Contents

Chapter 1

Lattice

秋风清，秋月明，落叶聚还散，寒鸦栖复惊，相思相见知何日，此时此夜难为情。
入我相思门，知我相思苦，长相思兮长相忆，短相思兮无穷极，早知如此绊人心，还如当初不相识

李白

1.1 Intro

Definition 1.1.1. Let $v_1, ..., v_n$ be linearly independent vectors in $\mathbb{R}^n$, then a ***lattice with basis*** $v_1, ..., v_n$ is defined to be

$$\Lambda = \{\sum_{i=1}^{n} a_i v_i : a_i \in \mathbb{Z}\}$$

We may write this as $\Lambda = \langle v_1, ..., v_n \rangle_{\mathbb{Z}}$.

Remark 1.1.2. Each element of the lattice has a unique representation as integer linear combination of its basis. In particular, basis of a lattice is not unique as we can multiply $A \in \mathrm{SL}_n(\mathbb{Z})$ to a basis and still get a basis, where $\mathrm{SL}_n(\mathbb{Z})$ means n by n special linear group over $\mathbb{Z}$.

Definition 1.1.3. Let $v_1, ..., v_n$ be a basis of lattice Λ, define $d(\Lambda)$, the ***determinant*** of Λ, to be $|\det(v_1, ..., v_n)|$.

Remark 1.1.4. It is not hard to see our definition does not depend on the choice of basis as they only differ by an invertible matrix over $\mathbb{Z}$ and hence if v_i and w_i are both basis for Λ so $A[v_1, ..., v_n]^T = [w_1, ..., w_n]^T$ then we have $|\det(w_1, .., w_n)| =$

$|\det(A \cdot [v_1, ..., v_n])| = |\det(A) \cdot \det(v_1, ..., v_n)| = |\pm 1| \cdot |\det(v_1, ..., v_n)|$. In addition we see $d(\Lambda) > 0$ as $v_1, ..., v_n$ are linearly independent.

Definition 1.1.5. Let Λ, Ω be two lattice, if $\Omega \subseteq \Lambda$ then we say Ω is a ***sublattice*** of Λ.

Remark 1.1.6. Note if $w_1, ..., w_n \in \Lambda$, then they generate a sublattice Λ_0. In particular, we have

$$A \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = \begin{bmatrix} w_1 \\ \vdots \\ w_n \end{bmatrix}$$

Let $D = |\det(A)|$, note this is a positive integer. Further, we have

$$D = \frac{|\det(w_1, ..., w_n)|}{|\det(v_1, ..., v_n)|} = \frac{d(\Lambda_0)}{\Lambda}$$

Definition 1.1.7. Let Λ_0 be a sublattice of Λ, then we define the ***index*** of Λ_0 in Λ to be $\frac{d(\Lambda_0)}{d(\Lambda)}$.

Remark 1.1.8. Suppose $\Lambda = \langle v_1, ..., v_n \rangle_{\mathbb{Z}}$ and $\Lambda_0 \subseteq \Lambda$ is a sublattice so $\Lambda_0 = \langle w_1, ..., w_n \rangle$. Then we have a matrix $A \in M_n(\mathbb{Z})$ and $D = |\det(A)|$ so

$$A \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} = \begin{bmatrix} w_1 \\ \vdots \\ w_n \end{bmatrix}$$

Thus

$$\begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} = \frac{1}{\det(A)} \operatorname{Adj}(A) \cdot \begin{bmatrix} w_1 \\ \cdots \\ w_n \end{bmatrix} \Rightarrow \begin{bmatrix} Dv_1 \\ \vdots \\ Dv_n \end{bmatrix} = \frac{D}{\det(A)} \operatorname{Adj}(A) \cdot \begin{bmatrix} w_1 \\ \vdots \\ w_n \end{bmatrix}$$

where $\operatorname{Adj}(A)$ is the adjugate or adjoint matrix with defining property to be $A^{-1} = \frac{1}{\det(A)} \operatorname{Adj}(A)$. This above calculation shows $D\Lambda \subseteq \Lambda_0$ as $Dv_i \in \Lambda_0$.

Theorem 1.1.9. *Let Λ_1 be a sublattice of Λ in $\mathbb{R}^n$, then*

1. *If $\Lambda = \langle v_1, ..., v_n \rangle_{\mathbb{Z}}$, then there exists a basis $w_1, ..., w_n$ of Λ_1 and a lower triangular matrix $A \in M_n(\mathbb{Z})$ such that $A[v_1, ..., v_n]^T = [w_1, ..., w_n]^T$, $a_{ii} > 0$ for all i, and $0 \leq a_{ij} < a_{jj}$ for all $1 \leq j < i \leq n$.*
2. *If $w_1, ..., w_n$ is a basis of Λ_1 then there exists a basis $v_1, ..., v_n$ for Λ and lower triangular such that $A[v_1, ..., v_n]^T = [w_1, ..., w_n]^T$, $a_{ii} > 0$, and $0 \leq a_{ij} < a_{ii}$ for all $1 \leq j < i \leq n$.*

Proof. We prove (1) first. Let D be the index of Λ_1 in Λ, for each i with $1 \leq i \leq n$ there exists $w_i = a_{i1}v_1 + ... + a_{ii}v_i$ in Λ_1 with $a_{ij} \in \mathbb{Z}$ and $a_{ii} > 0$ since $Dv_i \in \Lambda_1$. We choose w_i for $i = 1, ..., n$ in such a way that a_{ii} is positive and as small as possible. Since $w_1, ..., w_n \in \Lambda_1$, we have $\sum b_i w_i \in \Lambda_1$ for all $b_i \in \mathbb{Z}$. Now we only need to show $\Lambda_1 = \langle w_1, ..., w_n \rangle_{\mathbb{Z}}$. Suppose not, then we can find $z \in \Lambda_1$ so $z \notin \langle w_1, ..., w_n \rangle_{\mathbb{Z}}$.

Then we can find $c_1, ..., c_n$ so $z = c_1v_1 + ... + c_nv_n$. Now choose $z \in \Lambda_1$ for which $c_{i+1} = ... = c_n = 0$ with i minimal. Then we have $z = c_1v_1 + ... + c_iv_i$. Let $c_i = qa_{ii} + r$ with $0 \leq r < a_{ii}$, then

$$z - qw_i = (c_1 - qa_{ii})v_1 + ... + rv_i$$

where $z - qw_i \in \Lambda_1$ and is an integer linear combination of $v_1, ..., v_i$. Also note $r \neq 0$ since i is minimal. This contradicts the minimal choice of a_{ii}. Thus $\Lambda_1 = \langle w_1, ..., w_n \rangle$ as desired.

It remains to check $0 \leq a_{ij} < a_{jj}$ for $1 \leq j < i \leq n$. To do this, we replace w_i with w_i' for $= 1, ..., n$ where

$$w_i' = b_{i1}w_1 + ... + b_{i,i-1}w_{i-1} + w_i$$

with the b_{ij}'s integers to be chosen. Note that $w_1', ..., w_n'$ is a basis for Λ_1 and that $w_i' = a_{i1}'v_1 + ... + a_{ii}'v_i$ with $a_{ii} = a_{ii}'$. In addition, we have

$$a_{ij}' = b_{ij}a_{jj} + b_{i,j+1}a_{j+1,j} + ... + b_{i,i-1}a_{i-1,j} + a_{ij}$$

For each i now we choose $b_{i,i-1}, ..., b_{i,1}$ in that order so that $0 \leq a_{ij}' < a_{jj} = a_{jj}'$ as required.

Now we check (2). Let $w_1, ..., w_n$ be a basis for Λ_1. Let D be the index, then $D\Lambda \subseteq \Lambda_1$. Hence, by part (1) we have a basis $Dv_1, ..., Dv_n$ for $D\Lambda$ so

$$\begin{aligned} Dv_1 &= a_{11}w_1 \\ &\vdots \\ Dv_n &= a_{n1}w_1 + ... + a_{nn}w_n \end{aligned}$$

with $a_{ij} \in \mathbb{Z}$. Put

$$A = \begin{bmatrix} a_{11} & ... & 0 \\ \vdots & \ddots & \vdots \\ a_{n1} & ... & a_{nn} \end{bmatrix}$$

$$A \begin{bmatrix} w_1 \\ \vdots \\ w_n \end{bmatrix} = D \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}$$

and so we see

$$\begin{bmatrix} w_1 \\ \vdots \\ w_n \end{bmatrix} = D \cdot \frac{\mathrm{Adj}(A)}{\det(A)} \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}$$

where $\mathrm{Adj}(A) \in M_n(\mathbb{Z})$ is also lower triangular, say $\mathrm{Adj}(A) = (b_{ij})$. Note w_i can be expressed as a rational linear combination of $v_1, ..., v_n$ (we get this from $\frac{D}{\det(A)} \mathrm{Adj}(A)$) and hence it must be a integral linear combination of $v_1, ..., v_n$ (by the unique representation of w_i by $v_1, ..., v_n$). This means $\frac{D}{\det(A)} \mathrm{Adj}(A)$ is actually integer matrix. Thus we obtained our desired matrix in the claim. Next, we show $a_{ii} > 0$ but to get that all we need to do is flip the sign of a_{ii} if needed. Finally, for $1 \leq j < i \leq n$, we replace v_i by v_i' where $v_i' = c_{i1}v_1 + ... + c_{i,i-1}v_{i-1} + v_i$ where c_{ij}'s are integers chosen as in (1) to ensure the last assertion. ♡

Corollary 1.1.9.1. *Let Λ be a lattice in $\mathbb{R}^n$, and $w_1, ..., w_m$ be linearly independent vectors of Λ. Then there exists a basis $v_1, ..., v_n$ of Λ for which*

$$\begin{aligned} w_1 &= a_{11}v_1 \\ w_2 &= a_{21}v_1 + a_{22}v_2 \\ &\vdots \\ w_m &= a_{m1}v_1 + ... + a_{mm}v_m \end{aligned}$$

with $a_{ij} \in \mathbb{Z}$, $a_{ii} > 0$ and $0 \leq a_{ij} < a_{ii}$ for $1 \leq j < i \leq m$.

Proof. Extend $w_1, ..., w_m$ to a set of n linearly independent vectors $w_1, ..., w_n \in \Lambda$. Then consider the sublattice $\langle w_1, ..., w_n \rangle_{\mathbb{Z}}$ and apply Theorem 1.1.9. ♡

Corollary 1.1.9.2. *Let $w_1, ..., w_m$ be linearly independent vectors from a lattice Λ in $\mathbb{R}^n$ with $m < n$. There exist $w_{m+1}, ..., w_n$ in Λ such that $w_1, ..., w_n$ is a basis for Λ if and only if every vector $\sum_{i=1}^{m} a_i w_i \in \Lambda$ with $a_i \in \mathbb{R}$ for $i = 1, ..., m$ has in fact $a_i \in \mathbb{Z}$.*

Proof. $\Rightarrow$: Immediate.

$\Leftarrow$: We apply Corollary 1.1.9.1 to get a basis $v_1, ..., v_n$ of Λ with

$$\begin{aligned} w_1 &= a_{11}v_1 \\ &\vdots \\ w_m &= a_{m1}v_1 + ... + a_{mm}v_m \end{aligned}$$

Thus, $v_1 = \frac{1}{a_{11}}w_1$ and we get, by hypothesis, $\frac{1}{a_{11}} \in \mathbb{Z}$, i.e. $a_{11} = 1$ as $a_{11} \in \{1, -1\}$ and $a_{11} > 0$. Next, $w_2 = a_{21}v_1 + a_{22}v_2$ and so $\frac{1}{a_{22}}w_2 = \frac{a_{21}}{a_{22}}w_1 + v_2$ and so $a_{22} = 1$ as desired. In this way, we see $a_{11} = a_{22} = ... = a_{mm} = 1$ and $w_1, ..., w_m, v_{m+1}, ..., v_n$ is a basis of Λ. ♡

Corollary 1.1.9.3. *Let $v_1, ..., v_n$ be a basis of Λ and $w = a_1v_1 + ... + a_nv_n$ in Λ. Let $1 \leq m \leq n-1$, then $v_1, ..., v_{m-1}, w$ can be extended to a basis for Λ if and only if $\gcd(a_m, ..., a_n) = 1$.*

Proof. Let $g = \gcd(a_m, ..., a_n)$. If $v_1, ..., v_{m-1}, w$ can be extend to a basis, say $\langle v_1, ..., v_{m-1}, w, w_{m+1}, ..., w_n \rangle_{\mathbb{Z}} = \Lambda$, then

$$w - a_1v_1 - ... - a_{m-1}v_{m-1} = a_mv_m + ... + a_nv_n$$

therefore,

$$\frac{1}{g}(w - a_1v_1 - ... - a_{m-1}v_{m-1}) = \frac{a_m}{g}v_m + ... + \frac{a_n}{g}v_n$$

Now, $a_t/g \in \mathbb{Z}$ for $t = m, ..., n$. This means $\frac{1}{g}w - \frac{a_1}{g}v_1 - ... - \frac{a_{m-1}}{g}v_{m-1}$ is in Λ. Now apply Corollary 1.1.9.2 to conclude $1/g \in \mathbb{Z}$ and so $g = 1$.

Conversely, we wish to find $w_{m+1}, ..., w_n$ in Λ for which $v_1, ..., v_{m-1}, w, w_{m+1}, ..., w_n$ is a basis for Λ. Then

$$\begin{aligned} v_1 &= v_1 \\ &\vdots \\ v_{m-1} &= v_{m-1} \\ w &= a_{m,1}v_1 + ... + a_{m,m}v_m + ... + a_{m,n}v_n \\ w_{m+1} &= a_{m+1,1}v_1 + ... + a_{m+1,m}v_m + ... + a_{m+1,n}v_n \\ &\vdots \\ w_n &= a_{n,1}v_1 + ... + a_{n,m}v_m + ... + a_{n,n}v_n \end{aligned}$$

It suffice to show that we can choose coefficients $a_{j,k}$ with $m \leq j \leq n$ and $1 \leq k \leq n$ as integers in a way so the matrix associated with it has determinant ± 1. This is the same as to show the row $(a_m, ..., a_n)$ can be extented to an $(n-m+1) \times (n-m+1)$ matrix invertible in the ring $\mathbb{Z}$.

Consider the standard lattice $\Lambda_0 = \langle e_1, ..., e_{n-m+1} \rangle$ in $\mathbb{R}^{n-m+1}$, it suffice to show we can extend $(a_m, ..., a_n)$ to a basis for Λ_0. We appeal to Corollary 1.1.9.2. Notice if $\alpha \in \mathbb{R}$ with $\alpha \neq 0$ and $\alpha(a_m, ..., a_n) \in \Lambda_0$, then $\alpha \in \mathbb{Q}$. Say $\alpha = \frac{p}{q}$ with $\gcd(p, q) = 1$. Then $(\frac{pa_m}{q}, ..., \frac{pa_n}{q}) \in \Lambda_0$, hence $q \mid pa_m, ..., q \mid pa_n$ and since p, q are coprime we get $q \mid \gcd(a_m, ..., a_n)$. ♡

Remark 1.1.10. We use $\cdot$ to mean standard inner product on $\mathbb{R}^n$. Also recall the existence of dual basis for any given basis of $\mathbb{R}^n$, i.e. for basis $v_1, ..., v_n$ we have a v_i^* basis so $v_i^* \cdot v_j = \delta_{ij}$.

Definition 1.1.11. Let Λ be a lattice, define the ***polar lattice*** of Λ to be $\Lambda^* := \{v^* : v \in \Lambda\}$.

Theorem 1.1.12. *Let Λ be a lattice, then Λ^* contains exactly the set of vectors v so $v \cdot u \in \mathbb{Z}$ for all $u \in \Lambda$, $(\Lambda^*)^* = \Lambda$ and*

$$d(\Lambda)d(\Lambda^*) = 1$$

Proof. Let $\Lambda = \langle v_1, ..., v_n \rangle_{\mathbb{Z}}$ and let $v_1^*, ..., v_n^*$ be a basis for λ^*. If v is in Λ then we see $v = \sum a_i v_i$ while if $u \in \Lambda^*$ then we see $u = \sum b_i v_i^*$, hence we see

$$v \cdot u = \sum a_i b_i$$

as desired. Conversely, let w be a vector for which $w \cdot v \in \mathbb{Z}$ for all $v \in \Lambda$, let $c_i := w \cdot v_i$. Let $v^* = \sum c_i v_i^* \in \Lambda^*$, we will show $w = v^*$. To that end, note

$$(w - v^*) \cdot v_i = 0$$

for all i where $v_1, ..., v_n$ are linearly independent, i.e. we are forced to have $w = v^*$ as desired.

By what we proved we see $(\Lambda^*)^* = \Lambda$ as desired. Finally, note

$$\det(v_1^*, ..., v_n^*) \det(v_1, ..., v_n)$$

and hence $d(\Lambda)d(\Lambda^*) = 1$ as desired. ♡

Remark 1.1.13. Note if $w = (y_1, ..., y_n)$ is in $\mathbb{R}^n$, the set of x for which $x \cdot w = 0$ is given by $x_1y_1 + ... + x_ny_n = 0$, i.e. it is a hyperplane in $\mathbb{R}^n$.

Proposition 1.1.14. *Let Λ be a lattice in $\mathbb{R}^n$ and u be a vector in $\mathbb{R}^n$. There exist $n-1$ linearly independent vectors $w_1, ..., w_{n-1}$ in Λ with $u \cdot w_i = 0$ for $i = 1, ..., n-1$ if and only if $u = t \cdot w^*$ with $t \in \mathbb{R}$ and $w^* \in \Lambda^*$.*

Proof. $\Rightarrow$: By Corollary 1.1.9.1 there is a basis $v_1, ..., v_n$ of Λ such that $w_i = a_{i1}v_1 + ... + a_{ii}v_i$ with $a_{ij} \in \mathbb{Z}, a_{ii} \neq 0$ for $i = 1, ..., n-1$. Since $u \cdot w_i = 0$ for $i = 1, ..., n-1$ we see $u \cdot v_i = 0$ for $i = 1, ..., n-1$. Put $u \cdot v_n = t$ for some $t \in \mathbb{R}$. Note that if $v_1^*, ..., v_n^*$ is a polar basis for Λ^* then $u = tv_n^*$ as required.

