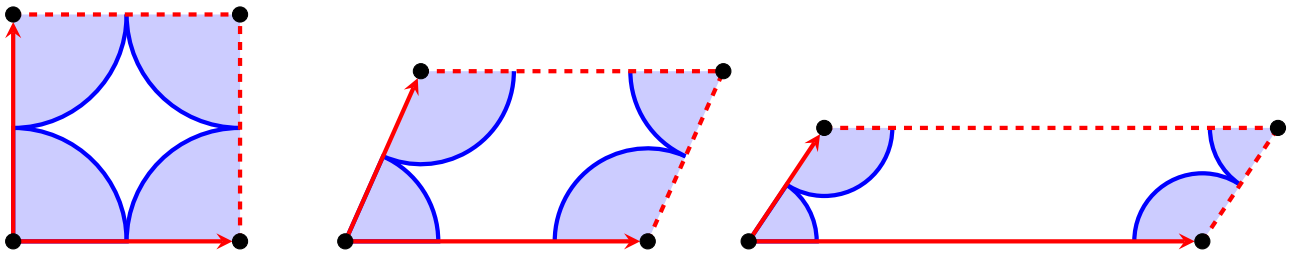


Figure 1.6: **Original lattice points with shifted parallelogram tiling.**

Lattices that have density 1 are called *unimodular lattices*, their generator matrices compose the *general linear group* of 2×2 matrices $GL_2(\mathbb{R})$. Along with density, lattices have another useful attribute called *minimal distance*.

Definition 1.2.5. The *minimal distance* of any lattice Λ is the minimum of $|\lambda|$ over non-zero $\lambda \in \Lambda$, i.e. the *minimal non-zero length of* Λ.



(a) $\Delta(S) = \frac{\pi}{4} \approx 0.7854$. (b) $\Delta(S) = \frac{97\pi}{576} \approx 0.5291$. (c) $\Delta(S) = \frac{13\pi}{144} \approx 0.2836$.

Figure 1.7: **Unimodular lattice tiles with varying minimal distances, yielding different packing densities.**

A lattice tile constructed by vectors v_1, v_2 depends on two lattice attributes: its determinant (corresponding to the area of the tile) and its minimal distance (corresponding to the shape of the tile). For any non-degenerate tile, one can place sphere sectors onto each tile vertex such that the area of each sphere is held within the tile (sphere packing clipped by the tile boundaries). With the interior angles of a parallelogram adding to 2π radians and a sphere sector of radius r placed at each parallelogram vertex, the combined area covered by the sphere sectors inside the tile sum to πr^2, where r is the common radius of each sphere sector. As such, we can derive an expression for the density of a sphere packing in terms of the minimal distance of a lattice Λ and the area of one of its tiles.

$$r = \frac{1}{2} \cdot \min_{\forall \lambda \in \Lambda}(\{|\lambda| : |\lambda| \neq 0\})$$

$$\Delta(S) = \frac{\pi r^2}{|\det(\Lambda)|}$$

We are interested in the packing of equal spheres, which means that we need to pick the maximum sphere radius such that there is no overlap of sphere sectors within the tile. This maximum radius length is *half of the minimal distance*, for should we attempt to use sphere sectors of a larger radius we result in overlap along the tile edges whose length are equal to the minimal distance of the lattice. If we can ensure that the density of a lattice is always within our control (using mathematics established in the next chapter), we can vary the packing density of a lattice by varying the minimal distance. In Figure 1.7, each lattice is unimodular but have different minimal distances. This results in sphere sectors of different radii and thus sphere packings of different densities. We see in Figure 1.7c that there is an abundance of horizontal empty space not covered by spheres while the tile sides whose length is equal to the minimal distance have touching spheres. In order to maximize packing density, we want there to be as little empty space as possible in the tile, achieved when all tile sides have length equal to the minimum distance of the lattice. This is also when the minimal distance is maximized. We thus establish the following lemma.

Lemma 1.2.3. In order to maximize the packing density of a lattice of density D, one needs to maximize the minimal distance.

1.3 Isometry and Homothecy

We've seen previously that multiple bases can generate the same lattice Λ. As such, it makes sense to consider certain generator matrices to be equivalent in the context of lattices. Similarly, it makes sense to consider certain lattices to be equal if they are transformations of each other such that the associated sphere packing density is unaffected by said transformations. These *density-preserving transformations* allow us to transform lattices in a certain manner that simplifies calculations for us or removes the need to concern oneself with certain variables. These transformations constitute two groups (although one is a subgroup of the other) called ***isometry*** and ***homothecy***.

Definition 1.3.1. A lattice Λ' is isometric with lattice Λ if it is obtained from Λ by an orthogonal linear transformation of $\mathbb{R}^2$.

Definition 1.3.2. A lattice Λ' is homothetic if it is isometric with $\alpha\Lambda$ for some $\alpha > 0$.

Should one set $\alpha = 1$ in the homothetic equivalence relation, one simply results with isometry. As such, isometry is a special case of homothecy wherein the scaling constant α is 1.

Isometry relies on orthogonal linear transformations of $\mathbb{R}^2$. These transformations form the *orthogonal group* $O_2(\mathbb{R})$. This is the group of 2×2 *orthonormal* matrices where for all $Q \in O_2(\mathbb{R})$,

$$Q^T Q = QQ^T = I_2$$

where Q^T is the transpose of matrix Q and I_2 is the 2×2 identity matrix. This leads to an alternative characterization of the orthogonal group: a matrix Q is orthogonal if its transpose is equal to its inverse.

$$Q^T = Q^{-1}.$$

For some 2×2 matrix $Q = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ to be orthogonal, the following mst hold true.

$$QQ^T = I_2$$
$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} a & c \\ b & d \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$
$$\begin{bmatrix} a^2 + b^2 & ac + bd \\ ac + bd & c^2 + d^2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

And so, orthogonality demands that the matrix Q satisfies the three equations

$$a^2 + b^2 = 1,$$
$$c^2 + d^2 = 1,$$
$$ac + bd = 0.$$

Recall that $Q \in \mathbb{R}^2$, wherein each column vector of Q corresponds to some vector coordinate in $\mathbb{R}^2$. Without loss of generality, one can let $a = \cos(\theta)$ and $c = \sin(\theta)$. Then:

$$\cos^2(\theta) + b^2 = 1 \Rightarrow b^2 = 1 - \cos^2(\theta), \tag{1.3}$$
$$\sin^2(\theta) + d^2 = 1 \Rightarrow d^2 = 1 - \sin^2(\theta), \tag{1.4}$$
$$\cos(\theta)\sin(\theta) + bd = 0 \Rightarrow bd = -\cos(\theta)\sin(\theta). \tag{1.5}$$

Recall the Pythagorean trigonometric identity: $\cos^2(\theta)+\sin^2(\theta)=1$. One can rearrange the identity to express $\cos(\theta)$ in terms of $\sin(\theta)$ and vice versa.

$$\sin^2(\theta) = 1-\cos^2(\theta),$$
$$\cos^2(\theta) = 1-\sin^2(\theta).$$

Applying these manipulations to equations (1.3) and (1.5) results in the following.

$$b^2=\sin^2(\theta) \Rightarrow b=\sin(\theta)$$
$$b=-\sin(\theta)$$

$$d^2=\cos^2(\theta) \Rightarrow d=\cos(\theta)$$
$$d=-\cos(\theta).$$

Equation (1.5) is used to determine the signs. The product bd must be negative, which is only possible if b and d have opposite signs. This results in two possibilities for orthogonal matrix transformations, outlined below.

$$\begin{bmatrix}\cos(\theta) & -\sin(\theta)\\ \sin(\theta) & \cos(\theta)\end{bmatrix} := \text{Rotation} \qquad \begin{bmatrix}\cos(\theta) & \sin(\theta)\\ \sin(\theta) & -\cos(\theta)\end{bmatrix} := \text{Reflection}$$

We can represent the orthogonal group $O_2(\mathbb{R})$ by the union of these two transformations types.

$$\boldsymbol{O_2(\mathbb{R}) := \left\{\begin{bmatrix}\cos(\theta) & -\sin(\theta)\\ \sin(\theta) & \cos(\theta)\end{bmatrix} \forall \theta \in [0,2\pi)\right\} \cup \left\{\begin{bmatrix}\cos(\theta) & \sin(\theta)\\ \sin(\theta) & -\cos(\theta)\end{bmatrix} \forall \theta \in [0,2\pi)\right\}}$$

The subgroup $SO_2(\mathbb{R})$ consisting of orthogonal matrices with determinant $+1$ is called the *special orthogonal group*. Each element in the special orthogonal group is called a *special orthogonal matrix*. Note that for reflective transformations Q of the form

$$Q = \begin{bmatrix}\cos(\theta) & \sin(\theta)\\ \sin(\theta) & -\cos(\theta)\end{bmatrix},$$

$|\det(Q)|$ is always negative. This means that every special orthogonal matrix acts on $\mathbb{R}^2$ via rotation as a reflective transformation would require a negative matrix determinant - not possible for special orthogonal matrices.

Lattices can be generated by matrices, their generator matrices can be multiplied by unimodular matrices $M \in GL_2(\mathbb{Z})$ in order to obtain a new basis, and they can be orthogonally transformed by matrices $U \in O_2(\mathbb{R})$. With these parallels being made between lattices and matrices, we would also like to define isometry and homothecy in terms of matrix transformations.

Definition 1.3.3. The lattices generated by A and A' are isometric if and only if

$$A' = UAM$$

for some $M \in GL_2(\mathbb{Z})$ and $U \in O_2(\mathbb{R})$.

Definition 1.3.4. The lattices generated by A and A' are homothetic if and only if

$$A' = \alpha UAM$$

for some $\alpha > 0$, $M \in GL_2(\mathbb{Z})$, and $U \in O_2(\mathbb{R})$.

We've established that isometry represents rotational and reflective transformations while homothecy adds on linear scaling. It is clear why these transformations preserve sphere packing density when we observe how they affect lattices in $\mathbb{R}^2$.
With isometric transformations, a lattice $\Lambda \subset \mathbb{R}^2$ experiences a change in its orientation relative to the origin and naught else. Such transformations do not affect the sphere packing density as each sphere experiences the same orthogonal transformation.
With homothecy, a lattice Λ can experience an isometric transformation but also linear scaling by some $\alpha > 0$. Should this occur, the radius of each sphere shall also experience the same linear scaling that is applied to the lattice itself, leaving the sphere packing density unaffected. For $A' = \alpha UAM$:

$$\begin{aligned}|\det(A')| &= |\det(\alpha UAM)|\\ &= |\det(U)\cdot\det(\alpha A)\cdot\det(M)|\\ &= 1\cdot|\det(\alpha A)|\cdot 1\\ &= \alpha^2|\det(A)|.\end{aligned}$$

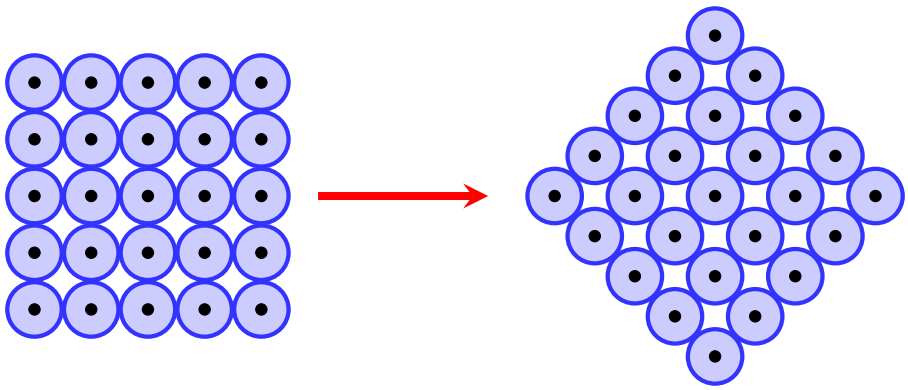


Figure 1.8: **Isometry - $\mathbb{Z}^2$ integer lattice packing under action by rotational transformation.**

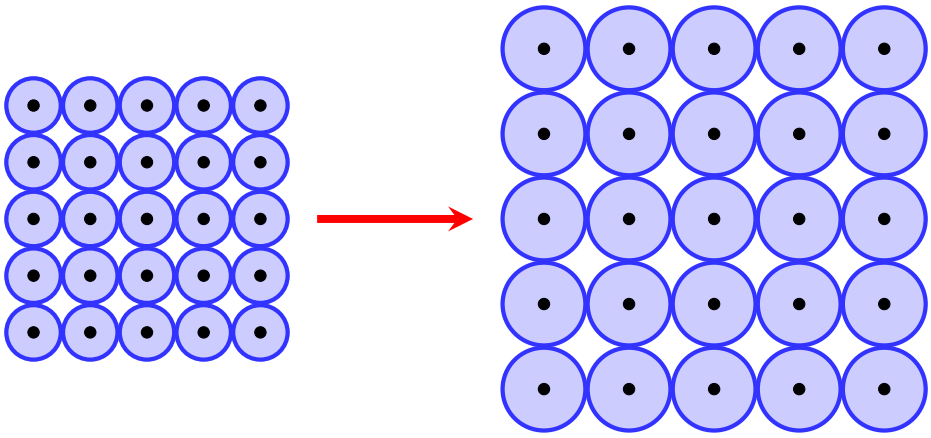


Figure 1.9: **Homothecy - $\mathbb{Z}^2$ integer lattice packing under action by scaling transformation.**

Thus, for all generator matrices $A \in M_2(\mathbb{R})$ there exists a unique α, specifically $\alpha = |\det(A)|^{-1/2}$, such that lattice associated to the basis $A' = \alpha UAM$ is unimodular, i.e.

$$\det(A') = \pm 1$$
$$|\det(A')| = 1.$$

Due to the distance-preserving feature of homothecy, the generator matrix A' produces a lattice which has the same sphere packing density as the lattice generated by A, but now each parallelogram tile mapped out by the lattice vectors has an area of 1. One can now choose A and U to have determinants $+1$, i.e. for A to lie in the *special linear group* $SL_2(\mathbb{R})$ and for U to lie in the *special orthogonal group* $SO_2(\mathbb{R})$. We then result in the *double coset space* $SO_2(\mathbb{R}) \setminus SL_2(\mathbb{R})/GL_2(\mathbb{Z})$, and a special set of lattices $\mathbf{\Lambda_2}$.

