

PMATH 944 Lecture 1: September 15, 2009

Cam Stewart MC 5051

Books:

Geometry of Numbers: J. W. S. Cassels

Geometry of Numbers: Lekkerkerker

Sphere Packings, Lattices and Groups: Conway & Sloane

Let v_1, \dots, v_n be linearly independent vectors in \mathbb{R}^n . The set

$$\Lambda = \{a_1v_1 + \dots + a_nv_n : (a_1, \dots, a_n) \in \mathbb{Z}^n\}$$

is said to be a lattice with basis v_1, \dots, v_n .

Note that since v_1, \dots, v_n are linearly independent, each element of Λ has a unique representation as a linear combination of v_1, \dots, v_n . Further, the coefficients in the representation are integer.

Observe that the basis v_1, \dots, v_n is not uniquely determined by Λ . In particular, let A be an $n \times n$ matrix with integer entries and determinant ± 1 and put

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

We claim that w_1, \dots, w_n is also a basis for Λ . Certainly w_1, \dots, w_n are linearly independent vectors in \mathbb{R}^n . Secondly note that

$$\begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = A^{-1} \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} \quad \text{and} \quad A^{-1} = \frac{1}{\det(A)} \text{adj } A$$

Recall that the i, j -th entry of $\text{adj } A$ is the cofactor of $a_{j,i}$. But the cofactor is an integer and $\det(A) = \pm 1$ so v_i can be expressed as an integer linear combination of w_1, \dots, w_n . Thus every element of Λ is an integer linear combination of w_1, \dots, w_n . Thus w_1, \dots, w_n is also a basis for Λ .

Suppose now that v_1, \dots, v_n is a basis for Λ and that w_1, \dots, w_n is also a basis for Λ . We'll now show that they are related as above. In particular since v_1, \dots, v_n is a basis we can express w_i for $i = 1, \dots, n$ as an integer linear combination of v_1, \dots, v_n . Thus there is an $n \times n$ matrix A with integer entries such that

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

Similarly, there is an $n \times n$ matrix B with integer entries such that

$$B \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$$

Therefore

$$AB \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} = \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix}$$

hence $AB = I$ so $\det A \cdot \det B = 1$. But $\det A$ and $\det B$ are integers so $\det A = \pm 1$.

We are now in a position to define the determinant $d(\Lambda)$ of Λ . Let v_1, \dots, v_n be a basis for Λ . We put

$$d(\Lambda) = |\det(v_1, \dots, v_n)|.$$

Here (v_1, \dots, v_n) represents the matrix obtained by writing the v_i s with respect to the standard basis $(1, 0, \dots, 0), (0, 1, 0, \dots, 0), \dots, (0, \dots, 0, 1)$ in \mathbb{R}^n .

Notice that $d(\Lambda)$ does not depend on the choice of basis v_1, \dots, v_n since if w_1, \dots, w_n is another basis for Λ then there is a matrix A with $\det(A) = \pm 1$ such that

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

and we see that

$$|\det(w_1, \dots, w_n)| = |\det(A)| \cdot |\det(v_1, \dots, v_n)| = |\det(v_1, \dots, v_n)|$$

Remark: Since v_1, \dots, v_n are linearly independent we see that $d(\Lambda) > 0$.

The simplest lattice in \mathbb{R}^n is Λ_0 where Λ_0 is generated by $(1, 0, \dots, 0), (0, 1, 0, \dots, 0), \dots, (0, \dots, 0, 1)$. Then $d(\Lambda_0) = 1$.

Let Λ and Λ_1 be lattices in \mathbb{R}^n . If $\Lambda_1 \subseteq \Lambda$ then Λ_1 is said to be a sublattice of Λ . Note that if w_1, \dots, w_n are linearly independent vectors in a lattice Λ in \mathbb{R}^n then they generate a sublattice Λ_1 of Λ and there is a matrix A with integer entries such that

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

Let $D = |\det(A)|$ and note that D is a positive integer. Further

$$D = \frac{|\det(w_1, \dots, w_n)|}{|\det(v_1, \dots, v_n)|} = \frac{|\det(w_1, \dots, w_n)|}{d(\Lambda)} = \frac{d(\Lambda_1)}{d(\Lambda)}$$

where Λ_1 is the lattice generated by w_1, \dots, w_n . D is known as the index of Λ_1 in Λ .

Suppose that Λ is a lattice in \mathbb{R}^n and Λ_1 is a sublattice of Λ of index D . Let v_1, \dots, v_n be a basis for Λ and w_1, \dots, w_n be a basis for Λ_1 . Then we have a matrix A with integer entries and $|\det(A)| = D$ such that

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

Thus

$$\begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = \frac{1}{\det A} \operatorname{adj} A \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix},$$

and so

$$\begin{pmatrix} Dv_1 \\ \vdots \\ Dv_n \end{pmatrix} = \frac{D}{\det A} \operatorname{adj} A \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix}$$

hence Dv_i is an integer linear combination of w_1, \dots, w_n for $i = 1, \dots, n$. In particular $Dv_i \in \Lambda_1$ for $i = 1, \dots, n$.

Theorem 1: Let Λ_1 be a sublattice of the lattice Λ in \mathbb{R}^n .

A) If v_1, \dots, v_n is a basis for Λ then there is a basis w_1, \dots, w_n of Λ_1 such that

$$\begin{aligned} w_1 &= a_{11}v_1 \\ w_2 &= a_{21}v_1 + a_{22}v_2 \\ &\vdots \\ w_n &= a_{n1}v_1 + \dots + a_{nn}v_n \end{aligned} \tag{1}$$

where

- i) the a_{ij} s are integers
 - ii) $a_{ii} > 0$ for $i = 1, \dots, n$
 - iii) $0 \leq a_{ij} < a_{jj}$ for $1 \leq j < i \leq n$.
- B) If w_1, \dots, w_n is a basis Λ_1 then there is a basis v_1, \dots, v_n for Λ such that (1) holds with
- i) the a_{ij} s are integers
 - ii) $a_{ii} > 0$ for $i = 1, \dots, n$
 - iii)' $0 \leq a_{ij} < a_{ii}$ for $1 \leq j < i \leq n$.

Proof:

- A) Let D be the index of Λ_1 in Λ . For each i with $1 \leq i \leq n$ there exist vectors

$$w_i = a_{i1}v_1 + \dots + a_{ii}v_i$$

in Λ_1 with $a_{ij} \in \mathbb{Z}$ and $a_{ii} > 0$ since $Dv_i \in \Lambda_1$. We choose w_i for $i = 1, \dots, n$ in such a way that a_{ii} is positive and as small as possible. Since w_1, \dots, w_n are in Λ_1 we have $b_1w_1 + \dots + b_nw_n$ in Λ_1 for any integers b_1, \dots, b_n .

We claim that w_1, \dots, w_n for a basis for Λ_1 .

If not then there is a vector z in Λ_1 which is not of the form $b_1w_1 + \dots + b_nw_n$ with b_1, \dots, b_n integers. Then there exist integers c_1, \dots, c_n such that $z = c_1v_1 + \dots + c_nv_n$. We now choose z in Λ_1 for which the representation has $c_{i+1} = \dots = c_n = 0$ with i minimal. In particular $z = c_1v_1 + \dots + c_iv_i$.

Let $c_i = qa_{ii} + r$ with $0 \leq r < a_{ii}$. Then

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

Note that $z - qw_i \in \Lambda_1$ and is an integer linear combination of v_1, \dots, v_i . Further note that $r \neq 0$ since i is minimal. But this contradicts the minimal choice of a_{ii} . Thus w_1, \dots, w_n forms a basis for Λ_1 .

It remains to check that iii) holds. To obtain iii) we replace w_i by w'_i for $i = 1, \dots, n$ where

$$w'_i = b_{i1}w_1 + \dots + b_{i,i-1}w_{i-1} + w_i,$$

with the b_{ij} s integers to be chosen. Note that w'_1, \dots, w'_n is a basis for Λ_1 and that

$$w'_i = a'_{i1}v_1 + \dots + a'_{ii}v_i$$

with $a'_{ii} = a_{ii}$ for $i = 1, \dots, n$. Further for $j < i$ we have

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

For each i we now choose $b_{i,i-1}, b_{i,i-2}, \dots, b_{i,1}$ in that order so that $0 \leq a'_{ij} < a_{jj} = a'_{jj}$ as required.

PMATH 944 Lecture 2: September 17, 2009

Theorem 1: Let Λ_1 be a sublattice of Λ in \mathbb{R}^n .

- A) ✓
- B) If w_1, \dots, w_n is a basis for Λ_1 , then there is a basis v_1, \dots, v_n of Λ such that (1) holds, i) and ii) hold, and
- iii)' $0 \leq a_{ij} < a_{ii}$, for $1 \leq j < i \leq n$.

Proof B): Let w_1, \dots, w_n be a basis for Λ_1 .

Let D be the index of Λ_1 in Λ . Recall that $D\Lambda$ is a sublattice of Λ_1 . In particular, by part A), there is a basis Dv_1, \dots, Dv_n of $D\Lambda$ such that

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

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

Put

$$A = \begin{pmatrix} a_{11} & & \\ \vdots & \ddots & \\ a_{n1} & \dots & a_{nn} \end{pmatrix}$$

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

and so

$$\begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} = D \cdot \frac{\text{adj } A}{\det(A)} \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$$

Further,

$$\text{adj } A = \begin{pmatrix} b_{11} & & \\ \vdots & \ddots & \\ b_{n1} & \dots & b_{nn} \end{pmatrix}$$

with $b_{ij} \in \mathbb{Z}$. Note that w_i can be expressed as a rational linear combination of v_1, \dots, v_n and that it is an integral linear combination of v_1, \dots, v_n .

Thus we obtain (1) with i) holding. To obtain ii), it suffices to change the sign of v_i if necessary, for $i = 1, \dots, n$.

Finally, to obtain iii)', we replace v_i by v'_i , where

$$v'_i = c_{i1}v_1 + \dots + c_{i,i-1}v_{i-1} + v_i$$

where the c_{ij} s are integers chosen as in A) to ensure iii)'.

Corollary 1: Let Λ be a lattice in \mathbb{R}^n , and let w_1, \dots, w_m be linearly independent vectors of Λ . Then there exists a basis v_1, \dots, v_n of Λ for which

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

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

Proof: Extend w_1, \dots, w_m to a set of n linearly independent vectors w_1, \dots, w_n of Λ . Consider the sublattice Λ_1 generated by the basis w_1, \dots, w_m and apply Theorem 1.

Corollary 2: Let w_1, \dots, w_m be linearly independent vectors from a lattice Λ in \mathbb{R}^n , with $m < n$. There exist w_{m+1}, \dots, w_n in Λ such that w_1, \dots, w_n is a basis for Λ , if and only if every vector $a_1w_1 + \dots + a_mw_m$ in Λ with $a_i \in \mathbb{R}$ for $i = 1, \dots, m$ has in fact $a_i \in \mathbb{Z}$ for $i = 1, \dots, m$.

Proof: \implies : immediate.

\Leftarrow : We apply Corollary 1 to get a basis v_1, \dots, v_n of Λ with

$$\begin{aligned} w_1 &= a_{11}v_1 \\ &\vdots \\ w_m &= a_{m1}v_1 + \dots + a_{mm}v_m \end{aligned} \quad a_{ij} \in \mathbb{Z}, a_{ii} > 0$$

Thus, $v_1 = \frac{1}{a_{11}}w_1$, and we get by hypothesis $\frac{1}{a_{11}} \in \mathbb{Z}$, hence $a_{11} = 1$ ¹⁾.

Next, $w_2 = a_{21}v_1 + a_{22}v_2$, hence $\frac{1}{a_{22}}w_2 = \frac{a_{21}}{a_{22}}w_1$ ²⁾ + v_2 , $\implies a_{22} = 1$.

In this way, we find $a_{11} = a_{22} = \dots = a_{mm} = 1$.

Then $w_1, \dots, w_m, v_{m+1}, \dots, v_n$ is a basis for Λ .

Corollary 3: Let v_1, \dots, v_n be a basis for Λ and let $w = a_1v_1 + \dots + a_nv_n$ be in Λ , so $a_i \in \mathbb{Z}$ for $i = 1, \dots, n$. Let m be an integer with $1 \leq m \leq n-1$.

Then

$$v_1, \dots, v_{m-1}, w \text{ can be extended to a basis for } \Lambda \iff \gcd(a_m, \dots, a_n) = 1.$$

Proof: \implies : Let $g = \gcd(a_m, \dots, a_n)$.

If v_1, \dots, v_{m-1}, w can be extended by say w_{m+1}, \dots, w_n to a basis for Λ , then

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

therefore

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

Now, $\frac{a_t}{g} \in \mathbb{Z}$, for $t = m, \dots, n$

Thus $\frac{1}{g}w - \frac{a_1}{g}v_1 - \dots - \frac{a_{m-1}}{g}v_{m-1}$ is in Λ . We now apply Corollary 2 to conclude $\frac{1}{g} \in \mathbb{Z}$, hence $g = 1$.

\Leftarrow : We wish to find w_{m+1}, \dots, w_n in Λ for which $v_1, \dots, v_{m-1}, w, w_{m+1}, \dots, w_n$ is a basis for Λ .

Then:

$$\begin{aligned} v_1 &= v_1 \\ &\vdots \\ v_{m-1} &= v_{m-1} \\ w &= a_1v_1 + \dots + a_mv_m + \dots + a_nv_n \\ w_{m+1} &= b_1v_1 + \dots + b_mv_m + \dots + b_nv_n \quad b_i \in \mathbb{Z} \\ &\vdots \\ w_n &= z_1v_1 + \dots + z_mv_m + \dots + z_nv_n \quad z_i \in \mathbb{Z} \end{aligned}$$

It suffices to show that we can choose the coefficients $b_1, \dots, b_n, \dots, z_1, \dots, z_n$ as integers in such a way that the associated coefficient matrix has determinant ± 1 . Notice that it is enough to show that the row (a_m, \dots, a_n) can be extended to an $(n-m+1) \times (n-m+1)$ matrix with integer entries and determinant ± 1 .