$\Leftarrow$: If $w^* = 0$ then $u = 0 \in \mathbb{R}^n$ and so $u \cdot w_i = 0$ for $i = 1, ..., n-1$. Thus suppose $w^* \neq 0$. Put $w^* = m \cdot v_1^*$ where m is a positive integer and v_1^* is such that $\frac{1}{k}v_1^*$ is not in Λ^* for any integer k with $k \geq 2$ (such v_1^* is said to be primitive[1]). By Corollary 1.1.9.2 we can extend v_1^* to a basis $v_1^*, ..., v_n^*$ of Λ^*. Let $v_1, ..., v_n$ be a basis for the polar lattice Λ of Λ^*. Then $v_1^* \cdot v_j = 0$ for $j = 2, ..., n$ and so $w^* \cdot v_j = 0$ for $j = 2, ..., n$ as required. ♡

Remark 1.1.15. It follows from the proof of Proposition 1.1.14 that if $w^* \in \Lambda^*$ then we can associate to it a lattice $\Lambda(w^*)$ in $\mathbb{R}^{n-1}$ with basis $v_2, ..., v_n$.

1.2 Minkowski's Theorems

Definition 1.2.1. Let U be the unit interval given by $U = \{t \in \mathbb{R} : 0 \leq t < 1\}$ and let U^n be the unit n-cube given by $U^n = \{(x_1, ..., x_n) : \forall 1 \leq i \leq n, 0 \leq x_i < 1\}$.

Definition 1.2.2. Let $x \in \mathbb{R}^n$ we define

$$\|x\|_\infty := \max_{1 \leq i \leq n} |x_i|$$

This is not known as the ***house*** of x and we shall just call it the L-infinity norm.

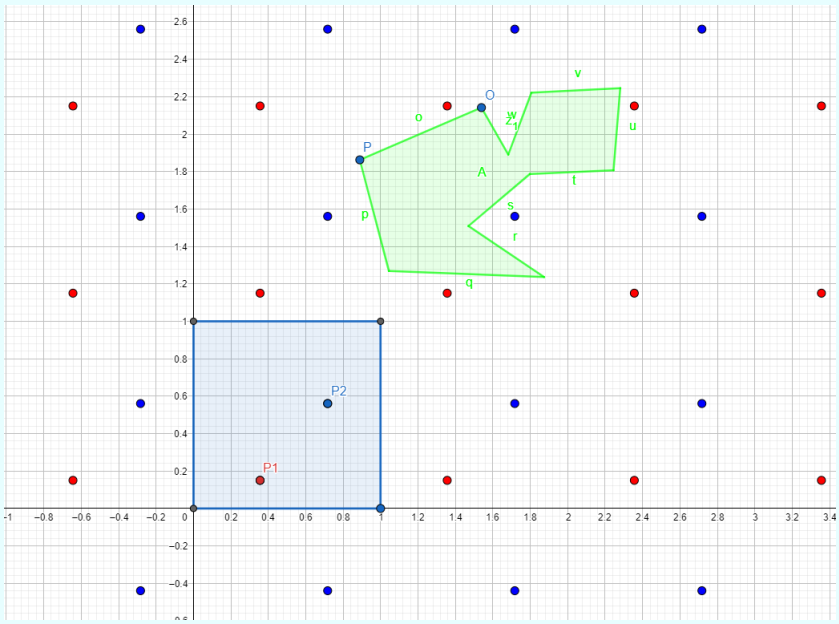
Definition 1.2.3. If $x \in \Lambda_0 = \langle e_1, ..., e_n \rangle_\mathbb{Z}$ we say x is an ***integer point***.

Definition 1.2.4. For any set $T \subseteq \mathbb{R}^n$, $x \in \mathbb{R}^n$ and $\lambda \in \mathbb{R}$ we define $T + x = x + T := \{y + x : y \in T\}$ and $\lambda T = \{\lambda x : x \in T\}$.

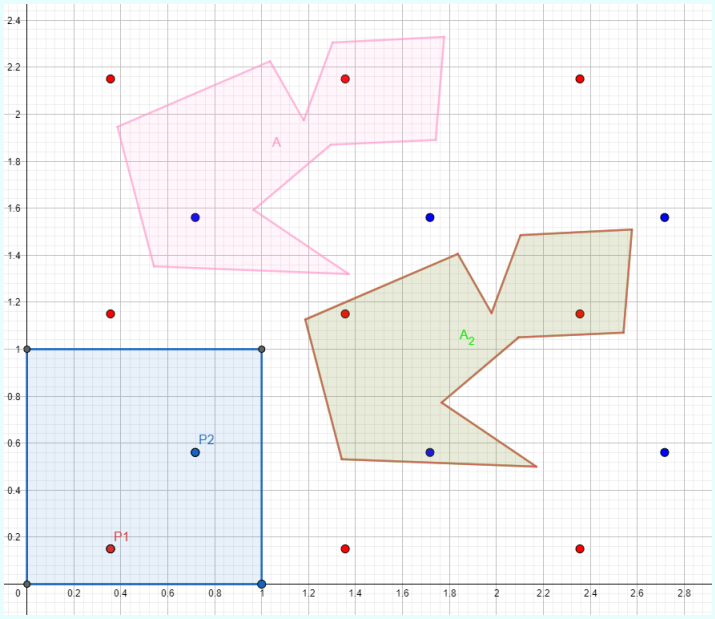
[1] Primitive element exists. Suppose for all $v \in \Lambda$ we can find k so $\frac{1}{k}v \in \Lambda$. Let $v_1, ..., v_n$ be a basis of Λ then we see $\frac{1}{k}v_1 \in \Lambda$ for some $k \geq 2$. In particular by Corollary 1.1.9.2 we see then v_1 cannot be extended to a basis as we can find rational value $\frac{1}{k} \notin \mathbb{Z}$ so $\frac{1}{k}v_1 \in \Lambda$, a contradiction.

Theorem 1.2.5 (Blichfeldt). *Let P be a non-empty set of points in $\mathbb{R}^n$ which is invariant by translation by integer points and has N points in U^n. Let A be a subset of $\mathbb{R}^n$ of positive Lebesgue measure $\mu(A)$, then there is an $x \in U^n$ such that $A + x$ contains at least $N \cdot \mu(A)$ points of P. Further, if A is compact then there is an $x \in U^n$ so $A + x$ contains more than $N \cdot \mu(A)$ points of P.*

Proof. Before we begin, let us picture one instance of this as follows: consider



where the green shape is our A with $\mu(A)$ approximately 0.76, blue square is U^2, and blue and red points are elements of P. Then, by translating A, we see we get



where the pink A on the left correspond to the general case (we find 2 points inside

it) and the brown A on the right correspond to the compact case (we even find more points inside it). Note in this case I did not translate them by a vector in U^2 (I just moved them around randomly until I find it contains two points) but I think we get the picture here.

Now we begin the proof. For any set S in $\mathbb{R}^n$ we let $v(S)$ be the number of points of P in S. Let $p_1, ..., p_N$ be the N points of P in U^n. We put $P_i := \{p_i + g : g \in A_0\}$ for $i = 1, ..., N$. Since P is invariant by translation by integer points, we see

$$P = \bigcup_{i=1}^{N} P_i$$

Further we have $P_i \cap P_j = \emptyset$ for $i \neq j$ (if $v \in P_i \cap P_j$ then we see we can find $g \in \Lambda_0$ so $p_i = p_j + g$ but both $p_i, p_j \in U^n$, which is absurd). Now for any $S \subseteq \mathbb{R}^n$ let $v_i(S)$ be the number of points of P_i in S for $i = 1, ..., N$. Let χ be the characteristic function of A, then

$$v_i(A + x) = \sum_{g \in \Lambda_0} \chi(p_i + g - x)$$

and we have

$$\begin{aligned} \int_{U^n} v_i(A + x)dx &= \int_{U^n} \sum_{g \in \Lambda_0} \chi(p_i + g - x)dx \\ &= \int_{\mathbb{R}^n} \chi(z)dz \\ &= \mu(A) \end{aligned}$$

Thus we see

$$\int_{U^n} v(A + x)dx = N\mu(A)$$

and so there is some element $x_1 \in U^n$ such that $v(A + x_1) \geq N\mu(A)$ and so $A + x_1$ contains at least $N\mu(A)$ points of P. Now suppose A is compact. If $N\mu(A)$ is not an integer there is nothing more to prove. Hence suppose $N\mu(A) = h$ for $h \in \mathbb{Z}^+$. For $k = 1, 2, ...$ we define A_k by $A_k = (1 + \frac{1}{k})A$, by what we just proved, for each k there is $x_k \in U^n$ so $v(A_k + x_k) \geq h + 1$. Since $x_k \in \overline{U^n}$ and $\overline{U^n}$ is compact we see we can find a subsequence x_{k_j} converges to a point x in $\overline{U^n}$. Since A is compact the sets $A_k + x_k$ are uniformly bounded and so contain only finitely many points of P.

Each of the sets $A_{k_j} + x_{k_j}$ contain at least $h + 1$ points of P and so we may assume by taking a further subsequence that there are $h + 1$ points, say $u_1, ..., u_{h+1}$, which occur in each set $A_{k_j} + x_{k_j}$ (suppose not, then $\bigcup A_{k_j} + x_{k_j}$ contains infinite many points of P, contradiction). $A + x$ is compact and in fact contains $u_1, ..., u_{h+1}$. Indeed, suppose that is not the case then $u_i \notin A + x$ for some i. However, then u_i is a positive distance away from $A + x$ and this cannot be the case, since $x_{k_j} \to x$, i.e. the distance from a point in A_{k_j} to the nearest point in A tends to 0 as $k_j \to \infty$. Hence $A + x$ must contain $h + 1$ of the points of P. We now choose g so $x - g \in U^n$ and then $A + x - g$ contains $h + 1$ points of P as required since P is invariant by translation by integer points. ♡

Definition 1.2.6. Let $S \subseteq \mathbb{R}^n$, we say S is ***symmetric about origin*** if $x \in S \Rightarrow -x \in S$.

Definition 1.2.7. Let $S \subseteq \mathbb{R}^n$, we say S is ***convex*** if $x, y \in S \Rightarrow \forall \lambda \in [0,1], \lambda x + (1-\lambda)y \in S$.

Theorem 1.2.8 (Minkowski's Convex Body Theorem)**.** *Let $A \subseteq \mathbb{R}^n$ be convex and symmetric about origin with volume $\mu(A)$. If $\mu(A) > 2^n$ or if A is compact and $\mu(A) \geq 2^n$, then A contains an integer point different from 0.*

Proof. Note $\mu(\frac{1}{2}A) > 1$, or if A is compact then $\mu(\frac{1}{2}A) \geq 1$. By Blichfeldt's Theorem 1.2.5 applied to $\frac{1}{2}A$ where $P = \Lambda_0$ we see we can find $x \in U^n$ so $\frac{1}{2}A + x$ contains two distinct integer points g_1, g_2. Then $g_1 - x, g_2 - x \in \frac{1}{2}A$ and so $g_1 - x = \frac{1}{2}x_1$ and $g_2 - x = \frac{1}{2}x_2$ with $x_1, x_2 \in A$. By symmetry we have $-(g_2 - x) = x - g_2 = -\frac{1}{2}x_2$ with $-x_2 \in A$. Since A is convex we have $\frac{1}{2}x_1 + \frac{1}{2}(-x_2) \in A$ and so $g_1 - g_2 = g_1 - x + x - g_2 \in A$. However $g_1 - g_2 \in \Lambda_0$ and since g_1, g_2 are distinct, the proof follows. ♡

Theorem 1.2.9 (Minkowski's Linear Forms Theorem)**.** *Let $B = (B_{ij})$ be an n by n matrix with real entries and non-zero determinant. Let $c_1, ..., c_n$ be positive real numbers with $c_1, ..., c_n \geq |\det(B)|$. Then there exists an integer point $x = (x_1, ..., x_n)$ different from 0 for which*

$$\forall i = 1, ..., n-1, |B_{i1}x_1 + ... + B_{in}x_n| < c_i$$

$$|B_{n1}x_1 + ... + B_{nn}x_n| \leq c_n$$

Proof. By dividing the ith row of B, say b_i, by $\frac{1}{c_i}$ we may assume $c_1 = ... = c_n = 1$ and $0 < |\det(B)| \leq 1$. Let $L_1(x), ..., L_n(x)$ be linear forms given by $L_i(x) = B_{i1}x_1 + ... + B_{in}x_n = b_i^T x$ for $i = 1, ..., n$. Then we wish to solve the system $|L_i(x)| < 1$ for $1 \leq i \leq n-1$ and $|L_n(x)| \leq 1$.

Let A be the set of $x \in \mathbb{R}^n$ for which $|L_i(x)| \leq 1$ for $i = 1, ..., n$, then A is symmetric about the origin. Also we claim A is convex. Indeed, let $x, y \in A$ and $0 \leq \lambda \leq 1$ then we have

$$\begin{aligned} |L_i(\lambda x + (1-\lambda)y)| &= |b_i^T(\lambda x + (1-\lambda)y)| \\ &= |\lambda \cdot b_i^T x + (1-\lambda) \cdot b_i^T y| \\ &\leq \lambda |b_i^T x| + (1-\lambda)|b_i^T y| \\ &\leq \lambda + 1 - \lambda = 1 \end{aligned}$$

In addition, we note

$$\mu(A) = \frac{1}{|\det(B)|}\mu(C^n)$$

where $C^n = \{(x_1, ..., x_n) \in \mathbb{R}^n : |x_i| \leq 1\}$. Hence $\mu(A) \geq 2^n$ and so by Minkowski's Convex Body theorem 1.2.8 we see there is an integer point x with $x \neq 0$ in A.

Finally, to get strict inequality in the first $n-1$ inequalities we introduce, for each $\epsilon > 0$ the set A_ϵ given by the inequalities $|L_i(x)| < 1$ for $i = 1, ..., n-1$ and

$|L_n(x)| < 1 + \epsilon$. Then $\mu(A_\epsilon) \geq (1+\epsilon)2^n > 2^n$ and sy by Minkowski's Convex Body Theorem again we get $x_\epsilon \in A_\epsilon \cap \Lambda_0$ with $x_\epsilon \neq 0$. Now take any sequence ϵ_k of positive reals which decreases to 0, we get a sequence x_{ϵ_k} associated with it. Since $\bigcup_{k=1}^{\infty} A_{\epsilon_k}$ is bounded we can find an integer point x in infinitely many of A_{ϵ_k} and hence we get our desired x. ♡

Theorem 1.2.10. *Let $a_{ij} \in \mathbb{R}$ with $1 \leq i \leq n$, $1 \leq j \leq m$, and let Q be a real number with $Q > 1$. Then there exists integers $q_1, ..., q_m$ and $p_1, ..., p_n$ with*

$$0 \leq \max_{1 \leq j \leq m} |q_j| < Q^{n/m}$$

and for all $i = 1, ..., n$ that

$$|a_{i1}q_1 + ... + a_{im}q_m - p_i| < \frac{1}{Q}$$

Proof. Let $l = m + n$ and consider the l linear forms $L_1, ..., L_l$ in $\mathbb{R}^l$ given by $L_i(x_1, ..., x_l) = x_i$ for $i = 1, ..., m$ and

$$L_{m+j}(x_1, ..., x_l) = a_{j1}x_1 + ... + a_{jm}x_m - x_{m-j}$$

Note that the determinants of the matrix associated with $L_1, ..., L_l$ is $(-1)^n$. Let $Q > 1$ and apply Minkowski's Linear Form Theorem 1.2.9 to the system of inequalities

$$\begin{cases} |L_i(x)| < Q^{n/m}, & \text{if } i = 1, ..., m \\ |L_{m+j}(x)| \leq \frac{1}{Q}, & \text{if } j = 1, ..., n \end{cases} \tag{eq 1}$$

to find a non-zero integer point x satisfying equation eq 1. We now put $q_i = x_i$ for $i = 1, ..., m$ and $p_j = x_{m+j}$ for $j = 1, ..., n$. Then

$$q = \max |q_i| < Q^{n/m}$$

and

$$|a_{j1}q_1 + ... + a_{jm}q_m - p_j| \leq \frac{1}{Q}$$

It remains to check $q \neq 0$. Suppose otherwise, then $q_1 = ... = q_m = 0$ and so

$$|p_j| \leq \frac{1}{Q}$$

for $j = 1, ..., n$. However $Q > 1$ so $p_1 = ... = p_n = 0$ and this contradicts the fact that x is a non-zero point. The result follows. ♡

Corollary 1.2.10.1. *Let a_{ij} be real numbers with $1 \leq i \leq n, 1 \leq j \leq m$. Suppose that for some t with $1 \leq t \leq n$, $1, a_{t1}, ..., a_{tm}$ are linearly independent over the rationals. Then there exists infinitely many coprime $m+n$-tuples of integers $(q_1, ..., q_m, p_1, ..., p_n)$ with $q = \max |q_j| > 0$ and*

$$|a_{i1}q_1 + ... + a_{im}q_m - p_i| < \frac{1}{q^{m/n}}$$

Proof. Take $Q = 2$ in the above theorem. By Theorem 7 there exists a solution $q_1, ..., q_m, p_1, ..., p_n$ satisfy the equations $|a_{i1}q_1 + ... + a_{im}q_m - p_i| < \frac{1}{q^{m/n}}$. We now divide through by the gcd of $q_1, ..., q_m$, $p_1, ..., p_n$ to give us a solution of the equation with a coprime $m+n$-tuples. Thus we may suppose, WLOG, that $q_1, ..., p_n$ are all coprime. Let

$$|q_1 a_{t1} + ... + q_m a_{tm}| = \delta_t$$

and we note $\delta_t > 0$ since $1, \alpha_{t1}, ..., \alpha_{tm}$ are linearly independent over $\mathbb{Q}$.

We now apply Theorem 1.2.10 with Q so that $\frac{1}{Q} < \delta_t$ to get a new $m+n$-tuple satisfy the equation we want to get. As before, we may divide by gcd to get coprime tuples. Repeat this process gives us infinitely many coprime tuples satisfying the equation. ♡

1.3 Sets Containing Lattice Points

Theorem 1.3.1. *Let Λ be a lattice in $\mathbb{R}^n$ and let A be a convex set in $\mathbb{R}^n$ which is symmetric about the origin and has volume greater than $2^n d(\Lambda)$, or if A is compact then $\geq 2^n d(\Lambda)$. Then A contains a point of Λ different from 0.*

Proof. Suppose $v_1, ..., v_n$ is a basis for Λ. Let $v_j = (a_{j1}, ..., a_{jn})$ for $j = 1, ..., n$. Let T be the linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^n$ associated with the matrix (a_{ij}). Then $\Lambda = T\Lambda_0$. Note $\mu(T^{-1}A) = d(\Lambda)^{-1}\mu(A)$ and that $T^{-1}A$ is convex and symmetric. Hence the result follows from Minkowski's Convex Body Theorem 1.2.8 ♡

Definition 1.3.2. We define Γ to be the function $\Gamma(z) = \int_0^\infty x^{z-1}e^{-x}dx$ if $\mathrm{Re}(z) > 0$.