Definition 1.3.5. Let G be a group, and let H and K be subgroups. Let H act on G by left multiplication and let K act on G by right multiplication. For each $x \in G$, the ***(H,K)*-double coset of x** is the set

$$HxK = \{hxk \mid h \in H, k \in K\}.$$

The set of all double cosets is then denoted by $H \setminus G/K$.

Definition 1.3.6. The set $\mathbf{\Lambda_2}$ is the set of homothecy classes of lattices in $\mathbb{R}^2$ with the double coset space $SO_2(\mathbb{R}) \setminus SL_2(\mathbb{R})/GL_2(\mathbb{Z})$.
This is also the space of isometry classes of lattices of density 1, i.e. *unimodular lattices.*

The "double coset space" is a condensed way of referring to all possible permutations of UAM for $U \in SO_2(\mathbb{R})$, $A \in SL_2(\mathbb{R})$, $M \in GL_2(\mathbb{Z})$. We can utilize isometry and homothecy to reduce the form that a generator matrix $A' \in SO_2(\mathbb{R}) \setminus SL_2(\mathbb{R})/GL_2(\mathbb{Z})$ takes. Consider an invertible matrix

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

The columns of A can be interpreted as coordinate vectors in the complex plane. For any pair of vectors w_1, w_2, it is always possible to apply orthogonal transformations (rotation and reflection) such that w_1 lies along the positive real axis.

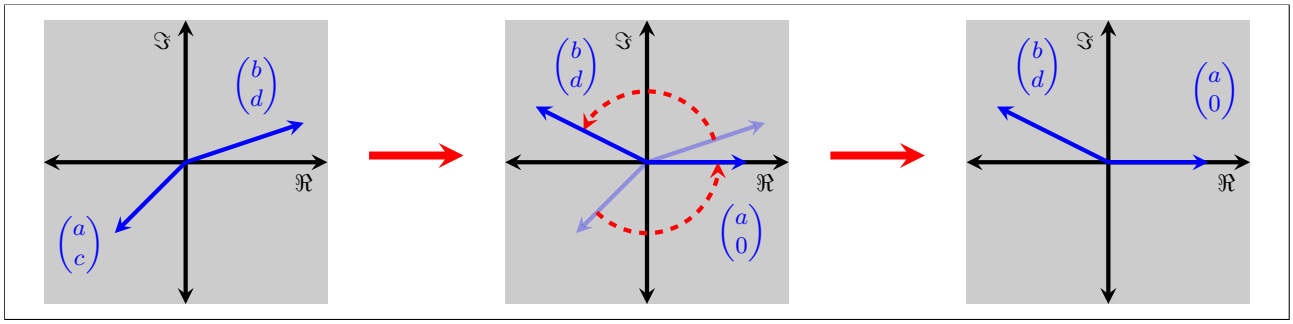


Figure 1.10: **Applying orthogonal rotation to arbitrary lattice basis.**

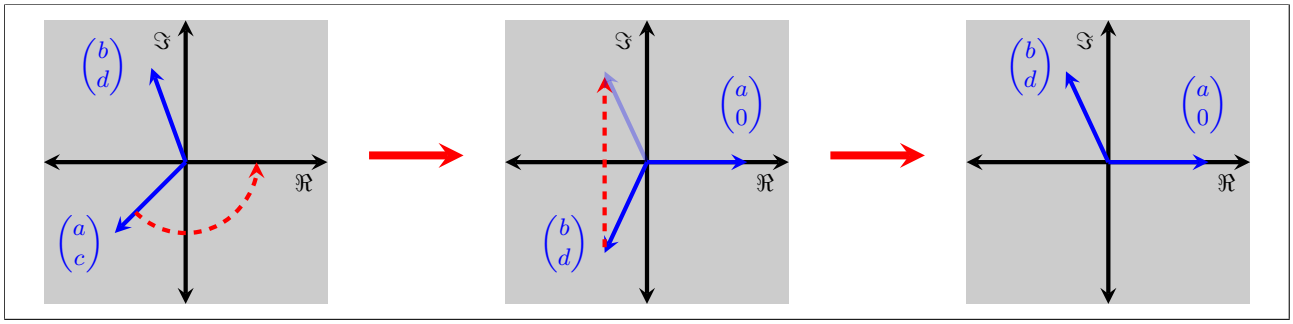


Figure 1.11: **Applying orthogonal rotation and reflection to arbitrary lattice basis.**

Upon applying such transformations, our column vectors are now of the following form.

$$w_1 = \begin{pmatrix} a \\ 0 \end{pmatrix} \qquad w_2 = \begin{pmatrix} b \\ d \end{pmatrix}$$
$$A = \begin{pmatrix} a & b \\ 0 & d \end{pmatrix}.$$

Recall that the basis lies in the special linear group; $A \in SL_2(\mathbb{R})$.

$$\det(A) = +1 \Rightarrow ad - b \cdot 0 = +1 \Rightarrow d = 1/a.$$

Our new basis is now of the form,

$$\boxed{A = \begin{pmatrix} a & b \\ 0 & 1/a \end{pmatrix}}$$

What was once a matrix form of four unknown variables is now a basis of two unknown variables. These two variables can be manipulated in a certain manner such that there is a bijection or an *isomorphism* between our new simplified basis form and the *complex upper half-plane* $\mathcal{H}$.

$$\mathcal{H} := \{\tau = x + iy \mid y > 0\}.$$

1.4 Isomorphism With Complex Upper Half-Plane

Consider a lattice $\Lambda \subset \mathbb{R}^2$ with basis

$$A = (w_1, w_2)$$

Theorem 1.4.1. *For each lattice $\Lambda \subset \mathbb{R}^2$, there exists a basis (w_1, w_2) such that the ratio $\tau = w_2/w_1$ lies in the region $\mathcal{F}$ satisfying the following conditions:*

1. $\Im(\tau) > 0$
2. $-\frac{1}{2} < \Re(\tau) \leq \frac{1}{2}$
3. $|\tau| \geq 1$
4. $\Re(\tau) \geq 0$ *if* $|\tau| = 1$.

The ratio τ is uniquely determined for each lattice by these conditions.

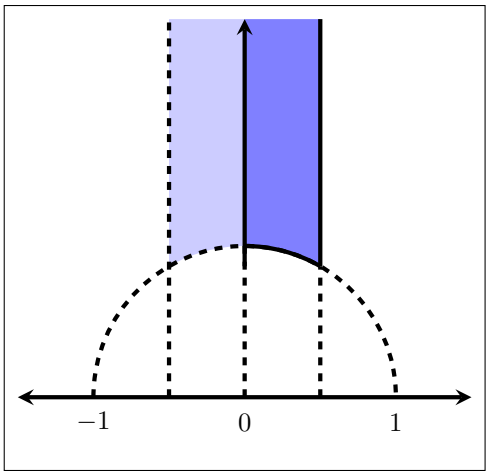


Figure 1.12: **Region defined by Theorem 1.4.1**

Proof. The lattices generated by bases (w_1, w_2) and (w_2, w_1) are identical. Thus, without loss of generality, one can say that

$$|w_1| \leq |w_2| \leq |w_1 \pm w_2|$$

i.e. the Euclidean length of one basis is greater than or equal to the length of the other, and the sum of the Euclidean lengths is greater than or equal to any one basis length. We then have two inequalities to work with:

$$\begin{aligned} &\text{Inequality (1)} : |w_1| \leq |w_2|, \\ &\text{Inequality (2)} : |w_2| \leq |w_1 \pm w_2|. \end{aligned}$$

One can manipulate Inequality (1) such that the vector moduli are grouped together.

$$\begin{aligned} |w_1| \leq |w_2| &\Rightarrow 1 \leq \frac{|w_2|}{|w_1|} \\ &\Rightarrow 1 \leq \left|\frac{w_2}{w_1}\right| \\ &\Rightarrow 1 \leq |\tau|. \end{aligned}$$

Condition 3 has been satisfied.
Should one divide both sides of Inequality (2) by $|w_1|$, one results in the following.

$$\begin{aligned} |w_2| \leq |w_1 \pm w_2| \Rightarrow \frac{|w_2|}{|w_1|} \leq \frac{|w_1 \pm w_2|}{|w_1|} &\Rightarrow \left|\frac{w_2}{w_1}\right| \leq \left|\frac{w_1 \pm w_2}{w_1}\right| \\ &\Rightarrow |\tau| \leq |1 \pm \tau| \end{aligned}$$

Splitting τ into its real and imaginary components allows us to simplify the above expression further.

$$\begin{aligned} |\tau| \leq |1 \pm \tau| &\Rightarrow \sqrt{\Re(\tau)^2 + \Im(\tau)^2} \leq \sqrt{(1 + \Re(\tau))^2 + \Im(\tau)^2} \\ &\Rightarrow \Re(\tau)^2 + \Im(\tau)^2 \leq 1 + 2\Re(\tau) + \Re(\tau)^2 + \Im(\tau)^2 \\ &\Rightarrow 0 \leq 1 + 2\Re(\tau) \Rightarrow 0 \leq \frac{1}{2} + \Re(\tau) \\ &\Rightarrow |\Re(\tau)| \leq \frac{1}{2}. \end{aligned}$$

Condition 2 has been satisfied.
Even with the above condition, it is possible that $\Re(\tau) = -\frac{1}{2}$. In such a case, one can replace (w_1, w_2) by $(w_1, w_1 + w_2)$. This, too, is an elementary column operation (addition of two rows), corresponding to the following matrix transformation.

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \in GL_2(\mathbb{Z}).$$

Condition 2 has been satisfied.
For $\tau = \frac{w_2}{w_1}$, it may be possible that $\Im(\tau)$ is negative. Then, one can replace (w_1, w_2) by $(-w_1, w_2)$, ensuring that $\Im(\tau) > 0$. This basis manipulation is an elementary column operation (multiplication by non-zero scalar), and corresponds to the following matrix transformation.

$$\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \in GL_2(\mathbb{Z})$$

Condition 1 has been satisfied.
Should it be the case that $|\tau| = 1$ with $\Re(\tau) < 0$, one can replace (w_1, w_2) with $(-w_2, w_1)$, yet again an elementary column operation corresponding to the following matrix transformation.

$$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \in GL_2(\mathbb{Z}).$$

Condition 4 has been satisfied. One is now required to prove that τ is uniquely determined for each lattice. Recall that a change of basis for lattices is carried out by applying a matrix transformation from the general linear group (also known as the unimodular group) $GL_2(\mathbb{Z})$. The theorem claims that if τ and τ' are in the region $\mathcal{F}$, and $\tau' = M\tau$ for some $M \in GL_2(\mathbb{Z})$ then $\tau = \tau'$, making τ unique for each lattice. The unimodular group $GL_2(\mathbb{Z})$ acts on the basis (w_1, w_2) via matrix multiplication as follows.

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} w_2 \\ w_1 \end{pmatrix} = \begin{pmatrix} aw_2 + bw_1 \\ cw_2 + dw_1 \end{pmatrix}.$$

Similarly, it acts on the quotient $\tau = w_2/w_1$ via a linear fractional transformation.