Consider the standard lattice Λ_0 in \mathbb{R}^{n-m+1} . It now suffices to show that we can extend (a_m, \dots, a_n) to a basis for Λ_0 . We appeal to Corollary 2. Notice that if $\alpha \in \mathbb{R}$ with $\alpha \neq 0$, and $\alpha(a_m, \dots, a_n)$ is in Λ_0 , then $\alpha \in \mathbb{Q}$, say $\alpha = \frac{p}{q}$ with p and q coprime non-zero integers.

Then

$$\left(\frac{pa_m}{q}, \dots, \frac{pa_n}{q} \right) \in \Lambda_0$$

hence $q \mid pa_m, \dots, q \mid pa_n$, and so, since p and q are coprime, $q \mid \gcd(a_m, \dots, a_n)$.

¹⁾ $a_{11} = \pm 1, a_{11} > 0$
²⁾ $= v_1$

Recall the standard dot product of two vectors $v = (a_1, \dots, a_n)$ and $w = (b_1, \dots, b_n)$ in \mathbb{R}^n , given by $v \cdot w = a_1 b_1 + \dots + a_n b_n$. Let v_1, \dots, v_n be a basis for a lattice Λ in \mathbb{R}^n . Since v_1, \dots, v_n are linearly independent, there exist vectors v_1^*, \dots, v_n^* such that

$$v_j^* \cdot v_i = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

v_1^*, \dots, v_n^* are linearly independent, and they generate a lattice Λ^* in \mathbb{R}^n . Λ^* is known as the *polar lattice* of Λ , and one can show that it does not depend on the choice of basis for Λ .

Theorem 2: Let Λ be a lattice in \mathbb{R}^n . The polar lattice Λ^* of Λ consists of all vectors v^* in \mathbb{R}^n for which $v^* \cdot v$ for all v in Λ . Further,

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

Proof: If v_1, \dots, v_n is...

PMATH 944 Lecture 3: September 22, 2009

Theorem 2: Let Λ be a lattice. The polar lattice of Λ consists of the vectors v^* such that $v^* \cdot v$ is an integer for all v in Λ . Λ is the polar lattice of Λ^* .

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

Proof: Let v_1, \dots, v_n be a basis for Λ and let v_1^*, \dots, v_n^* be a basis for Λ^* . If v is in Λ then there exist integers a_1, \dots, a_n such that

$$v = a_1 v_1 + \dots + a_n v_n$$

while if v^* is in Λ^* then there exist integers b_1, \dots, b_n such that

$$v^* = b_1 v_1^* + \dots + b_n v_n^*.$$

In particular

$$v^* \cdot v = \sum_{i=1}^n a_i b_i$$

which is an integer.

Now let w be a vector for which $w \cdot v$ is an integer for all v in Λ . Then there exist integers c_1, \dots, c_n such that $w \cdot v_i = c_i$ for $i = 1, \dots, n$.

Put $v^* = c_1 v_1^* + \dots + c_n v_n^*$ so $v^* \in \Lambda^*$. But then

$$(w - v^*) \cdot v_i = 0 \quad \text{for } i = 1, \dots, n.$$

But v_1, \dots, v_n are linearly independent in \mathbb{R}^n and so $w = v^*$ hence $w \in \Lambda^*$.

By what we have just proved we now see that Λ is the polar lattice of Λ^* . Finally,

$$\det(v_1^*, \dots, v_n^*) \cdot \det(v_1, \dots, v_n) = 1,$$

and so

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

Notice that if $w = (y_1, \dots, y_n)$ is in \mathbb{R}^n the set of $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ for which $x \cdot w = 0$ is given by (x_1, \dots, x_n) for which

$$x_1 y_1 + \dots + x_n y_n = 0$$

and so it determines a hyperplane in \mathbb{R}^n .

Proposition 3: Let Λ be a lattice in \mathbb{R}^n and let u be a vector in \mathbb{R}^n . There exist $n - 1$ linearly independent vectors w_1, \dots, w_{n-1} in Λ with $u \cdot w_i = 0$ for $i = 1, \dots, n - 1$ if and only if $u = t \cdot w^*$ with $t \in \mathbb{R}$ and $w^* \in \Lambda^*$.

Proof: \implies : By Corollary 1 of Theorem 1 there is a basis v_1, \dots, v_n of Λ such that

$$w_i = a_{i1}v_1 + \dots + a_{ii}v_i \quad \text{with} \quad a_{ij} \in \mathbb{Z} \quad \text{and} \quad a_{ii} \neq 0$$

for $i = 1, \dots, n-1$. Since $u \cdot w_i = 0$ for $i = 1, \dots, n$ we see that $u \cdot v_i = 0$ for $i = 1, \dots, n-1$. Put $u \cdot v_n = t$, for some $t \in \mathbb{R}$. Observe that if v_1^*, \dots, v_n^* is a polar basis for Λ^* then $u = tv_n^*$ as required.

\Leftarrow : If $w^* = \mathbf{0}$ then $u = \mathbf{0}$ and so $u \cdot w_i = 0$ for $i = 1, \dots, n-1$. Suppose $w^* \neq (0, \dots, 0)$. Put $w^* = m \cdot v_1^*$ where m is a positive integer and v_1^* is such that $\frac{1}{k} \cdot v_1^*$ is not in Λ^* for any integer k with $k \geq 2$. (v_1^* is said to be primitive for Λ^* .) By Corollary 2 of Theorem 1 we can extend v_1^* to a basis $v_1^*, v_2^*, \dots, v_n^*$ of Λ^* . Let v_1, \dots, v_n be a basis for the polar lattice Λ of Λ^* . Then $v_1^* \cdot v_j = 0$ for $j = 2, \dots, n$ and so $w^* \cdot v_j = 0$ for $j = 2, \dots, n$ as required.

Remark: It follows from the proof of Proposition 3 that if $w^* \in \Lambda^*$ then we can associate to it a lattice $\Lambda(w^*)$ in \mathbb{R}^{n-1} (with basis v_2, \dots, v_n).

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

$$U^n = \{(x_1, \dots, x_n) \in \mathbb{R}^n : 0 \leq x_i < 1 \text{ for } i = 1, \dots, n\}.$$

Let \overline{U}^n denote the closure of U^n . For $\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{R}^n$ we denote

$$|\overline{\mathbf{x}}| = \max_{i=1, \dots, n} |x_i|.$$

This is known as the house of \mathbf{x} . If $\mathbf{x} = (x_1, \dots, x_n) \in \Lambda_0$ then we say that \mathbf{x} is an integer point. For any set T in \mathbb{R}^n and \mathbf{x} in \mathbb{R}^n we define $T + \mathbf{x}$ by

$$T + \mathbf{x} = \{\mathbf{y} + \mathbf{x} : \mathbf{y} \in T\}.$$

Further for any $\lambda \in \mathbb{R}$ we define λT by

$$\lambda T = \{\lambda \mathbf{y} : \mathbf{y} \in T\}.$$

Theorem 4 (Blichfeldt, 1914): Let P be a non-empty set of points in \mathbb{R}^n which is invariant by translation by integer points and has precisely N points in U^n .

Let A be a subset of \mathbb{R}^n of positive Lebesgue measure $\mu(A)$. Then there is an \mathbf{x} in U^n such that $A + \mathbf{x}$ contains at least $N \cdot \mu(A)$ points of P . Further if A is compact then there is an \mathbf{x} in U^n such that $A + \mathbf{x}$ contains more than $N \cdot \mu(A)$ points of P .

Proof: For any set S in \mathbb{R}^n we let $v(S)$ be the number of points of P in S . Let $\mathbf{p}_1, \dots, \mathbf{p}_N$ be the N points of P in U^n . We put

$$P_i = \{\mathbf{p}_i + \mathbf{g} : \mathbf{g} \in \Lambda_0\}$$

for $i = 1, \dots, N$. Since P is invariant by translation by integer points, or equivalently by elements of Λ_0 ,

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

Further we have $P_i \cap P_j = \emptyset$ for $i \neq j$. Now for any $S \subseteq \mathbb{R}^n$ let $v_i(S)$ denote the number of points of P_i in S for $i = 1, \dots, N$.

Let χ be the characteristic function of A^3). Then

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

³⁾is 1 if argument is in A

we have

$$\begin{aligned} \int_{U^n} v_i(A + \mathbf{x}) \, d\mathbf{x} &= \int_{U^n} \sum_{\mathbf{g} \in \Lambda_0} \chi(\mathbf{p}_i + \mathbf{g} - \mathbf{x}) \, d\mathbf{x} \\ &= \int_{\mathbb{R}^n} \chi(\mathbf{z}) \, d\mathbf{z} \\ &= \mu(A) \end{aligned}$$

Thus

$$\int_{U^n} v(A + \mathbf{x}) \, d\mathbf{x} = N\mu(A).$$

Therefore there is some element \mathbf{x} in U^n such that $v(A + \mathbf{x}_1) \geq N\mu(A)$ and so $A + \mathbf{x}_1$ contains at least $N\mu(A)$ points of P .

If A is compact and $N\mu(A)$ is not an integer there is nothing more to prove. Suppose $N\mu(A) = h$ for $h \in \mathbb{Z}^+$. For $k = 1, 2, \dots$ we define A_k by

$$A_k = \left(1 + \frac{1}{k}\right)A.$$

By what we have just proved for each positive integer k there is an \mathbf{x}_k in U^n such that

$$v(A_k + \mathbf{x}_k) \geq h + 1.$$

PMATH 944 Lecture 4: September 24, 2009

Blichfeldt's theorem

It remains to consider the case when A is compact and $N\mu(A)$ is an integer h . For $k = 1, 2, \dots$ we put $A_k = \left(1 + \frac{1}{k}\right)A$. By what we have just proved there is a sequence of points $\mathbf{x}_k \in U^n$, $k = 1, 2, \dots$ for which

$$v(A_k + \mathbf{x}_k) \geq h + 1$$

Since $\mathbf{x}_k \in \overline{U}^n$ and \overline{U}^n is compact there is a subsequence \mathbf{x}_{k_j} , $j = 1, 2, \dots$ which converges to a point \mathbf{x} in \overline{U}^n . Since A is compact the sets $A_k + \mathbf{x}_k$ are uniformly bounded and so contain only finitely many points of P .

Each of the sets $A_{k_j} + \mathbf{x}_{k_j}$ contain at least $h + 1$ points of P and so we may assume by taking a further subsequence that there are $h + 1$ points of P say $\mathbf{u}_1, \dots, \mathbf{u}_{h+1}$ which occur in each set $A_{k_j} + \mathbf{x}_{k_j}$. $A + \mathbf{x}$ is compact and in fact contains $\mathbf{u}_1, \dots, \mathbf{u}_{h+1}$ for if not then $\mathbf{u}_i \notin A + \mathbf{x}$ for some i with $1 \leq i \leq h + 1$. But then \mathbf{u}_i is a positive distance from \mathbf{u}_i to $A + \mathbf{x}$ and this can't be since $\mathbf{x}_{k_j} \rightarrow \mathbf{x}$ and the distance from a point in A_{k_j} to the nearest point in A tends to zero as $k_j \rightarrow \infty$. Thus $A + \mathbf{x}$ contains $h + 1$ of the points of P . We now choose \mathbf{g} so that $\mathbf{x} - \mathbf{g} \in U^n$ and then $A + \mathbf{x} - \mathbf{g}$ then has $h + 1$ points of P as required since P is invariant by translation by integer points.

Let $S \subseteq \mathbb{R}^n$. S is said to be symmetric about the origin (or symmetric) if whenever $\mathbf{x} \in S$ then $-\mathbf{x} \in S$. S is said to be convex if whenever \mathbf{x}, \mathbf{y} are in S and $\lambda \in \mathbb{R}$ with $0 \leq \lambda \leq 1$ then $\lambda\mathbf{x} + (1 - \lambda)\mathbf{y} \in S$. In other words S is convex if whenever \mathbf{x} and \mathbf{y} are in S the line segment joining them is also in S .

Theorem 5 (Minkowski's Convex Body Theorem, 1896) Let A be a convex subset of \mathbb{R}^n which is symmetric about the origin and has 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 the $\mathbf{0}$.

Proof: Notice that $\mu\left(\frac{1}{2}A\right) > 1$ or if A is compact $\mu\left(\frac{1}{2}A\right) \geq 1$. By Blichfeldt's Theorem applied to $\frac{1}{2}A$ where $P = \Lambda_0$, there exists an \mathbf{x} in U^n for which $\frac{1}{2}A + \mathbf{x}$ contains two distinct integer points \mathbf{g}_1 and \mathbf{g}_2 . Notice that $\mathbf{g}_1 - \mathbf{x}$ and $\mathbf{g}_2 - \mathbf{x}$ are in $\frac{1}{2}A$ and so $\mathbf{g}_1 - \mathbf{x} = \frac{1}{2}\mathbf{x}_1$ and $\mathbf{g}_2 - \mathbf{x} = \frac{1}{2}\mathbf{x}_2$ for $\mathbf{x}_1, \mathbf{x}_2 \in A$. By symmetry, $-(\mathbf{g}_2 - \mathbf{x}) = \mathbf{x} - \mathbf{g}_2 = \frac{1}{2}(-\mathbf{x}_2)$ with $-\mathbf{x}_2 \in A$. Since A is convex $\frac{1}{2}\mathbf{x}_1 + \frac{1}{2}(-\mathbf{x}_2)$ is in A thus

$$\mathbf{g}_1 - \mathbf{x} + \mathbf{x} - \mathbf{g}_2 = \mathbf{g}_1 - \mathbf{g}_2 \in A.$$

But $\mathbf{g}_1 - \mathbf{g}_2 \in \Lambda_0$ and since \mathbf{g}_1 and \mathbf{g}_2 are distinct $\mathbf{g}_1 - \mathbf{g}_2 \neq \mathbf{0}$.