Proposition 1.3.3. *Let R be a positive real number and n be a positive integer. The volume of the sphere of radius R in $\mathbb{R}^n$ is $\omega_n R^n$ where $\omega_n = \frac{\pi^{n/2}}{\Gamma(1+\frac{n}{2})}$*

Proof. It suffice to show ω_n is the volume of the unit sphere given by $\{(x_1, ..., x_n) : \sum x_i^2 \leq 1\}$. We have $\omega_1 = 1$ and $\omega_2 = \pi$ and we proceed inductively. Suppose $n \geq 3$ then

$$\omega_n = \int_{\sum x_i^2 \leq 1} dx_1 ... dx_n = \int_{-1}^{1}\int_{-1}^{1}\left(\int_{\mathbb{R}^{n-2}} g(x_1, ..., x_n)\right) dx_{n-1}dx_n$$

where g is the characteristic function of the unit sphere. Thus

$$\begin{aligned}\omega_n &= \int_{x_{n-1}^2+x_n^2\leq 1} \omega_{n-2}(1-x_{n-1}^2-x_n^2)^{(n-2)/2}dx_{n-1}dx_n \\ &= \omega_{n-2}\int_{x_{n-1}^2+x_n^2\leq 1}(1-x_{n-1}^2-x_n^2)^{(n-2)/2}dx_{n-1}dx_n \\ &= \omega_{n-2}\int_0^{2\pi}\int_0^1 (1-r^2)^{(n-2)/2}drd\theta \\ &= 2\pi\omega_{n-2}(-\frac{1}{n}(1-r^2)^{n/2})|_0^1 \\ &= \frac{2\pi}{n}\omega_{n-2}\end{aligned}$$

Thus $\omega_{2n} = \frac{2\pi}{2n}\cdot\frac{2\pi}{2(n-1)}...\cdot\frac{2\pi}{4}\cdot\frac{2\pi}{2} = \frac{\pi^n}{n!}$ and

$$\omega_{2n+1} = \frac{2\pi}{2n+1}\cdot\frac{2\pi}{2n-1}...\frac{2\pi}{3}\cdot 2 = \frac{\pi^n}{(n+\frac{1}{2})(n-\frac{1}{2})...(\frac{3}{2})(\frac{1}{2})}$$

The result now follows from that fact $\Gamma(x+1) = x\Gamma(x)$ for $x > 0$ and $\Gamma(\frac{1}{2}) = \sqrt{\pi}$ ♡

Theorem 1.3.4. *Let Λ be a lattice in $\mathbb{R}^n$, there is a non-zero element $x \in \Lambda$ which*

$$0 \leq x\cdot x = \sum x_i^2 \leq 4\cdot(\omega_n^{-1}d(\Lambda))^{2/n}$$

Proof. We apply Theorem 1.3.1 to the set $A = \{x \in \mathbb{R}^n : \sum_{i=1}^n x_i^2 \leq t\}$ with $t = 4\cdot(\omega_n^{-1}d(\Lambda))^{2/n}$. Then $\mu(A) = \omega_n t^{n/2} = 2^n d(\Lambda)$. Since A is convex, symmetric, the result follows. ♡

Remark 1.3.5. Note Theorem 1.3.4 is close to the truth since Minkowski constructed for each $n \geq 1$ a lattice Λ for which

$$\min_{x\in\Lambda\setminus\{0\}} x\cdot x \geq (\omega_n^{-1}d(\Lambda))^{2/n}$$

In particular, Theorem 1.3.4 cannot be improved by more than a factor of 4. Rogers was able to Theorem 1.3.4 somewhat. He replaced $4\omega_n^{-2/n}$ by $4(\frac{\sigma_n}{\omega_n})^{2/n}$ where σ_n is the quotient of two geometrical figures with the property that $\sigma_n \sim \frac{n}{e2^{n/2}}$ as $n \to \infty$. We have

$$\omega_n^{-2/n} \sim \frac{n}{2\pi e}, 4(\frac{\sigma_n}{\omega_n})^{2/n} \sim \frac{n}{\pi e}, 4\omega_n^{-2/n} \sim \frac{2n}{\pi e}$$

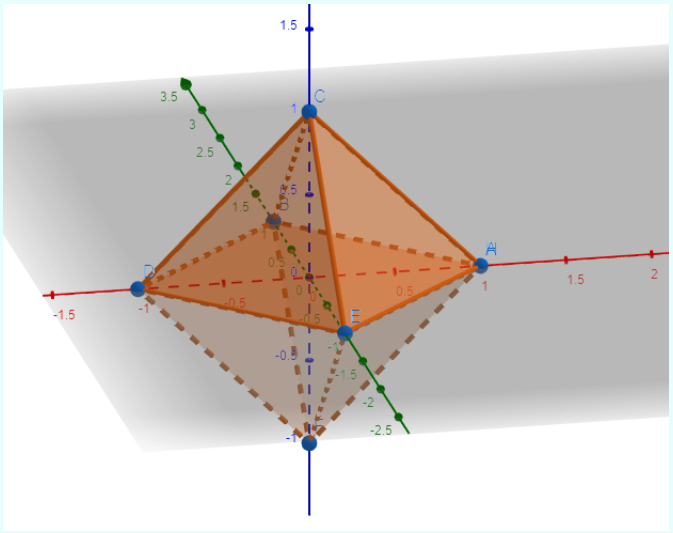
Definition 1.3.6. For each $\lambda \in \mathbb{R}$ with $\lambda > 0$, define

$$A_\lambda^{(n)} = A_\lambda = \{(x_1,...,x) \in \mathbb{R}^n : |x_1| + ... + |x_n| \leq \lambda\}$$

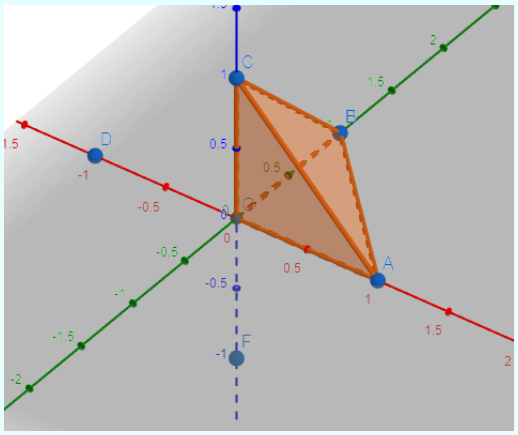
Definition 1.3.7. For $n \geq 1$, define

$$A_\lambda^{(n)^+} = A_\lambda^+ = \{(x_1,...,x_n) : \sum x_i \leq \lambda, 0 \leq x_i \leq \lambda\}$$

Example 1.3.8. The following is A_1^3:



The following is A_1^{3+}:



Remark 1.3.9. A easy computation shows

$$\begin{aligned}
\mu(A_1^+) &= \int_0^1 \int_0^{1-x_1} \dots \int_0^{1-x_1-x_2-\dots-x_{n-1}} dx_n \dots dx_1 \\
&= \int_0^1 \int_0^{1-x_1} \dots \int_0^{1-x_1-\dots-x_{n-2}} (1-x_1-\dots-x_{n-1}) dx_{n-1} \dots dx_1 \\
&= \int_0^1 \int_0^{1-x_1} \dots \int_0^{1-x_1-\dots-x_{n-3}} \frac{(1-x_1-\dots-x_{n-2})^2}{2} dx_{n-2} \dots dx_1 \\
&= \frac{1}{n!}
\end{aligned}$$

Hence $\mu(A_\lambda^{(n)}) = \frac{2^n \lambda^n}{n!}$.

Also, it is not hard to see A_λ is symmetric and convex.

Theorem 1.3.10. *Let Λ be a lattice of $\mathbb{R}^n$. Then there is a non-zero point $x \in \Lambda$ with*

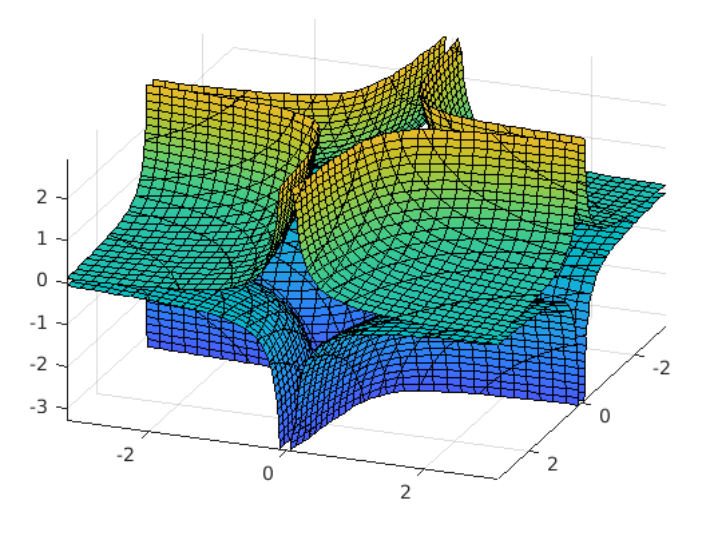
$$|x_1| + ... + |x_n| \leq (n!d(\Lambda))^{1/n}$$

Proof. Apply Theorem 1.3.1 to the set A_λ where $\lambda = (n!d(\Lambda))^{1/n}$. Then $\mu(A_\lambda) = 2^n d(\Lambda)$ and hence the result follows. ♡

Definition 1.3.11. Let $n \geq 1$ and $\lambda > 0$ with $\lambda \in \mathbb{R}$. Define

$$B_\lambda^{(n)} = B_\lambda = \{(x_1, ..., x_n) : |x_1...x_n| \leq \lambda^n\}$$

Example 1.3.12. The following is the surface of $B_1^{(3)}$ (to think $B_1^{(3)}$ just fill the space between those surfaces):



which is generated using MATLAB with the code

```
f = @(x,y,z) abs(x.*y.*z)-1;
fimplicit3(f)
```

Remark 1.3.13. Note B_λ is not convex. However, by arithmetic-geometric mean inequality, we see given non-negative x_i we have

$$(x_1...x_n)^{1/n} \leq \frac{x_1 + ... + x_n}{n}$$

and hence B_λ contains $A_{n\lambda}^{(n)}$ where $A_{n\lambda}^{(n)}$ is convex.

Theorem 1.3.14. *Let C be non-singular n by n matrix with entries from $\mathbb{R}$ and let $L_i(x) = \sum_j c_{ij}x_j$. Then there exists an integer point x different from 0 for which*

$$|L_1(x)...L_n(x)| \leq \frac{n!}{n^n}|\det(C)|$$

Proof. We apply Theorem 1.3.10 with the lattice $\Lambda := C\Lambda_0$ determined by C and $\lambda := \frac{(n!\det(C))^{1/n}}{n}$. Note since $B_\lambda^{(n)}$ contains $A_{n\lambda}^{(n)}$ we see the proof follows. ♡

Remark 1.3.15. If Λ_1 is a sublattice of Λ in $\mathbb{R}^n$, then clearly we can define Λ/Λ_1 as the quotient of $\mathbb{Z}$-module. Note this is finite and it has order equal the index of $[\Lambda : \Lambda_1]$, as we will prove next.

Proposition 1.3.16. *Let Λ_1 be sublattice of $\Lambda \subseteq \mathbb{R}^n$, then the index of Λ_1 in Λ is equal $|\Lambda/\Lambda_1|$.*

Proof. We have $\Lambda = \langle v_1, ..., v_n\rangle$ and $\Lambda_1 = \langle w_1, ..., w_n\rangle$ where the vectors are of the form of Theorem 1.1.9. Then the index is $\prod a_{ii}$. We claim that every vector u in Λ is equivalent to precisely one of $\sum q_i v_i$ with $0 \leq q_i < a_{ii}$ for $i = 1, ..., n$.

Let $u = \sum u_i v_i \in \Lambda$. First we shift u by a multiple w_n to find an equivalent vector with nth coordinate in the range $0 \leq q_n < a_{nn}$. Next we subtract a multiple w_{n-1} from this vector to get q_{n-1} in the range $0 \leq q_{n-1} < a_{n-1,n-1}$. Continuing this we see u is equivalent to a vector of the form $\sum q_i v_i$ with $0 \leq q_i < a_{ii}$. It remains to show no two vectors of the form $\sum q_i v_i$ are equivalent.

Suppose $\sum q_i v_i \sim \sum p_i v_i$ in Λ/Λ_1. Then consider $\sum r_i v_i$ with $r_i := q_i - p_i$. We see $|r_i| < a_{ii}$ and let j be the largest integer so $r_j \neq 0$. Then we may replace w_j in the basis $w_1, ..., w_n$ of Λ_1 with $w_j - \lambda \sum r_i v_i$ where λ is some multiple that makes the resulting basis lower triangular form but with a_{jj} replaced by a smaller non-negative integer. The final reduction does not change the diagonal, but the resulting determinant is different, which gives a contradiction. Hence the index is $\prod a_{ii}$. ♡

Remark 1.3.17. Let A be a convex subset of $\mathbb{R}^n$ which is symmetric about origin and of finite volume. Let Λ be a lattice in $\mathbb{R}^n$. Minkowski introduced minima $\lambda_1, ..., \lambda_n$ associated with Λ and A by putting

$$\lambda_j(A, \Lambda) = \lambda_j = \inf\{\lambda \in \mathbb{R} : \lambda A \text{ contains } j \text{ linearly independent vectors of } \Lambda\}$$

Then we see $0 < \lambda_1 \leq \lambda_2 \leq ... \leq \lambda_n < \infty$. Minkowski, in what is known as Minkowski's Second Theorem on Convex Bodies proved that

$$\frac{2^n d(\Lambda)}{n!} \leq \lambda_1 ... \lambda_n \mu(A) \leq 2^n d(\Lambda)$$

We will not give a proof in this note, but will remark the upper bound is tricky.

This result is sharp, in the sense that neither the upper bound or the lower bound can be improved in general, as we will show next.

Take any positive reals $\gamma_1, ..., \gamma_n$ with $\gamma_1 \leq \gamma_2 \leq ... \leq \gamma_n < \infty$. Consider the lattice $\Lambda = \langle \gamma_1 e_1, ..., \gamma_n e_n\rangle \subseteq \mathbb{R}^n$ and the set $A = \{(x_1, ..., x_n) \in \mathbb{R}^n : |x_i| \leq 1\}$. Plainly we see $\lambda_i(A, \Lambda) = \gamma_i$. Further, $d(\Lambda) = \gamma_1 ... \gamma_n$ and thus

$$\lambda_1 ... \lambda_n \mu(A) = \gamma_1 ... \gamma_n 2^n = 2^n d(\Lambda)$$

so the upper bound is sharp.

Now take $A = A_1^{(n)}$, then $\lambda_i(A, \Lambda) = \gamma_i$ but this time $\lambda_1 ... \lambda_n \mu(A) = \frac{2^n}{n!} d(\Lambda)$ so the lower bound is sharp as well.

Theorem 1.3.18. *A subset Λ of $\mathbb{R}^n$ is a lattice if and only if:*

1. *If $a, b \in \Lambda$ then $a + b, a - b \in \Lambda$.*
2. *Λ contains n linearly independent points $a_1, ..., a_n$.*
3. *Λ is discrete set, i.e. it has no limit points.*

Proof. ($\Rightarrow$) is immediate.

($\Leftarrow$): We prove this by induction. For $n = 1$ we note by (2) we see Λ is not empty with $a \in \Lambda$. By (1) we see it contains $0, -a$. Since Λ is discrete we can find a smallest positive real number $a \in \Lambda$. Then by (1) we see $\Lambda = \{ga : g \in \mathbb{Z}\}$ as desired.

Suppose it holds for $n - 1$ and assume the conditions $(1), (2), (3)$ for $\Lambda \subseteq \mathbb{R}^n$. We may choose our coordinate system so that $n - 1$ linearly independent points of Λ lie in a subspace of the form $\mathbb{R}^{n-1} \times \{0\}$, so $\pi_n(a_i) = 0$ with $\pi_n : \mathbb{R}^n \to \mathbb{R}$ the nth coordinate projection. Then $\Lambda' = \Lambda \cap \mathbb{R}^{n-1} \times \{0\}$ projects down to a subset of $\mathbb{R}^{n-1}$ which is a lattice by our inductive hypothesis. Let $b_1, ..., b_{n-1}$ be a basis for Λ', then Λ contains a point of the form $b_n = (b_{1n}, ..., b_{nn})$ with $b_{nn} > 0$. In fact there is a point b_n of this form with b_{nn} minimal. Suppose otherwise, then we can find a sequence $b_n^{(j)} = (b_{1n}^{(j)}, ..., b_{nn}^{(j)})$ in Λ with $b_{nn}^{(j)} > 0$ and

$$\lim_{j \to \infty} b_{nn}^{(j)} \to 0$$

However we can translate $b_n^{(j)}$ by some linear combination of $b_1, ..., b_{n-1}$ so that $(b_{1,b}^{(j)}, ..., b_{n-1,n}^{(j)}, 0)$ are in the compact set $\{\sum \lambda_i b_i : |\lambda_i| \leq 1\}$ so thus $b_n^{(j)}$ are all in a compact set and have a limit point, contradicting the fact Λ is discrete. We now claim every element of Λ is an integer linearly combination of $b_1, ..., b_n$. Let $d \in \Lambda$ with $d = (d_1, ..., d_n)$. Then

$$d' = d - [\frac{d_n}{b_{nn}}] b_n \in \Lambda$$

where $[x]$ is the closest integer to x. Then the nth coordinate of d' is non-negative and smaller than b_{nn}. Therefore it is 0. Thus $d' \in \Lambda'$ and so it is integer linear combination of $b_1, ..., b_{n-1}$. Thus d is integer linear combination of $b_1, ..., b_n$ and Λ is a lattice as desired. ♡

Proposition 1.3.19. *Let $n, m, k_1, ..., k_m$ be positive integers and a_{ij}, $i = 1, ..., m$ and $j = 1, ..., n$ be integers. The set Λ of points $\mu = (u_1, ..., u_n)$ in $\mathbb{R}^n$ with integral coordinates satisfying*

$$\sum_{j=1}^{n} a_{ij} u_j \equiv 0 \pmod{k_i}$$

for $i = 1, ..., m$ is a lattice in $\mathbb{R}^n$ with $d(\Lambda) \leq k_1 ... k_m$.