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot \tau = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot \frac{w_2}{w_1} = \frac{aw_2 + bw_1}{cw_2 + dw_1} = \frac{a\frac{w_2}{w_1} + b\frac{w_1}{w_1}}{c\frac{w_2}{w_1} + d\frac{w_1}{w_1}}$$
$$\Rightarrow \frac{a\tau + b}{c\tau + d}.$$

Suppose that $\tau' = \frac{a\tau+b}{c\tau+d}$ for some $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL_2(\mathbb{Z})$. Then:

$$\begin{aligned}
\Im(\tau') = \Im\left(\frac{a\tau + b}{c\tau + d}\right) &= \Im\left(\frac{(a\tau + b)(\overline{c\tau + d})}{|c\tau + d|^2}\right) \\
&= \frac{\Im([(ax + b) + aiy][(cx + d) - ciy])}{|c\tau + d|^2} \\
&= \frac{\Im((ax + b)(-ciy) + (aiy)(cx + d))}{|c\tau + d|^2} \\
&= \frac{\Im(-aiycx + aiycx - bciy + daiy)}{|c\tau + d|^2} \\
&= \frac{\Im(iy(ad - bc))}{|c\tau + d|^2} \\
&= \frac{\Im(iy)}{|c\tau + d|^2} \\
&= \frac{\Im(\tau)}{|c\tau + d|^2}.
\end{aligned}$$

Without loss of generality, one can say that $\Im(\tau') \geq \Im(\tau)$. It follows that $|c\tau + d| \leq 1$. Because c and d are integers, the number of possible cases for c, d such that the inequality holds true are finite.
Take $c = 0$. Then:

$$|c\tau + d| \leq 1 \Rightarrow |d| \leq 1 \Rightarrow d = \pm 1 \text{ or } d = 0.$$

Recall that $ad - bc = 1$. With $c = 0$, $ad = 1 \Rightarrow d \neq 0$. This leaves only two possibilities.

$$a = d = 1, \quad \text{or}$$
$$a = d = -1.$$

This corresponds to the following forms for τ'.

$$\text{If } a = d = 1: \quad \begin{pmatrix} 1 & \pm b \\ 0 & 1 \end{pmatrix} \cdot \tau = \frac{\tau \pm b}{1} = \tau \pm b$$

$$\text{If } a = d = -1: \quad \begin{pmatrix} -1 & \pm b \\ 0 & -1 \end{pmatrix} \cdot \tau = \frac{-\tau \pm b}{-1} = \tau \pm b.$$

In both cases, $\tau' = \tau \pm b$, implying that

$$\pm b = \tau' - \tau \Rightarrow |b| = |\Re(\tau') - \Re(\tau)|$$

Recall that the region $\mathcal{F}$ is only defined in the interval $-\frac{1}{2} < \Re(\tau) \leq \frac{1}{2}$, i.e. the region's width parallel to the real axis is less than 1. If $|\Re(\tau') - \Re(\tau)| \geq 1$, then the displacement of τ' relative to τ extends beyond the width of region $\mathcal{F}$. Recall also that $b \in \mathbb{Z}$. Then,

$$|b| = |\Re(\tau') - \Re(\tau)| < 1 \Rightarrow b \neq 1 \Rightarrow b = 0$$
$$\Rightarrow \boxed{\tau' = \tau}$$

Consider now the case wherein $c \neq 0$,

$$|c\tau + d| \leq 1 \Rightarrow \left|\tau + \frac{d}{c}\right| \leq \frac{1}{|c|}.$$

This implies that $|c| = 1$, for if $|c|$ where ≥ 2 then,

$$\left|\tau + \frac{d}{2}\right| \leq \frac{1}{2} \Rightarrow \Im(\tau) \leq \frac{1}{2}.$$

This is not possible if $\tau \in \mathcal{F}$ as the lowest possible height for an element $\tau \in \mathcal{F}$ occurs when $|\tau| = 1$, where $\Im(\tau) = \frac{\sqrt{3}}{2} > \frac{1}{2}$. The inequality now becomes

$$\left|\tau + \frac{d}{c}\right| \leq \frac{1}{|c|} \Rightarrow |\tau + d| \leq 1. \tag{1.6}$$

This inequality is only satisfied when $d = \pm 1$ or $d = 0$.
Assume $|c| = 1,\ d = +1:\ |\tau + d| \leq 1 \Rightarrow |\tau + 1| \leq 1$. Recall that $|\tau|$ cannot be less than 1, therefore $|\tau + 1| = 1$. This implies that $|\tau| = 1$, and thus in order to satisfy the condition we require some τ such that $|\tau| = |\tau + 1| = 1$. This is only satisfied when $\tau = e^{2\pi i/3}$ which is not in the region $\mathcal{F}$. Thus, the condition is never satisfied and $d \neq +1$.
Assume $|c| = 1,\ d = -1:\ |\tau + d| \leq 1 \Rightarrow |\tau - 1| \leq 1$. With similar reasoning to the case wherein $d = +1$, we require some τ such that $|\tau| = |\tau - 1| = 1$. This is only satisfied when $\tau = e^{\pi i/3}$, which is in the region $\mathcal{F}$. Recall the relation between $\Im(\tau')$ and $\Im(\tau)$.

$$\Im(\tau') = \frac{\Im(\tau)}{|c\tau + d|^2}$$

Here, $|c\tau + d|^2 = |e^{\pi i/3} - 1|^2 = |1| = 1$. Thus:

$$\Im(\tau') = \Im(\tau).$$

No other $\tau \in \mathcal{F}$ exists such that $|\tau| = |\tau - 1| = 1$, thus:

$$\boxed{\tau' = \tau}$$

There remains only the case wherein $|c| = 1,\ d = 0$.

$$|c\tau + d| \leq 1 \Rightarrow |\tau| \leq 1.$$

Recall that for $\tau \in \mathcal{F}$, $|\tau| \geq 1$. Therefore, $1 \leq |\tau| \leq 1$, i.e. $|\tau| = 1$. Recall also that $ad - bc = 1$. It follows that,

$$bc = -1 \Rightarrow \frac{b}{c} = -1$$

as $c = \pm 1$. Then,

$$\tau' = \frac{a\tau + b}{c\tau + d} \Longrightarrow \tau' = \frac{a\tau + b}{c\tau} = \frac{a\tau}{c\tau} + \frac{b}{c\tau} \Longrightarrow \tau' = \frac{a}{c} - \frac{1}{\tau} \Longrightarrow \tau' = \pm a - \frac{1}{\tau} \Longrightarrow \tau' = \pm a - \overline{\tau}.$$

Hence, $\tau' + \overline{\tau} = \pm a \Longrightarrow \Re(\tau' + \tau) = \pm a$ where $a \in \mathbb{Z}$. Recall that $-\frac{1}{2} < \Re(\tau) \leq \frac{1}{2}$. There exist no two real numbers $a, b \in \left(-\frac{1}{2}, \frac{1}{2}\right]$ such that $|a + b| \geq 1$. Thus, the condition is only satisfied when $a = 0$, in which case

$$\tau' = \pm a - \overline{\tau} \Longrightarrow \tau' = -\overline{\tau} \Longrightarrow \Re(\tau') = -\Re(\tau).$$

Recall that for $\tau \in \mathcal{F}$, if $|\tau| = 1$ then $\Re(\tau) \geq 0$. This means that

$$\Re(\tau') = -\Re(\tau) \leq 0$$

In this case, $|\tau'| = |-\overline{\tau}| = |\tau| = 1$. Therefore:

$$\Re(\tau') \geq 0$$

These two inequalities are only satisfied when $\Re(\tau') = \Re(\tau) = 0$ and, as a consequence of the fact that $|\tau| = 1$, the condition is only fulfilled when $\tau' = \tau = +i$. Therefore, τ is unique. This completes the proof. □

The region defined by Theorem 1.4.1 (see Figure 1.12) is called the *fundamental region of unimodular lattices*. This region converges at $+i\infty$ making it closed and bounded, i.e. *compact*. It is defined by

$$\mathcal{F} := \{\infty\} \cup \left\{\tau \in \mathcal{H} : -\frac{1}{2} < \Re(\tau) \leq \frac{1}{2} \text{ if } |\tau| > 1;\ \Re(\tau) > 0 \text{ if } |\tau| = 1\right\}.$$

Theorem 1.4.1 allows us to send an arbitrary lattice bases $A = (w_1, w_2)$ to a unique element in the compact region $\mathcal{F}$. We can construct a more specific isomorphism by applying isometry and homothecy to said arbitrary basis and feeding it to the theorem.

Key Result 1. There exists an isomorphism between unimodular lattices generated by bases in the double coset space $SO_2(\mathbb{R}) \setminus SL_2(\mathbb{R})/GL_2(\mathbb{R})$ and the compact region in the complex upper half-plane $\mathcal{H}$ defined by

$$\mathcal{F} = \{\infty\} \cup \left\{\tau \in \mathcal{H} : -\frac{1}{2} < \Re(\tau) \leq \frac{1}{2} \text{ if } |\tau| > 1;\ \Re(\tau) > 0 \text{ if } |\tau| = 1\right\}.$$

Proof. All lattice-generating bases in the double coset space $SO_2(\mathbb{R}) \setminus SL_2(\mathbb{R})/GL_2(\mathbb{R})$ have a matrix representation of the following form:

$$A = (w_1, w_2) = \begin{pmatrix} a & b \\ 0 & 1/a \end{pmatrix}$$

By Theorem 1.4.1, the ratio $\tau = w_2/w_1$ is in the compact region $\mathcal{F}$.

$$w_1 = \begin{pmatrix} a \\ 0 \end{pmatrix} = a + 0 \cdot i,$$
$$w_2 = \begin{pmatrix} b \\ 1/a \end{pmatrix} = b + \frac{1}{a} \cdot i$$
$$\tau = \frac{w_2}{w_1} = \frac{b + \frac{1}{2} \cdot i}{a} = \frac{b}{a} + \frac{1}{a^2} \cdot i$$

Recall that $\tau = x + iy \in \mathcal{H}$. By matching the real and complex components of the ratio $\frac{w_2}{w_1}$ and the arbitrary form of τ, one results in the following.

$$\frac{1}{a^2} = y \Rightarrow \boxed{a = y^{-1/2}}$$
$$\frac{b}{a} = x \Rightarrow b = ax \Rightarrow \boxed{b = xy^{-1/2}}.$$

Thus, all $\tau \in \mathcal{F} \subset \mathcal{H}$ have an associated lattice generated by the basis

$$\boxed{A = \begin{pmatrix} y^{-1/2} & xy^{-1/2} \\ 0 & y^{1/2} \end{pmatrix} = y^{-1/2} \begin{pmatrix} 1 & x \\ 0 & y \end{pmatrix}}$$

□

1.5 Packing Optimality Using Fundamental Region Isomorphism

We've seen previously that in order to maximize the packing density of a lattice, one ought to maximize its minimal distance. For a basis obtained from Key Result 1,

$$A = \begin{pmatrix} y^{-1/2} & xy^{-1/2} \\ 0 & y^{1/2} \end{pmatrix},$$

one can generate a unimodular lattice using the generator columns.

$$v_1 = \begin{pmatrix} y^{-1/2} \\ 0 \end{pmatrix} \qquad v_2 = \begin{pmatrix} xy^{-1/2} \\ y^{1/2} \end{pmatrix}$$

One can then identify the lattice distances obtained from these generator columns and in doing so obtain the sphere packing density associated with the lattice. Figure 1.13 outlines the distances which are candidates for the *minimal distance* of the lattice.

Consider the Euclidean lengths of each of the generator vectors.

$$|v_1| = \sqrt{\left(y^{-1/2}\right)^2 + 0^2} = y^{-1/2},$$
$$|v_2| = \sqrt{\left(xy^{-1/2}\right)^2 + \left(y^{1/2}\right)^2} = \sqrt{x^2y^{-1} + y}.$$

One ought to also consider the distance between the coordinates $(y^{-1/2}, 0)$ and $(xy^{-1/2}, y^{1/2})$.

$$c^2 = \left[y^{-1/2} - xy^{-1/2}\right]^2 + \left[y^{1/2}\right]^2 \Longrightarrow c = \sqrt{y^{-1}(1-x)^2 + y}.$$

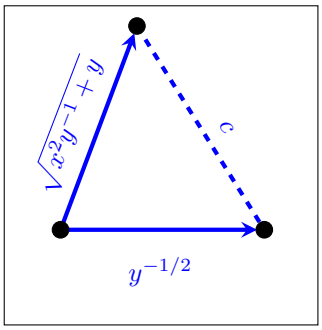


Figure 1.13: **Distances in a unimodular lattice obtained from basis utilizing Key Result 1.**

We have three possible (squared) distances that could be the minimal distance of the lattice.

$$d_1 = \sqrt{x^2y^{-1} + y},$$
$$d_2 = y^{-1/2},$$
$$d_3 = \sqrt{y^{-1}(x-1)^2 + y}.$$

Note that all d_1, d_2, d_3 for all $(x, y) \in \mathcal{F}$. Note also that $d_1 \geq y^{1/2}$ and $d_3 \geq y^{1/2}$, while $d_2 \leq y^{1/2}$. Therefore, the shortest distance is always d_2, although d_1 and d_3 may be equal to d_2 for certain values of x and y. Therefore, $d_2 = y^{-1/2}$ is the *minimal distance* of the lattice associated with $\tau = x + iy \in \mathcal{F}$. One ought to maximize the minimal distance in order to find the optimal packing density. To maximize $y^{-1/2}$, one ought to minimize y. The minimal value of y is found at

$$\tau = e^{\pi i/3} = \cos\left(\frac{\pi}{3}\right) + i\sin\left(\frac{\pi}{3}\right) = \frac{1}{2} + i\frac{\sqrt{3}}{2},$$

where $y = \frac{\sqrt{3}}{2}$. The minimal distance of the associated unimodular lattice is $y^{-1/2} = \left(\frac{2}{\sqrt{3}}\right)^{1/2}$.
Recall that the density of a sphere packing S in terms of the minimal distance of a lattice Λ is

$$\Delta(S) = \frac{\pi r^2}{|\det(\Lambda)|}$$

where r is half the minimal distance of the lattice and $|\det(\Lambda)|$ is the determinant of the generator basis of the lattice, also corresponding to the area of the parallelogram tile constructed by the column vectors of the basis. Recall that the lattice associated with $\tau = x + iy \in \mathcal{F}$ is *unimodular*, i.e. $|\det(\Lambda)| = 1$. Therefore, the density of the sphere packing is

$$\Delta(S) = \frac{\pi r^2}{1} = \pi r^2 = \pi \cdot \left[\frac{1}{2}\left(\frac{2}{\sqrt{3}}\right)^{1/2}\right]^2 = \pi \cdot \frac{2}{4\sqrt{3}}$$

$$\boxed{\Delta(S) = \frac{\pi}{2\sqrt{3}} \approx 0.9068996821}$$

One can use $\tau = e^{\pi i/3}$ to obtain the generator matrix for this optimal sphere packing.