Remark: Note that Minkowski's Convex Body Theorem is best possible in the sense that the conclusion does not hold with 2^n replaced by a smaller number as the example

$$A = \{ (t_1, \dots, t_n) \in \mathbb{R}^n : |t_i| < 1, i = 1, \dots, n \}.$$

One can also check that the hypothesis of symmetry and convexity can't be omitted.

Theorem 6 (Minkowski's Linear Forms Theorem): Let $B = (B_{ij})$ be an $n \times n$ matrix with real entries and non-zero determinant. Let c_1, \dots, c_n be positive real numbers with $c_1 \cdots c_n \geq |\det B|$. Then there exists an integer point $\mathbf{x} = (x_1, \dots, x_n)$ different from $\mathbf{0}$ for which

$$|B_{i,1}x_1 + \cdots + B_{i,n}x_n| < c_i \quad \text{for } i = 1, \dots, n-1$$

and

$$|B_{n,1}x_1 + \cdots + B_{n,n}x_n| \leq c_n.$$

Proof: Let $L_1(\mathbf{x}), \dots, L_n(\mathbf{x})$ be linear forms given by

$$L_i(\mathbf{x}) = B_{i,1}x_1 + \cdots + B_{i,n}x_n \quad \text{for } i = 1, \dots, n.$$

Next put,

$$L'_i(\mathbf{x}) = \frac{1}{c_i}L_i(\mathbf{x}) \quad \text{for } i = 1, \dots, n.$$

Then we wish to solve the system

$$|L'_i(\mathbf{x})| < 1 \quad \text{for } i = 1, \dots, n-1$$

and

$$|L'_n(\mathbf{x})| \leq 1.$$

The absolute value of the determinant of the matrix determined by the coefficients of L'_1, \dots, L'_n is at most 1. Thus we may assume that $c_1 = \cdots = c_n = 1$ and $0 < |\det B| \leq 1$.

Let A be the set of $\mathbf{x} \in \mathbb{R}^n$ for which

$$|L_i(\mathbf{x})| \leq 1 \quad \text{for } i = 1, \dots, n.$$

Certainly A is symmetric about the origin. Also A is convex since if \mathbf{x} and \mathbf{y} are in A then for any λ with $0 \leq \lambda \leq 1$

$$\begin{aligned} |L_i(\lambda\mathbf{x} + (1-\lambda)\mathbf{y})| &= |\lambda(B_{i,1}x_1 + \cdots + B_{i,n}x_n) + (1-\lambda)(B_{i,1}y_1 + \cdots + B_{i,n}y_n)| \\ &\leq \lambda|B_{i,1}x_1 + \cdots + B_{i,n}x_n| + (1-\lambda)|B_{i,1}y_1 + \cdots + B_{i,n}y_n| \\ &\leq \lambda + 1 - \lambda = 1 \end{aligned}$$

Further we remark that

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

where $B = (B_{ij})$ and $\tilde{U}^n = \{ (t_1, \dots, t_n) \in \mathbb{R}^n : |t_i| \leq 1 \}$. Therefore $\mu(A) \geq 2^n$. By Minkowski's Convex Body Theorem there is an integer point \mathbf{x} with $\mathbf{x} \neq \mathbf{0}$ 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

$$|L_i(\mathbf{x})| < 1 \quad \text{for } i = 1, \dots, n-1$$

and

$$|L_n(\mathbf{x})| < 1 + \epsilon.$$

Then $\mu(A_\epsilon) \geq (1 + \epsilon)2^n > 2^n$ and so we may apply Minkowski's Convex Body Theorem to find an integer point \mathbf{x}_ϵ in A_ϵ with $\mathbf{x}_\epsilon \neq \mathbf{0}$. Now take any sequence ϵ_k of positive reals which decreases to 0. Associated to

it we get a sequence \mathbf{x}_{ϵ_k} of integer points different from $\mathbf{0}$. Since $\bigcup_{k=1}^{\infty} A_{\epsilon_k}$ is bounded there exists an integer point \mathbf{x} in infinitely many of the sets A_{ϵ} hence \mathbf{x} satisfies

$$|L_i(\mathbf{x})| < 1 \quad \text{for } i = 1, \dots, n-1$$

and

$$|L_n(\mathbf{x})| \leq 1.$$

PMATH 944 Lecture 5: September 29, 2009

Theorem 7: Let α_{ij} be real numbers, with $1 \leq i \leq n$, $1 \leq j \leq m$, and let Q be a real number with $Q > 1$. Then there exist integers q_1, \dots, q_m and p_1, \dots, p_n with

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

and

$$|\alpha_{i1}q_1 + \dots + \alpha_{im}q_m - p_i| \leq \frac{1}{Q} \quad \text{for } i = 1, \dots, n.$$

(This was proved in 1842 by Dirichlet under the assumption that Q is an integer.)

We have the following consequence of Theorem 7:

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

$$|\alpha_{i1}q_1 + \dots + \alpha_{im}q_m - p_i| < \frac{1}{q^{m/n}} \quad \text{for } i = 1, \dots, n. \quad (2)$$

Proof: Take $Q = 2$. By Theorem 7 there exists a solution $q_1, \dots, q_m, p_1, \dots, p_n$ of (2). We now divide through by the gcd of $q_1, \dots, q_m, p_1, \dots, p_n$ to give us a solution of (2) with a coprime $m+n$ -tuple. Thus we may suppose, without loss of generality, $\gcd(q_1, \dots, q_m, p_1, \dots, p_n) = 1$. Let

$$|q_1\alpha_{t1} + \dots + q_m\alpha_{tm}| = \delta_t$$

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

We now apply Theorem 7 with Q so that $\frac{1}{Q} < \delta_t$ to get a new $m+n$ -tuple satisfying (2). We remove the gcd to make the $m+n$ -tuple coprime. Repeating this process gives us infinitely many coprime $m+n$ -tuples satisfying (2).

Proof of Theorem 7: Put $l = m+n$ and consider the l linear forms L_1, \dots, L_l in $\mathbf{x} = (x_1, \dots, x_l)$ given by

$$L_i(\mathbf{x}) = x_i \quad \text{for } i = 1, \dots, m$$

and

$$L_{m+j}(\mathbf{x}) = \alpha_{j1}x_1 + \dots + \alpha_{jm}x_m - x_{m+j} \quad \text{for } j = 1, \dots, n.$$

Note that the determinant of the matrix associated with L_1, \dots, L_l is $(-1)^n$.

Let $Q > 1$ and apply Minkowski's Linear Forms Theorem to the system of inequalities:

$$|L_i(\mathbf{x})| < Q^{n/m} \quad \text{for } i = 1, \dots, m \quad (3)$$

and

$$|L_{m+j}(\mathbf{x})| \leq \frac{1}{Q} \quad \text{for } j = 1, \dots, n \quad (4)$$

to find a non-zero integer point \mathbf{x} satisfying (3) and (4). We now put $q_i = x_i$ for $i = 1, \dots, m$ and $p_j = x_{m+j}$ for $j = 1, \dots, n$. Then

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

and

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

It remains to check that $q \neq 0$. Suppose otherwise. Then $q_1 = \cdots = q_m = 0$ so

$$|p_j| \leq \frac{1}{Q} \quad \text{for } j = 1, \dots, n.$$

But $Q > 1$ so $p_1 = \cdots = p_n = 0$ and this contradicts the fact that \mathbf{x} is a non-zero point. The result follows.

Theorem 8: 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 has volume $\geq 2^n d(\Lambda)$. Then A contains a point of Λ different from $\mathbf{0}$.

Proof: Suppose v_1, \dots, v_n is a basis for Λ . Let $v_j = (\alpha_{j1}, \dots, \alpha_{jn})$ for $j = 1, \dots, n$. Let T be the linear transformation from \mathbb{R}^n to \mathbb{R}^n associated with the matrix (α_{ij}) . Then $\Lambda = T\Lambda_0$. Notice that $\mu(T^{-1}A) = d(\Lambda)^{-1}\mu(A)$ and that $T^{-1}A$ is a convex set which is symmetric about the origin. The result now follows from Minkowski's Convex Body Theorem.

Proposition 9: Let R be a positive real number and let n be a positive integer. The volume of the sphere of radius R in \mathbb{R}^n is $\omega_n R^n$ where $\omega_n = \frac{\pi^{n/2}}{\Gamma(1+n/2)}$.

Proof: It suffices to prove that ω_n is the volume of the unit sphere given by

$$\{(x_1, \dots, x_n) \in \mathbb{R}^n : x_1^2 + \cdots + x_n^2 \leq 1\}.$$

We have $\omega_1 = 2$ and $\omega_2 = \pi$. We now proceed inductively. Suppose $n \geq 3$. Then

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

where g is the characteristic function of the unit sphere.⁴⁾ Thus

$$\begin{aligned} \omega_n &= \int_{x_{n-1}^2 + x_n^2 \leq 1} \omega_{n-2} (1 - x_{n-1}^2 - x_n^2)^{(n-2)/2} dx_{n-1} dx_n \\ &= \omega_{n-2} \int_{x_{n-1}^2 + x_n^2 \leq 1} (1 - x_{n-1}^2 - x_n^2)^{(n-2)/2} dx_{n-1} dx_n \end{aligned}$$

Change to polar coordinates (r, θ) . Thus

$$\begin{aligned} \omega_n &= \omega_{n-2} \int_0^{2\pi} \int_0^1 (1 - r^2)^{(n-2)/2} r dr d\theta \\ &= 2\pi \omega_{n-2} \int_0^1 (1 - r^2)^{(n-2)/2} r dr \\ &= 2\pi \omega_{n-2} \left[-\frac{1}{n} (1 - r^2)^{n/2} \right]_0^1 \\ &= \frac{2\pi}{n} \omega_{n-2} \end{aligned}$$

Thus

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

while

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

⁴⁾This not necessary; can ignore.

The result follows on noting that $\Gamma(x+1) = x\Gamma(x)$ for $x > 0$ and that $\Gamma(\frac{1}{2}) = \sqrt{\pi}$.

Theorem 10: Let Λ be a lattice in \mathbb{R}^n . There is a non-zero element $\mathbf{x} \in \Lambda$ for which

$$0 < \mathbf{x} \cdot \mathbf{x} = x_1^2 + \cdots + x_n^2 \leq 4(\omega_n^{-1}d(\Lambda))^{2/n}.$$

Proof: We apply Theorem 8 to the set

$$A = \{ \mathbf{x} \in \mathbb{R}^n : x_1^2 + \cdots + x_n^2 \leq t \}$$

with $t = 4(\omega_n^{-1}d(\Lambda))^{2/n}$. Then

$$\begin{aligned} \mu(A) &= \omega_n t^{n/2} = \omega_n 2^n \omega_n^{-1} d(\Lambda) \\ &= 2^n d(\Lambda) \end{aligned}$$

A is convex, symmetric about the origin and compact and the result now follows from Theorem 8.

PMATH 944 Lecture 6: October 1, 2009

Theorem 10 is close to the truth since Minkowski constructed for each $n \in \mathbb{Z}^+$ a lattice Λ in \mathbb{R}^n for which

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

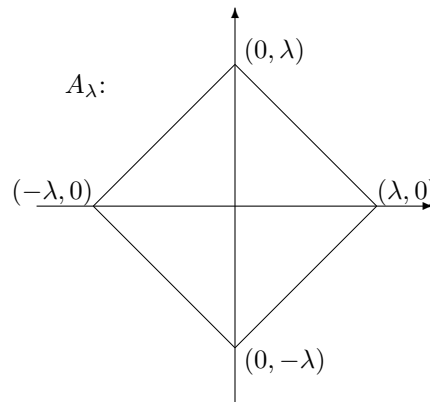
In particular Theorem 10 can't be improved by more than a factor of 4. Rogers was able to improve on Theorem 10 somewhat. He replaced $4\omega_n^{-2/n}$ in Theorem 10 by $4\left(\frac{\sigma_n}{\omega_n}\right)^{2/n}$ where σ_n is the quotient of two geometrical figures with the property that $\sigma_n \sim \frac{n}{e^{2n/2}}$ as $n \rightarrow \infty$. We have

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

How about other convex bodies of interest? For each $\lambda \in \mathbb{R}$ with $\lambda > 0$ define

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

Thus in \mathbb{R}^2 ,



Define for $n \in \mathbb{Z}^+$

$$A_\lambda^{(n)+} = A_\lambda^+ = \{ (x_1, \dots, x_n) \in \mathbb{R}^n : \lambda \geq x_i \geq 0 \text{ for } i = 1, \dots, n \}$$

The volume of A_λ is $2^n \lambda^n$ times the volume of A_1^+ .

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

⁵⁾Also require $x_1 + \cdots + x_n \leq \lambda$ (correction from next class).