Proof. First we remark that Λ is a subset of Λ_0 and so is discrete. Next we observe that, for $k = \prod k_i$ we have

$$ke_1, ..., ke_n$$

are n linearly independent points in Λ. Finally $u = (u_1, ..., u_n), v = (v_1, ..., v_n) \in \Lambda$ then $u + v$ are in Λ since

$$\sum_j a_{ij}(u_j \pm v_j) \equiv 0 \pm 0 \equiv 0 \pmod{k_i}$$

Thus by Theorem 1.3.18 we have Λ is a lattice and so a sublattice of Λ_0.

Let I be the index of Λ in Λ_0, then $I = \frac{d(\Lambda)}{d(\Lambda_0)} = \frac{d(\Lambda)}{1}$. By Proposition 1.3.16 this is the order of Λ_0/Λ. However, that's immediate that this number is bounded by $k_1...k_m$. ♡

Theorem 1.3.20 (Lagrange's Theorem)**.** *Every positive integer m can be expressed as the sum of four squares of integers.*

Proof. We may assume $m > 1$ and m is square-free. Let $m = p_1...p_r$ with $p_1, ..., p_r$ all distinct primes. Note for every prime we can find a_p, b_p so

$$a_p^2 + b_p^2 + 1 \equiv 0 \pmod{p}$$

Indeed, if $p = 2$ we take $a_p = 1$, $b_p = 0$. If p is odd then the integer a^2 with $0 \leq a \leq \frac{1}{2}p$ are distinct mod p (consider $a_1^2 - a_2^2 \equiv (a_1 - a_2)(a_1 + a_2) \pmod{p}$). Similarly the integer $-1 - b^2$ with $0 \leq b < \frac{1}{2}p$ are distinct mod p. Since

$$\frac{1}{2}(p+1) + \frac{1}{2}(p+1) > p$$

there exist integers a_p, b_p with $a_p^2 \equiv -1 - b_p^2 \pmod{p}$ as required.

We define the lattice $\Lambda \subseteq \mathbb{R}^4$ as the set of points $(u_1, ..., u_4)$ so

$$\begin{cases} u_1 \equiv a_{p_i}u_3 + b_{p_i}u_4 \pmod{p_i} \\ u_2 \equiv b_{p_i}u_2 + a_{p_i}u_4 \pmod{p_i} \end{cases}$$

for $i = 1, ..., r$. Further, $d(\Lambda) \leq (p_1...p_r)^2$. Let $A = \{(x_1, ..., x_4) : \sum x_i^2 < 2m\}$, i.e. it is the sphere of radius $\sqrt{2m}$ in $\mathbb{R}^4$. We see $\mu(A) = \frac{\pi^2}{2}(\sqrt{2m})^4 = 2\pi^2 m^2$. Since $2\pi^2 m^2 > 2^4 m^2 \geq 2^4 d(\Lambda)$ there is a non-zero point $(u_1, ..., u_4)$ of Λ in A by Theorem 1.3.1. In particular we see

$$0 < \sum u_i^2 < 2m$$

where

$$\begin{aligned} u_1^2 + ... + u_4^2 &\equiv (a_{p_i}u_3 + b_{p_i}u_4)^2 + (b_{p_i}u_3 - a_{p_i}u_4)^2 + u_3^2 + u_4^2 \pmod{p_i} \\ &\equiv (a_{p_i}^2 + b_{p_i}^2 + 1)u_3^2 + (a_{p_i}^2 + b_{p_i}^2 + 1)u_4^2 \pmod{p_i} \\ &\equiv (a_{p_i}^2 + b_{p_i}^2 + 1)(u_3^2 + u_4^2) \pmod{p_i} \\ &\equiv 0 \pmod{p_i} \end{aligned}$$

for all $i = 1, ..., r$. Now by Chinese remainder theorem, we see

$$u_1^2 + u_2^2 + u_3^2 + u_4^2 \equiv 0 \pmod{m}$$

as desired. ♡

Remark 1.3.21. In many combinatorial settings, it is important to find short vectors in a lattice in an efficient way. Finding the shortest vector in a given lattice, with the usual distance, is NP-hard as shown by Ajtai. However, if we only look for one short vector, we can do so efficiently. This algorithm we use is the L^3-algorithm, where L^3 stands for Lenstra, Lenstra, and Lovasz.

1.4 Lenstra-Lenstra-Lovasz

Definition 1.4.1. Let $b_1, ..., b_n$ be linearly independent vectors for $\mathbb{R}^n$ and let $\beta_1, ..., \beta_n$ be the vectors obtained by Gram-Schmidt, with $\mu_{ij} := \frac{\langle b_i, \beta_j \rangle}{\langle \beta_j, \beta_j \rangle}$ where we recall $\langle \cdot, \cdot \rangle$ is the standard inner product (dot product). Then we say the basis $b_1, ..., b_n$ for a lattice Λ is ***reduced*** if $|\mu_{ij}| < \frac{1}{2}$ for all i, j and $|\beta_i + \mu_{i,i-1}\beta_{i-1}|^2 \geq \frac{3}{4}|\beta_{i-1}|^2$.

Remark 1.4.2. Note that the vectors $\beta_i + \mu_{i,i-1}\beta_{i-1}$ and β_{i-1} are the projections of b_i, b_{i-1} respectively on the orthogonal complement of the span of $b_1, ..., b_{i-2}$.

We also remark that the constant $\frac{3}{4}$ is arbitrary in the sense that it can be replaced by δ for any $\frac{1}{4} < \delta < 1$.

Finally, a remark on the condition $|\beta_i + \mu_{i,i-1}\beta_{i-1}|^2 \geq \delta|\beta_{i-1}|^2$, which you will find it justified once you read the actual LLL algorithm. As Gram-Schmidt orthogonalization depends on the order of the vectors, its vectors change if b_{i-1} and b_i are swapped (this will happen in the LLL algorithm). In fact, only β_{i-1} and β_i will be changed, and the new β'_{i-1} is just $\beta_i + \mu_{i,i-1}\beta_{i-1}$. Therefore, this condition $|\beta_i + \mu_{i,i-1}\beta_{i-1}|^2 \geq \delta|\beta_{i-1}|^2$ means by swapping b_i and b_{i-1}, the norm of β_{i-1} does not decrease too much, where the loss is quantified by δ.

The most natural value of δ should actually be $\delta = 1$, which in dimension 2 give us a algorithm called Lagrange's reduction. However, if $\delta = 1$ then it is unknown if such a reduced basis can be computed in polynomial time or not and so we need to adjust the value of δ and $\delta = \frac{3}{4}$ was used in the paper where L^3 was introduced, and hence the convention.

Proposition 1.4.3. *Let $b_1, ..., b_n$ be a reduced basis for Λ in $\mathbb{R}^n$ and $\beta_1, ..., \beta_n$ be the vectors obtained by Gram-Schmidt on b_i. Then:*

1. $|b_j|^2 \leq 2^{i-1}|\beta_i|^2$ *for* $1 \leq j \leq i \leq n$.
2. $d(\Lambda) \leq |b_1| \cdot ... \cdot |b_n| \leq 2^{n(n-1)/4} d(\Lambda)$
3. $|b_1| \leq 2^{(n-1)/4} d(\Lambda)^{1/n}$

Proof. By the definition of a reduced basis, we see

$$|\beta_i + \mu_{i,i-1}\beta_{i-1}|^2 \geq \frac{3}{4}|\beta_{i-1}|^2$$

with $|\mu_{i,i-1}| \leq \frac{1}{2}$. Thus

$$\begin{aligned}|\beta_i + \mu_{i,i-1}\beta_{i-1}|^2 &= \langle \beta_i + \mu_{i,i-1}\beta_{i-1}, \beta_i + \mu_{i,i-1}\beta_{i-1}\rangle \\ &= |\beta_i|^2 + \mu_{i,i-1}^2|\beta_{i-1}|\end{aligned}$$

for $i = 2, ..., n$ Thus

$$\begin{aligned}|\beta_i|^2 &= |\beta_i + \mu_{i,i-1}\beta_{i-1}|^2 - \mu_{i,i-1}^2|\beta_{i-1}|^2 \\ &\geq (\frac{3}{4} - \mu_{i,i-1}^2)|\beta_{i-1}|^2 \\ &\geq \frac{1}{2}|\beta_{i-1}|^2\end{aligned}$$

or equivalently $|\beta_{i-1}|^2 \leq 2|\beta_i|^2$. Thus inductively we see

$$|\beta_j|^2 \leq 2^{i-j}|\beta_i|^2$$

for $1 \leq j \leq i \leq n$. Now note

$$\begin{aligned}|b_i|^2 &= |\beta_i|^2 + \sum_{j=1}^{i-1} \mu_{ij}^2|\beta_j|^2 \\ &\leq |\beta_i|^2(1 + \sum_{j=1}^{i-1} \frac{1}{4}2^{i-j}) \\ &\leq |\beta_i|^2(1 + \frac{1}{4}(2^i - 2)) \\ &\leq 2^{i-1}|\beta_i|^2\end{aligned}$$

Hence we see

$$|b_j|^2 \leq 2^{j-1}|\beta_j|^2 \leq 2^{j-1} \cdot 2^{i-j}|\beta_i|^2 = 2^{i-1}|\beta_i|^2$$

This concludes (1).

Now note that $d(\Lambda) = |\det(\beta_1, ..., \beta_n)|$ and so by Hadamard's inequality we have

$$d(\Lambda) = |\beta_1|...|\beta_n|$$

Hence, we see

$$|b_i| \leq 2^{(i-1)/2}|\beta_i|$$

for $1 \leq i \leq n$ and so

$$|b_1|...|b_n| \leq 2^0 \cdot 2^{1/2} \cdot ... \cdot 2^{(n-1)/2}|\beta_1| \cdot ... \cdot |\beta_n| = 2^{n(n-1)/4}d(\Lambda)$$

This proves (2).

To prove (3), we apply (1) to $j = 1$ and see

$$|b_1| \leq 2^{(i-1)/2}|\beta_i|$$

and so $|b_1| \leq 2^{(n-1)/4}d(\Lambda)^{1/n}$ as desired. ♡

Proposition 1.4.4. *Let $b_1, ..., b_n$ be reduced basis for a lattice Λ in $\mathbb{R}^n$. Then for any vector $0 \neq x \in \Lambda$ we have*

$$|b_1| \leq 2^{n-1}|x|^2$$

Proof. Write $x = \sum g_i b_i = \sum \lambda_i \beta_i$ with $g_i \in \mathbb{Z}$ and $\lambda_i \in \mathbb{R}$. Let i be the largest index for which $g_i \neq 0$. Then by construction $\lambda_i = g_i$. Thus

$$|x|^2 \geq \lambda_i^2 |\beta_i|^2 \geq |\beta_i|^2$$

and by Proposition 1.4.3.1 we see

$$2^{i-1}|x|^2 \geq 2^{i-1}|\beta_i| \geq |b_1|^2$$

♡

Proposition 1.4.5. *Let $b_1, ..., b_n$ be a reduced basis for a lattice Λ in $\mathbb{R}^n$. Let $x_1, ..., x_t$ be t linearly independent vectors from Λ. Then*

$$|b_j|^2 \leq 2^{n-1} \max\{|x_1|^2, ..., |x_t|^2\}$$

for $j = 1, ..., t$.

Proof. Write $x_j = g_{1j}b_1 + ... + g_{nj}b_n$ with $g_{ij} \in \mathbb{Z}$ for $1 \leq j \leq t$, $1 \leq i \leq n$. For each j let $i(j)$ be the largest index for which g_{ij} is non-zero. Just as in the proof of Proposition 1.4.4 we see

$$|x_j|^2 \geq |\beta_{i(j)}|^2$$

Renumber the x_j's so that $i(1) \leq i(2) \leq ... \leq i(t)$. Observe $j \leq i(j)$ since otherwise $x_1, ..., x_j$ would be in the span of $b_1, ..., b_{j-1}$ which contradicts the assumption that $x_1, ..., x_j$ are linearly independent. Thus by Proposition 1.4.3.1, we see

$$|b_j|^2 \leq 2^{i(j)-1}|\beta_{i(j)}|^2 \leq 2^{i(j)-1}|x_j|^2$$

for $j = 1, ..., t$. Since $i(j) \leq n$ our result follows. ♡

Remark 1.4.6 (Lenstra-Lenstra-Lovasz Algorithm)**.** We now describe the L^3-algorithm, which take input as a basis for a lattice and output a reduced basis. Let $b_1, ..., b_n$ be a basis for a lattice Λ in $\mathbb{R}^n$.

First, apply Gram-Schmidt to $\{b_1, ..., b_n\}$ to get $\beta_1, ..., \beta_n$ and $\mu_{ij} := \frac{\langle b_i, \beta_j \rangle}{\langle \beta_j, \beta_j \rangle}$.

Define integer $k = 2$ and we shall iterate on k. Define boolean functions based on k to check the following conditions:

1. $|\mu_{ij}| \leq \frac{1}{2}$ for $1 \leq j < i < k$.
2. $|\beta_i + \mu_{i,i-1}\beta_{i-1}|^2 \geq \frac{3}{4}|\beta_{i-1}|^2$ for $1 < i < k$.
3. $|\mu_{k,k-1}| \leq \frac{1}{2}$
4. $|\beta_k + \mu_{k,k-1}\beta_{k-1}|^2 < \frac{3}{4}|\beta_{k-1}|^2$

For example, we say $(1)_k$ and $(2)_k$ are true if the two claims above holds about that particular k. In particular, right now $(1)_k = (1)_2$ and $(2)_k = (2)_2$ holds as we defined $k = 2$.

If $1 < k \leq n$ and $(3)_k$ holds then we do nothing, otherwise replace b_k by $b_k - r_{k-1}b_{k-1}$ in our basis where r is the closest integer to $\mu_{k,k-1}$. This has the effect of replacing $\mu_{k,k-1}$ with $\mu_{k,k-1} - r$ where $|\mu_{k,k-1} - r| \leq \frac{1}{2}$. The numbers μ_{kj} with $j < k - 1$ are replaced by $\mu_{kj} - r\mu_{k-1,j}$ and re-compute the Gram-Schmidt. The other μ_{ij} and b_i with $i \notin \{k, k-1\}$ and $i \leq k$ are not changed. At this point, we may assume $(3)_k$ holds.

Now consider two cases.

Case 1: If $k \geq 2$ and $(4)_k$ holds, then we interchange b_k and b_{k-1} in our basis (so $i \neq k, k-1$). We leave the other b_i's unchanged. Notice that β_k, β_{k-1} and the numbering $\mu_{k,k-1}, \mu_{k-1,j}, \mu_{kj}, \mu_{ik}, \mu_{i,k-1}$ for $j < k-1$ and $i > k$ changed as we swap b_k and b_{k-1}. Let us call our new basis $c_1, ..., c_n$ (with Gram-Schmidt basis $\theta_1, ..., \theta_n$) so that $c_i = b_i$ for $i \neq k, k-1$ and $c_{k-1} = b_k$ and $c_k = b_{k-1}$. Note that θ_{k-1} is the projection of b_k on the orthogonal complement of the span of $\{b_1, ..., b_{k-2}\}$ and so $\theta_{k-1} = \beta_k + \mu_{k,k-1}\beta_{k-1}$. Therefore,

$$|\theta_{k-1}|^2 < \frac{3}{4}|\beta_{k-1}|^2$$

In particular the new $|\beta_{k-1}|^2$ is less than $\frac{3}{4}$ of the old $|\beta_{k-1}|^2$. We now replace k by $k-1$ and return to the algorithm.

Case 2: If $k = 1$ or $(4)_k$ is false, then we achieve $|\mu_{kj}| \leq \frac{1}{2}$ for $1 \leq j \leq k-1$ and we may replace k with $k+1$ and return to the algorithm.

To achieve $|\mu_{kj}| \leq \frac{1}{2}$ for $1 \leq j \leq k-1$ we do the following. First note $\mu_{k,k-1} \leq \frac{1}{2}$. Then let l be the largest integer with $1 \leq l < k-1$ for which $|\mu_{kl}| > \frac{1}{2}$. Let r be the integer closest to μ_{kl} and replace b_k by $b_k - rb_l$. Note that μ_{kl} is then replaced by $\mu_{kl} - r$ and $|\mu_{kl} - r| \leq \frac{1}{2}$. We now recalculate $\mu_{k,j}$ for $j < l$. We then repeat the process until we have

$$|\mu_{kj}| \leq \frac{1}{2}$$

for $1 \leq j \leq k-1$.

Example 1.4.7. Before we show this algorithm terminates, let us compute an easy example. Consider the lattice generated by $(1, 2)$ and $(2, 1)$ with a basis $b_1 = (5, 4)$ and $b_2 = (2, 1)$. Then we get $\beta_1 = b_1$ and

$$\beta_2 = b_2 - \frac{\langle b_2, \beta_1 \rangle}{\langle \beta_1, \beta_1 \rangle}\beta_1 = b_2 - \frac{14}{41}\beta_1 = (12/41, -15/41)$$

Next, we have $\mu_{11} = 1$, $\mu_{12} = 0$, $\mu_{21} = \frac{14}{41}$ and $\mu_{22} = 1$.

Run the algorithm start with $k = 2$ we get $\mu_{2,1} \leq \frac{1}{2}$ already, so we got nothing to do. Next, we check the value

$$|\beta_k + \mu_{k,k-1}\beta_{k-1}|^2 = |\beta_2 + \frac{14}{41}\beta_1|^2 = |(2, 1)|^2 = 5$$

while on the other hand we get $\frac{3}{4}|\beta_1|^2 = \frac{123}{4} > 5$. Hence we need to interchange b_k and b_{k-1} in our basis and we need to run Gram-Schmidt again with this new basis $b_1 = (2,1)$ and $b_2 = (5,4)$. This time we get $\beta_1 = (2,1)$ and $\beta_2 = (-18/5, -9/5)$ and we set $k = k-1$, i.e. now $k = 1$ and we are back to the algorithm.