$$\tau = x + iy = \frac{1}{2} + i\frac{\sqrt{3}}{2}$$

$$x = \frac{1}{2} \qquad y = \frac{\sqrt{3}}{2}$$

$$A = \begin{pmatrix} y^{-1/2} & xy^{-1/2} \\ 0 & y^{1/2} \end{pmatrix} \approx \begin{pmatrix} 1.07457 & 0.53729 \\ 0 & 0.93060 \end{pmatrix}$$

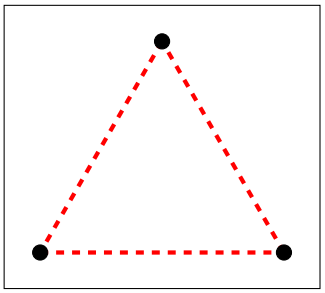


Figure 1.14: **Fundamental lattice points of optimal lattice basis in $\mathbb{R}^2$.**

Recall that for a basis A, a lattice $\Lambda \subset \mathbb{R}^2$ is defined by

$$\Lambda := \{A\gamma \mid \gamma \in \mathbb{Z}^2\}.$$

By taking all the integer combinations of the column vectors of the basis A, one can extend the fundamental lattice points to a full lattice and observe the resulting sphere packing.

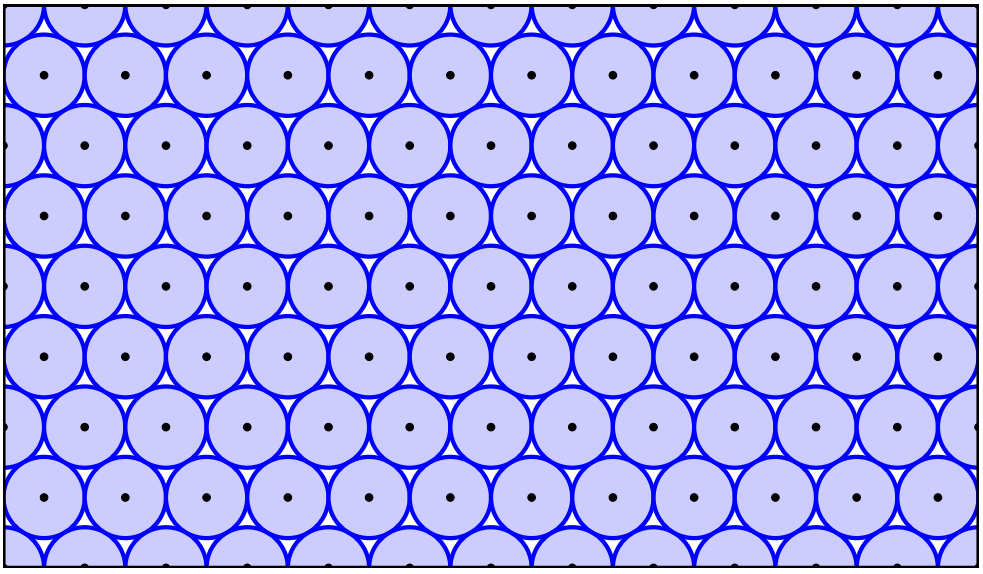


Figure 1.15: **Optimal lattice packing in $\mathbb{R}^2$. $\Delta(S) \approx 0.9069$.**

Chapter 2

Modular Forms

2.1 Theta Functions

The problem of sphere packing can be reduced to a problem of minimizing lattice distances by restricting the spheres in a manner such that their centres map out a lattice in $\mathbb{R}^2$. We've established a method of identifying the *minimal distance* of a lattice, and in doing so obtained an optimal solution for two-dimensional sphere packing. Should one wish to investigate lattice lengths beyond the minimal distance, then the outlined method falls short with its lack of generality. If one is to ask for *all* lengths of vectors of Λ and their multiplicities, one may wish to turn to *theta functions*. These are generating functions of the squared lengths of a lattice, also known as *theta series*, defined by

$$\Theta_\Lambda(z) := \sum_{x\in\Lambda} z^{(x,x)},$$

where (x,x) is the dot product of x with itself, yielding the square length of the vector $x \in \Lambda$. For any lattice $\Lambda \subset \mathbb{R}^n$, there is just one vector of square length 0. Thus:

$$\Theta_\Lambda(z) = 1 + \sum_{m>0} N_m(\Lambda) z^m,$$

where $N_m(\Lambda) = \#\{x \in \Lambda \mid (x,x) = m\}$ is the number of lattice vectors of length $\sqrt{m}$.

While the definition is plagued by obscurity, theta functions become easier to comprehend upon the exploration of an example. Consider the lattice $\mathbb{Z} \subset \mathbb{R}$. There is one lattice vector of length 0: the vector from 0 to itself. There are two vectors of length 1: $0 \to -1$ and $0 \to 1$. There are two vectors of length 2: $0 \to -2$ and $0 \to 2$. In fact, for each $m \in \mathbb{N}$, there are exactly two vectors of length m in the lattice $\mathbb{Z} \subset \mathbb{R}$. Thus, our theta function for $\mathbb{Z}$ is

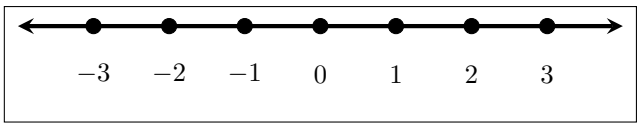


Figure 2.1: **Lattice $\mathbb{Z} \subset \mathbb{R}$.**

$$\begin{aligned}
\Theta_\mathbb{Z}(z) &= 1 + 2z^{(1,1)} + 2z^{(2,2)} + 2z^{(3,3)} + ... \\
\Rightarrow \Theta_\mathbb{Z}(z) &= 1 + 2(z + z^4 + z^9 + ...) \\
\Rightarrow \Theta_\mathbb{Z}(z) &= \sum_{k=-\infty}^{\infty} z^{k^2}.
\end{aligned}$$

One can perform the substitution $z = e^{ci\tau}$ (for some positive constant $c > 0$) to yield an analytic function of $\tau \in \mathcal{H}$, i.e. an infinitely differentiable function on $\mathcal{H}$ where each $f(\tau)$ is single-valued.
Naturally, there is no known way to compute $\Theta_\Lambda(z)$ for $\Lambda \subset \mathbb{R}^n$ once n is large enough, but that does not veil theta functions in irrelevance. Theta functions, relations among them, and their properties as analytic functions yield a substantial amount of information concerning lattices and their vector lengths. One such identity relates the theta functions of any two lattices $\Lambda_1, \Lambda_2 \subset \mathbb{R}^n$ and the theta function of their *direct sum* $\Lambda_1 \oplus \Lambda_2 \subset \mathbb{R}^{n+n}$.

Definition 2.1.1. The *direct sum* of two lattices $M \in \mathbb{R}^m$ and $N \in \mathbb{R}^n$ is $X \in \mathbb{R}^{m+n}$ such that

$$\begin{aligned}
M \oplus N &\longrightarrow X \\
(m,n) &\longrightarrow (m+n).
\end{aligned}$$

for each $m \in M,\ n \in N$.

Lemma 2.1.1. For lattices $\Lambda_1 \in \mathbb{R}^m$ and $\Lambda_2 \in \mathbb{R}^n$:

$$\Theta_{\Lambda_1 \oplus \Lambda_2}(z) = \Theta_{\Lambda_1}(z)\Theta_{\Lambda_2}(z).$$

Proof. Expand theta functions of singular lattices (right-hand side).

$$\Theta_{\Lambda_1}(z) = \sum_{x_1 \in \Lambda_1} z^{(x_1, x_1)},$$
$$\Theta_{\Lambda_2}(z) = \sum_{x_2 \in \Lambda_2} z^{(x_2, x_2)}.$$

Expand theta function of direct sum of lattices (left-hand side).

$$\Theta_{\Lambda_1 \oplus \Lambda_2}(z) = \sum_{\substack{x_1 \in \Lambda_1 \\ x_2 \in \Lambda_2}} z^{(x_1 + x_2, x_1 + x_2)} = \sum_{\substack{x_1 \in \Lambda_1 \\ x_2 \in \Lambda_2}} z^{(x_1, x_1)} z^{(x_2, x_2)} = \sum_{x_1 \in \Lambda_1} z^{(x_1, x_1)} \sum_{x_2 \in \Lambda_2} z^{(x_2, x_2)}.$$

Compare expansions of theta functions of singular lattices to the expansion of the theta function of the direct sum of lattices and observe that they are identical. Thus:

$$\boxed{\Theta_{\Lambda_1 \oplus \Lambda_2} = \Theta_{\Lambda_1}(z)\Theta_{\Lambda_2}(z)}$$

□

2.2 Poisson Summation Formula

Another, perhaps more prolific and crucial, identity arises from the *Poisson summation formula*. It incorporates something called the *dual* of a lattice, defined below.

Definition 2.2.1. The dual lattice Λ^* is defined by

$$\Lambda^* := \{y \in \mathbb{R}^n \mid \forall x \in \Lambda : (y, x) \in \mathbb{Z}\}.$$

More simply, the dual of a lattice Λ are all points in the lattice Λ^* such that the dot product of any point $x \in \Lambda$ with any point in $y \in \Lambda^*$ is an integer. For $n = 2$, i.e. two dimensions, the dual lattice Λ^* can be constructed by placing a lattice point at the centre of each lattice tile, see Figure 2.2

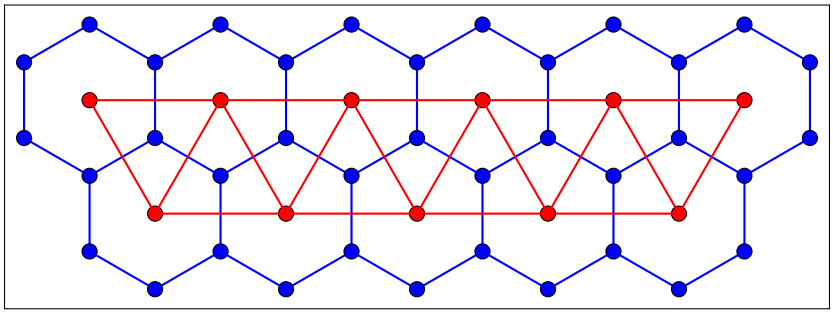


Figure 2.2: **Hexagonal lattice (blue) beneath the overlay of its dual (red).**

If matrix A generates lattice Λ then the transpose of its inverse generates the dual.

Lemma 2.2.1. If the lattice $\Lambda \subset \mathbb{R}^n$ is defined to be

$$\Lambda := \{Mc \mid c \in \mathbb{Z}^n\}$$

then its dual Λ^* is

$$\Lambda^* = \left\{M^{-T}c \mid c \in \mathbb{Z}^n\right\}.$$

Proof. Recall the definition of the dual lattice.

$$\Lambda^* := \{y \in \mathbb{R}^n \mid \forall x \in \Lambda : (y, x) \in \mathbb{Z}\}.$$

Consider Λ to be generated by the basis M. Incorporating M into the definition of the dual lattice yields the following.

$$\Lambda^* = \{y \in \mathbb{R}^n \mid \forall x \in \mathbb{Z}^n : (Mx, y) \in \mathbb{Z}\}.$$

where $\Lambda = M\mathbb{Z}^n \ \forall \ M \in GL_2(\mathbb{R})$. One can manipulate the dot product (Mx, y) as follows.

$$\begin{aligned}(Mx, y) &= (Mx)^T y ... \text{ (dot product to scalar product transformation property)} \\ &= M^T x^T y \\ &= M^T x^T y \\ &= (x^T)(M^T y) \\ &= (x, M^T y).\end{aligned}$$

Then,

$$\Lambda^* = \left\{y \in \mathbb{R}^n \mid \forall x \in \mathbb{Z}^n : (x, M^T y) \in \mathbb{Z}\right\}$$

If $(x, M^T y) \in \mathbb{Z}$ and $x \in \mathbb{Z}^n$, then $M^T y \in \mathbb{Z}^n \implies y \in M^{-T}\mathbb{Z}^n$. And so, if $\Lambda = M\mathbb{Z}^n$ then $\Lambda^* = M^{-T}\mathbb{Z}^n$. One can then define the dual lattice by

$$\boxed{\Lambda^* := \left\{M^{-T}c \mid c \in \mathbb{Z}^n\right\}.}$$

□

Another axiom of duals is that the dual of a dual lattice is the original lattice, i.e. $(\Lambda^*)^* = \Lambda$.