Notice that

$$\int_0^u \frac{(u-x)^n}{n!} dx = \left[-\frac{(u-x)^{n+1}}{(n+1)!} \right]_0^u = \frac{u^{n+1}}{(n+1)!}.$$

Therefore

$$\mu(A_1^+) = \frac{1}{n!}$$

so

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

Further observe that $A_\lambda^{(n)}$ is symmetric about the origin. Furthermore it is convex since if γ is a real number with $0 \leq \gamma \leq 1$ and $\mathbf{x} = (x_1, \dots, x_n)$, $\mathbf{y} = (y_1, \dots, y_n) \in A_\lambda$ then

$$\gamma \mathbf{x} + (1-\gamma) \mathbf{y} = (\gamma x_1 + (1-\gamma)y_1, \dots, \gamma x_n + (1-\gamma)y_n) \in A_\lambda$$

since

$$\begin{aligned} |\gamma x_1 + (1-\gamma)y_1| + \dots + |\gamma x_n + (1-\gamma)y_n| &\leq \gamma(|x_1| + \dots + |x_n|) + (1-\gamma)(|y_1| + \dots + |y_n|) \\ &\leq \gamma \lambda + (1-\gamma) \lambda = \lambda. \end{aligned}$$

Theorem 11: Let Λ be a lattice in \mathbb{R}^n . Then there is a non-zero point $\mathbf{x} = (x_1, \dots, x_n)$ in Λ with

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

Proof: We apply Theorem 8 to the set $A_\lambda^{(n)}$ where $\lambda = (n! d(\Lambda))^{1/n}$. Then the volume of $A_\lambda^{(n)}$ is $2^n d(\Lambda)$. The set is convex, symmetric about $\mathbf{0}$ and compact and so the result follows.

We may apply Theorem 8 to sets which contain sets which are convex, symmetric, and of large enough volume. In this connection we introduce for each $n \in \mathbb{Z}^+$ and $\lambda \in \mathbb{R}$, $\lambda \geq 0$,

$$B_\lambda^{(n)} = \{ (x_1, \dots, x_n) \in \mathbb{R}^n : |x_1 \cdots x_n| \leq \lambda^n \}.$$

$B_\lambda^{(n)}$ is not convex. However we can appeal to the arithmetic-geometric mean inequality: Given non-negative real numbers x_1, \dots, x_n we have

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

Thus $B_\lambda^{(n)}$ contains $A_{n\lambda}^{(n)}$ and $A_{n\lambda}^{(n)}$ is convex.

Theorem 12: Let $C = (c_{ij})$ be a non-singular $n \times n$ matrix with entries from \mathbb{R} and put

$$L_i(\mathbf{x}) = c_{i1}x_1 + \dots + c_{in}x_n \quad \text{for } i = 1, \dots, n.$$

Then there exists an integer point \mathbf{x} different from $\mathbf{0}$ for which

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

Proof: We apply Theorem 11 with the lattice Λ determined by the row vectors of C and the region $B_\lambda^{(n)}$ where $\lambda = \frac{(n! \det(C))^{1/n}}{n}$. Since $B_\lambda^{(n)}$ contains $A_{n\lambda}^{(n)}$ the result follows.

Let A_1 be a sublattice of a lattice Λ in \mathbb{R}^n . We can put an equivalence relation $\sim (\sim_{A_1})$ on Λ by the rule $\mathbf{x}_1 \sim \mathbf{x}_2$ if and only if $\mathbf{x}_1 - \mathbf{x}_2 \in A_1$. \sim is an equivalence relation on Λ and it partitions Λ into a finite set of equivalence classes.

Proposition 13: Let A_1 be a sublattice of a lattice Λ in \mathbb{R}^n . The index of A_1 in Λ is the number of equivalence classes of Λ under \sim_{A_1} .

Proof: By Theorem 1 we can find bases v_1, \dots, v_n for Λ and w_1, \dots, w_n for A_1 of the form given in Theorem 1. Then the index is $\prod_{i=1}^n a_{ii}$. We claim that every vector \mathbf{u} in Λ is equivalent to precisely one of $q_1 \mathbf{v}_1 + \dots + q_n \mathbf{v}_n$ with $0 \leq q_i < a_{ii}$ for $i = 1, \dots, n$. This will prove the result.

Let $\mathbf{u} = u_1\mathbf{v}_1 + \cdots + u_n\mathbf{v}_n \in \Lambda$. First we shift \mathbf{u} by a multiple \mathbf{w}_n to find an equivalent vector with n th coordinate in the range $0 \leq q_n < a_{n,n}$. Next we subtract a multiple \mathbf{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 in this way we see that \mathbf{u} is equivalent to a vector of the form $q_1\mathbf{v}_1 + \cdots + q_n\mathbf{v}_n$ with $0 \leq q_i < a_{ii}$ for $i = 1, \dots, n$. It remains to show that no two vectors of the form $q_1\mathbf{v}_1 + \cdots + q_n\mathbf{v}_n$ with $0 \leq q_i < a_{ii}$ for $i = 1, \dots, n$ are equivalent under \sim .

PMATH 944 Lecture 7: October 6, 2009

Corrections: Addition of absolute values to $|\det(C)|$ in Theorem 12.

$$(A_\lambda^{(n)})^+ = \{ (x_1, \dots, x_n) \in \mathbb{R}^n : x_i \geq 0, \quad i = 1, \dots, n, \quad x_1 + \cdots + x_n \leq \lambda \}$$

Proposition 13: ...

Every vector in Λ is equivalent to a vector of the form $q_1\mathbf{v}_1 + \cdots + q_n\mathbf{v}_n$ with $0 \leq q_i < a_{ii}$ for $i = 1, \dots, n$.

Finally we should show that all vectors of the above form are inequivalent. So suppose two are equivalent, their difference $r_1\mathbf{v}_1 + \cdots + r_n\mathbf{v}_n$ is in Λ_1 with $|r_i| < a_{ii}$ for $i = 1, \dots, n$. Let j be the largest integer for which $r_j \neq 0$. Then we replace w_j in the basis w_1, \dots, w_n of Λ_1 by w_j minus a multiple of $r_1\mathbf{v}_1 + \cdots + r_n\mathbf{v}_n$ so that the resulting basis is in lower triangular form but with a_{jj} replaced by a smaller non-negative integer. The final reduction (to Hermite normal form) doesn't change the diagonal. But the resulting determinant is different which gives a contradiction. Therefore the index is $\prod_{i=1}^n a_{ii}$.

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

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

Then

$$0 < \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n < \infty.$$

Minkowski, in what is known as Minkowski's Second Theorem on Convex Bodies proved that

$$\frac{2^n d(A)}{n!} \leq \lambda_1 \cdots \lambda_n \mu(A) \leq 2^n d(A).^{6)}$$

We won't give a proof: the upper bound is tricky.

The result is sharp in the sense that neither the upper bound or the lower bound can be improved in general.

Take any positive real numbers $\gamma_1, \dots, \gamma_n$ with $0 < \gamma_1 \leq \gamma_2 \leq \cdots \leq \gamma_n < \infty$. Consider the lattice generated by

$$(\gamma_1, 0, \dots, 0), (0, \gamma_2, 0, \dots, 0), \dots, (0, \dots, 0, \gamma_n).$$

Let A be the cube $A = \{ (x_1, \dots, x_n) \in \mathbb{R}^n : |x_i| \leq 1, \quad i = 1, \dots, n \}$. "Plainly" $\lambda_i(A, \Lambda) = \lambda_i = \gamma_i$ for $i = 1, \dots, n$. Further $d(A) = \gamma_1 \cdots \gamma_n$. Thus

$$\lambda_1 \cdots \lambda_n \mu(A) = \gamma_1 \cdots \gamma_n 2^n = 2^n d(A),$$

so the upper bound is sharp.

If we now take $A = A_1^{(n)} = \{ \mathbf{x} \in \mathbb{R}^n : |x_1| + \cdots + |x_n| \leq 1 \}$ then

$$\lambda_i(A, \Lambda) = \lambda_i = \gamma_i \text{ for } i = 1, \dots, n \text{ as before.}$$

We have

$$\lambda_1 \cdots \lambda_n \mu(A) = \gamma_1 \cdots \gamma_n \frac{2^n}{n!} = \frac{2^n}{n!} d(A)$$

and so the lower bound is sharp.

Sometimes it is useful to have another characterization of a lattice.

⁶⁾upper bound implies $\lambda_1^n \mu(A) \leq 2^n d(A) \implies$ Minkowski's Convex Body Theorem

Theorem 14: A subset Λ of \mathbb{R}^n is a lattice in \mathbb{R}^n if and only if

- i) If \mathbf{a}, \mathbf{b} are in Λ the $\mathbf{a} + \mathbf{b}$ and $\mathbf{a} - \mathbf{b}$ are in Λ .
- ii) Λ contains n linearly independent points $\mathbf{a}_1, \dots, \mathbf{a}_n$.
- iii) Λ is a discrete set, in other words it has no limit points.

Proof: (\implies) Follows immediately from the definition of a lattice.

(\impliedby) We prove this by induction on n . For $n = 1$ we note by ii) that Λ contains a non-zero point a . By i) Λ contains 0 and $-a$. Further since Λ is discrete there is a smallest positive real number a in Λ . Then by i)

$$\Lambda = \{ga : g \in \mathbb{Z}\}$$

as required.

Suppose the result holds for dimension $n - 1$ with $n \geq 2$. We may choose our coordinate system in \mathbb{R}^n so that $n - 1$ linearly independent points of Λ lie in a subspace of the form $\mathbb{R}^{n-1} \times \{0\}$ so $x_n = 0$. 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 $\mathbf{b}_1, \dots, \mathbf{b}_{n-1}$ be a basis for Λ' . Then Λ contains a point of the form $\mathbf{b}_n = (b_{1n}, \dots, b_{nn})$ with $b_{nn} > 0$. In fact there is a point \mathbf{b}_n of this form with b_{nn} minimal. Suppose otherwise. Then we can find a sequence $\mathbf{b}_n^{(j)} = (b_{1n}^{(j)}, \dots, b_{nn}^{(j)})$ in Λ with $b_{nn}^{(j)} > 0$ and

$$b_{nn}^{(j)} \rightarrow 0 \text{ as } j \rightarrow \infty.$$

But we can translate $\mathbf{b}_n^{(j)}$ by some linear combination of $\mathbf{b}_1, \dots, \mathbf{b}_{n-1}$ so that $(b_{1n}^{(j)}, \dots, b_{n-1,n}^{(j)}, 0)$ are in the compact set

$$\{\lambda_1 \mathbf{b}_1 + \dots + \lambda_{n-1} \mathbf{b}_{n-1} : |\lambda_i| \leq 1\}$$

so thus the $\mathbf{b}_n^{(j)}$ s are all in a compact set and so have a limit point contradicting the fact that Λ is discrete. We now claim that every element of Λ is an integer linear combination of $\mathbf{b}_1, \dots, \mathbf{b}_n$. Let $\mathbf{d} \in \Lambda$ with $\mathbf{d} = (d_1, \dots, d_n)$. Then

$$\mathbf{d}' = \mathbf{d} - \left\lfloor \frac{d_n}{b_{nn}} \right\rfloor \mathbf{b}_n \in \Lambda.$$

The n th coordinate of \mathbf{d}' is non-negative and smaller than b_{nn} . Therefore it is 0. Thus $\mathbf{d}' \in \Lambda'$ and so is an integer linear combination of $\mathbf{b}_1, \dots, \mathbf{b}_{n-1}$. Therefore \mathbf{d} is an integer linear combination of $\mathbf{b}_1, \dots, \mathbf{b}_n$. Thus Λ is a lattice basis $\mathbf{b}_1, \dots, \mathbf{b}_n$ and the result follows.

Proposition 15: Let n, m, k_1, \dots, k_m be positive integers and let $a_{ij}, i = 1, \dots, m, j = 1, \dots, n$ be integers. The set Λ of points $\mathbf{u} = (u_1, \dots, u_n)$ in \mathbb{R}^n with integral coordinates satisfying

$$\sum_{j=1}^n a_{ij} u_j \equiv 0 \pmod{k_i} \text{ for } i = 1, \dots, m$$

is a lattice in \mathbb{R}^n with $d(\Lambda) \leq k_1 \cdots k_m$.

PMATH 944 Lecture 8: October 8, 2009

Proposition 15: Let n, m, k_1, \dots, k_m be positive integers and let a_{ij} ($1 \leq i \leq m, 1 \leq j \leq n$) be integers. The set Λ of points $\mathbf{u} = (u_1, \dots, u_n)$ with integer coordinates satisfying

$$\sum_{j=1}^n a_{ij} u_j \equiv 0 \pmod{k_i} \text{ for } i = 1, \dots, m$$

is a lattice with determinant $d(\Lambda) \leq k_1 \cdots k_m$.

Proof: First we remark that Λ is a subset of Λ_0 and so is discrete. Next we observe that

$$(k_1 \cdots k_m, 0, \dots, 0), (0, k_1 \cdots k_m, 0, \dots, 0), \dots, (0, \dots, 0, k_1 \cdots k_m)$$

are n linearly independent points in Λ . Finally we have that if $\mathbf{u} = (u_1, \dots, u_n)$ and $\mathbf{v} = (v_1, \dots, v_n)$ are in Λ then $\mathbf{u} + \mathbf{v}$ are in Λ since

$$\sum_{j=1}^n a_{ij}(u_j \pm v_j) \equiv \left(\sum_{j=1}^n a_{ij}u_j \right) \pm \left(\sum_{j=1}^n a_{ij}v_j \right) \equiv 0 \pm 0 \equiv 0 \pmod{k_i} \quad \text{for } i = 1, \dots, m.$$

Thus by Theorem 14, Λ is a lattice in \mathbb{R}^n and so is a sublattice of Λ_0 .