Now $k = 1$, we don't need to check $\mu_{k,k-1}$ as this is meaningless. Hence we are at Case 2 now and we may replace k with $k+1$ and back to the algorithm.

Now $k = 2$, we need to check $\mu_{2,1}$, which is equal to $\frac{\langle b_2, \beta_1 \rangle}{\langle \beta_1, \beta_1 \rangle} = \frac{14}{5}$ and this is greater than $1/2$. Thus, let $r = 3$, which is the closest integer to $14/5$ and we need to replace b_2 with $b_2 - 3b_1 = (-1,1)$. Thus our basis become $b_1 = (2,1)$ and $b_2 = (-1,1)$ and we need to re-compute Gram-Schmidt and get $\beta_1 = (2,1)$ and $\beta_2 = (-3/5, 6/5)$. Then we see

$$|\beta_2 + \mu_{21}\beta_1|^2 = 2 < \frac{15}{4} = \frac{3}{4}|\beta_1|^2$$

Hence, we need to interchange b_1 and b_2 in our basis, i.e. our basis now is $b_1 = (-1,1)$ and $b_2 = (2,1)$. We need to re-compute the Gram-Schmidt and get $\beta_1 = (-1,1)$ and $\beta_2 = (3/2, 3/2)$. Now replace k with $k - 1 = 1$ and return to the algorithm.

Since $k = 1$ we don't need to check $\mu_{k,k-1}$ and we just increase k by 1 and get $k = 2$.

Now $k = 2$, we need to check $|\mu_{21}| = \frac{1}{2} \leq \frac{1}{2}$, hence we don't need to do anything right now. Next, compute

$$|\beta_2 + \mu_{21}\beta_1|^2 = |(2,1)|^2 = 5 \geq \frac{3}{4} \cdot 2 = |\beta_1|^2$$

This tell us we are in Case 2 so just increase k by 1, i.e. $k = 3$ and we are done.

Hence, we get $b_1 = (-1,1)$ and $b_2 = (2,1)$ (with $\beta_1 = (-1,1)$ and $\beta_2 = (3/2, 3/2)$) is our reduced basis for $w_1 = (5,4)$ and $w_2 = (2,1)$. One should try to do a computation to show this is indeed reduced basis.

Remark 1.4.8 (Termination of The L^3 Algorithm)**.** We shall now show the algorithm terminates after finitely many steps. Do that end, we define

$$d_i := \det(A_i)$$

where

$$A_i = \begin{bmatrix} \langle b_1, b_1 \rangle & \dots & \langle b_1, b_i \rangle \\ \vdots & \ddots & \vdots \\ \langle b_n, b_1 \rangle & \dots & \langle b_i, b_i \rangle \end{bmatrix} = \begin{bmatrix} b_1^T \\ \vdots \\ b_i^T \end{bmatrix} \begin{bmatrix} b_1 & \dots & b_i \end{bmatrix} = [\beta_1, \dots, \beta_i]^T [\beta_1, \dots, \beta_i]$$

and define $D = \prod_{i=1}^n d_i$. Note $d_n = d(\Lambda^2)$ and

$$d_i = (|\beta_1| \dots |\beta_i|)^2 = d(\Lambda_i)^2$$

where Λ_i is the lattice generated by $b_1, \dots, b_i$ in the i-dimensional subspace of $\mathbb{R}^n$. Note that D changes only if one of β_i changes and this only occurs in case 1 during

the algorithm. Further, in case 1 we interchanged b_k and b_{k-1}. Since $d_i = (|\beta_1|...|\beta_i|)^2$ we see d_i only changes when $i = k-1$ in which case it gets smaller by a factor of at least $3/4$. Viz, D is getting smaller by a factor fo at least $3/4$. To complete our argument, we will show D is bounded from below in terms of A.

Put $m(\Lambda) = \min\{x \cdot x : x \in \Lambda, x \neq 0\}$, by Theorem 1.3.4 we see

$$m(\Lambda) \leq 4(\omega_i^{-1} d(\Lambda_i))^{2/i}$$

and hence $d_i \geq m(\Lambda_i)^i 4^{-i} \omega_i^2$ and since $m(\Lambda_i) \geq m(\Lambda)$ we see

$$d_i \geq m(\Lambda)^i 4^{-i} \omega_i^2$$

for $i = 1, ..., n$. Thus we see

$$D = d_1 d_2 ... d_n \geq \left(\frac{m(\Lambda)}{4}\right)^{n(n+1)/2} (\omega_1 ... \omega_n)^2$$

as desired.

Hence, we can pass through case 1 only infinitely many times. In case 1, k decreases by 1. In case 2, k increases by 1 and so after finitely many steps our algorithm terminates as desired.

Remark 1.4.9 (Efficiency of L^3 Algorithm)**.** We remark that Lenstra, Lenstra, and Lovasz proved that if Λ is a sublattice of Λ_0 with basis $b_1, ..., b_n$ and if B is a real number with $B \geq 2$ and $|b_i|^2 \leq B$ for $i = 1, ..., n$, then the number of arithmetical operations needed for the L^3 algorithm is $O(n^4 \log B)$ and the integers on which these operations are performed have binary length $O(n \log B)$. The algorithm runs in polynomial time in terms of B.

Remark 1.4.10. The L^3-algorithm can be used to find a short vector in a lattice Λ. We just put a basis for the lattice in reduced form $b_1, ..., b_n$. Then b_1 is a short vector in Λ.

Example 1.4.11 (Application: Approximate Real Numbers)**.** Let $\alpha_1, ..., \alpha_n \in \mathbb{R}$ and let $0 < \epsilon < 1$. The goal of this example is to produce efficiently a positive integer q and integers $p_1, ..., p_n$ for which

$$|q\alpha_i - p_i| < \epsilon, i = 1, 2, ..., n$$

with $1 \leq q \leq 2^{n(n+1)/4} \epsilon^{-n}$.

First recall, by Theorem 1.2.10, on taking $\epsilon = \frac{1}{Q}$, $n = n$ and $m = 1$, that such a q exists with $1 \leq q \leq \epsilon^{-n}$ (it may not be efficient!).

Now consider the lattice Λ generated by the row of the $n+1$ by $n+1$ matrix

$$A = \begin{bmatrix} 1 & 0 & ... & 0 & 0 \\ 0 & 1 & ... & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & ... & 1 & 0 \\ \alpha_1 & \alpha_2 & ... & \alpha_n & \delta \end{bmatrix}, \quad \delta := 2^{-n(n+1)/4} \epsilon^{n+1}$$

Note $d(\Lambda) = \delta$ and by LLL algorithm we can find a small non-zero vector b in the lattice (i.e. run the algorithm and get reduced basis $b_1, b_2, ..., b_n$, then let b equal b_1). Since $b \in \langle e_1, ..., e_n, \begin{bmatrix} \alpha_1 \\ \vdots \\ \delta \end{bmatrix} \rangle$ we see there exists $p_1, ..., p_n, q \in \mathbb{Z}$ so

$$b = -p_1 e_1 - ... - p_n e_n + q \begin{bmatrix} \alpha_1 \\ \vdots \\ \delta \end{bmatrix}$$

and hence we get

$$b = (q\alpha_1 - p_1, q\alpha_2 - p_2, ..., q\alpha_n - p_n, q\delta)$$

Note we may suppose $q \geq 0$ by replace b with $-b$ if needed. Further by Proposition 1.4.3.3 we can find b so

$$|b| \leq 2^{n/4} d(\Lambda)^{1/(n+1)} = 2^{n/4} \delta^{1/(n+1)} = \epsilon$$

Since $|b| \leq \epsilon$ and $\epsilon < 1$ we see $q \neq 0$ since in that case $|b| = |(p_1, ..., p_n, 0)| \geq 1$ since $p_1, ..., p_n$ are not all zero as we have assumed $b \neq 0$. Thus we get

$$1 \leq q \leq 2^{n(n+1)/4} \epsilon^{-n}$$

as desired. In particular, if $\alpha_1, ..., \alpha_n, \epsilon \in \mathbb{Q}$ then L^3 algorithm runs in polynomial time and hence efficient.

Example 1.4.12 (Application: Approximate Linear Forms)**.** In this example, we want to find a small linear form with integer coefficients $\alpha_1, ..., \alpha_n$. In particular, g iven $0 < \epsilon < 1$, how do we find, efficiently, integers $q_1, ..., q_n$ and p so

$$|L(q_1, ..., q_n) - p| < \epsilon$$

with $L(x_1, ..., x_n) = \sum_{i=1}^n \alpha_i x_i$ and

$$1 \leq \max_i |q_i| \leq 2^{(n+1)/4} \epsilon^{-1/n}$$

Note the existence of such numbers are asserted by Corollary 1.2.10.1.

Now consider a the lattice Λ generated by the rows of the matrix

$$A = \begin{bmatrix} 1 & 0 & 0 & ... & 0 \\ \alpha_1 & \delta & 0 & ... & 0 \\ \alpha_2 & 0 & \delta & ... & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \alpha_n & 0 & 0 & ... & \delta \end{bmatrix}, \quad \delta := \left(\frac{\epsilon^{1/n}}{2^{1/4}} \right)^{n+1}$$

with a particular element b of this lattice to be

$$b = (L(q_1, ..., q_n) - p, q_1\delta, q_2\delta, ..., q_n\delta)$$

with $q_i, p \in \mathbb{Z}$. By LLL we can find a non-zero vector b in Λ of this form with $|b| \leq 2^{n/4} d(\Lambda)^{1/(n+1)}$ where $d(\Lambda) = (\frac{\epsilon}{2^{n/4}})^{n+1}$. Thus we see

$$|b| \leq 2^{n/4} \cdot \frac{\epsilon}{2^{n/4}} = \epsilon$$

Further, since $b \neq 0$ and $\epsilon < 1$ we have $0 < |b| < 1$ hence $q_1, ..., q_n$ are not all zero and so $0 < \max_i |q_i|$.

Example 1.4.13 (Application: Approximate Matrix)**.** Finally, let

$$A = \begin{bmatrix} \alpha_{11} & ... & \alpha_{1m} \\ \vdots & \ddots & \vdots \\ \alpha_{n1} & ... & \alpha_{nm} \end{bmatrix}$$

be a matrix with real entries and $0 < \epsilon < 1$. Consider the lattice generated by the rows of the $m + n$ by $m + n$ matrix

$$B = \begin{bmatrix} I & 0 \\ A^T & \delta I \end{bmatrix}, \quad \delta = (2^{-(n+m-1)/4} \cdot \epsilon)^{\frac{n}{m}+1}$$

We note $d(\Lambda) = \delta^m = (2^{-(n+m-1)/4} \cdot \epsilon)^{n+m}$ and by LLL we can find $b \in \Lambda$ with

$$|b| \leq \delta^{m/(n+m)} \cdot 2^{(n+m-1)/4} = \epsilon$$

Also, note using the basis we have, we see b is of the form

$$b = \begin{bmatrix} L_1(q_1, ..., q_m) - p_1 \\ L_2(q_1, ..., q_m) - p_2 \\ \vdots \\ L_n(q_1, ..., q_m) - p_n \\ q_1\delta \\ q_2\delta \\ \vdots \\ q_m\delta \end{bmatrix}$$

where L_i is the linear form defined by the ith row of A, i.e. $L_i(x_1, ..., x_m) = \sum_{j=1}^{m} \alpha_{ij} x_j$ and $q_i, p_i \in \mathbb{Z}$. In particular this means

$$|L_i(q_1, ..., q_m) - p_i| < \epsilon, i = 1, 2, ..., n$$

and $|q_j\delta| < \epsilon$ for $j = 1, ..., m$, hence

$$|q_j| < \delta^{-1}\epsilon = 2^{(\frac{n+m-1}{4})(\frac{n+m}{m})}\epsilon^{-n/m}$$

where not all q_i's are zero.

Remark 1.4.14. Theorem 1.2.10 tell us we can make linear forms in a_{ij}'s with integer coefficients which are simultaneously close to integers. LLL gives us an efficient method for finding the associated integers coefficients. Can we do better than Theorem 1.2.10? Not for real numbers in general, but for algebraic numbers α_{ij} we can say more, as it follows from the work of Schmidt.

1.5 Schmidt's Subspace Theorem

Remark 1.5.1. In this section, we will first state two theorems as an motivation for the Schmidt's subspace theorem and assume the subspace theorem to prove them. Finally, Schmidt's subspace theorem will not be proved in this section.

Definition 1.5.2. Let x be a real number, then $\|x\|_{\mathbb{Z}}$ is defined to be the distance from x to the nearest integer.

Theorem 1.5.3. *Let $1, \alpha_1, ..., \alpha_n$ be real algebraic numbers which are linearly independent over $\mathbb{Q}$. Let $\delta > 0$, then there are only finitely many n-tuples of non-zero integers $q_1, ..., q_n$ with*

$$|q_1 ... q_n|^{1+\delta} \cdot \|q_1\alpha_1 + ... + q_n\alpha_n\|_{\mathbb{Z}} < 1$$

Corollary 1.5.3.1. *Let $1, \alpha_1, ..., \alpha_n$ be real algebraic numbers which are $\mathbb{Q}$-linearly independent. Let $\delta > 0$, then there are only finitely many $n+1$-tuples of integers $q_1, ..., q_n, p$ with $q = \max_i |q_i| > 0$ for which*

$$|\alpha_1 q_1 + ... + \alpha_n q_n - p| < \frac{1}{q^{n+\delta}}$$

Proof. Apply Theorem 1.5.3 to all finite subsets of $\{\alpha_1, ..., \alpha_n\}$ and the result follows. ♡

Corollary 1.5.3.2 (Roth's Theorem)**.** *Let $\delta > 0$, if α is an algebraic number, then there are only finitely many rationals p/q with $\gcd(p, q) = 1$ and $q > 0$ for which*

$$|\alpha - \frac{p}{q}| < \frac{1}{q^{2+\delta}}$$

Proof. Let $n = 1$ in the above corollary and we are done. ♡

Theorem 1.5.4. *Suppose $\alpha_1, ..., \alpha_n$ be real algebraic numbers with $1, \alpha_1, ..., \alpha_n$ be linearly independent over $\mathbb{Q}$. Let $\delta > 0$, then there are only finitely many positive integer q with*

$$q^{1+\delta} \|\alpha_1 q\|_{\mathbb{Z}} \cdot ... \cdot \|\alpha_n q\|_{\mathbb{Z}} < 1$$

Corollary 1.5.4.1. *Let $1, \alpha_1, ..., \alpha_n$ be real algebraic numbers which are $\mathbb{Q}$-linearly independent. Let $\delta > 0$, then there are only finitely many n-tuples of rationals $(p_1/q, ..., p_n/q)$ with $q > 0$ for which*

$$|\alpha_i - \frac{p_i}{q}| < \frac{1}{q^{1+\frac{1}{n+\delta}}}$$

Remark 1.5.5. Note Theorem 1.5.3 and Theorem 1.5.4 are both proven by Schmidt and are consequences of the Schmidt's subspace theorem.

Definition 1.5.6. Let $x \in \mathbb{R}^n$ we define

$$\|x\|_\infty := \max_{1 \le i \le n} |x_i|$$

This is not known as the ***house*** of x (to me) and we shall just call it the L-infinity norm.

Theorem 1.5.7 (Schmidt's Subspace Theorem)**.** *Suppose $L_1(x), ..., L_n(x)$ are linearly independent linear forms in $x = (x_1, ..., x_n)$ with algebraic number coefficients. Let $\delta > 0$, then there are finitely many proper subspaces $T_1, ..., T_w$ of $\mathbb{R}^n$ such that for every integer point x with the property $x \neq 0$ and*

$$|L_1(x)...L_n(x)| < \frac{1}{\|x\|_\infty^\delta}$$

we have $x \in T_i$ for some $i = 1, 2, ..., w$.

Remark 1.5.8. We remark that:

1. The result is not effective in the sense that the proof does not yield a procedure for determining the subspaces $T_1, ..., T_w$.
2. The integer points in a proper subspace of $\mathbb{R}^n$ lies in a rational subspace of $\mathbb{R}^n$, in other words in a subspace determined by a linear form with rational coefficients.
3. The proof generalizes Roth's theorem, uses ideas from the geometry of numbers and is difficult.

Proof of Theorem 1.5.4. Let q be a positive integer satisfying

$$q^{1+\delta} \|\alpha_1 q\|_\mathbb{Z} ... \|\alpha_n q\|_\mathbb{Z} < 1$$

Choose integers $p_1, ..., p_n$ such that $\|\alpha_i q\|_\mathbb{Z} = |\alpha_i q - p_i|$ for $i = 1, ..., n$, then put $x = (p_1, ..., p_n, q)$. Let K_1, K_2 be positive numbers depend on only $\alpha_1, ..., \alpha_n$ and n such that[1] $\|x\|_\infty \le K_1 q$ and the choice of K_2 will be given later.

We consider the linear forms

$$L_i(x_1, ..., x_{n+1}) = \alpha_i x_{n+1} - x_i, \quad i = 1, ..., n$$

$$L_{n+1}(x_1, ..., x_{n+1}) = x_{n+1}$$

where we see they are $n+1$-linearly independent linear forms with algebraic coefficients. We note that

$$|L_1(x)...L_{n+1}(x)| = \|\alpha_1 q\|_\mathbb{Z} \cdot ... \cdot \|\alpha_n q\|_\mathbb{Z} \cdot q \Rightarrow |L_1(x)...L_{n+1}(x)| < \frac{1}{q^\delta} < \frac{1}{\|x\|_\infty^{\delta/2}}$$

for q sufficiently large.

[1] We can find such K_1 since $p_1, ..., p_n$ are very close to $\alpha_i q$. If $\alpha_i < 1$ then p_i is close to a number less than q and hence $p_i \le q$. If $\alpha_i > 1$ then p_i is close to a multiple of q and we just take K_1 to be a integer greater than the product of $|\alpha_i|$ where $|\alpha_i| > 1$ and we see $K_1 q$ must be greater than $\|x\|_\infty$ as desired.