Lemma 2.2.2. For any lattice $\Lambda \subset \mathbb{R}^n$, $(\Lambda^*)^* = \Lambda$.

Proof. Consider lattice $\Lambda = M\mathbb{Z}^n$ and its dual defined by $\Lambda^* := \{y \in \mathbb{R}^n \mid \forall x \in \Lambda : (x, y) \in \mathbb{Z}\}$. Then,

$$\begin{aligned}\Lambda &= \{y \in \mathbb{R}^n \mid \forall x \in \mathbb{Z}^n : (Mx, y) \in \mathbb{Z}\} \\ &= \left\{y \in \mathbb{R}^n \mid \forall x \in \mathbb{Z}^n : (x, M^{-T}y) \in \mathbb{Z}\right\} \\ &= \left\{yM^{-T} \in \mathbb{Z}^n \mid y \in \mathbb{R}^n, M \in GL_n(\mathbb{R})\right\}\end{aligned}$$

Recall that $\Lambda = \{xM \in \mathbb{R}^n \mid x \in \mathbb{Z}^n, M \in GL_n(\mathbb{R})\}$. And so:

$$\begin{aligned}(\Lambda^*)^* &= \{z(M^{-T})^{-T} \in \mathbb{R}^n \mid z \in \mathbb{Z}^n, M \in GL_n(\mathbb{R})\} \\ &= \{xM \in \mathbb{R}^n \mid x \in \mathbb{Z}^n, M \in GL_n(\mathbb{R})\} = \Lambda\end{aligned}$$

Thus:

$$\boxed{(\Lambda^*)^* = \Lambda}$$

□

The densities of a lattice Λ and its dual Λ^* are closely related, with the relationship outlined in the below lemma.

Lemma 2.2.3. Consider lattice Λ where $\det(\Lambda)$ corresponds to the determinant of its generator basis. Then,

$$\det(\Lambda^*) = \frac{1}{\det(\Lambda)}$$

Proof. Consider lattice $\Lambda = M\mathbb{Z}^n$. Recall that its dual is defined by $\Lambda^* = M^{-T}\mathbb{Z}^n$. Upon evaluation of corresponding determinants, one finds that

$$\Lambda = M\mathbb{Z}^n \longrightarrow \det(\Lambda) = \det(M)$$
$$\Lambda^* = M^{-T}\mathbb{Z}^n \longrightarrow \det(\Lambda^*) = \det(M^{-T})$$

The determinants of the dual lattice can be further simplified:

$$\det(M^{-T}) = \det(M^T)^{-1} = \frac{1}{\det(M^T)} = \frac{1}{\det(M)}.$$

Thus:

$$\boxed{\det(\Lambda^*) = \frac{1}{\det(\Lambda)}}$$

□

With the concept of a dual established, one can proceed to explore the Poisson summation formula. There are different variations which can be applied to different spaces. We begin by stating and proving the Poisson summation formula functions over $\mathbb{R}$. This standard formula shall then be employed when we move on to establish a Poisson summation formula for lattices. In the next few theorems, we shall see the phrase "sufficiently nice" being used a descriptor for some function f. Initially, one may find the usage of this phrase to be a glaring hole in the weave of mathematical rigour, a pestilence that one ought to avoid. The exact requirements underlying the ambiguity are as follows.

1. $\int_{\mathbb{R}^n} \big|f(x)\big|dx < \infty$.
2. The series $\sum_{\lambda\in\Lambda} \big|f(\lambda + u)\big|$ converges uniformly for all u belonging to a compact subset of $\mathbb{R}^n$.
3. The series $\sum_{\lambda^*\in\Lambda^*} \hat{f}(y)$ is absolutely convergent.

Theta functions naturally satisfy these conditions and so exploring the restrictions yields little contextual value here, hence the free use of "sufficiently nice". The weave of rigour has been repaired and the pestilence is naught but a figment of a pedant's dream.

Theorem 2.2.1. ***Poisson Summation Formula for*** $\mathbb{R}$***:*** *Given a function f which is "sufficiently nice" with Fourier transform*

$$\hat{f}(y) = \int_{\mathbb{R}} f(x)e^{2\pi i(x,y)}dy,$$

then:

$$\sum_{n=-\infty}^{\infty} f(n) = \sum_{n=-\infty}^{\infty} \hat{f}(n).$$

Proof. One requires a compact domain in order to employ Fourier analysis. f may not be defined on a compact domain, and so we define a new function as follows.

$$F(x) := \sum_{n=-\infty}^{\infty} f(x+n).$$

This function is sufficiently nice if f is sufficiently nice and is periodic with period 1, i.e. for $x \in \mathbb{Z} : f(x+n) = f(n)$. We can consider F to be a function on the compact domain $\mathbb{R}/\mathbb{Z}$, and so it has a Fourier expansion such that

$$F(x) = \sum_{n=-\infty}^{\infty} a_m e^{2\pi mx},$$

where

$$a_m = \int_0^1 F(x)e^{-2\pi mx}dx = \int_0^1 \sum_{n=-\infty}^{\infty} f(x+n)e^{-2\pi mx}dx.$$

Since $e^{-2\pi mx} = e^{-2\pi m(x+n)}$, we may interchange the order of summation and integration, giving:

$$a_m = \sum_{n=-\infty}^{\infty} \int_0^1 f(x+n)e^{-2\pi xm(x+n)}dx = \int_{i\infty}^{+\infty} f(x)e^{-2\pi mx}dx = \hat{f}(m).$$

Now, take $x = 0$ and one results in

$$\sum_{n=-\infty}^{\infty} f(n) = F(0) = \sum_{n=-\infty}^{\infty} a_n = \sum_{n=-\infty}^{\infty} \hat{f}^n.$$

And so:

$$\boxed{\sum_{n=-\infty}^{\infty} f(n) = \sum_{n=-\infty}^{\infty} \hat{f}(n)}$$

□

The Poisson summation formula can then be applied to lattices, with the details outlined in the proof below.

Theorem 2.2.2. ***Poisson Summation Formula for lattices:*** *Given a lattice Λ, its dual Λ^*, and a sufficiently nice function f,*

$$\sum_{\lambda\in\Lambda} f(\lambda) = \frac{1}{\det(\Lambda)} \sum_{\lambda^*\in\Lambda^*} \hat{f}(\lambda^*).$$

Proof. Let Λ be a lattice with a corresponding basis $A \in GL_n(\mathbb{R})$. Then:

$$\begin{aligned}\sum_{\lambda\in\Lambda} f(\lambda) &= \sum_{m_1\in\mathbb{Z}} \sum_{m_2\in\mathbb{Z}} ... \sum_{m_n\in\mathbb{Z}} f(m_1\lambda_1 + m_2\lambda_2 + ... + m_n\lambda_n) \\ &= \sum_{m_1\in\mathbb{Z}} ... \sum_{m_n\in\mathbb{Z}} f\circ A \begin{pmatrix} m_1 \\ \vdots \\ m_n \end{pmatrix}.\end{aligned}$$

Apply Poisson summation formula (Theorem 2.2.1) n times.

$$\sum_{\lambda\in\Lambda} f(\lambda) = \sum_{m_1\in\mathbb{Z}} ... \sum_{m_n\in\mathbb{Z}} f\circ A \begin{pmatrix} m_1 \\ \vdots \\ m_n \end{pmatrix} = \sum_{k_1,...,k_n} \widehat{f\circ A} \begin{pmatrix} k_1 \\ \vdots \\ k_n \end{pmatrix}, \tag{†}$$

where $\widehat{\bullet}$ is the n-dimensional Euclidean Fourier transform. Now apply change of variable $y = Ax$ in the Fourier transform.

$$\widehat{f\circ A}(\epsilon) = \int_{\mathbb{R}^n} e^{-2\pi i(\epsilon,x)} f(Ax)dx = \frac{1}{|A|}\int_{\mathbb{R}^n} e^{-2\pi i(\epsilon,A^{-1}y)} f(y)dy,$$

where we introduce $|A| = \det(A) = \det(\Lambda)$ due to the scaling property of the Fourier transform. Then,

$$\implies \frac{1}{\det(\Lambda)} \int_{\mathbb{R}^n} e^{-2\pi i(A^{-T}\epsilon,y)f(y)dy} = \frac{1}{\det(\Lambda)} \hat{f}(A^{-T}\epsilon).$$

Returning to (†), one sees that

$$\sum_{\lambda\in\Lambda} f(\lambda) = \sum_{k_1,...,k_n} \widehat{f\circ A} \begin{pmatrix} k_1 \\ \vdots \\ k_n \end{pmatrix} = \frac{1}{\det(\Lambda)} \sum_{k_1,...,k_n} \hat{f}\left(A^{-T}\begin{pmatrix} k_1 \\ \vdots \\ k_n \end{pmatrix}\right) = \frac{1}{\det(\Lambda)} \sum_{\lambda^*\in\Lambda^*} \hat{f}(\lambda^*).$$

Thus:

$$\boxed{\sum_{\lambda\in\Lambda} f(\lambda) = \frac{1}{\det(\Lambda)} \sum_{\lambda^*\in\Lambda^*} \hat{f}(\lambda^*)}$$

□

One can utilize the Poisson summation formula for lattices to obtain an equation relating the theta function of a lattice Λ and the theta function of its dual Λ^*.

Theorem 2.2.3. *For lattice Λ, with density $D = 1/\det(\Lambda)$, and its dual $\Lambda*$,*

$$\Theta_\Lambda\left(e^{-\pi t}\right) = Dt^{-n/2}\Theta_{\Lambda^*}\left(e^{-\pi/t}\right).$$

Proof. Take the sufficiently nice function $f(x) = e^{-\pi(x,x)}$. We calculate this function's Fourier transform.

$$\begin{aligned}\hat{f}(y) &= \int_{\mathbb{R}^n} e^{-2\pi i(x,y)} f(x)dx \\ &= \int_{\mathbb{R}^n} e^{-2\pi i(x,y)} e^{-\pi(x,x)} dx \\ &= \int_{\mathbb{R}^n} e^{-2\pi i(x_1y_1+\ldots+x_ny_n)} e^{-\pi\left(x_1^2+\ldots+x_n^2\right)} dx\end{aligned}$$

Split integral over $\mathbb{R}^n$ into n integrals over $\mathbb{R}$.

$$\begin{aligned}\hat{f}(y) &= \underbrace{\int_{\mathbb{R}} \ldots \int_{\mathbb{R}}}_{n \text{ integrals}} e^{-2\pi i(x_1y_1+\ldots+x_ny_n)} e^{-\pi\left(x_1^2+\ldots x_n^2\right)} dx_1\ dx_2\ \ldots\ dx_n \\ &= \int_{\mathbb{R}} e^{-2\pi i x_n y_n} e^{-\pi x_n^2} \ldots \int_{\mathbb{R}} e^{-2\pi i x_1 y_1} e^{-\pi x_1^2} dx_1\ \ldots\ dx_n.\end{aligned}$$

This is an iterated integral that is identical at each coordinate (x_i, y_i). By evaluating the integral at one such coordinate, we can synthesize the result of the iterated integral by multiplying the identical isolated integrals together.

$$\begin{aligned}\int_{\mathbb{R}} e^{-2\pi ixy} e^{-\pi x^2} dx &= \int_{\mathbb{R}} e^{-\pi x^2 - 2\pi ixy} dx = \int_{\mathbb{R}} e^{-\pi((x+iy)^2+y^2)} dx = e^{-\pi y^2} \int_{\mathbb{R}} e^{-\pi(x+iy)^2} dx \\ &= e^{-\pi y^2} \int_{\mathbb{R}+iy} e^{-\pi t^2} dt = e^{-\pi y^2} \int_{\mathbb{R}} e^{-\pi x^2} dx = e^{-\pi y^2} \cdot 1 = e^{-\pi y^2}\end{aligned}$$

Apply isolated coordinate integral result into iterated integral:

$$\begin{aligned}\hat{f}(y) &= \int_{\mathbb{R}} e^{-2\pi i x_n y_n} e^{-\pi x_n^2} \ldots \int_{\mathbb{R}} e^{-2\pi i x_1 y_1} e^{-\pi x_1^2} dx_1\ \ldots\ dx_n. \\ &= e^{-\pi y_n^2} \int_{\mathbb{R}} e^{-\pi x_n^2}\ \ldots\ e^{-\pi y_1^2} \int_{\mathbb{R}} e^{-\pi x^2} dx_1\ \ldots\ dx_n \\ &= e^{-\pi y_1^2} e^{-\pi y_2^2}\ \ldots\ e^{-\pi y_n^2} \cdot \int_{\mathbb{R}} e^{-\pi x_n^2} \ldots \int_{\mathbb{R}} e^{-\pi x_1^2} dx_1\ \ldots\ dx_n \\ &= e^{-\pi y_1^2} \ldots e^{-\pi y_n^2} \cdot 1 \\ &= e^{-\pi\left(y_1^2+\ldots+y_n^2\right)} \\ &= f(y)\end{aligned}$$