Let I denote the index of Λ in Λ_0 . Then $I = \frac{d(\Lambda)}{d(\Lambda_0)}$. But $d(\Lambda_0) = 1$ and so $I = d(\Lambda)$. It remains to estimate the index of Λ in Λ_0 . By Proposition 13 this is the number of equivalence classes of Λ_0 under \sim_Λ . Notice that $\mathbf{u}, \mathbf{v} \in \Lambda_0$ are equivalent if $\mathbf{u} - \mathbf{v} \in \Lambda$ hence, with $\mathbf{u} = (u_1, \dots, u_n)$ and $\mathbf{v} = (v_1, \dots, v_n)$, if

$$\sum_{j=1}^n a_{ij}(u_j - v_j) \equiv 0 \pmod{k_i} \quad \text{for } i = 1, \dots, m.$$

Thus $I = d(\Lambda) \leq k_1 \cdots k_m$.

Theorem 16: (Lagrange's Theorem). Every positive integer can be expressed as the sum of four squares of integers.

Proof: We may restrict our attention, without loss of generality, to integers m with $m > 1$ which are squarefree. Let $m = p_1 \cdots p_r$ with p_1, \dots, p_r distinct primes.

We now remark that for every prime p there exist integers a_p and b_p for which

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

If $p = 2$ we take $a_p = 1, b_p = 0$. If p is odd then the integers a^2 with $0 \leq a < \frac{1}{2}p$ are distinct mod p . (Consider $a_1^2 - a_2^2 = (a_1 - a_2)(a_1 + a_2) \pmod{p}$.) Similarly the integers $-1 - b^2$ with $0 \leq b < \frac{1}{2}p$ are distinct mod p . Therefore, since $\frac{1}{2}(p+1) + \frac{1}{2}(p+1) > p$ there must exist integers a_p and b_p with $a_p^2 \equiv -1 - b_p^2 \pmod{p}$ as required.

We define the lattice Λ in \mathbb{R}^4 as the set of points (u_1, u_2, u_3, u_4) with integer coordinates satisfying

$$\begin{aligned} & u_1 \equiv a_{p_i}u_3 + b_{p_i}u_4 \pmod{p_i} \\ \text{and} & & & \text{for } i = 1, \dots, r \\ & u_2 \equiv b_{p_i}u_3 - a_{p_i}u_4 \pmod{p_i}. \end{aligned}$$

Further $d(\Lambda) \leq (p_1 \cdots p_r)^2 = m^2$.

Let $A = \{ (x_1, x_2, x_3, x_4) \in \mathbb{R}^4 : x_1^2 + x_2^2 + x_3^2 + x_4^2 < 2m \}$. A is the sphere of radius $\sqrt{2m}$ in \mathbb{R}^4 . Thus it is a convex set which is symmetric about the origin and it has volume $\frac{\pi^2}{2}(\sqrt{2m})^4 = 2\pi^2 m^2$. Since $2\pi^2 m^2 > 2^4 m^2 \geq 2^4 d(\Lambda)$ there is a non-zero point (u_1, u_2, u_3, u_4) of Λ in A by Theorem 8. In particular

$$0 < u_1^2 + u_2^2 + u_3^2 + u_4^2 < 2m. \tag{5}$$

But

$$\begin{aligned} u_1^2 + u_2^2 + u_3^2 + u_4^2 &\equiv (a_{p_i}u_3 + b_{p_i}u_4)^2 + (b_{p_i}u_3 - a_{p_i}u_4)^2 + u_3^2 + u_4^2 \pmod{p_i} \\ &\equiv (a_{p_i}^2 + b_{p_i}^2 + 1)u_3^2 + (a_{p_i}^2 + b_{p_i}^2 + 1)u_4^2 \pmod{p_i} \\ &\equiv (a_{p_i}^2 + b_{p_i}^2 + 1)(u_3^2 + u_4^2) \pmod{p_i} \\ &\equiv 0 \pmod{p_i} \quad \text{for } i = 1, \dots, r \end{aligned}$$

By the Chinese Remainder Theorem

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

By (5), $u_1^2 + u_2^2 + u_3^2 + u_4^2 = m$ as required.

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 respect to the usual Euclidean distance, is a difficult problem, it is NP-hard as shown by Ajtai. However, if we look for only a “short” vector in the lattice we can do so efficiently. The algorithm we use is the L^3 -algorithm. Here L^3 stands for Lenstra, Lenstra, and Lovász.

Let $\mathbf{b}_1, \dots, \mathbf{b}_n$ be a basis for a lattice Λ in \mathbb{R}^n . Let (\cdot, \cdot) denote the usual inner product in \mathbb{R}^n . The Gram–Schmidt orthogonalization produces orthogonal vectors $\tilde{\mathbf{b}}_1, \dots, \tilde{\mathbf{b}}_n$ and real numbers μ_{ij} with $(1 \leq j < i \leq n)$ inductively by

$$\tilde{\mathbf{b}}_i = \mathbf{b}_i - \sum_{j=1}^{i-1} \mu_{ij} \tilde{\mathbf{b}}_j \quad \text{and} \quad \mu_{ij} = \frac{(\mathbf{b}_i, \tilde{\mathbf{b}}_j)}{(\tilde{\mathbf{b}}_j, \tilde{\mathbf{b}}_j)}.$$

Note that $\tilde{\mathbf{b}}_i$ is the projection of \mathbf{b}_i on the orthogonal complement of $\text{Sp}\{\tilde{\mathbf{b}}_1, \dots, \tilde{\mathbf{b}}_{i-1}\}$. Further $\text{Sp}\{\mathbf{b}_1, \dots, \mathbf{b}_i\} = \text{Sp}\{\tilde{\mathbf{b}}_1, \dots, \tilde{\mathbf{b}}_i\}$ for $i = 1, \dots, n$.

PMATH 944 Lecture 9: October 13, 2009

Let $\mathbf{b}_1, \dots, \mathbf{b}_n$ be linearly independent vectors in \mathbb{R}^n . Apply Gram–Schmidt to get

$$\tilde{\mathbf{b}}_1, \dots, \tilde{\mathbf{b}}_n \quad \text{orthogonal linearly independent vectors in } \mathbb{R}^n.$$

$\tilde{\mathbf{b}}_i$ is the projection of \mathbf{b}_i on the orthogonal complement of $\text{Sp}\{\mathbf{b}_1, \dots, \mathbf{b}_{i-1}\}$. Further

$$\text{Sp}\{\mathbf{b}_1, \dots, \mathbf{b}_i\} = \text{Sp}\{\tilde{\mathbf{b}}_1, \dots, \tilde{\mathbf{b}}_i\} \quad \text{for } i = 1, \dots, n.$$

Definition: A basis $\mathbf{b}_1, \dots, \mathbf{b}_n$ for a lattice Λ in \mathbb{R}^n is said to be reduced if

- i) $|\mu_{ij}| \leq \frac{1}{2}$ for $1 \leq j < i \leq n$
- ii) $|\tilde{\mathbf{b}}_i + \mu_{i,i-1} \tilde{\mathbf{b}}_{i-1}|^2 \geq \frac{3}{4} |\tilde{\mathbf{b}}_{i-1}|^2$ for $2 \leq i \leq n$.

Here $|\mathbf{x}|$ is the Euclidean length of \mathbf{x} , so $|\mathbf{x}|^2 = \mathbf{x} \cdot \mathbf{x}$.

Remarks

1. The vectors $\tilde{\mathbf{b}}_i + \mu_{i,i-1} \tilde{\mathbf{b}}_{i-1}$ and $\tilde{\mathbf{b}}_{i-1}$ are the projections of \mathbf{b}_i and \mathbf{b}_{i-1} respectively on the orthogonal complement of $\text{Sp}\{\mathbf{b}_1, \dots, \mathbf{b}_{i-2}\}$.
2. The constant $\frac{3}{4}$ is somewhat arbitrary, it could have been replaced by y for any y with $\frac{1}{4} < y < 1$.

Objective:

1. Describe properties of a reduced basis for a lattice Λ .
2. Give an algorithm (the L^3 -algorithm) for efficiently transforming a basis to a reduced basis.

Proposition 17: Let $\mathbf{b}_1, \dots, \mathbf{b}_n$ be a reduced basis for a lattice Λ in \mathbb{R}^n and let $\tilde{\mathbf{b}}_1, \dots, \tilde{\mathbf{b}}_n$ be the vectors obtained by applying the Gram–Schmidt process. Then

- i) $|\mathbf{b}_j|^2 \leq 2^{i-1} |\tilde{\mathbf{b}}_i|^2$ for $1 \leq j \leq i \leq n$
- ii) $d(\Lambda) \leq |\mathbf{b}_1| \cdots |\mathbf{b}_n| \leq 2^{n(n-1)/4} d(\Lambda)$
- iii) $|\mathbf{b}_1| \leq 2^{(n-1)/4} d(\Lambda)^{1/n}$

Proof: By the definition of a reduced basis

$$|\tilde{\mathbf{b}}_i + \mu_{i,i-1} \tilde{\mathbf{b}}_{i-1}|^2 \geq \frac{3}{4} |\tilde{\mathbf{b}}_{i-1}|^2 \quad \text{with } |\mu_{i,i-1}| \leq \frac{1}{2}.$$

Thus

$$\begin{aligned} |\tilde{\mathbf{b}}_i + \mu_{i,i-1} \tilde{\mathbf{b}}_{i-1}|^2 &= (\tilde{\mathbf{b}}_i + \mu_{i,i-1} \tilde{\mathbf{b}}_{i-1}, \tilde{\mathbf{b}}_i + \mu_{i,i-1} \tilde{\mathbf{b}}_{i-1}) \\ &= |\tilde{\mathbf{b}}_i|^2 + \mu_{i,i-1}^2 |\tilde{\mathbf{b}}_{i-1}|^2, \quad \text{for } i = 2, \dots, n. \end{aligned}$$

Thus

$$\begin{aligned} |\tilde{\mathbf{b}}_i|^2 &= |\tilde{\mathbf{b}}_i + \mu_{i,i-1}\tilde{\mathbf{b}}_{i-1}|^2 - \mu_{i,i-1}^2|\tilde{\mathbf{b}}_{i-1}|^2 \\ &\geq \left(\frac{3}{4} - \mu_{i,i-1}^2\right)|\tilde{\mathbf{b}}_{i-1}|^2 \\ &\geq \frac{1}{2}|\tilde{\mathbf{b}}_{i-1}|^2 \end{aligned}$$

or equivalently $|\tilde{\mathbf{b}}_{i-1}|^2 \leq 2|\tilde{\mathbf{b}}_i|^2$.

Thus, by induction,

$$|\tilde{\mathbf{b}}_j|^2 \leq 2^{i-j}|\tilde{\mathbf{b}}_i|^2 \quad \text{for } 1 \leq j \leq i \leq n. \quad (6)$$

Now

$$\begin{aligned} |\mathbf{b}_i|^2 &= |\tilde{\mathbf{b}}_i|^2 + \sum_{j=1}^{i-1} \mu_{ij}^2 |\tilde{\mathbf{b}}_j|^2 \\ &\leq |\tilde{\mathbf{b}}_i|^2 \left(1 + \sum_{j=1}^{i-1} \frac{1}{4} 2^{i-j}\right) \\ &\leq |\tilde{\mathbf{b}}_i|^2 (1 + \frac{1}{4}(2^i - 2)) \end{aligned}$$

so

$$|\mathbf{b}_i|^2 \leq 2^{i-1}|\tilde{\mathbf{b}}_i|^2 \quad \text{for } i = 1, \dots, n. \quad (7)$$

Thus, by (6) and (7),

$$|\mathbf{b}_j|^2 \leq 2^{j-1}|\tilde{\mathbf{b}}_j|^2 \leq 2^{j-1} \cdot 2^{i-j}|\tilde{\mathbf{b}}_i|^2 = 2^{i-1}|\tilde{\mathbf{b}}_i|^2 \quad \text{for } 1 \leq j \leq i \leq n$$

and this proves i).

Note that $d(\Lambda) = |\det(\mathbf{b}_1, \dots, \mathbf{b}_n)|$ and so by Hadamard's inequality,

$$d(\Lambda) \leq |\mathbf{b}_1| \cdots |\mathbf{b}_n|.$$

By construction

$$d(\Lambda) = |\det(\mathbf{b}_1, \dots, \mathbf{b}_n)| = |\det(\tilde{\mathbf{b}}_1, \dots, \tilde{\mathbf{b}}_n)|.$$

But $\tilde{\mathbf{b}}_1, \dots, \tilde{\mathbf{b}}_n$ are orthogonal and so

$$d(\Lambda) = |\det(\tilde{\mathbf{b}}_1, \dots, \tilde{\mathbf{b}}_n)| = |\tilde{\mathbf{b}}_1| \cdots |\tilde{\mathbf{b}}_n|.$$

By i)

$$|\mathbf{b}_i| \leq 2^{(i-1)/2}|\tilde{\mathbf{b}}_i| \quad \text{for } 1 \leq i \leq n.$$

and so

$$|\mathbf{b}_1| \cdots |\mathbf{b}_n| \leq 2^0 \cdot 2^{1/2} \cdots 2^{(n-1)/2} |\tilde{\mathbf{b}}_1| \cdots |\tilde{\mathbf{b}}_n| = 2^{n(n-1)/4} d(\Lambda)$$

and this proves ii).