Thus, by the subspace theorem we have x lies in one of finitely many proper subspaces $T_1, ..., T_w$ of $\mathbb{R}^{n+1}$ and since x is integer point, it lies in a proper rational subspace T. We can find $c_1, ..., c_{n+1}$ in $\mathbb{Q}$ so T is determined by the vanishing of $c_1x_1 + ... + c_{n+1}x_{n+1}$, i.e. we have

$$c_1p_1 + ... + c_np_n + c_{n+1}q = 0$$

In particular then we see

$$\begin{aligned} |c_1(\alpha_1 q - p_1) + ... + c_n(\alpha_n q - p_n)| &= |c_1\alpha_1 q + ... + c_n\alpha_n q - c_1p_1 - ... - c_np_n| \\ &= |c_1\alpha_1 q + ... + c_n\alpha_n q + c_{n+1}q| \\ &= |c_1\alpha_1 + ... + c_n\alpha_n + c_{n+1}|q \\ &> K_2 q \end{aligned}$$

for fixed positive number K_2 since $1, \alpha_1, ..., \alpha_n$ are linearly independent over $\mathbb{Q}$ and hence $|c_1\alpha_1 + ... + c_n\alpha_n + c_{n+1}|$ is non-zero (not all c_i's are zero since that will imply T is not a proper subspace). Thus we see

$$K_2 q < |c_1| + ... + |c_n|$$

Since K_2 is fixed, this concludes the possible values of q are bounded, i.e. we have only finitely many possibilities. ♡

Proof of Theorem 1.5.3. We will prove the claim by induction on n. For $n = 1$ the result holds by Theorem 1.5.4. Now assume $n > 1$ and say we have integers $q_1, ..., q_n$, not all zero, for which

$$\|\alpha_1 q_1 + ... + \alpha_n q_n\|_{\mathbb{Z}} \cdot |q_1...q_n|^{1+\delta} < 1$$

Choose p to be the closest integer to $\alpha_1 q_1 + ... + \alpha_n q_n$ so $\alpha_1 q_1 + ... + \alpha_n q_n - p < 1$. Write $x = (q_1, ..., q_n, p)$ and let

$$L_i(x_1, ..., x_{n+1}) = x_i, \quad i = 1, 2, ..., n$$

and

$$L_{n+1}(x_1, ..., x_{n+1}) = \sum_{i=1}^{n} \alpha_i x_i - x_{n+1}$$

Then we have $n + 1$ linearly independent linear forms with algebraic coefficients. Note that

$$|L_1(x)...L_{n+1}x| = |q_1...q_n| \cdot \|\alpha_1 q_1 + ... + \alpha_n q_n\|_{\mathbb{Z}}$$

We have $\|x\|_\infty < K_1 q$ where $q = \max_i |q_i|$ as before and let $K_1, K_2, ...$ be positive numbers depends on $\alpha_1, ..., \alpha_n$ and n. Observe that

$$|L_1(x)...L_{n+1}(x)| < \frac{1}{|q_1...q_n|^\delta} < \frac{1}{\|x\|_\infty^{\delta/2}}$$

for q sufficiently large, as we may assume. Then by subspace theorem we see x lies in one of a finite collection of proper rational subspaces of $\mathbb{R}^{n+1}$. Let T be

such a subspace and hence we see we can find $c_1, ..., c_{n+1}$ so $T = \{(y_1, ..., y_{n+1}) : c_1y_1 + ... + c_{n+1}y_{n+1} = 0\}$ with $c_i \in \mathbb{Q}$ and not all c_i's are zero.

Now we have two cases, where in the first case, one of $c_1, ..., c_n$ is not zero, and the second case where c_{n+1} is not zero.

In the first case, WLOG we may suppose c_n is not zero and hence we get

$$c_1q_1 + ... + c_nq_n + c_{n+1}p = 0$$

so

$$c_n\alpha_nq_n = -\alpha_n \cdot (\sum_{i=1}^{n-1} c_iq_i) - c_{n+1}\alpha_np$$

Thus we get

$$\begin{aligned} |c_n| \cdot |\alpha_1q_1 + ... + \alpha_nq_n - p| &= |\sum_{i=1}^{n-1}(c_n\alpha_i - c_i\alpha_n)q_i - (c_n + c_{n+1})p| \\ &= |c_n + c_{n+1}\alpha_n| \cdot \left|\sum_{i=1}^{n-1}\left(\frac{c_n\alpha_i - c_i\alpha_n}{c_n + c_{n+1}\alpha_n}\right)q_i - p\right| \end{aligned}$$

where $c_n + c_{n+1}\alpha_n \neq 0$ since $1, \alpha_1, ..., \alpha_n$ are linearly independent over $\mathbb{Q}$. Now put

$$\alpha_i' = \frac{c_n\alpha_i - c_i\alpha_n}{c_n + c_{n+1}\alpha_n}, \quad i = 1, ..., n-1$$

we get

$$|c_n| \cdot |\alpha_1q_1 + ... + \alpha_nq_n - p| = |c_n + c_{n+1}\alpha_n| \cdot |\sum_{i=1}^{n-1}\alpha_i'q_i - p|$$

Therefore, we get

$$\left\|\sum_{i=1}^{n-1}\alpha_i'q_i\right\|_{\mathbb{Z}} < \frac{K_2}{|q_1...q_n|^{1+\delta}} < \frac{1}{|q_1...q_{n-1}|^{1+\frac{\delta}{2}}}$$

for q sufficiently large.

We remark $1, \alpha_1', ..., \alpha_{n-1}'$ are linearly independent over $\mathbb{Q}$. To see this, suppose

$$\sum_{i=1}^{n-1}\lambda_i\alpha_i' + \lambda_n = 0$$

with $\lambda_i \in \mathbb{Q}$, then we get

$$\begin{gathered} \lambda_1(c_n\alpha_1 - c_1\alpha_n) + \cdots + \lambda_{n-1}(c_n\alpha_{n-1} - c_{n-1}\alpha_n) + \lambda_n(c_n + c_{n+1}\alpha_n) = 0 \\ \lambda_1c_n\alpha_1 + \cdots + \lambda_{n-1}c_n\alpha_{n-1} - (\lambda_1c_1 + \cdots + \lambda_{n-1}c_{n-1} + \lambda_nc_{n+1})\alpha_n + \lambda_nc_n = 0 \end{gathered}$$

but since we have $1, \alpha_1, ..., \alpha_n$ are linearly independent over $\mathbb{Q}$ and $c_n \neq 0$, we must have $\lambda_1 = ... = \lambda_n = 0$ as desired.

Then by induction, we see $|q_1|, ..., |q_n|$ are bounded in the case where not all $c_1, ..., c_n$'s are zero.

Now it remains to consider the case $c_1 = ... = c_n = 0$ and $c_{n+1} \neq 0$. Then we have

$$c_{n+1}p = 0 \Rightarrow p = 0$$

In this case, we get

$$|q_1...q_n|^{1+\delta}|\alpha_1 q_1 + ... + \alpha_n q_n| < 1$$

and hence

$$|q_1...q_n|^{1+\delta}|\alpha_n| \cdot \left|\frac{\alpha_1}{\alpha_n}q_1 + ... + \frac{\alpha_{n-1}}{\alpha_n}q_{n-1} + q_n\right| < 1$$

Now let $\alpha_i' = \frac{\alpha_i}{\alpha_n}$ for $i = 1, ..., n-1$, we see $1, \alpha_1', ..., \alpha_{n-1}'$ are linearly independent over $\mathbb{Q}$ so

$$|q_1...q_{n-1}|^{1+\frac{\delta}{2}} \cdot \left\|q_1\alpha_1' + ... + q_{n-1}\alpha_{n-1}'\right\| < 1$$

and so $\max_i |q_i|$ is bounded by induction and the result follows. ♡

Theorem 1.5.9. *Let α_{ij} with $1 \leq i \leq n, 1 \leq j \leq m$ be real algebraic numbers and suppose $1, \alpha_{i1}, ..., \alpha_{im}$ are linearly independent over $\mathbb{Q}$ for $i = 1, ..., n$. Let $\delta > 0$, then there are only finitely many m-tuples of non-zero integers $(q_1, ..., q_m)$ for which*

$$|q_1...q_m|^{1+\delta} \prod_{i=1}^{n} \|\alpha_{i1}q_1 + ... + \alpha_{im}q_m\|_{\mathbb{Z}} < 1$$

Chapter 2

Lattice II

2.1 Extremal Lattice

Remark 2.1.1. For each n, let us consider those latice of dimension n with $d(\Lambda) = 1$, say this collection is $\mathcal{D}_1^n$. In this collection, let us look for lattices Λ for which the minimal non-zero distance between lattice points $\mu(\Lambda)$ is large (recall $\mu(\Lambda)$ is Lebesgue measure).

In particular, for each $n = 1, 2, 3, 4, ...$ define

$$\mu_n = \sup_{\Lambda \in \mathcal{D}_1^n} \left(\min_{x,y \in \Lambda, x \neq y} |x - y| \right) = \sup_{\Lambda \in \mathcal{D}_1^n} \mu(\Lambda)$$

However, followed by Mahler's Compactness Theorem, this sup above is actually a maximum. Those lattices which maximum is attained are known as extremal lattices. The value of μ_n have been determined for $n = 1, ..., 8$ and they are

$$\mu_1 = 1, \mu_2 = \sqrt[4]{4/3}, \mu_3 = \sqrt[6]{2}, \mu_4 = \sqrt[8]{4}, \mu_5 = \sqrt[10]{8}, \mu_6 = \sqrt[12]{64/3}, \mu_7 = \sqrt[14]{64}, \mu_8 = \sqrt{2}$$

Example 2.1.2. For $n = 2$ we have:

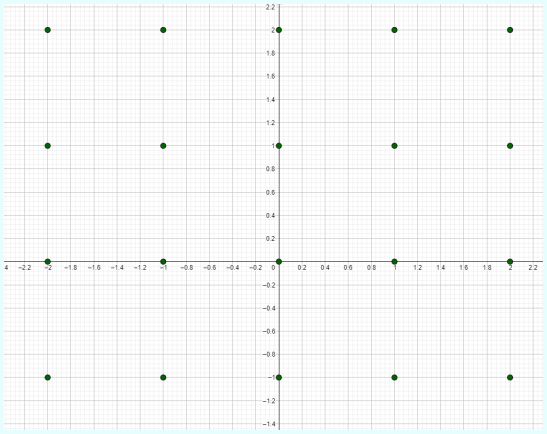


Figure 2.1: Not extremal

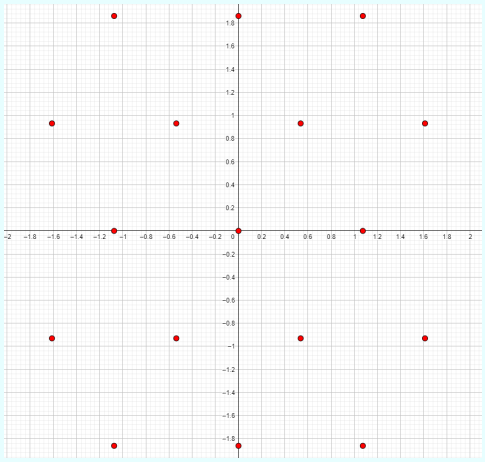


Figure 2.2: Extremal

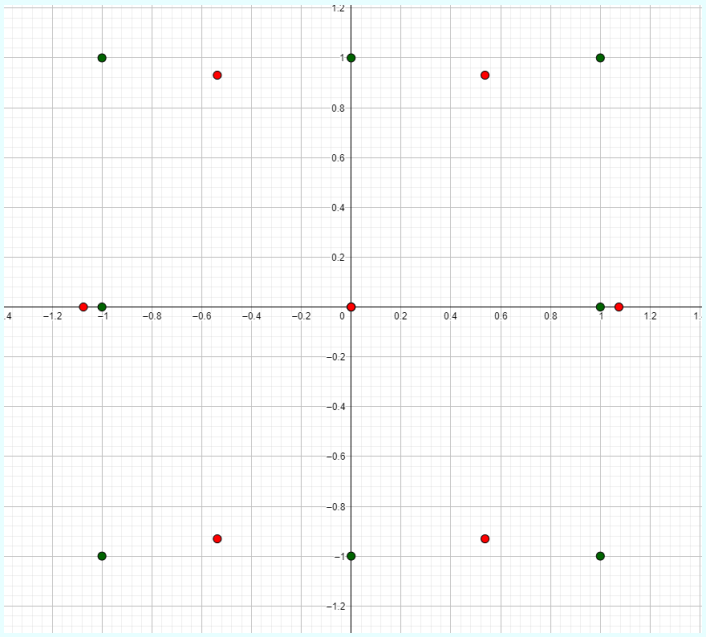


Figure 2.3: The two put together

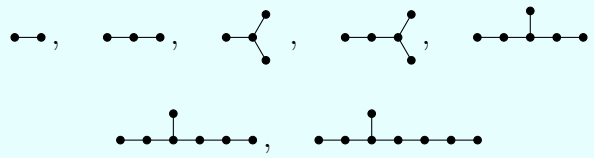
Example 2.1.3. In above example and remark we claimed and drawed the extremal lattice for $n = 2$. Now we prove it. Let Λ be the lattice generated by $(\sqrt[4]{(4/3)}, 0)$ and $(\frac{1}{2}\sqrt[4]{4/3}, \sqrt[4]{3/4})$, clearly $d(\Lambda) = 1$ so $\Lambda \in \mathcal{D}_1^2$.

Next, suppose we have $\Delta \in \mathcal{D}_1^2$ so that $\mu(\Delta) > \sqrt[4]{4/3}$. Then WLOG we may assume Δ has basis $(a, 0)$ and $(b, 1/a)$ for some $a > 0$. Further, by adding an appropriate multiple we may assume $|b| \leq \frac{a}{2}$. Furthermore we may suppose $a = \mu(\Delta)$. If $a > \sqrt[4]{4/3}$ then $3a^4 > 4$ so $\frac{3}{4}a^2 > \frac{1}{a^2}$. However then $(b, \frac{1}{a})$ is closer to the origin than $(a, 0)$ since $b^2 + \frac{1}{a^2} < \frac{a^2}{4} + \frac{3}{4}a^2 = a^2$, which is a contradiction.

Remark 2.1.4. The first few extremal lattices can be represented by Dynkin diagrams which arise in the study of Lie groups. A graph will consist of n points which correspond to generators of the lattice. Each of the generators will be of the same length. If two generators are not connected by an edge they are orthogonal. If they

are connected by an edge then the angle between them is 60 degree or $\pi/3$. Finally we normalize the length of the generators so the determinant of the lattice is 1.

Here are the graphs associated with extremal lattices for $n = 2, 3, ..., 8$.



These lattices give the values of μ which we claimed were the extremal values. The difficult task is to prove they are extremal.

Remark 2.1.5. We look more closely at the lattices associated with these diagrams in the above remark. Let $b_1, ..., b_n$ be basis vectors in such a lattice (satisfying the dynkin graph conditions). We assume initially that each vector is of length $\sqrt{2}$. Notice that the inner product $b_i \cdot b_j = |b_i| \cdot |b_j| \cdot \cos(\theta_{ij})$ where θ_{ij} is the angle between the two vectors, i.e. if the angle is $\pi/3$ then $\cos\theta_{ij} = 1$.

Notice that if B is the matrix with ij-entry $b_{ij} := b_i \cdot b_j$, then $\det(B) = d(\Lambda)^2$. Next, we observe that each non-zero vector in Λ has length at least $\sqrt{2}$. To see this, suppose $b = \sum k_i b_i$ with not all k_i equal zero. Then

$$b \cdot b = \sum_{i,j} k_i k_j (b_i \cdot b_j) = 2(\sum k_i^2) + 2 \sum_{i \prec j} k_i k_j$$

where $\prec$ means $i < j$ and b_i and b_j are connected by an edge in the graph. This quantity is an even integer and so the length of b is at least $\sqrt{2}$.

Therefore, to determine $\mu(\Lambda)$ in each example it suffice to compute $\det(B)$ and then normalize the length of the vectors so $d(\Lambda) = 1$. For example, for $n = 2$ we have

$$A_2 = \text{}, \quad B = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}, \quad \text{and } \det(B) = 3$$

Thus it suffice to divide our basis vector by $\sqrt[4]{3}$ and then $\mu(A_2) = \frac{\sqrt{2}}{\sqrt[4]{3}} = \sqrt[4]{4/3}$ as desired. For $n = 3$ we get

$$A_3 = \text{}, \quad B = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 2 \end{bmatrix}, \quad \text{and } \det(B) = 4$$

Thus we must divide b_1, b_2, b_3 by $\sqrt[6]{4}$ and so the minimal length of a vector in Λ_3 is $\frac{\sqrt{2}}{\sqrt[6]{4}} = 2^{1/6}$Similarly we get

$$D_4 = \text{  }, \quad B = \begin{bmatrix} 2 & 1 & 0 & 0 \\ 1 & 2 & 1 & 1 \\ 0 & 1 & 2 & 0 \\ 0 & 1 & 0 & 2 \end{bmatrix}, \quad \text{and } \det(B) = 4$$

and so $\mu(\Lambda_4) = 2^{1/4}$

Remark 2.1.6. Now, note D_4 is our first distinct graph (compare to A_2 and A_3, which are just lines), we take a closer look at D_4 with vector length $\sqrt{2}$, say $D_4(\sqrt{2})$. We claim this lattice is the same as the lattice Λ_1 of vectors in $\mathbb{R}^4$ of the form $(u_1, ..., u_4)$ with the u_i's integers and $u_1 + ... + u_4 \equiv 0 \pmod 2$. What are the shortest vectors in the above lattice? They are

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

One can check the lattice is generated by

$$(1, 0, 0, 1), (1, 0, 1, 0), (1, 0, 0, -1), \quad \text{and } (0, 1, 1, 0)$$

Observe that

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

Thus Λ_1 is a representation for the diagram $D_4(\sqrt{2})$.

Now put a sphere of radius $\frac{1}{\sqrt{2}}$ around each lattice point in Λ_1. Notice that any two lattice points in Λ_1 differ by a vector of length at least $\sqrt{2}$. Thus the spheres may touch but they do not overlap in a set of positive volume. Consider the sphere around $(0, 0, 0, 0)$.

It is surrounded by several spheres which touch it. They are $(\pm 1, \pm 1, 0, 0)$, $(\pm 1, 0, \pm 1, 0)$, $(\pm 1, 0, 0, \pm 1)$, $(0, \pm, \pm 1, 0)$, $(0, \pm 1, 0, \pm 1)$ and $(0, 0, \pm 1, \pm 1)$. Thus the central sphere is surrounded by $\binom{4}{2} \cdot 4 = 24$ spheres which touch it. Recently in year 2003, Oleg Musin proved that there is no configuration of 25 spheres of equal radius which touch a central sphere of the same radius without overlap in $\mathbb{R}^4$.