We have established that for $f(x) = e^{-\pi(x,x)}$, $f = \hat{f}$. Consider the lattice $t^{1/2}\Lambda$ for some constant t. The density of this lattice is

$$\frac{1}{\det(t^{1/2}\Lambda)} = \frac{1}{t^{n/2}\det(\Lambda)} = Dt^{-n/2}$$

where D is the density of the lattice Λ and n is the dimension of the vector space within which the lattice Λ lives. Recall that for lattice Λ, the dual lattice is Λ^{-T}. Applying this to the lattice $t^{1/2}\Lambda$, one results in

$$(t^{1/2}\Lambda)^{-T} = t^{-1/2}\Lambda^{-T} = t^{-1/2}\Lambda^*$$

where Λ^* is the dual of Λ. Now apply the lattice Poisson summation formula for the function $f(x) = e^{-\pi(x,x)}$

and the lattice $t^{1/2}\Lambda$.

$$\sum_{x\in t^{1/2}\Lambda} e^{-\pi(x,x)} = Dt^{-n/2} \sum_{x\in t^{-1/2}\Lambda^*} \hat{f}(x)$$
$$\implies \sum_{x\in t^{1/2}\Lambda} e^{-\pi(x,x)} = Dt^{-n/2} \sum_{x\in t^{-1/2}\Lambda^*} e^{-\pi(x,x)}$$
$$\implies \sum_{x\in\Lambda} e^{-\pi\left(xt^{1/2},xt^{1/2}\right)} = Dt^{-n/2} \sum_{x\in\Lambda^*} e^{-\pi\left(xt^{-1/2},xt^{-1/2}\right)}$$
$$\implies \sum_{x\in\Lambda} e^{\pi t(x,x)} = Dt^{-n/2} \sum_{x\in\Lambda^*} e^{-\pi t^{-1}(x,x)}$$

Consider theta functions $\Theta_\Lambda\left(e^{-\pi t}\right)$ and $\Theta_{\Lambda^*}\left(e^{-\pi/t}\right)$:

$$\Theta_\Lambda\left(e^{-\pi t}\right) = \sum_{x\in\Lambda} e^{\pi t(x,x)}$$
$$\Theta_\Lambda\left(e^{-\pi/t}\right) = \sum_{x\in\Lambda} e^{\pi/t(x,x)}$$

We note the presence of expanded theta functions in the equation and conclude the following.

$$\boxed{\Theta_\Lambda\left(e^{-\pi t}\right) = Dt^{-n/2}\Theta_{\Lambda^*}\left(e^{-\pi/t}\right)}$$

□

Note that $e^{-\pi}$ is a common term in the input for both theta functions present in the above equation. One can simplify the equation (syntactically) by defining a slightly more specific theta function as follows.

$$\Theta_\Lambda(t) = \sum_{x\in\Lambda} e^{-\pi t(x,x)}$$

The derived equation then becomes

$$\Theta_\Lambda(t) = Dt^{-n/2}\Theta_{\Lambda^*}\left(\frac{1}{t}\right) \tag{†}$$

As we progress with our approach to exploring modular forms through theta functions, we shall define a specific form of the generalized theta function as follows (note the lowercase theta symbol).

$$\theta_\Lambda(\tau) = \sum_{x\in\Lambda} e^{\pi i\tau(x,x)}.$$

where $\tau \in \mathcal{H}$, the complex upper half-plane. One can then compare the new function to the old generalized theta function, specifically how to move from $\Theta_\Lambda(\tau)$ to $\theta_\Lambda(\tau)$.

$$\text{Old Theta Function} : \Theta_\Lambda(t) = \sum_{x\in\Lambda} e^{-\pi t(x,x)}$$
$$\text{New Theta Function} : \theta_\Lambda(\tau) = \sum_{x\in\Lambda} e^{\pi i\tau(x,x)}$$

$$\theta_\Lambda(it) = \sum_{x\in\Lambda} e^{\pi i(it)(x,x)} = \sum_{x\in\Lambda} e^{-\pi t(x,x)} = \Theta_\Lambda(t).$$
$$\theta_\Lambda\left(-\frac{1}{it}\right) = \sum_{x\in\Lambda} e^{\pi i\left(-\frac{1}{it}\right)(x,x)} = \sum_{x\in\Lambda} e^{-\pi/t(x,x)} = \Theta_\Lambda\left(\frac{1}{t}\right)$$

And so, we establish that:

$$\Theta_\Lambda(t) = \theta_\Lambda(it) \tag{2.1}$$
$$\Theta_\Lambda\left(\frac{1}{t}\right) = \theta_\Lambda\left(-\frac{1}{it}\right) \tag{2.2}$$

2.3 Functional Equation For Self-Dual Lattices

Lattices are able to be *self-dual*. This is when a lattice Λ is its own dual, i.e. $\Lambda = \Lambda^*$. This occurs when Λ is unimodular (i.e. density = 1) and it is *integral*, meaning that $(x, y) \in \mathbb{Z}$ for all $x, y \in \Lambda$. In the world of matrices, self-duality is present when the lattice basis A is unimodular, i.e. $\det(A) = \pm 1$, and when the matrix product $A^T A$ has integer entries. This product is called the *Gram matrix*. Examples of self-dual lattices are $\mathbb{Z}$ and the E_8 lattice, where the E_8 lattice is defined by $E_8 := \mathbb{Z}^8 \cup (\mathbb{Z} + \frac{1}{2})^8$. For any two self-dual lattice Λ_1 and Λ_2, the direct sum $\Lambda_1 \oplus \Lambda_2$ is also self-dual. Suppose that lattice Λ is self-dual. Then:

$$\det(\Lambda^*) = \frac{1}{\det(\Lambda)} = \frac{1}{\det(\Lambda)} \Longrightarrow \det(\Lambda) = \det(\Lambda^*) = 1 = D$$

where D is the density of lattices Λ and Λ^*. Thus, for a self-dual lattice Λ, equation (†) becomes

$$\begin{aligned} &\Theta_\Lambda(t) = Dt^{-n/2}\Theta_{\Lambda^*}\left(\frac{1}{t}\right) \\ \Longrightarrow\ &\Theta_\Lambda(t) = t^{-n/2}\Theta_\Lambda\left(\frac{1}{t}\right). \\ \Longrightarrow\ &\Theta_\Lambda\left(\frac{1}{t}\right) = t^{n/2}\Theta_\Lambda(t) \end{aligned}$$

This is a *functional equation* of Θ_Λ for self-dual Λ. One can then replace the generalized theta function Θ with the newly constructed, more specific θ using the identities outlined in (2.1) and (2.2).

$$\theta_\Lambda\left(-\frac{1}{it}\right) = t^{n/2}\theta_\Lambda(it)$$

Take $\tau = it \in \mathcal{H}$. Then the above functional equation becomes

$$\boxed{\theta_\Lambda\left(-\frac{1}{\tau}\right) = \left(\frac{\tau}{i}\right)^{n/2}\theta_\Lambda(\tau)} \qquad (*)$$

The transformation property of θ_Λ is similar but not equal to that of a modular form. What conditions must the lattice Λ satisfy in order for θ_Λ to be a modular form?
If (x, x) is an even integer $\forall x \in \Lambda$, i.e. $\exists c \in \mathbb{Z} :\ (x, x) = 2c$, then:

$$e^{\pi i(\tau+1)(x,x)} = e^{2\pi i(\tau+1)c} = e^{2\pi i\tau c} = e^{\pi i\tau(x,x)}.$$

Thus, θ would have period 1:

$$\boxed{\theta_\Lambda(\tau + 1) = \theta_\Lambda(\tau)} \qquad (**)$$

In the context of matrices, this means that the diagonal entries of lattice basis A are even. Such matrices are called *even matrices*. Modular forms lie at the core of eight-dimensional sphere packing, and we see the number 8 and its multiples appear in other areas as if by magical convenience. The following lemma establishes the connection between lattice self-duality, the state of being even, and the enigma that is the number 8.

Lemma 2.3.1. If a lattice Λ is self-dual and even then its ambient dimension is a multiple of 8, i.e. for self-dual lattice even lattice $\Lambda \subset \mathbb{R}^n$: $8|n$.

Proof. The modular group $SL_2(\mathbb{Z})$ is generated by matrices

$$T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$
$$S = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

This is the group of all 2×2 integer matrices with determinant $+1$. Matrices S and T are incorporated in identities $(*)$ and $(**)$ as shown below.

$$\begin{aligned} \theta_\Lambda(S \cdot \tau) &= \theta_\Lambda\left(-\frac{1}{\tau}\right) = \left(\frac{\tau}{i}\right)^{n/2}\theta_\Lambda(\tau) \\ \theta_\Lambda(T \cdot \tau) &= \theta_\Lambda(\tau + 1) = \theta_\Lambda(\tau) \end{aligned}$$

Choose $\tau = i$ and apply alternating $(*)$ and $(**)$:

$$\theta_\Lambda(i) \overset{(**)}{=\!=} \theta_\Lambda(i+1) \overset{(**)}{=\!=} \left(\frac{i+1}{i}\right)^{n/2} \theta_\Lambda\left(-\frac{1}{i+1}\right) = (1-i)^{n/2}\theta_\Lambda\left(\frac{i-2}{2}\right) \overset{(**)}{=\!=} (1-i)^{n/2}\theta_\Lambda(\frac{i+1}{2})$$
$$\overset{(*)}{=\!=} \left(\frac{1}{i}\frac{i+1}{2}(1-i)\right)^{n/2} \theta_\Lambda\left(-\frac{2}{i+1}\right) = (-i)^{n/2}\theta_\Lambda(i-1) \overset{(**)}{=\!=} (-i)^{n/2}\theta_\Lambda(i)$$

And so, one results in $\theta_\Lambda(i) = (-i)^{n/2}\theta_\Lambda(i) \implies (-i)^{n/2} = 1$. Recall that $i = \sqrt{-1}$. $i^2 = -1$, $i^4 = 1$. One would require for $n/2$ to be a multiple of 4 in order for the equation to hold true, and so n would need to be a multiple of 8, i.e. $8|n$. □

For a self-dual, even lattice Λ, our functional equation becomes

$$\boxed{\theta_\Lambda\left(-\frac{1}{\tau}\right) = \tau^{n/2}\theta_\Lambda(\tau)}$$

We now have two identities that hold true for self-dual, even lattices.

1. $\theta_\Lambda(\tau+1) = \theta_\Lambda(\tau)$
2. $\theta_\Lambda\left(-\frac{1}{\tau}\right) = \tau^{n/2}\theta_\Lambda(\tau)$

2.4 Modular Forms Manifested By Theta Functions

What follows is a theorem that finally reaches the ever veiled concept of modular forms.

Theorem 2.4.1. *If Λ is a self-dual, even lattice with $\dim(\Lambda) = 8|n$ then $\theta_\Lambda(\tau)$ is a modular form of weight $n/2$.*

Proof. Recall the definition of a modular form: A sufficiently nice analytic function ϕ on $\mathcal{H}$ satisfying the identity

$$\phi\left(\frac{a\tau+b}{c\tau+d}\right) = (\epsilon_g\sqrt{c\tau+d})^n\phi(\tau)$$

is called a *modular form of weight* $(n/2)$ where ϵ_g is an eight root of unity. For the first root of unity, the identity becomes

$$\phi\left(\frac{a\tau+b}{c\tau+d}\right) = (c\tau+d)^{n/2}\phi(\tau)$$

The modular group $SL_2(\mathbb{Z})$ is generated by matrices $S = \begin{pmatrix}0 & -1\\ 1 & 0\end{pmatrix}$ and $T = \begin{pmatrix}1 & 1\\ 0 & 1\end{pmatrix}$. The theta series expansion shows invariance under T while the functional equation shows invariance under S for $8|n$. There exists an integer combination of S and T for all transformations $A \in SL_2(\mathbb{Z})$, which means that for all transformations A, the theta function $\theta_\Lambda(A\tau)$ is invariant, i.e. satisfying the modular form requirements. □

Chapter 3

Visualizing the E_8 Lattice

The esotericism of modular forms may initially seem to be a deterrent, for on paper they seem to be naught more than some complex functions with a quaint symmetry. However, satisfying oneself with such a conclusion would be underselling the significance of modular forms. In July 2022, Maryna Viazovska was awarded the prestigious Fields Medal in part for her proof that "no packing of unit balls in Euclidean space $\mathbb{R}^8$ has density greater than that of the E_8 lattice packing"[5], with her paper concerning itself with *sphere packing*, *modular forms*, and *Fourier analysis*, concepts all present in our journey from lattices to modular forms. The E_8 lattice has been popularized by the American Institute of Mathematics (AIM) as part of their ambitious project known as *Atlas of Lie Groups and Representations*, see `https://aimath.org/E8/` for more information.