To prove iii) we apply i) with $j = 1$. Then

$$|\mathbf{b}_1| \leq 2^{(i-1)/2}|\tilde{\mathbf{b}}_i|, \quad \text{for } i = 1, \dots, n.$$

Thus

$$|\mathbf{b}_1| \leq 2^{(n-1)/4} d(\Lambda)^{1/n}.$$

Proposition 18: Let $\mathbf{b}_1, \dots, \mathbf{b}_n$ be a reduced basis for a lattice Λ in \mathbb{R}^n . Then for any vector \mathbf{x} in Λ with $\mathbf{x} \neq \mathbf{0}$ we have

$$|\mathbf{b}_1|^2 \leq 2^{n-1}|\mathbf{x}|^2$$

Proof: Write $\mathbf{x} = g_1\mathbf{b}_1 + \cdots + g_n\mathbf{b}_n$ with g_1, \dots, g_n integers and

$$\mathbf{x} = \lambda_1\tilde{\mathbf{b}}_1 + \cdots + \lambda_n\tilde{\mathbf{b}}_n$$

with $\lambda_1, \dots, \lambda_n$ real numbers. Let i be the largest index for which $g_i \neq 0$. Then by construction $\lambda_i = g_i$. Thus

$$|\mathbf{x}|^2 \geq \lambda_i^2 |\tilde{\mathbf{b}}_i|^2 \geq |\tilde{\mathbf{b}}_i|^2$$

and by Proposition 17 i),

$$2^{i-1}|\mathbf{x}|^2 \geq 2^{i-1}|\tilde{\mathbf{b}}_i|^2 \geq |\mathbf{b}_1|^2$$

are required.

Proposition 19: Let $\mathbf{b}_1, \dots, \mathbf{b}_n$ be a reduced basis for a lattice Λ in \mathbb{R}^n . Let $\mathbf{x}_1, \dots, \mathbf{x}_t$ be t linearly independent vectors from Λ . Then

$$|\mathbf{b}_j|^2 \leq 2^{n-1} \max\{|\mathbf{x}_1|^2, \dots, |\mathbf{x}_t|^2\} \quad \text{for } j = 1, \dots, t.$$

Proof: Write $\mathbf{x}_j = g_{1j}\mathbf{b}_1 + \cdots + g_{nj}\mathbf{b}_n$ 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 18

$$|\mathbf{x}_j|^2 \geq |\tilde{\mathbf{b}}_{i(j)}|^2.$$

Renumber the \mathbf{x}_j s so that $i(1) \leq i(2) \leq \cdots \leq i(t)$. Observe that $j \leq i(j)$ since otherwise $\mathbf{x}_1, \dots, \mathbf{x}_j$ would be in $\text{Sp}\{\mathbf{b}_1, \dots, \mathbf{b}_{j-1}\}$ which contradicts the assumption that $\mathbf{x}_1, \dots, \mathbf{x}_j$ are linearly independent. Thus by Proposition 17 i),

$$|\mathbf{b}_j|^2 \leq 2^{i(j)-1} |\tilde{\mathbf{b}}_{i(j)}|^2 \leq 2^{i(j)-1} |\mathbf{x}_j|^2 \quad \text{for } j = 1, \dots, t.$$

Since $i(j) \leq n$ our result follows.

We now describe the L^3 -algorithm for transforming a basis $\mathbf{b}_1, \dots, \mathbf{b}_n$ for a lattice Λ in \mathbb{R}^n to a reduced basis for Λ . The first step is to apply Gram-Schmidt and compute $\tilde{\mathbf{b}}_1, \dots, \tilde{\mathbf{b}}_n$ and the μ_{ij} s. During the course of the algorithm we will change the \mathbf{b}_j s and each time we recompute the $\tilde{\mathbf{b}}_j$ s and the μ_{ij} s.

At each step of the algorithm there is a current subscript k with k in $\{1, \dots, n+1\}$. We start with $k = 2$.

We shall now iterate a sequence of steps which starts from and returns to a situation where the following conditions are satisfied

$$1) \mu_{ij} \leq \frac{1}{2} \text{ for } 1 \leq j < i < k$$

and

$$2) |\tilde{\mathbf{b}}_i + \mu_{i,i-1}\tilde{\mathbf{b}}_{i-1}| \geq \frac{3}{4}|\tilde{\mathbf{b}}_{i-1}|^2 \text{ for } 1 < i < k$$

Note that 1) and 2) hold for $k = 2$.

PMATH 944 Lecture 10: October 15, 2009

$k \in \{1, \dots, n+1\}$. Start with $k = 2$.

Return to situation:

$$1) |\mu_{ij}| \leq \frac{1}{2} \text{ for } 1 \leq j < i < k$$

and

$$2) |\tilde{\mathbf{b}}_i + \mu_{i,i-1}\tilde{\mathbf{b}}_{i-1}|^2 \geq \frac{3}{4}|\tilde{\mathbf{b}}_{i-1}|^2 \text{ for } 1 < i < k.$$

Plainly 1) and 2) hold when $k = 2$.

If $k = n+1$ then the basis is reduced and the algorithm terminates. If $1 < k \leq n$ then we first achieve

$$|\mu_{k,k-1}| \leq \frac{1}{2}. \tag{8}$$

If (8) does not hold then let r be the closest integer to $\mu_{k,k-1}$ and replace \mathbf{b}_k by $\mathbf{b}_k - r\mathbf{b}_{k-1}$ in our basis.

This has the effect of replacing $\mu_{k,k-1}$ by $\mu_{k,k-1} - r$ and $|\mu_{k,k-1} - r| \leq \frac{1}{2}$. The numbers μ_{kj} with $j < k - 1$ are replaced by $\mu_{kj} - r\mu_{k-1,j}$. The other μ_{ij} s and \mathbf{b}_i s with i different from k and $k - 1$ with $i \leq k$ are not changed. We may now assume that (8) holds.

We now distinguish two cases:

Case 1: If $k \geq 2$ and

$$|\tilde{\mathbf{b}}_k + \mu_{k,k-1}\tilde{\mathbf{b}}_{k-1}|^2 < \frac{3}{4}|\tilde{\mathbf{b}}_{k-1}|^2$$

then we interchange \mathbf{b}_k and \mathbf{b}_{k-1} in our basis (so $i \neq k, k - 1$). We leave the other \mathbf{b}_i s unchanged. Notice that $\tilde{\mathbf{b}}_k, \tilde{\mathbf{b}}_{k-1}$ and the numbers $\mu_{k,k-1}, \mu_{k-1,j}, \mu_{kj}, \mu_{ik}, \mu_{i,k-1}$ for $j < k - 1$ and $i > k$ are changed. Let us call our new basis $\mathbf{c}_1, \dots, \mathbf{c}_n$ so that $\mathbf{c}_i = \mathbf{b}_i$ for $i \neq k, k - 1$ and $\mathbf{c}_{k-1} = \mathbf{b}_k, \mathbf{c}_k = \mathbf{b}_{k-1}$. Note that $\tilde{\mathbf{c}}_{k-1}$ is the projection of \mathbf{b}_k on the orthogonal complement of the span of $\{\mathbf{b}_1, \dots, \mathbf{b}_{k-2}\}$ and so $\tilde{\mathbf{c}}_{k-1} = \tilde{\mathbf{b}}_k + \mu_{k,k-1}\tilde{\mathbf{b}}_{k-1}$. Therefore

$$|\tilde{\mathbf{c}}_{k-1}|^2 < \frac{3}{4}|\tilde{\mathbf{b}}_{k-1}|^2.$$

In particular the “new” $|\tilde{\mathbf{b}}_{k-1}|^2$ is less than $\frac{3}{4}$ of the “old” $|\tilde{\mathbf{b}}_{k-1}|^2$. We now replace k by $k - 1$ and return to the start of the algorithm.

Case 2: If $k = 1$ or

$$|\tilde{\mathbf{b}}_k + \mu_{k,k-1}\tilde{\mathbf{b}}_{k-1}|^2 \geq \frac{3}{4}|\tilde{\mathbf{b}}_{k-1}|^2$$

then we achieve $|\mu_{kj}| \leq \frac{1}{2}$ for $1 \leq j \leq k - 1$; we replace k by $k + 1$ and we return to the start of the algorithm.

To achieve $|\mu_{kj}| \leq \frac{1}{2}$ for $1 \leq j \leq k - 1$ we do the following. First note that $\mu_{k,k-1} \leq \frac{1}{2}$. 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 \mathbf{b}_k by $\mathbf{b}_k - r\mathbf{b}_l$. 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

$$|\mu_{kj}| \leq \frac{1}{2} \quad \text{for } 1 \leq j \leq k - 1.$$

We shall now show that the algorithm terminates after only finitely many steps. We introduce the quantities for $1 \leq i \leq n$.

$$\begin{aligned} d_i &= \det((\mathbf{b}_j, \mathbf{b}_i)) \text{ for } 1 \leq j \leq i, 1 \leq i \leq n \\ &= \det((\mathbf{b}_1, \dots, \mathbf{b}_i) \cdot (\mathbf{b}_1, \dots, \mathbf{b}_i)^{\text{tr}}) \\ &= \det((\tilde{\mathbf{b}}_1, \dots, \tilde{\mathbf{b}}_i) \cdot (\tilde{\mathbf{b}}_1, \dots, \tilde{\mathbf{b}}_i)^{\text{tr}}) \end{aligned}$$

since the determinant does not change if we add a multiple of one row to another. We put $D = \prod_{i=1}^n d_i$. Note that $d_n = d(A)^2$. Further,

$$\begin{aligned} d_i &= (|\tilde{\mathbf{b}}_1| \cdots |\tilde{\mathbf{b}}_i|)^2 = \det((\tilde{\mathbf{b}}_1, \dots, \tilde{\mathbf{b}}_i) \cdot (\tilde{\mathbf{b}}_1, \dots, \tilde{\mathbf{b}}_i)^{\text{tr}}) \\ &= \det((\mathbf{b}_1, \dots, \mathbf{b}_i) \cdot (\mathbf{b}_1, \dots, \mathbf{b}_i)^{\text{tr}}) \\ &= d(A_i)^2 \end{aligned}$$

where A_i is the lattice generated by $\mathbf{b}_1, \dots, \mathbf{b}_i$ in the i -dimensional subspace of \mathbb{R}^n spanned by these vectors. Note that D changes only if one of the $\tilde{\mathbf{b}}_i$ s changes and this only occurs in case 1. Further in case 1 we interchange \mathbf{b}_k and \mathbf{b}_{k-1} . Since $d_i = (|\tilde{\mathbf{b}}_1| \cdots |\tilde{\mathbf{b}}_i|)^2$ we see that d_i only changes when $i = k - 1$ in which case it gets smaller by a factor of at least $3/4$. Further D is smaller by a factor of at least $3/4$. To complete our argument we'll show that D is bounded from below in terms of A .

Put $m(A) = \min\{\mathbf{x} \cdot \mathbf{x} : \mathbf{x} \in A, \mathbf{x} \neq \mathbf{0}\}$. By Theorem 10,

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

hence

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

Since $m(A_i) \geq m(A)$,

$$d_i \geq m(A)^i 4^{-i} \omega_i^2, \text{ for } i = 1, \dots, n.$$

Thus

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

as required.

Therefore we can pass through case 1 only finitely many times. In case 1, k decreases by 1. In case 2, k increases by 1 and so after finitely many steps $k = n + 1$ and our algorithm terminates.

In fact the algorithm is efficient. Lenstra, Lenstra, and Lovász proved that if Λ is a sublattice of Λ_0 with basis $\mathbf{b}_1, \dots, \mathbf{b}_n$ and if B is a real number with $B \geq 2$ and

$$|\mathbf{b}_i|^2 \leq B \quad \text{for } i = 1, \dots, n$$

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 $O(n \log B)$. By an arithmetical operation I mean an addition, subtraction, multiplication, or division, and by the binary length of an integer n , I mean the length or number of digits in the base 2 expansion of n . The algorithm runs in polynomial time in terms of B .

PMATH 944 Lecture 11: October 20, 2009

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 $\mathbf{b}_1, \dots, \mathbf{b}_n$. Then \mathbf{b}_1 is a short vector in Λ .

Let $\alpha_1, \dots, \alpha_n$ be in \mathbb{R} and let ϵ be a real number with $0 < \epsilon < 1$. How do we produce *efficiently* a positive integer q and integers p_1, \dots, p_n for which

$$|q\alpha_i - p_i| < \epsilon \quad \text{for } i = 1, \dots, n,$$

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

If $\alpha_1, \dots, \alpha_n$ and ϵ are in \mathbb{Q} then we can use L^3 to find q in polynomial time in terms of the input. First recall, by Theorem 7, on taking $\epsilon = \frac{1}{Q}$, that such a q exists with

$$1 \leq q \leq \epsilon^{-n}.$$

We consider the lattice Λ generated by the rows of the matrix

$${}_{n+1} \left\{ \overbrace{\begin{pmatrix} 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & & 0 & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & 0 & & 1 & 0 \\ \alpha_1 & \alpha_2 & \cdots & \alpha_n & \delta \end{pmatrix}}^{n+1} \right. \quad \text{where } \delta = 2^{-n(n+1)/4} \epsilon^{n+1}.$$

Note that $d(\Lambda) = \delta$. By L^3 we can find a small non-zero vector \mathbf{b} ($= \mathbf{b}_1$) in the lattice with

$$\mathbf{b} = (q\alpha_1 - p_1, q\alpha_2 - p_2, \dots, q\alpha_n - p_n, q\delta)$$

where q and p_1, \dots, p_n are integers. Note that we may suppose that $q \geq 0$ by replacing \mathbf{b} by $-\mathbf{b}$ if necessary. Further by Proposition 17 iii), we can find \mathbf{b} with

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

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

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

What if we want to find a small linear form with integer coefficients in $\alpha_1, \dots, \alpha_n$? Given ϵ with $0 < \epsilon < 1$ how do we find efficiently integers q_1, \dots, q_n and p such that

$$|q_1\alpha_1 + \dots + q_n\alpha_n - p| < \epsilon$$

and with

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

Again by Theorem 7 the objective is best possible up to the factor $2^{(n+1)/4}$.