2.2 Kissing Numbers and Packing

Definition 2.2.1. The ***kissing number*** τ_n for $n = 1, 2, 3, ...$ is defined to be the maximum number of unit spheres in $\mathbb{R}^n$ which can touch a central sphere so their interiors do not overlap.

Example 2.2.2. We have $\tau_4 \geq 24$ by the example in last section and we have $\tau_4 \leq 24$ by Musin. Plainly $\tau_1 = 2$ and the hexagonal packing gives $\tau_2 = 6$.

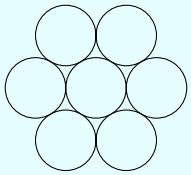


Figure 2.4: Hexagonal Packing

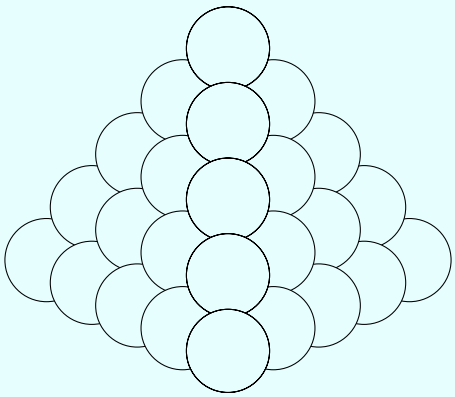
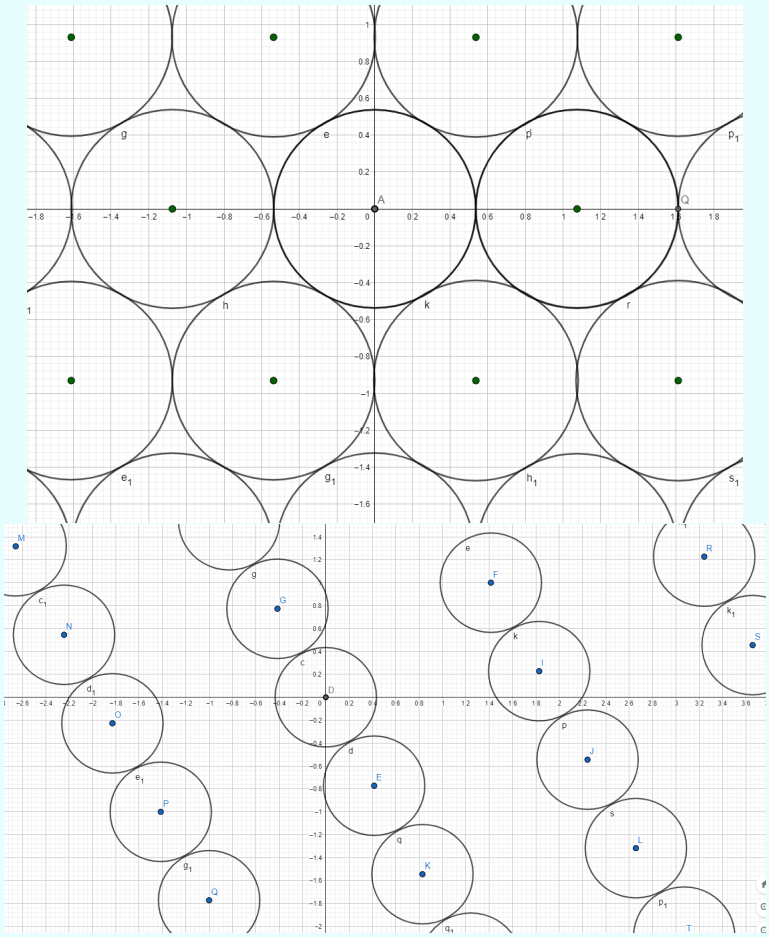


Figure 2.5: Cannonball Packing

For $n = 3$, the standard cannonball packing gives $\tau_3 \geq 12$. It is not clear what τ_3 should be at first glance and there was a dispute between Newton and Gregory as to whether $\tau_3 = 12$ or $\tau_3 = 13$. The first correct proof that $\tau_3 = 12$ is due to Schutte and van der Waerden in 1953.

Definition 2.2.3. A ***sphere packing*** of $\mathbb{R}^n$ is a collection of spheres in $\mathbb{R}^n$ of equal radius whose interiors do not overlap. If the centres of the spheres occur at the points of a lattice we say the packing is a ***lattice packing*** of spheres.

Example 2.2.4. We have the following two distinct lattice packing:



Definition 2.2.5. Let P be a sphere packing with radius ρ, define the ***packing density*** $\Delta = \Delta(P)$ as follows: define S_x^0 to mean a sphere of radius x around 0 in $\mathbb{R}^n$ and $\delta(P \cap S_R^0)$ equal the number of spheres in P inside S_R^0, then

$$\Delta = \Delta(P) = \limsup_{R\to\infty} \frac{\delta(P \cap S_R^0) \cdot \mu(S_\rho^0)}{\mu(S_R^0)}$$

Definition 2.2.6. Now we define Δ_n for $n = 1, 2, 3, \ldots$ as follows

$$\Delta_n = \sup_P \Delta(P)$$

where P run over all possible sphere packings in $\mathbb{R}^n$. Similarly we define

$$\Delta_n(L) = \sup_P \Delta(P)$$

where this time P run over all possible lattice packing in $\mathbb{R}^n$.

Example 2.2.7. Note if L is a lattice then the largest radius of spheres in a sphere packing associated with the lattice is $\frac{1}{2}$ times the minimal non-zero distance between points in the lattice.

If we consider the lattice packing of spheres of radius ρ_0 around each lattice point of Λ then

$$\Delta(\Lambda) = \frac{\mu(S_{\rho_0})}{d(\Lambda)}$$

Certainly $\Delta_n \geq \Delta_n(L)$ for $n = 1, 2, 3, \ldots$. In fact, $\Delta_i = \Delta_i(L)$ for $i = 1, 2$. Indeed, for $n = 2$ the hexagonal lattice yields Δ_2 and we have

$$\Delta_2 = \frac{\pi(\frac{1}{2}\sqrt[4]{4/3})^2}{1} = \frac{\pi}{12}$$

Now let us compute packing density of D_4. Since the minimal non-zero distance between two lattice points in $D_4(\sqrt{2})$ is $\sqrt{2}$ we may take $\rho = \frac{1}{2}\sqrt{2} = \frac{1}{\sqrt{2}}$ and we have

$$\Delta(D_4) = \frac{\frac{\pi^2}{2} \cdot (\frac{1}{\sqrt{2}})^4}{2} = \frac{\pi^2}{16}$$

This is the largest known lattice packing in $\mathbb{R}^4$ and it was proved by Korkine and Zolotareff in 1872 that $\Delta_4(L) = \Delta_4(D_4)$.

Example 2.2.8. Let us consider the lattice of integer points in $\mathbb{R}^n$ denoted by Λ_0, i.e. the diagram is

• • ... •

with n disconnected dots. We have $d(\Lambda_0) = 1$ and the vectors of minimal non-zero length in Λ_0 are $\pm e_i$ with each vector of length 1. Thus the lattice packing associated with Λ_0 consists of spheres of radius $\frac{1}{2}$ around each integer points. Hence the packing density is

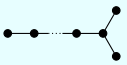
$$\frac{\pi^{n/2}}{\Gamma(1 + n/2)} \left(\frac{1}{2}\right)^n$$

In $\mathbb{R}^2$ we have this is equal $\pi/4$, in $\mathbb{R}^3$ it is $\pi/6$, and in $\mathbb{R}^4$ we have $\pi^2/32$. In particular, the kissing numbers are bounded by Λ_0 in the sense that using sphere packing given by Λ_0 we have kissing numbers are always at least $2n$.

Example 2.2.9. The lattice A_3 associated with

may also be associated with what we call D_3. Since the package I'm using don't have the shape for D_n when $n = 3$, so if you must see what it looks like, consider the "tail" of the following diagram (the diagram you see is D_4)

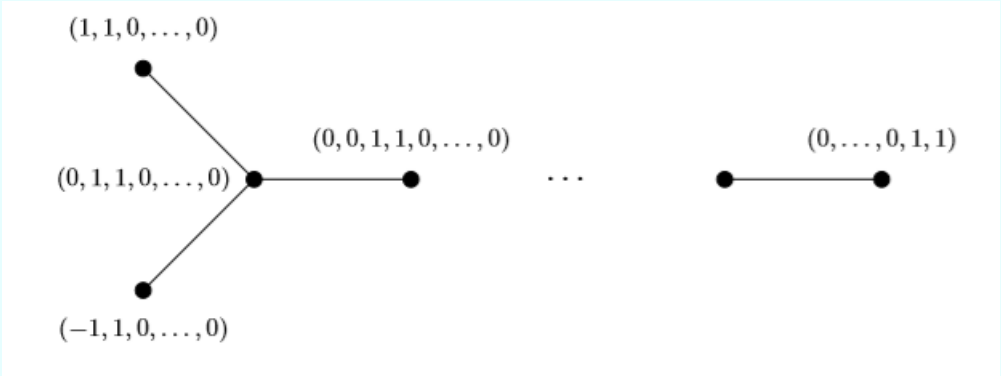
. For $n = 3, 4, 5, 6, ...$ we denote by D_n the lattice associated with



We can represent D_n as the sublattice of Λ_0 given by

$$\{(x_1, ..., x_n) \in \mathbb{Z}^n : x_1 + ... + x_n \equiv 0 \pmod 2\}$$

The lattice is generated by elements of length $\sqrt{2}$, which is the minimal non-zero distance between vectors in the lattice. We take



Thus, we get

$$d\left(D_n(\sqrt{2})\right) = \left|\det \begin{pmatrix} -1 & 1 & 0 & 0 & \cdots & 0 \\ 1 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 1 & & 0 \\ \vdots & \vdots & & \ddots & \ddots & \\ 0 & 0 & \cdots & 0 & 1 & 1 \end{pmatrix}\right| = |-1 \cdot 1 - 1 \cdot 1| = 2$$

The kissing number associated with the lattice $D_n(\sqrt{2})$ corresponds to the number of non-zero vectors of minimal length, so it is $4\binom{n}{2} = 2n(n-1)$. We have a central sphere around $(0, ..., 0)$ of radius $\frac{1}{2}\sqrt{2}$ and it is touched by the $2n(n-1)$ non-overlapping spheres of radius $\frac{\sqrt{2}}{2}$ around $(\pm 1, \pm 1, 0, ..., 0)$,..., $(0, ..., 0, \pm 1, \pm 1)$. Put spheres of radius $\frac{\sqrt{2}}{2}$ around each lattice point to give a sphere packing with density

$$\Delta(D_n(\sqrt{2})) = \frac{\pi^{n/2}}{2\Gamma(1 + n/2)2^n/2}$$

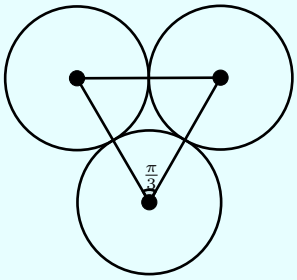
Note that

$$\Delta(D_3) = \frac{\pi}{\sqrt{18}} \approx 0.7405...$$

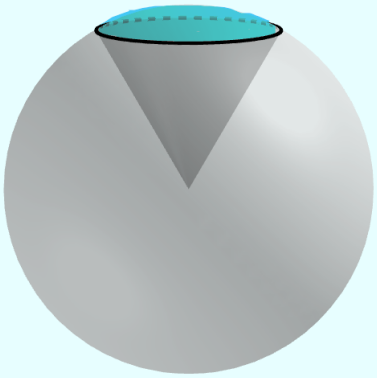
The sphere packing of D_3 corresponds to the cannonball packing. In 1831 Gauss proved that $\Delta_3(L) = \Delta(D_3)$, i.e. D_3 gives the maximal density among all sphere packings induced by lattice in $\mathbb{R}^3$.

Remark 2.2.10. Kepler conjectured that $\Delta_3 = \Delta_3(L) = \Delta(D_3)$, i.e. the cannonball packing gives t he most efficient packing of spheres in $\mathbb{R}^3$. In 1958 Rogus proved $\Delta_3 \leq 0.7796$ and in 1983 Lindsay proved $\Delta_3 \leq 0.7784$. In 2005 Hales gave a proof of Kepler's conjecture which is about 120 pages. It depended on a massive amount of computation and this part of the argument is hard to check.

Example 2.2.11. Consider the kissing number problem in $\mathbb{R}^3$. The following is three spheres touching in $\mathbb{R}^3$:



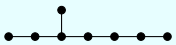
The center of the spheres form an equilateral triangle. Given a configuration of spheres of radius 1 touching a central sphere of radius 1 we can associate to each sphere touching the central sphere a shadow or spherical cap determined by a cone of radius $\pi/3$ from the origin. Such a cap in $\mathbb{R}^3$ is indicated by the cyan highlighter as follows:



The surface area of such a cap is $2\pi h$ where h is the height of the spherical cap and in our case $h = 1 - \frac{\sqrt{3}}{2}$ and so the area is $(2 - \sqrt{3})\pi$. The total surface area of the sphere is 4π and so the kissing number τ_3 is at most $\frac{4\pi}{(2-\sqrt{3})\pi} < 15$, i.e. $\tau_3 \leq 14$. The packing associated with D_3 gives $\tau_3 \geq 12$.

In fact, the following kissing numbers are known:

$$\tau_1 = 2, \tau_2 = 6, \tau_3 = 12, \tau_4 = 24, \tau_8 = 240, \tau_{24} = 196560$$

where τ_8 is given by E_8 with diagram  and τ_{25} is given by what's called the Leech lattice.

Remark 2.2.12. The results above about kissing numbers depend on linear programming and the study of positive semidefinite functions on the sphere S^{n-1} in $\mathbb{R}^n$.

Let $\{x_1, ..., x_m\}$ be points in S^{n-1}, then $x_i \cdot x_i = 1$ for $i = 1, ..., m$. Let θ_{ij} be the distance between x_i and x_j on the surface of S^{n-1}, then this is just the angle in radians determined by the points (those are called spherical codes). Notice that for any real numbers $t_1, ..., t_m$ we have

$$\|t_1 x_1 + ... + t_m x_m\|^2 = \sum_{i,j} t_i t_j \cos(\theta_{ij}) \geq 0$$

or equivalently the matrix

$$[\cos(\theta_{ij})] := \begin{bmatrix} \cos(\theta_{11}) & ... & \cos(\theta_{1m}) \\ \vdots & \ddots & \vdots \\ \cos(\theta_{m1}) & ... & \cos(\theta_{mm}) \end{bmatrix}$$

is positive semidefinite.

Define the Gegenbauer polynomials $G_k^{(n)}(t)$ as follows: for $n = 3, 4, 5, ...$ we have

$$\sum_{k=0}^{\infty} r^k C_k^{(n)}(t) = (1 - 2rt + r^2)^{(2-n)/2}$$

then define $G_k^{(n)}(t) = \frac{C_k^{(n)}(t)}{C_k^{(n)}(1)}$ so that $G_k^{(n)}(1) = 1$. Similarly we may define $G_k^{(n)}(t)$ for $n = 1, 2, 3, ...$ recursively by the rules

$$G_0^{(n)}(t) = 1, G_1^{(n)}(t) = t, G_{k+1}^{n}(t) = \frac{(2k + n - 4)tG_k^{(n)}(t) - (k - 1)G_{k-1}^{(n)}(t)}{k + n - 3}$$

In the case $n = 3$ the polynomials are known as the Legendre polynomials.

In 1943, Schoenberg proved that if we apply $G_k^{(n)}$ to each entry in the above matrix, we still get positive semidefinite matrix, where $G_k^{(n)}$'s are Gegenbauer polynomials defined above. He also proved that if $[f(\cos\theta_{ij})]$ is positive semidefinite for all choices of $x_1, ..., x_m$ in S^{n-1}, then f can be expressed as a linear combination (perhaps infinite) with non-negative coefficients of Gegenbauer polynomials.

Since $[G_k^{(n)}(\cos\theta_{ij})]$ is still positive semidefinite, if $a_0, ..., a_n$ are non-negative real numbers then the matrix with entry $T(\cos\theta_{ij})$ would still be positive semidefinite where $T = a_0 G_0^{(n)} + ... + a_d G_d^{(n)}$. We put

$$f(n, a_0, ..., a_d)(t) = T(t)$$

where T as defined above, and we define $S_f(x_1, ..., x_m)$ by

$$\begin{aligned} S_f(x_1, ..., x_m) &= \sum_{i,j} f(\cos\theta_{ij}) \\ &= \sum_{k=0}^{d} a_k \sum_{i,j} G_k^{(n)}(\cos\theta_{ij}) \end{aligned}$$

Thus, since $a_0, ..., a_d$ are non-negative and $\sum_{i,j} G_k^{(n)}(\cos\theta_{ij}) \geq 0$ for $k = 0, ..., d$ we get

$$S_f(x_1, ..., x_m) \geq a_0 \sum_{i,j} G_0^{(n)}(\cos\theta_{ij}) = a_0 m^2$$

Now suppose $x_1, ..., x_m$ is a configuration of m points on S^{n-1} which correspond to the m points of contacts by some m many kissing spheres with the center sphere. Then $\theta_{ij} \geq \frac{\pi}{3}$ provided $i \neq j$ hence $\cos\theta_{ij} \leq \frac{1}{2}$ for $i \neq j$.

Suppose $a_0, ..., a_d$ are non-negative real numbers for which $f(t) \leq 0$ for t in the range $[-1, \frac{1}{2}]$. Then

$$S_f(x_1, ..., x_m) \leq m f(1)$$

and if $a_0 > 0$ we get

$$m \leq \frac{f(1)}{a_0}$$

as $S_f(x_1, ..., x_m) \geq a_0 m^2$. The strategy now is to choose $a_0, ..., a_d$ so $f(t) \leq 0$ for $[-1, \frac{1}{2}]$ and such that a_0 is large and $f(1)$ is small.