There exists something called the E_8 *root system*, the set of all *minimal length* points in the E_8 lattice. Should one take all the possible linear combinations of the vectors in the root system, one indeed obtains the entirety of the E_8 lattice. The theta function $\theta_{E_8}(\tau)$ associated with the E_8 lattice is a modular form.

$$\theta_{E_8}(\tau) = 1 + 240e^{2\pi i\tau} + 2160e^{4\pi i\tau} + 6720e^{6\pi i\tau} + ...$$

We see (as is standard) 1 vector of length zero, and then 240 vectors of length $\sqrt{2}$. These 240 vectors can be projected onto the two-dimensional space to produce the two-dimensional representation of an eight-dimensional object called the *Gosset polytope* 4_{21}. The first of such representation was hand-drawn in the 1960s by the British mathematician Peter McMullen, see Figure 3.1

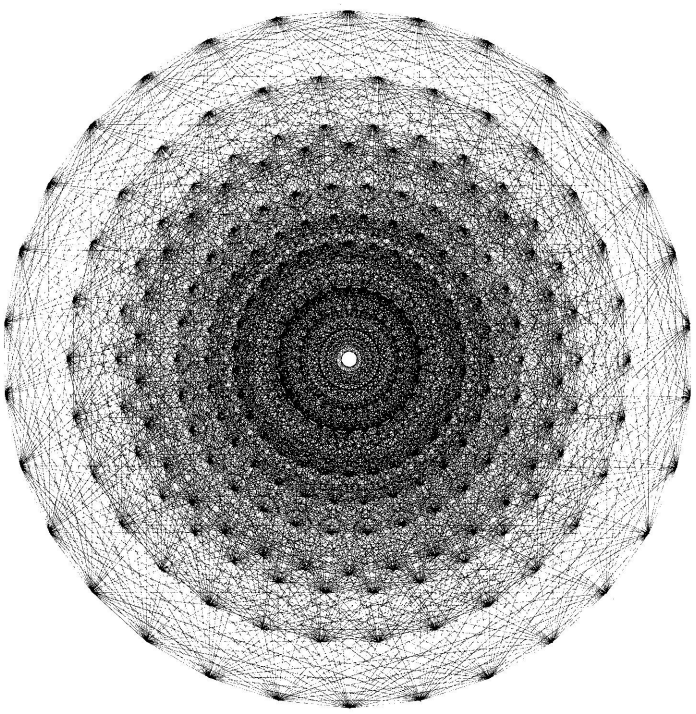


Figure 3.1: **2D representation of the E_8 root system, hand-drawn by Peter McMullen in the 1960s.**

This paper shall conclude by exploring the visualization of the E_8 root system using some Python packages for computation and graphics.

To begin, recall the definition of the E_8 lattice.

$$E_8 := \mathbb{Z}^8 \cup \left(\mathbb{Z} + \frac{1}{2}\right)^8.$$

The E_8 *root system* is then the subset of the lattice consisting only of the *minimal length* vectors. The minimal length for the E_8 lattice is $\sqrt{2}$. Vectors of this length take the form

$$F_1 = (\pm 1, \pm 1, 0, 0, 0, 0, 0, 0)$$

where the signs are independent and the non-zero entries can be in any position, and

$$F_2 = \frac{1}{2} \cdot (\pm 1, \pm 1, \pm 1, \pm 1, \pm 1, \pm 1, \pm 1, \pm 1)$$

where the signs are independent provided the condition $\prod_{i=1}^{8} a_i = 1$ for all entries a_i in the vector form F_2. The are 112 vector of the form F_1 and 128 vectors of the form F_2, totalling to a sum of 240. To begin our visualization, we ought to first obtain the root system and store it appropriately. Import relevant packages for vector computation.

```
import numpy as np
from itertools import combinations, product
```

We first compute roots of the form outlined by F_1. Two ± 1's with independent signs can take up any vector index. We iterate over all index pair combinations in an eight-dimensional vector, then iterate over all 4 combinations the numbers $(\pm 2, \pm 2)$ can take, before finally storing each root computed within the inner loop. We'll be using ± 2 instead of ± 1 however, this scaling factor makes no difference in the visualization but makes certain computations more convenient.

```
roots = []
for i, j in combinations(range(8), 2):
    for x, y in product([-2, 2], repeat=2):
        v = np.zeros(8)
        v[i] = x
        v[j] = y
        roots.append(v)
```

Now we compute roots of the form outlined by F_2. Here, too, we multiply each vector entry by 2 for convenience in computation, so we are in fact dealing with the placement of ± 1's in an eight-dimensional vector. We require that the product of the vector entries be equal to 1; the `numpy` library makes the checking for this requirement trivial.

```
for v in product([-1, 1], repeat=8):
    if np.prod(v) == 1:
        roots.append(v)
```

Finally, we convert the standard Python list to a `numpy` array which allows us to perform convenient vector computation later on.

```
roots = np.array(roots).astype(int)
```

At this stage, we can go about visualizing the lattice root system by projecting each eight-dimensional vector onto the two-dimensional plane, i.e. plotting the (x, y) coordinate of each vector. We'll need to import some additional libraries to support graph plotting and define our canvas parameters.

```
import matplotlib.pyplot as plt
import seaborn as sns
import pandas as pd

plt.rcParams.update(
    {
        "figure.figsize": (8, 8),
        "font.size": 16,
    }
)
plt.axis("off")
```

Now we will create a scatter plot and plot our root system, or at least one of its two-dimensional perspectives using coordinates from the first and second dimensions.

```
plt.scatter(roots[:, 0], roots[:, 1], s=10, color="black", zorder=1)
plt.show()
```

The resulting plot may give the impression that the computed root system is extremely primitive. In fact, one can plot the root system onto any possible plane (not just the one spanned by the first and second coordinates) and one observes identical graphs.

```
g = sns.PairGrid(pd.DataFrame(roots))
g.map_offdiag(plt.scatter, s=100, alpha=0.7)
g.map_diag(plt.hist, bins=21)

for ax in g.axes.ravel():
    ax.set_xticks([])
    ax.set_yticks([])
    ax.set_xlim(-2.5, 2.5)
    ax.set_ylim(-2.5, 2.5)

plt.show(block=True)
```

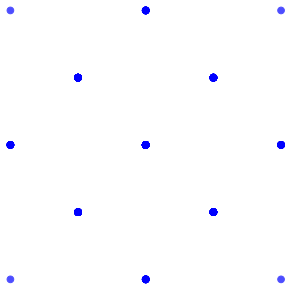


Figure 3.2: **Projection of E_8 root system as seen from a primitive perspective.**

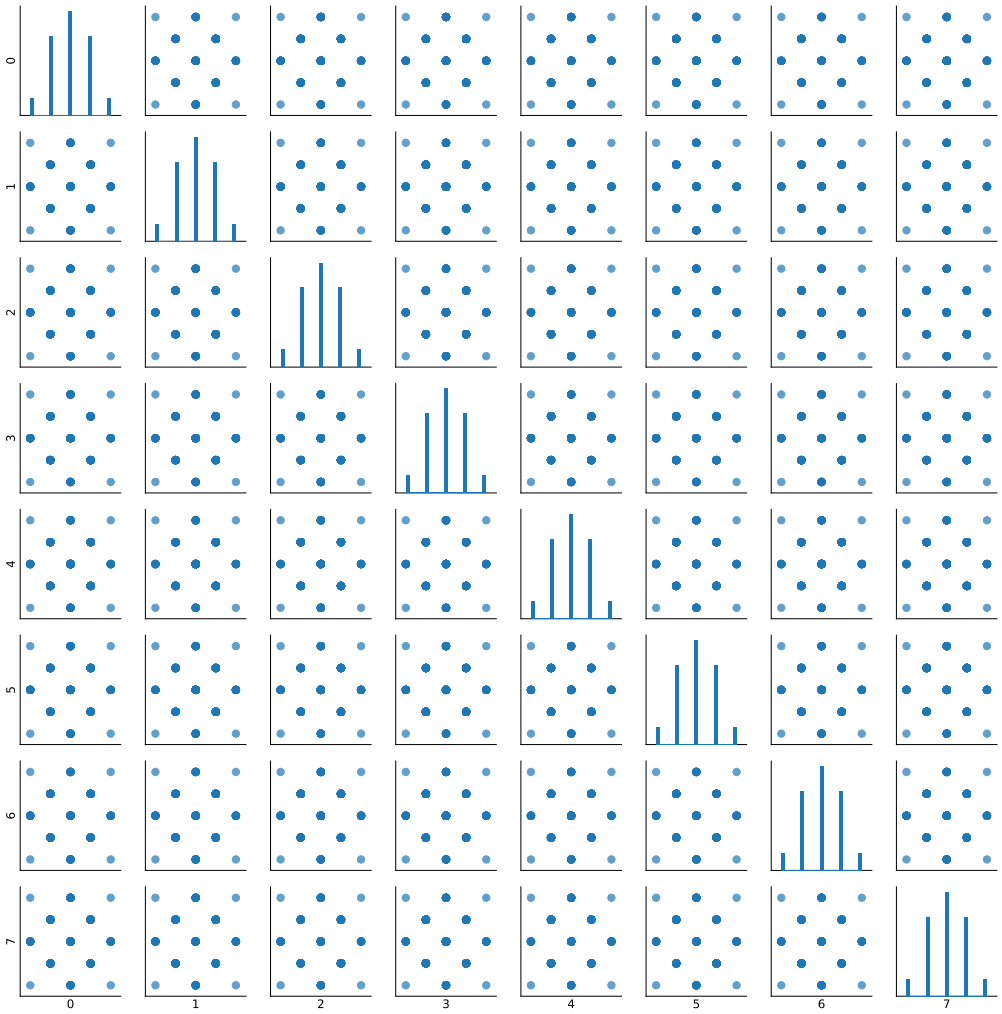


Figure 3.3: **Pairwise scatter plot visualizing E_8 root system from different two-dimensional perspectives.**

Should one presume that the E_8 root system looks the same (in its two-dimensional projection) regardless of which perspective one might take, one would be misled. The perspectives displayed by taking the roots in their standard form simply happen to be quite mundane. In order to get a different perspective, we want to

rotate this eight-dimensional object about some axis or a combination of axes. Although one cannot pick up this eight-dimensional object and rotate it in one's hand as one might do with an object living in tangible dimensions, we can still apply rotation transformations using matrix multiplication and in doing so observe a different perspective of the object. We define a small function to generate a random rotation matrix, apply the rotation transformation to the roots array, and plot the new rotated perspectives on a pairwise scatter plot as before.

```
38 def random_rotation_matrix(D):
39     q, _ = np.linalg.qr(np.random.normal(size=(D, D), ))
40     return q
41
42 rotated_roots = np.dot(roots, random_rotation_matrix(8))
43 g = sns.PairGrid(pd.DataFrame(rotated_roots))
44 g.map_offdiag(plt.scatter, s=40, alpha=0.7)
45 g.map_diag(plt.hist, bins=21)
46
47 for ax in g.axes.ravel():
48     ax.set_xticks([])
49     ax.set_yticks([])
50     ax.set_xlim(-2.5, 2.5)
51     ax.set_ylim(-2.5, 2.5)
52
53 plt.show(block=True)
```

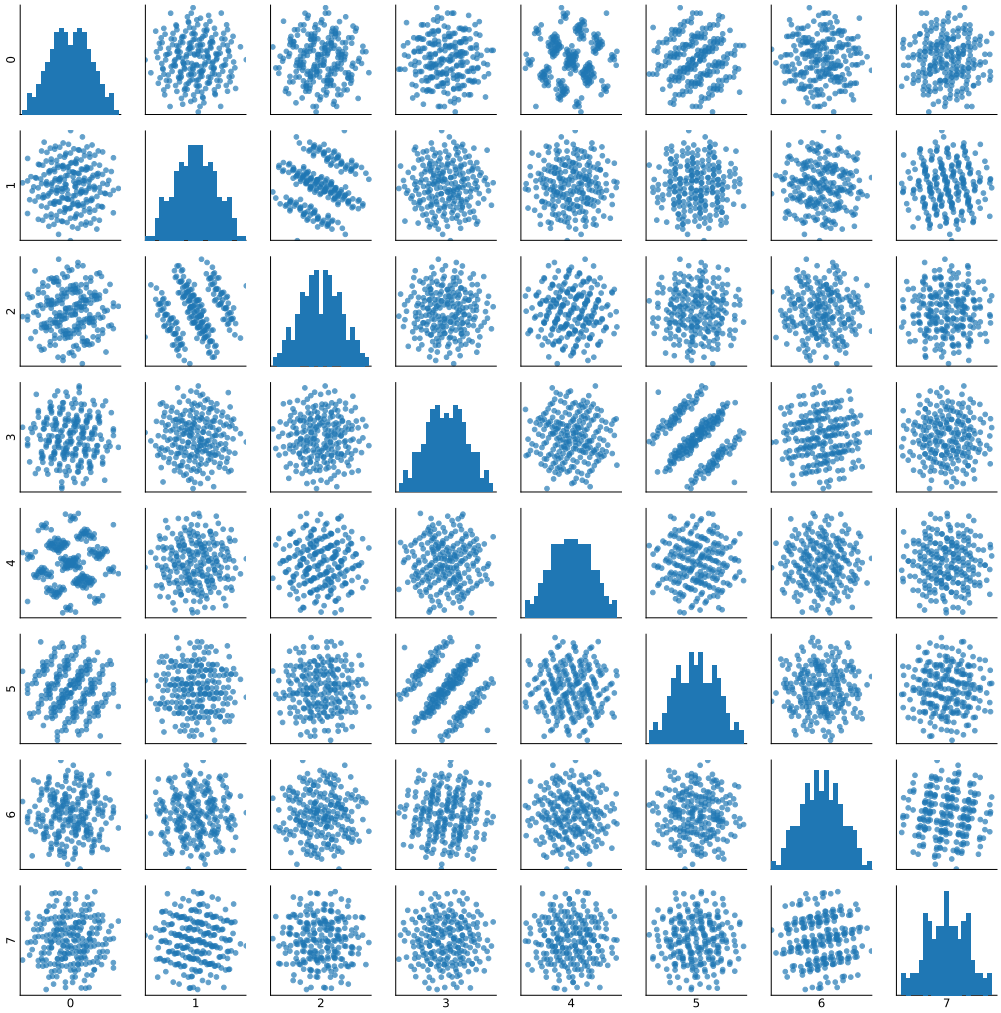


Figure 3.4: **Two-dimensional perspectives of a rotated E_8 root system.**

Here we see the apparent disorder one would expect when ones projects a high-dimensional object onto a two-dimensional space. The famous picture drawn by Peter McMullen is one such projection under a very specific rotation. The matrix that orients the roots in the manner visualized by McMullen is defined below. Under this rotation, the coordinates in the first two entries of each eight-dimensional vector form 8 distinct