We consider the lattice Λ generated by the rows of the matrix

$$\begin{pmatrix} 1 & 0 & \cdots & 0 \\ \alpha_1 & \delta & & 0 \\ \vdots & & \ddots & \\ \alpha_n & 0 & & \delta \end{pmatrix} \quad \text{where } \delta = \left(\frac{\epsilon^{1/n}}{2^{1/4}} \right)^{n+1}.$$

A typical vector \mathbf{b} in Λ is of the form

$$(q_1\alpha_1 + \dots + q_n\alpha_n - p, q_1\delta, q_2\delta, \dots, q_n\delta)$$

with q_1, \dots, q_n and p integers. By L^3 we can find a non-zero vector \mathbf{b} in Λ of this form with $|\mathbf{b}| \leq 2^{n/4} d(\Lambda)^{1/(n+1)}$, and since $d(\Lambda) = \delta^n = \left(\frac{\epsilon}{2^{n/4}} \right)^{n+1}$ we see that

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

Further since $\mathbf{b} \neq \mathbf{0}$ and $\epsilon < 1$ we have $0 < |\mathbf{b}| < 1$ hence q_1, \dots, q_n are not all zero and so

$$0 < \max_i |q_i|.$$

Finally suppose that α_{ij} ($1 \leq i \leq n$, $1 \leq j \leq m$) are all real numbers and that ϵ is a real number with $0 < \epsilon < 1$. Consider the lattice Λ generated by the rows of the matrix

$$\begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ \alpha_{11} & \cdots & \alpha_{n1} & \delta & & \\ \vdots & & \vdots & & \ddots & \\ \alpha_{1m} & \cdots & \alpha_{nm} & & & \delta \end{pmatrix} \quad (7).$$

Note that $d(\Lambda) = \delta^m = (2^{-(n+m-1)/4} \cdot \epsilon)^{n+m}$.

By L^3 there is a non-zero vector \mathbf{b} in Λ with

$$\begin{aligned} |\mathbf{b}| &\leq \delta^{m/(m+n)} 2^{(n+m-1)/4} \\ &= 2^{-(n+m-1)/4} \cdot \epsilon \cdot 2^{(n+m-1)/4} = \epsilon. \end{aligned}$$

We have

$$\mathbf{b} = (q_1\alpha_{11} + q_2\alpha_{12} + \dots + q_m\alpha_{1m} - p_1, \\ q_1\alpha_{21} + q_2\alpha_{22} + \dots + q_m\alpha_{2m} - p_2, \dots, q_1\alpha_{n1} + \dots + q_m\alpha_{nm} - p_n, q_1\delta, q_2\delta, \dots, q_m\delta)$$

with q_1, \dots, q_m and p_1, \dots, p_n integers. Then

$$|q_1\alpha_{i1} + \dots + q_m\alpha_{im} - p_i| < \epsilon \quad \text{for } i = 1, \dots, n$$

⁷⁾ an $m+n \times m+n$ matrix where $\delta = (2^{-(n+m-1)/4} \cdot \epsilon)^{n/m+1}$

and $|q_j \delta| < \epsilon$ for $j = 1, \dots, m$ so $|q_j| < \delta^{-1} \epsilon = 2^{\binom{n+m-1}{4} \binom{n+m}{m}} \epsilon^{-n/m}$. Further, as before the q_i s are not all zero.

Theorem 7 tells us that we can make linear forms in the α_{ij} s with integer coefficients which are simultaneously close to integers. L^3 gives us an efficient method for finding the associated integer coefficients. Can we do better than Theorem 7? Not for real numbers in general, but for algebraic numbers α_{ij} we can say more. It follows from work of Schmidt that:

Theorem 20: Let $1, \alpha_1, \dots, \alpha_n$ be real algebraic numbers which are linearly independent over \mathbb{Q} . Let $\delta > 0$. There are only finitely many n -tuples of non-zero integers q_1, \dots, q_n with

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

where for any real number x , $\|x\|$ denotes the distance from x to the nearest integer.

Applying Theorem 20 to all finite subsets of $\{\alpha_1, \dots, \alpha_n\}$ we deduce the following:

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

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

Note the special case $n = 1$ is Roth's Theorem.

PMATH 944 Lecture 12: October 22, 2009

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

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

The special case when $n = 1$ is due to Roth. In particular, let $\delta > 0$, if α is an algebraic number then there are only finitely many rationals p/q with $q > 0$ for which

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

\implies Thue equations such as $x^3 - 2y^3 = 6$, have only finitely many solutions in integers x and y . Roth obtained the Fields Medal in 1958. Schmidt also proved:

Theorem 21: Suppose that $\alpha_1, \dots, \alpha_n$ are real algebraic numbers with $1, \alpha_1, \dots, \alpha_n$ \mathbb{Q} -linearly independent. Let $\delta > 0$. There are only finitely many positive integers q with

$$q^{1+\delta} \|\alpha_1 q\| \cdots \|\alpha_n q\| < 1.$$

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

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

Theorems 20 and 21 are consequences of the following result.

Theorem 22: (Schmidt's Subspace Theorem). Suppose $L_1(\mathbf{x}), \dots, L_n(\mathbf{x})$ are *linearly independent* linear forms in $\mathbf{x} = (x_1, \dots, x_n)$ with algebraic coefficients. Let $\delta > 0$. There are *finitely* many *proper* subspaces T_1, \dots, T_w of \mathbb{R}^n such that every integer point $\mathbf{x} = (x_1, \dots, x_n)$ with $\mathbf{x} \neq \mathbf{0}$ for which

$$|L_1(\mathbf{x}) \cdots L_n(\mathbf{x})| < \frac{1}{|\mathbf{x}|^\delta}$$

lies in (at least) one of the subspaces.

Remarks:

1. The result is not effective in the sense that the proof does not yield a procedure for determining the subspaces T_1, \dots, T_w .
2. The integer points in a proper subspace of \mathbb{R}^n lie 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.

Let us now deduce Theorem 21 from the Subspace Theorem. Let q be a positive integer satisfying

$$q^{1+\delta} \|\alpha_1 q\| \cdots \|\alpha_n q\| < 1.$$

Choose integers p_1, \dots, p_n such that $\|\alpha_i q\| = |\alpha_i q - p_i|$, for $i = 1, \dots, n$. Then put $\mathbf{x} = (p_1, \dots, p_n, q)$. Let K_1, K_2 denote positive numbers which depend on $\alpha_1, \dots, \alpha_n$ and n only. Note that

$$\overline{\mathbf{x}} \leq K_1 q.$$

We consider the linear forms

$$\begin{aligned} L_i(\mathbf{X}) &= \alpha_i X_{n+1} - X_i \quad \text{for } i = 1, \dots, n \\ L_{n+1}(\mathbf{X}) &= X_{n+1} \end{aligned}$$

L_1, \dots, L_{n+1} are $n+1$ -linearly independent linear forms with algebraic coefficients.

We have

$$|L_1(\mathbf{x}) \cdots L_{n+1}(\mathbf{x})| = \|\alpha_1 q\| \cdots \|\alpha_n q\| \cdot q$$

so

$$|L_1(\mathbf{x}) \cdots L_{n+1}(\mathbf{x})| < \frac{1}{q^\delta} < \frac{1}{\overline{\mathbf{x}}^{\delta/2}},$$

for q sufficiently large, as we may assume.

By the Subspace Theorem \mathbf{x} lies in one of finitely many proper subspaces T_1, \dots, T_w of \mathbb{R}^{n+1} . Since \mathbf{x} has integer coordinates it lies in a proper rational subspace T . We can find c_1, \dots, c_{n+1} in \mathbb{Q} such that T is determined by $c_1 X_1 + \cdots + c_{n+1} X_{n+1}$. Then

$$c_1 x_1 + \cdots + c_{n+1} x_{n+1} = 0. \tag{9}$$

Since $\mathbf{x} \in T$,

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

since $1, \alpha_1, \dots, \alpha_n$ are linearly independent over \mathbb{Q} . Thus

$$\begin{aligned} K_2 q &< |c_1(\alpha_1 q - p_1) + \cdots + c_n(\alpha_n q - p_n)| \\ &\leq |c_1| + \cdots + |c_n| \end{aligned}$$

which implies q is bounded as required.

We shall now deduce Theorem 20 from the Subspace Theorem.

Proof: By induction on n . For $n = 1$ the result holds by Theorem 21, say. Suppose $n > 1$. Assume that q_1, \dots, q_n are non-zero integers with

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

We now choose p , an integer, so that

$$\|\alpha_1 q_1 + \cdots + \alpha_n q_n\| = |\alpha_1 q_1 + \cdots + \alpha_n q_n - p|.$$

Write $\mathbf{x} = (x_1, \dots, x_{n+1}) = (q_1, \dots, q_n, p)$.

Let K_3, K_4 be positive numbers which depend on $\alpha_1, \dots, \alpha_n$. Then

$$|\mathbf{x}| = \max(|q_1|, \dots, |q_n|, |p|) \leq K_3 q$$

where $q = \max_i |q_i|$. Put

$$L_i(\mathbf{X}) = X_i \quad \text{for } i = 1, \dots, n$$

and

$$L_{n+1}(\mathbf{X}) = \alpha_1 X_1 + \dots + \alpha_n X_n - X_{n+1}.$$

Then

$$|L_1(\mathbf{x}) \cdots L_{n+1}(\mathbf{x})| = |q_1 \cdots q_n| |\alpha_1 q_1 + \dots + \alpha_n q_n| < \frac{1}{|q_1 \cdots q_n|^\delta} < \frac{1}{|\mathbf{x}|^{\delta/2}},$$

for q sufficiently large. Then by the Subspace Theorem \mathbf{x} lies in one of finitely many proper rational subspaces of \mathbb{R}^{n+1} .

PMATH 944 Lecture 13: October 27, 2009

We deduce Theorem 20 from the Subspace Theorem by induction on n . $n = 1$ ✓. Assume we have integers q_1, \dots, q_n , not all zero, for which

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

Choose p to be the closest integer to $\alpha_1 q_1 + \dots + \alpha_n q_n$ so that $\alpha_1 q_1 + \dots + \alpha_n q_n - p < 1$. Write

$$\begin{aligned} \mathbf{x} &= (q_1, \dots, q_n, p) \text{ and put} \\ L_i(\mathbf{X}) &= X_i \text{ for } i = 1, \dots, n \end{aligned}$$

and

$$L_{n+1}(\mathbf{X}) = \alpha_1 X_1 + \dots + \alpha_n X_n - X_{n+1}.$$

We have $n + 1$ linearly independent forms with algebraic coefficients.

Note that

$$|L_1(\mathbf{x}) \cdots L_{n+1}(\mathbf{x})| = |q_1 \cdots q_n| |\alpha_1 q_1 + \dots + \alpha_n q_n|.$$

We have $|\mathbf{x}| < K_1 q$ where $q = \max_i |q_i|$ and K_1, K_2, \dots denote positive numbers which depend on $\alpha_1, \dots, \alpha_n$ and n . Observe that

$$|L_1(\mathbf{x}) \cdots L_{n+1}(\mathbf{x})| < \frac{1}{|q_1 \cdots q_n|^\delta} < \frac{1}{|\mathbf{x}|^{\delta/2}},$$

for q sufficiently large, as we may assume. Then by the Subspace Theorem \mathbf{x} lies in one of a finite collection of proper rational subspaces of \mathbb{R}^{n+1} . Let T be such a subspace. Then T can be expressed as the set of rational points $(y_1, \dots, y_{n+1}) \in \mathbb{R}^{n+1}$ for which $c_1 y_1 + \dots + c_{n+1} y_{n+1} = 0$ ⁸⁾ with $c_1, \dots, c_{n+1} \in \mathbb{Q}$ and not all the c_i are zero.

Suppose first that $c_i \neq 0$ for some i with $1 \leq i \leq n$. Without loss of generality we may suppose $c_n \neq 0$. Then

$$c_1 q_1 + \dots + c_n q_n + c_{n+1} p = 0$$

so

$$c_n \alpha_n q_n = -c_1 \alpha_n q_1 - \dots - c_{n-1} \alpha_n q_{n-1} - c_{n+1} \alpha_n p$$

Thus

$$\begin{aligned} |c_n| |\alpha_1 q_1 + \dots + \alpha_n q_n - p| &= |(c_n \alpha_1 - c_1 \alpha_n) q_1 + \dots + (c_n \alpha_{n-1} - c_{n-1} \alpha_n) q_{n-1} - (c_n + c_{n+1}) p| \\ &= |c_n + c_{n+1} \alpha_n| \left| \left(\frac{c_n \alpha_1 - c_1 \alpha_n}{c_n + c_{n+1} \alpha_n} \right) q_1 + \dots + \frac{(c_n \alpha_{n-1} + c_{n-1} \alpha_n)}{(c_n + c_{n+1} \alpha_n)} q_{n-1} - p \right| \end{aligned}$$

⁸⁾ * defining equation of subspace

Note that $c_n + c_{n+1}\alpha_n \neq 0$ since $1, \alpha_1, \dots, \alpha_n$ are linearly independent over \mathbb{Q} . Put

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

Then

$$|c_n|\alpha_1q_1 + \dots + \alpha_nq_n - p| = |c_n + c_{n+1}\alpha_n|\alpha'_1q_1 + \dots + \alpha'_{n-1}q_{n-1} - p|.$$

Therefore

$$\|\alpha'_1q_1 + \dots + \alpha'_{n-1}q_{n-1}\| < \frac{K_2}{|q_1 \dots q_n|^{1+\delta}} < \frac{1}{|q_1 \dots q_{n-1}|^{1+\delta/2}}$$

for q sufficiently large.