For $n = 8$ we consider

$$f(t) = G_0^{(8)} + \frac{16}{7} G_1^{(8)} + \frac{200}{63} G_2^{(8)} + \frac{832}{231} G_3^{(8)} + \frac{1216}{429} G_4^{(8)} + \frac{5120}{3003} G_5^{(8)} + \frac{2560}{4641} G_6^{(8)}$$

then $f(t) = \frac{320}{3}(t+1)(t+\frac{1}{2})^2 t^2 (t-\frac{1}{2})$ and so one can check $f(t) \leq 0$ for $[-1, \frac{1}{2}]$. Thus

$$\tau_8 \leq \frac{320}{3} \cdot 2 \cdot \frac{3^2}{2^2} \cdot \frac{1}{2} = 240$$

For $n = 24$ we can find a non-negative linear combination of $G_k^{(24)}$ to give

$$f(t) = \frac{1490944}{15}(t+1)\left(t+\frac{1}{2}\right)^2 \left(t+\frac{1}{4}\right)^2 t^2 \left(t-\frac{1}{4}\right)^2 \left(t-\frac{1}{2}\right)$$

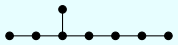
and $f(t) \leq 0$ for $t \in [-1, \frac{1}{2}]$. Thus

$$\tau_{24} \leq 196560$$

We note in the case of $n = 8$ and 24, the combination is obtained by running linear programming packages (i.e. not out of the blue).

Now it remains to show E_8 lattice give us lower bound of $240 \leq \tau_8$ and the Leech lattice give us lower bound of $196560 \leq \tau_{24}$.

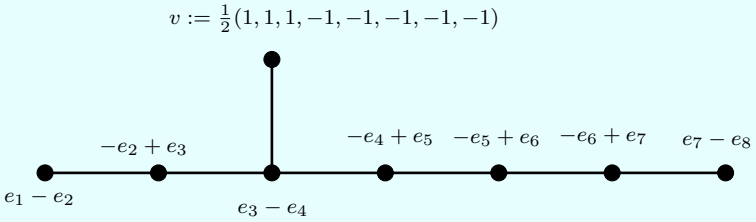
Example 2.2.13. Let us now look at E_8 more closely. Recall the diagram for E_8 is



Now let each vector be normalized to have length $\sqrt{2}$ we have the matrix B of inner product is

$$B = B\left(E_8(\sqrt{2})\right) = \begin{pmatrix} 2 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 2 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 2 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 2 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 2 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 2 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 2 \end{pmatrix}$$

One should check $\det B = 1$. Also note we may realize $E_8(\sqrt{2})$ in the following way:



Again, since the minimal distance between distinct points in $E_8(\sqrt{2})$ is $\sqrt{2}$ we may put sphere of radius $\frac{\sqrt{2}}{2}$ around each vector in the lattice. Then the number of vectors in $E_8(\sqrt{2})$ of length $\sqrt{2}$ will be the kissing number of the lattice.

Now note $2 \cdot v = (1,1,1,-1,-1,-1,-1,-1)$ is in the lattice. However, what about $e_1 + e_2$? Well, we just note the integral span of the basis vectors on the bottom row consists of all integer vectors whose sum of coordinates is zero. The sum of coordinates of $2v$ is -2, hence all vectors whose sum is congruent to 0 mod 2 are in $E_8(\sqrt{2})$, hence $e_1 + e_2$.

In particular, by the above argument we see any vector which has two coordinates from $\{1, -1\}$ and others 0 are in $E_8(\sqrt{2})$. This g ives us $4 \cdot \binom{8}{2} = 112$ vectors of length $\sqrt{2}$. These vectors together with v allow us to show the vectors $(\frac{\delta_1}{2}, ..., \frac{\delta_8}{2})$ are in the lattice where $\delta_i \in \{1, -1\}$ with $\prod_{i=1}^{8} \delta_i = -1$. There are 2^7 of these vectors of length $\sqrt{2}$. In total we get $112 + 128 = 240$ such vectors.

Note there are no other vectors of length $\sqrt{2}$ in the lattice, since if one coordinate is $\frac{3}{2}$ or larger in absolute value, the vector is of length greater than $\sqrt{2}$, and if there are more than 2 coordinates of absolute value at least 1, then the length again exceeds $\sqrt{2}$.

Therefore $\tau_8(E_8(\sqrt{2})) = 240$, i.e. $\tau_8 \geq 240$ as desired. With the last remark we get

$$\tau_8 = 240$$

The packing density associated to $E_8(\sqrt{2})$ is

$$\frac{\frac{\pi^4}{\Gamma(5)} \cdot \left(\frac{1}{\sqrt{2}}\right)^8}{1} = \frac{\pi^4}{24 \cdot 16} = \frac{\pi^4}{384} = 0.2537...$$

This is the largest lattice packing density in $\mathbb{R}^8$ and it is the largest packing density known in $\mathbb{R}^8$.

There are 240 vectors x in $E_8(\sqrt{2})$ for which $x \cdot x = 2$. The next smallest norm in the lattice is 4 and there are 2160 vectors x in $E_8(\sqrt{2})$ for which $x \cdot x = 4$. These are of the form

$$\pm 2e_1, ..., \pm 2e_8$$
$$\pm\Big(\sum_{\substack{I \subseteq \{1,...,8\}, |I|=4}} \sum_{i \in I} e_i\Big)$$

and

$$(\epsilon_1 \frac{3}{2}, \epsilon_2 \frac{1}{2}, ..., \epsilon_8 \frac{1}{2}), \quad \text{where } \epsilon_i \in \{1, -1\} \wedge \prod_{i=1}^{8} \epsilon_i = 1$$

and all permutations of the coordinates are allowed. There are 6720 elements of norm 6 and 17520 elements of norm 8, and 30240 elements of norm 10.

In fact, for each positive integer m the number $N(m)$ of x in $E_9(\sqrt{2})$ for which $x \cdot x = 2m$ is given by

$$240 \cdot \sigma_3(m), \quad \text{where } \sigma_3(m) = \sum_{d|m, d>0} d^3$$

How do we get such a result?

Let Λ be a lattice in $\mathbb{R}^n$ with $x \cdot y \in \mathbb{Z}$ for any $x, y \in \Lambda$. Suppose $b_1, ..., b_n$ is a basis for Λ and, as before, put B be the matrix with ij-entry given by $b_i \cdot b_j$. Then for any $x \in \Lambda$ we can find integer $k_1, ..., k_n$ so

$$x = \sum k_i b_i$$

and so

$$x \cdot x = 2 \sum_{\substack{i,j \\ i<j}} k_i k_j (b_i \cdot b_j)$$

and so this is a quadratic form in $(k_1, ..., k_n)$ given by

$$\begin{bmatrix} k_1 & ... & k_n \end{bmatrix} B \begin{bmatrix} k_1 \\ \vdots \\ k_n \end{bmatrix}$$

Let $q = e^{2\pi i z}$ for $z \in \mathbb{C}$, we now can define the theta function of the lattice Λ, denoted $\theta_\Lambda(z)$ by

$$\theta_\Lambda(z) = \sum_{x \in \Lambda} q^{\frac{x \cdot x}{2}} = \sum_{x \in \Lambda} e^{(x \cdot x)\pi i z}$$

If B has integer entries and determinant 1 and $x \cdot x \equiv 0 \pmod 2$ for all $x \in \Lambda$, then it can be proved that $\theta_\Lambda(z)$ is a modular form of weight $n/2$.

Remark 2.2.14. Note the group $\mathrm{SL}_2(\mathbb{Z})$ acts on the half-plane $H = \{z \in \mathbb{C} : \mathrm{Im}(z) > 0\}$ as follows $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathrm{SL}_2(\mathbb{Z})$ then we get $Az = \frac{az+b}{cz+d}$.

Definition 2.2.15. Let k be an integer, we say a meromorphic function $f : H \to \mathbb{C}$ is ***weakly modular of weight*** $2k$ if

$$f(z) = (cz+d)^{-2k} f(Az), \quad \text{for all } A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathrm{SL}_2(\mathbb{Z})$$

Remark 2.2.16. If $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ then $Az = z+1$ and if f is weakly modular of weight $2k$ then we must have $f(z+1) = f(z)$. In other word, f can be expressed in terms of $q = e^{2\pi i z}$. In particular, f determines a function $\tilde{f}(q)$ where

$$\tilde{f} : \{q \in \mathbb{C} : 0 < |q| < 1\} \to \mathbb{C}$$

and $\tilde{f}$ is meromorphic on the punctured disk $\{q \in \mathbb{C} : 0 < |q| < 1\}$ and if it extends to a meromorphic function on all of the disc then we say f is a ***modular function***. If $\tilde{f}$ is holomorphic on the punctured disk $\{q \in \mathbb{C} : 0 < |q| < 1\}$ and extends to a holomorphic function on $\{q \in \mathbb{C} : |q| < 1\}$ then we say f is a ***modular form***.

Remark 2.2.17. The space of modular forms of weight $2k$ forms a vector space M_{2k} over $\mathbb{C}$ of dimension d_k where if $k \geq 0$ then

$$d_k = \begin{cases} \lfloor \frac{k}{6} \rfloor, & k \equiv 1 \pmod 6 \\ \lfloor \frac{k}{6} \rfloor + 1, & k \not\equiv 1 \pmod 6 \end{cases}$$

Remark 2.2.18. The lattice $\Lambda = E_8(\sqrt{2})$ determines $\theta_{E_8(\sqrt{2})}(z)$ which is a modular form of weight 4. Thus $\theta_{E_8(\sqrt{2})}(z)$ lies in M_4, i.e. a vector space of dimension 1 over $\mathbb{C}$. Now $E_2(z) = 1 + 240 \sum_{n=1}^{\infty} \sigma_3(n) q^n$ is in M_4 and we have

$$\theta_{E_8(\sqrt{2})}(z) = \sum_{m=0}^{\infty} r_\Lambda(m) q^m$$

where $r_\Lambda(m)$ counts the number of vectors x in $\Lambda = E_8(\sqrt{2})$ for which $x \cdot x = 2m$. Thus $E_2(z) = \theta_{E_8(\sqrt{2})}(z)$.

2.3 Leech Lattice

Remark 2.3.1. Associated to each lattice Λ in $\mathbb{R}^n$ is $\mathrm{Aut}(\Lambda)$, the group of symmetries of the lattice which fix the origin, or equivalently the set of isometries of $\mathbb{R}^n$ which fix the origin and take the lattice to itself.

For example, the automorphism group of the hexagonal lattice A_2 is the Dihedral group D_6. On the other hand, the automorphism group of $E_8(\sqrt{2})$ is a group of order $2^{14} \cdot 3^2 \cdot 5^2 \cdot 7$ and it permutes the 240 vectors of minimal length transitively.

Remark 2.3.2. In this section we will construct the Leech lattice L, which is a lattice in $\mathbb{R}^{24}$. It has determinant 1, the dual of L is L, the associated inner product matrix B has integer entries, and $x \in L$ and $x \neq 0$ then

$$x \cdot x \geq 4$$

Proposition 2.3.3. *There exists a* 12 *dimensional subspace* S *of* $\mathbb{F}_2^{24}$ *with the following property: for every non-zero v ector* $s \in S$ *the number of coordinates which are* 1 *is at least* 8 *and is congruent to* 0 *mod* 4. *Further,* $(1, ..., 1) \in S$.

Proof. To prove this, we will realize S as the span of the row of a 12 by 24 matrix over $\mathbb{F}_2$ which we shall construct. Let A be a symmetric 11 by 11 matrix whose first row is

$$1, 1, 1, 0, 1, 1, 0, 1, 0, 0, 0$$

and whose remaining rows are obtained by permuting the rows cyclically to the left, i.e. we get

$$A = \begin{pmatrix} 1 & 1 & 1 & 0 & 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 1 & 0 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 1 & 0 & 1 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 & 1 & 0 & 0 & 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & 1 & 1 & 0 & 1 & 0 & 0 \end{pmatrix}$$

Now let B be the symmetric 12 by 12 matrix obtained by adjoining a first row of the form $(0, 1, 1, ..., 1)$ and a column of the form $(0, 1, 1, ..., 1)$, i.e. we get

$$B = B = \begin{pmatrix} 0 & 1 & 1 & \cdots & 1 \\ 1 & & & & \\ \vdots & & & A & \\ 1 & & & & \end{pmatrix}$$

Since any two rows of A have exactly three columns of the form $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ we see

$$B^2 = BB^T = I$$

over $\mathbb{F}_2$. Now put

$$C = (I_{12}|B)$$

to be a 12 by 24 matrix. We claim that S is the subspace of $\mathbb{F}_2^{24}$ generated by the rows of C. Denote the subspace generated by the rows of C to be S_1.

First note $(1, ..., 1) \in S_1$ since we can obtain it by adding the rows of C over $\mathbb{F}_2$. Next we note the number of 1 in any row of C is either 8 or 12 and any two

rows of C are orthogonal since any two rows of A have precisely three 1s in common columns.

For any 24-tuple $s = (s_1, ..., s_{24})$ in $\mathbb{F}_2^{24}$ we put $\|s\|$ equal the number of coordinates of s which are 1. We note if s is a linear combination of rows of C then $\|s\| = 0 \pmod 4$. To see this, note any two rows of C are orthogonal so if we add one row of C to another to get a matrix C' then the rows of C' will be orthogonal. If a row q_1 is obtained by adding a row s of C to a row r of C then

$$\|q_1\| = \|r\| + \|s\| - 2n$$

where n is the number of columns for which both entries are 1. Since r and s are orthogonal n is even and since r and s are in C, $\|r\|$ and $\|s\|$ are in $\{8, 12\}$. Thus $\|r\| \equiv 0 \pmod 4$, $\|s\| \equiv 0 \pmod 4$ and so we get $\|q_1\| \equiv 0 \pmod 4$ as desired. The result now follows by induction.

Now we can prove if s is a non-zero linear combination of the rows of C then $\|s\| \geq 8$. Since $\|s\| \equiv 0 \pmod 4$ it suffice to show $\|s\| \geq 5$.

Suppose s is a linear combination of k elements of C. If $k = 1$ then we are done. If $k = 2$ then since the rows of A have exactly three columns with two 1s and each row of A has six 1s we find $\|s\|$ is either 8 or 12.

If $k = 3$ then s is the sum of the first row and two other rows then since the rows of A have exactly three columns with two 1s we see $\|s\| = 8$. On the other hand, if the three rows do not include the first row then first 13 coordinates of s contain four 1s. If there are no other 1s in s the sum of three rows of A give the zero vector $(0, ..., 0)$ in $\mathbb{F}_2^{11}$ which does not happen. Thus $\|s\| \geq 5$ and so $\|s\| \geq 8$.

If $k = 4$ then the first 12 coordinates of s have four 1s. If the remaining coordinates are 0 then the sum of 4 rows of B are $(0, ..., 0)$ which contradicts the fact B is non-singular. Now recall $B^2 = I$, i.e. $\|s\| \geq 5$.

If $k \geq 5$ then we get at least five 1s in the first 12 coordiantes and the result follows. This complete the proof. ♡

Remark 2.3.4 (Construction of Leech Lattice)**.** Let $e_1, ..., e_{24}$ be the standard basis of $\mathbb{R}^{24}$. We put $b_i = \frac{1}{\sqrt{8}} e_i$ for all $1 \leq i \leq 24$. Let L_0 be the lattice generated by $b_1, ..., b_{24}$ in $\mathbb{R}^{24}$. Let L be the sublattice of L_0 whose elements are of the form

$$t_1 b_1 + ... + t_{24} b_{24}$$

where $t_1, ..., t_{24} \in \mathbb{Z}$ and satisfying either one of the following:

1. $t_1, ..., t_{24} \in 2\mathbb{Z}$ and $\sum t_i \equiv 0 \pmod 8$ and $\frac{1}{2}(t_1, ..., t_{24})$ reduced mod 2 lies in the subspace S as in the above proposition.
2. $t_1, ..., t_{24}$ are odd, $\sum t_i \equiv 4 \pmod 8$ and $\frac{1}{2}(1 + t_1 + ... + t_{24})$ reduced mod 2 lies in the subspace S of proposition above.

Notice L is a lattice and hence a sublattice of L_0. To see this note L contains 24 linearly independent vectors since it contains $8b_1, ..., 8b_{24}$. Further it is discrete

since it is contained in L_0. Next note if $x \in L$ then $-x \in L$ and if $x, y \in L$ then $x + y \in L$ since S is a subspace of $\mathbb{F}_2^{24}$.

This L is called the ***Leech lattice***.

Remark 2.3.5. Now we show $x \in L$ then $x \cdot x \equiv 0 \pmod 2$. If $x = \sum t_i b_i$ then

$$x \cdot x = \frac{1}{8}(\sum t_i^2)$$

and hence it suffice to prove

$$\sum t_i^2 \equiv 0 \pmod{16}$$

Consider first that $t_1, ..., t_{24}$ are all even. Then if $t_i \equiv 0 \pmod 4$ we get $t_i^2 \equiv 0 \pmod{16}$. If $t_i \equiv 2 \pmod 4$ then $t_i^2 \equiv 4 \pmod{16}$. Recall if $s \in S$ then $\|s\| \equiv 0 \pmod 4$ so the number of indices i for which $t_i \equiv 2 \pmod 4$ is a multiple of 4 and hence

$$\sum t_i^2 \equiv 0 \pmod{16}$$

as desired. On the other hand, if $t_1, ..., t_{24}$ are all odd then $t_i \equiv \pm 1 \pmod 8$ imply $t_i^2 \equiv 1 \pmod{16}$ whilte if $t_i \equiv \pm 3 \pmod 8$ we have $t_i^2 \equiv 8 \pmod{16}$. Let α_j be the number of t_is with $t_i \equiv j \pmod 8$ we get

$$\sum t_i^2 \equiv \alpha_1 + 9\alpha_3 + 9\alpha_5 + \alpha_7 \pmod{16}$$

We also have

$$24 = \alpha + \alpha_3 + \alpha_5 + \alpha_7 \equiv 0 \pmod 8$$

and, by the definition of L,

$$\alpha_1 + 3\alpha_3 + 5\alpha_5 + 7\alpha_7 \equiv 4 \pmod 8$$

Further, by the Proposition of this section, we get

$$\alpha_1 + \alpha_5 \equiv 0 \pmod 4$$

so

$$2(\alpha_1 + \alpha_5) \equiv 0 \pmod 8$$

Thus $\alpha_3 + \alpha_5$ is odd and so we see

$$\sum t_i^2 \equiv 24 + 8(\alpha_3 + \alpha_5) \equiv 0 \pmod{16}$$

as desired.