rings around the origin.

```
rotation_matrix = np.array(
    [
        [-0.4253, -0.0568, 0.0444, -0.6299, -0.0245, -0.4093, 0.4535, -0.2079],
        [-0.2080, -0.3752, 0.2248, 0.4525, 0.3674, 0.1566, 0.2612, -0.5767],
        [-0.2655, 0.3893, -0.0384, -0.2431, 0.7515, 0.1513, -0.3598,-0.0063],
        [-0.0409, -0.0279, 0.6295, 0.2028, 0.2559, -0.3084, 0.1899, 0.6028],
        [0.1313, -0.1925, 0.2703, -0.4232, -0.0058, 0.7712, 0.2608, 0.1750],
        [0.7039, -0.2095, 0.2762, -0.3195, 0.2130, -0.2936, -0.2328, -0.3102],
        [-0.4323, -0.2949, 0.4125, -0.1062, -0.3332, 0.0322, -0.6533, -0.0742],
        [0.0618, 0.7318, 0.4796, 0.0745, -0.2783, 0.0880, 0.1064, -0.3581],
    ]
)
```

To rotate the root vectors we find the dot product of `roots` and `rotation_matrix`.

```
roots_rotated = np.dot(roots, rotation_matrix)
```

Creating a scatter plot by using the first and second coordinates of each vector yields the ring pattern we are searching for.

```
plt.scatter(roots_rotated[:, 0], roots_rotated[:, 1], s=10, color="black")
plt.show()
```

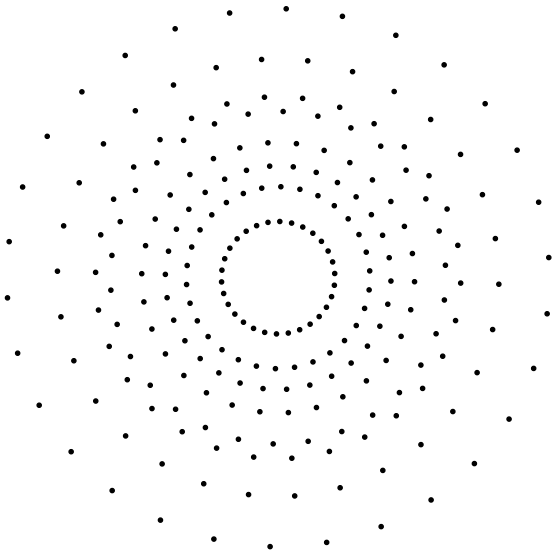


Figure 3.5: E_8 **roots from a particular perspective.**

Two roots are adjacent they form an angle of $\pi/3$ between one another. Thus, in order to identify all the edges, we will iterate over all possible root combinations and evaluate the angle between them. To do so, we define two new functions; one to convert an arbitrary n-dimensional vector into a unit vector, another to calculate the angle two n-dimensional vectors make between one another.

```
def unit_vector(vector:
    return vector / np.linalg.norm(vector)

def angle_between(v1, v2):
    v1_u = unit_vector(v1)
    v2_u = unit_vector(v2)
    return np.arccos(np.clip(np.dot(v1_u, v2_u), -1.0, 1.0))
```

Now we construct an array and fill it with all edge vertices.

```
edges = []
for i, j in combinations(roots, 2):
    if abs(np.pi/3 - angle_between(i, j)) < 0.00000001:
        edges.append([i, j])

```

```
edges = np.array(edges).astype(int)
edges_rotated = np.dot(edges, rotation_matrix)
```

Recall that the root vectors form 8 rings around the origin when transformed in the manner seen in Figure 3.5. Each ring has a unique radius r. One way to differentiate between different edges is to ensure that the colour of each edge is determined only by the outermost ring it connects. First we define a unique colour for each unique ring radius using a dictionary wherein each key corresponds to a radius.

```
colors = {
    0.48: "#fcc200",
    0.77: "#0058fc",
    0.95: "#135ebf",
    1.15: "#00ddff",
    1.41: "#ad0000",
    1.53: "#ff1443",
    1.85: "#00ffa6",
    2.28: "#5400ad"
}
```

We can now plot the edges onto the same diagram. Note the negative `zorder` parameter in the `plt.plot()` function. This ensures that the edges do not cover the nodes (who have a `zorder` of 1), we'd like for them to remain visible.

```
for edge in edges_rotated:
    x_values = [edge[0][0], edge[1][0]]
    y_values = [edge[0][1], edge[1][1]]
    ring1 = round(np.linalg.norm(np.asarray(edge[0][:2]) - (0,0)), 2)
    ring2 = round(np.linalg.norm(np.asarray(edge[1][:2]) - (0,0)), 2)
    dominant_ring = max(ring1, ring2)
    plt.plot(x_values, y_values, colors[dominant_ring], "o", linestyle="-", linewidth=0.5,
    ↪  markersize=0, alpha=0.5, zorder=-1)

plt.show()
```

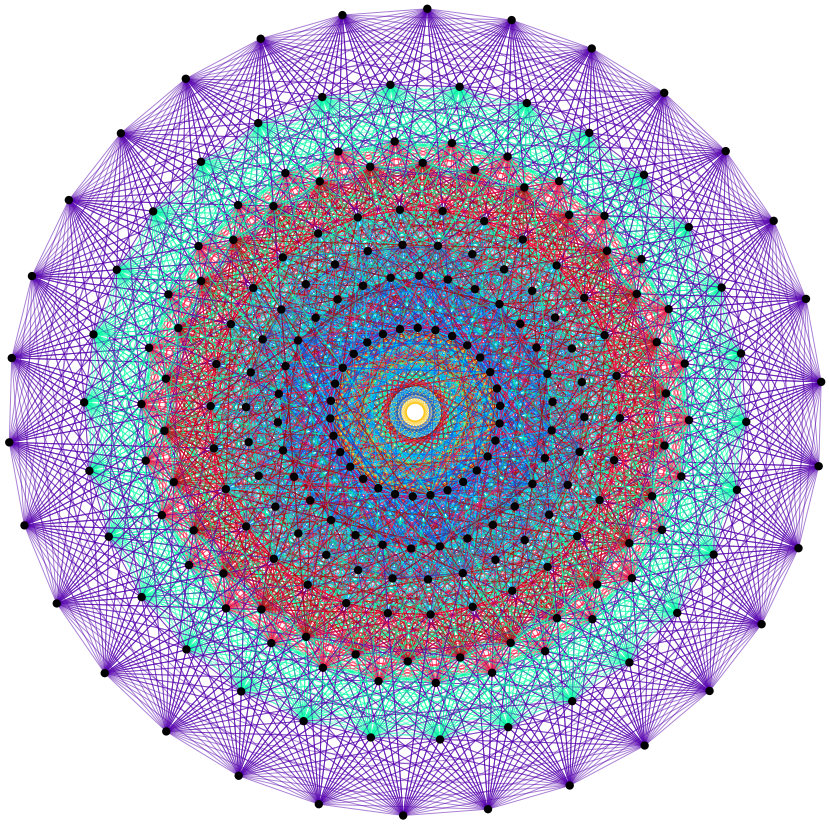


Figure 3.6: **The famous E_8 root system diagram recreated.**

Full Visualization Code

```python
import numpy as np
import matplotlib.pyplot as plt
from itertools import combinations, product

# pyplot parameters. See plt.rcParams.keys() for list of valid parameters
plt.rcParams.update(
    {
        "figure.figsize": (8,8),
        "font.size": 16,
    }
)
plt.axis("off")


#--- E8 root calculation ---#
# All root entries are multiplied by 2 to remove floats.

roots = []

# Computes roots of the form (+-1, +-1, 0, 0, 0, 0, 0, 0) where signs and non-zero entry placements are independent
for i, j in combinations(range(8), 2):
    for x, y in product([-2, 2], repeat=2):
        v = np.zeros(8)
        v[i] = x
        v[j] = y
        roots.append(v)

# Computes roots of the form 0.5*(+-1, +-1, ..., +-1) where signs are independent with the sole restriction that the product of all root entries must be equal to 1.
for v in product([-1, 1], repeat=8):
    if np.prod(v) == 1:
        roots.append(v)

roots = np.array(roots).astype(int)
print("Roots computed")


#--- Edge calculation ---#
def unit_vector(vector):
    """Converts vector to a unit vector in the same direction"""
    return vector / np.linalg.norm(vector)

def angle_beween(v1, v2):
    """Computes the angle between two vectors"""
    v1_u = unit_vector(v1)
    v2_u = unit_vector(v2)
    return np.arccos(np.clip(np.dot(v1_u, v2_u), -1.0, 1.0))

edges = []

for i, j in combinations(roots, 2): # tests each root vector pair
    # Two root vectors are connected by an edge if and only if they form an angle of pi/3
    if abs(np.pi/3 - angle_between(i, j)) < 0.00000001:
        edges.append([i, j])

edges = np.array(edges).astype(int)
print("Edges computed")


#--- E8 visualization ---#

# Rotation matrix that yields a symmetric two-dimensional perspective of root system
rotation_matrix = np.array(
```

```
    [
        [-0.4253, -0.0568, 0.0444, -0.6299, -0.0245, -0.4093, 0.4535, -0.2079],
        [-0.2080, -0.3752, 0.2248, 0.4525, 0.3674, 0.1566, 0.2612, -0.5767],
        [-0.2655, 0.3893, -0.0384, -0.2431, 0.7515, 0.1513, -0.3598,-0.0063],
        [-0.0409, -0.0279, 0.6295, 0.2028, 0.2559, -0.3084, 0.1899, 0.6028],
        [0.1313, -0.1925, 0.2703, -0.4232, -0.0058, 0.7712, 0.2608, 0.1750],
        [0.7039, -0.2095, 0.2762, -0.3195, 0.2130, -0.2936, -0.2328, -0.3102],
        [-0.4323, -0.2949, 0.4125, -0.1062, -0.3332, 0.0322, -0.6533, -0.0742],
        [0.0618, 0.7318, 0.4796, 0.0745, -0.2783, 0.0880, 0.1064, -0.3581],
    ]
)

# Apply rotation transformation to roots and edges
roots_rotated = np.dot(roots, rotation_matrix)
edges_rotated = np.dot(edges, rotation_matrix)

# Plot the first two coordinates of each vector onto the two-dimensional plane
plt.scatter(roots_rotated[:, 0], roots_rotated[:, 1], s=100, color="blue", alpha=0.7, zorder=1)
print("Roots plotted")

# Different root rings correspond to different edges colors. Dictionary keys are ring radii
colors = {
    0.48: "#fcc200",
    0.77: "#0058fc",
    0.95: "#135ebf",
    1.15: "#00ddff",
    1.41: "#ad0000",
    1.53: "#ff1443",
    1.85: "#00ffa6",
    2.28: "#5400ad"
}

# Plot edges. Note zorder=-1 to ensure that roots are always visible
for edge in edges_rotated:
    x_values = [edge[0][0], edge[1][0]]
    y_values = [edge[0][1], edge[1][1]]
    ring1 = round(np.linalg.norm(np.asarray(edge[0][:2]) - (0,0)), 2)
    ring2 = round(np.linalg.norm(np.asarray(edge[1][:2]) - (0,0)), 2)
    dominant_ring = max(ring1, ring2)
    plt.plot(x_values, y_values, colors[dominant_ring], "o", linestyle="-", linewidth=0.5,
    ↪  markersize=0, alpha=0.5, zorder=-1)

print("Edges plotted. Saving...")

plt.savefig("plot.pdf")
# plt.show
```

References

[1] Lars V Ahlfors. "Complex analysis: an introduction to the theory of analytic functions of one complex variable". In: *New York, London* 177 (1953).

[2] Tim Anderton. *Visualizing Lattices*. 2020. URL: `http://asymptoticlabs.com/blog/posts/visualizing-lattices.html`.

[3] Noam D Elkies. "Lattices, Linear Codes, and Invariants, Part I". In: *Notices of the AMS* 47.10 (2000).

[4] Jean-Pierre Serre. *A Course in Arithmetic*. Vol. 7. Springer Science & Business Media, 2012.

[5] Maryna Viazovska. "The sphere packing problem in dimension 8". In: *Annals of Mathematics* 185.3 (May 2017).