We remark that $1, \alpha'_1, \dots, \alpha'_{n-1}$ are linearly independent over \mathbb{Q} . To see this suppose that

$$\lambda_1\alpha'_1 + \dots + \lambda_{n-1}\alpha'_{n-1} + \lambda_n = 0$$

with $\lambda_1, \dots, \lambda_n$ in \mathbb{Q} . Then

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

But $1, \alpha_1, \dots, \alpha_n$ are linearly independent over \mathbb{Q} and so, since $c_n \neq 0$, $\lambda_1 = \dots = \lambda_n = 0$. Then by induction $|q_1|, \dots, |q_n|$ are bounded.

It remains to consider the case when $c_1 = \dots = c_n = 0$ and $c_{n+1} \neq 0$. Then

$$c_{n+1}p = 0 \quad \text{so} \quad p = 0.$$

In this case

$$|q_1 \dots q_n|^{1+\delta} |\alpha_1q_1 + \dots + \alpha_nq_n| < 1$$

so

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

Put $\alpha'_i = \frac{\alpha_i}{\alpha_n}$ for $i = 1, \dots, n-1$.

Then $1, \alpha'_1, \dots, \alpha'_{n-1}$ are linearly independent over \mathbb{Q} and so

$$|q_1 \dots q_{n-1}|^{1+\delta/2} \|q_1\alpha'_1 + \dots + q_{n-1}\alpha'_{n-1}\| < 1.$$

Therefore $\max_i |q_i|$ is bounded by induction. The result follows.

In a similar way we can deduce the following consequences of the Subspace Theorem.

Theorem 23: Let α_{ij} ($1 \leq i \leq n, 1 \leq j \leq m$) be real algebraic numbers and suppose that $1, \alpha_{i1}, \dots, \alpha_{im}$ are linearly independent over \mathbb{Q} , for $i = 1, \dots, n$. Let $\delta > 0$. Then there are only finitely many m -tuples of non-zero integers (q_1, \dots, q_m) for which

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

Results of this sort have application to the study of Diophantine equations such as Norm form equations.

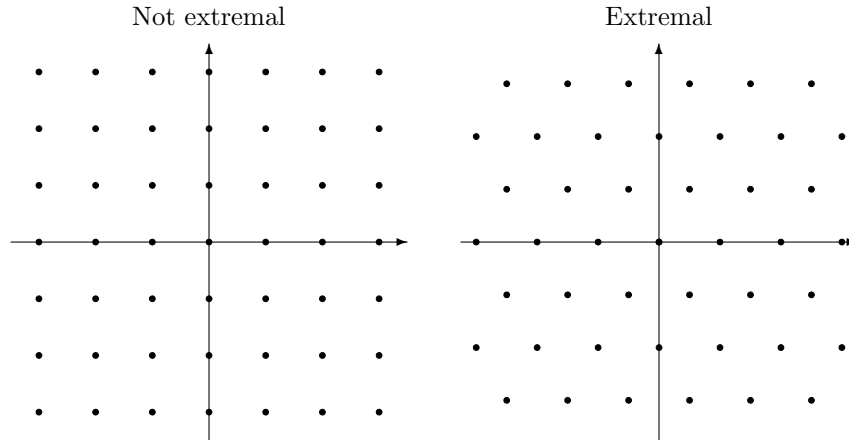
For each dimension n let us consider those lattices with $d(\Lambda) = 1$. In this collection let us look for lattices Λ for which the minimal non-zero distance between lattice points $\mu(\Lambda)$ is large.

We define μ_n for $n = 1, 2, \dots$ by

$$\begin{aligned} \mu_n &= \sup_{\substack{\text{lattices } \Lambda \text{ in } \mathbb{R}^n \\ \text{with } d(\Lambda) = 1}} \left(\min_{\substack{\mathbf{x}, \mathbf{y} \in \Lambda \\ \mathbf{x} \neq \mathbf{y}}} |\mathbf{x} - \mathbf{y}| \right) \\ &= \sup_{\substack{\Lambda \text{ in } \mathbb{R}^n \\ d(\Lambda) = 1}} (\mu(\Lambda)) \end{aligned}$$

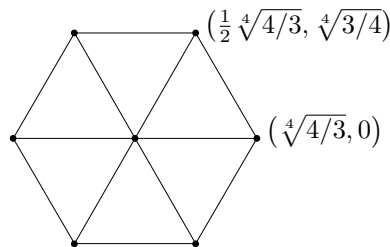
It follows from Mahler's Compactness Theorem that the supremum is actually a maximum. Lattices for which the maximum is attained are known as extremal lattices. The values of μ_n have been determined for $n = 1, \dots, 8$ and they are

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



PMATH 944 Lecture 14: October 29, 2009

We'll prove that $\mu_2 = \sqrt[4]{4/3}$. We first note that this is a lattice Λ in \mathbb{R}^2 with $d(\Lambda) = 1$ and $\mu(\Lambda) = \sqrt[4]{4/3}$. We take the basis vectors for Λ to be $(\sqrt[4]{4/3}, 0)$, $(\frac{1}{2}\sqrt[4]{4/3}, \sqrt[4]{3/4})$. Observe that $d(\Lambda) = 1$ and that both generating vectors have length $\sqrt[4]{4/3}$ and that this is the minimal distance between two distinct vectors in Λ :

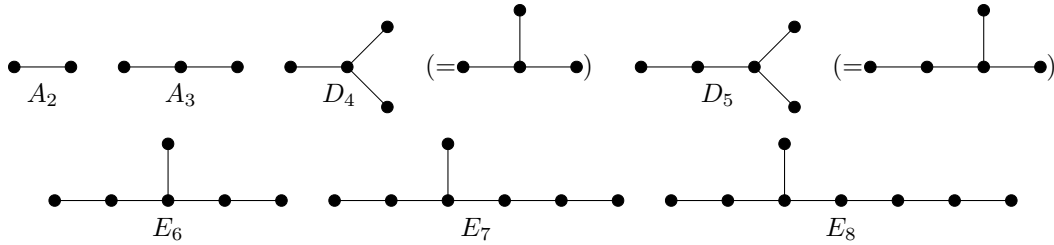


This is the maximum for suppose that Λ' is a lattice in \mathbb{R}^2 with $d(\Lambda') = 1$ for which $\mu(\Lambda') > \sqrt[4]{4/3}$. Then, without loss of generality, we may suppose that a basis for Λ' is of the form $(a, 0)$, $(b, 1/a)$ with $a > 0$. Further, by adding an appropriate multiple of $(a, 0)$ to $(b, 1/a)$ we may suppose that $|b| \leq \frac{a}{2}$.

Furthermore we may suppose that $a = \mu(\Lambda')$. If $a > \sqrt[4]{4/3}$ then $3a^4 > 4$ so $\frac{3}{4}a^2 > \frac{1}{a^2}$. But then $(b, \frac{1}{a})$ is closer to the origin than $(a, 0)$ since $b^2 + \frac{1}{a^2} < \frac{a^2}{4} + \frac{3}{4}a^2 = a^2$, and this is a contradiction.

The first few extremal lattices can be represented by graphs. The graphs are 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° or $\frac{\pi}{3}$. Finally we normalize the length of the generators so that the determinant of the lattice is 1.

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



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

We'll look more closely at the lattices associated with these diagrams. Let $\mathbf{b}_1, \dots, \mathbf{b}_n$ be basis vectors in such a lattice. We'll assume initially that each vector is of length $\sqrt{2}$. Notice that the inner product $\mathbf{b}_i \cdot \mathbf{b}_j = |\mathbf{b}_i||\mathbf{b}_j| \cos \theta_{ij}$ where θ_{ij} is the angle between the vectors \mathbf{b}_i and \mathbf{b}_j . Thus if the angle is 60° then $\mathbf{b}_i \cdot \mathbf{b}_j = 2 \cos \frac{\pi}{3} = 1$.

Notice that if

$$B = (\mathbf{b}_i \cdot \mathbf{b}_j)_{\substack{i=1, \dots, n \\ j=1, \dots, n}} \quad \text{then the} \quad \det(B) = d(\Lambda)^2.$$

To see this let e_1, \dots, e_n be the standard basis in \mathbb{R}^n and put $\mathbf{b}_i = \sum_{j=1}^n B_{ij} e_j$ with $B_{ij} \in \mathbb{R}$. Then

$$B = ((B_{ij})^{\text{tr}}(B_{ij}))$$

and so

$$\det(B) = (\det(B_{ij}))^2 = d(\Lambda)^2.$$

Next we observe that each non-zero vector in Λ has length at least $\sqrt{2}$. To see this suppose that $\mathbf{b} = k_1 \mathbf{b}_1 + \dots + k_n \mathbf{b}_n$ is in Λ with k_1, \dots, k_n integers, not all zero. Then

$$\begin{aligned} \mathbf{b} \cdot \mathbf{b} &= (k_1 \mathbf{b}_1 + \dots + k_n \mathbf{b}_n) \cdot (k_1 \mathbf{b}_1 + \dots + k_n \mathbf{b}_n) \\ &= \sum_{i=1}^n \sum_{j=1}^n k_i k_j (\mathbf{b}_i \cdot \mathbf{b}_j) \\ &= 2(k_1^2 + \dots + k_n^2) + 2 \sum_{\substack{i < j \\ i \text{ and } j \text{ connected} \\ \text{by an edge}}} k_i k_j. \end{aligned}$$

This quantity is an even integer and so the length of \mathbf{b} is at least $\sqrt{2}$.

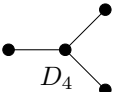
Therefore to determine $\mu(\Lambda)$ in each example it suffices to compute $\det B$ and then normalize the length of the vectors so that $d(\Lambda) = 1$.

$$\begin{array}{c} \bullet \text{---} \bullet \\ A_2 \end{array} \quad B = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \quad \text{and} \quad \det B = 3.$$

Thus it suffices to divide our basis vectors by $\sqrt[4]{3}$ and then $\mu(\Lambda_2) = \frac{\sqrt{2}}{\sqrt[4]{3}} = \sqrt[4]{4/3}$. \checkmark

$$\begin{array}{c} \bullet \text{---} \bullet \text{---} \bullet \\ A_3 \end{array} \quad B(A_3(\sqrt{2})) = \begin{pmatrix} 2 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 2 \end{pmatrix} \quad \text{and so } \det B = 6 - 2 = 4.$$

We must then divide $\mathbf{b}_1, \mathbf{b}_2$ and \mathbf{b}_3 by $\sqrt[6]{4}$ and so the minimal length of a vector in Λ_3 is $\frac{\sqrt{2}}{\sqrt[6]{4}} = 2^{1/6}$.



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

Thus we must divide each vector \mathbf{b}_i by $4^{1/8} = 2^{1/4}$ and so $\mu(\Lambda_4) = \frac{\sqrt{2}}{2^{1/4}} = 2^{1/4}$.

Let us look more closely at $D_4(\sqrt{2})$. We claim that the lattice is the same as the lattice of vectors in \mathbb{R}^4 of the form (u_1, u_2, u_3, u_4) with the u_i s integers and $u_1 + u_2 + u_3 + u_4 \equiv 0 \pmod{2}$. 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 that the lattice is generated by

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

Notice that

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

PMATH 944 Lecture 15: November 3, 2009

Consider the lattice $D_4(\sqrt{2})$ in \mathbb{R}^4 . It has diagram  and each basis vector has length $\sqrt{2}$. In fact

$D_4(\sqrt{2})$ can be represented as the lattice Λ_1 in \mathbb{R}^4 which is the sublattice of Λ_0 given by the congruence condition: (u_1, u_2, u_3, u_4) is in the lattice $\iff u_1 + u_2 + u_3 + u_4 \equiv 0 \pmod{2}$. One can check that this lattice is generated by

$$(2, 0, 0, 0), (1, 1, 0, 0), (1, 0, 1, 0), (1, 0, 0, 1).$$

As a consequence $d(\Lambda_1) = \left| \det \begin{pmatrix} 2 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix} \right| = 2$. Equivalently it is generated by $(1, 0, 0, 1), (1, 0, 1, 0), (1, 0, 0, -1)$ and $(0, 1, 1, 0)$. Notice that

$$\overbrace{\begin{pmatrix} 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & -1 \\ 0 & 1 & 1 & 0 \end{pmatrix}}^B \cdot \overbrace{\begin{pmatrix} 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & -1 & 0 \end{pmatrix}}^{B^{\text{tr}}} = \begin{pmatrix} 2 & 1 & 0 & 0 \\ 1 & 2 & 1 & 1 \\ 0 & 1 & 2 & 0 \\ 0 & 1 & 0 & 2 \end{pmatrix}$$

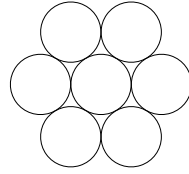
Thus Λ_1 is a representation for $D_4(\sqrt{2})$.

We now put a sphere of radius $\frac{\sqrt{2}}{2} = \frac{1}{\sqrt{2}}$ around each lattice point in $\Lambda_1(D_4(\sqrt{2}))$. 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 $(\pm 1, \pm 1, 0, 0), (\pm 1, 0, \pm 1, 0), (\pm 1, 0, 0, \pm 1), (0, \pm 1, \pm 1, 0), (0, \pm 1, 0, \pm 1), (0, 0, \pm 1, \pm 1)$. Thus the central sphere is surrounded by $\binom{4}{2} \cdot 4 = 24$ spheres which touch it. Recently (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 .

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

Thus $\tau_4 \geq 24$ by the example and $\tau_4 \leq 24$ by the result of Musin. Plainly $\tau_1 = 2$ and the hexagonal packing



gives $\tau_2 = 6$. It is not so clear what τ_3 is at first glance. The standard cannonball packing gives $\tau_3 \geq 12$. There was a dispute between Newton and Gregory as to whether τ_3 was 12 or 13. The first correct proof that $\tau_3 = 12$ is due to Schütte and van der Waerden in 1953.

Definition: 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 that the packing is a lattice packing (of spheres).

Given a sphere packing in \mathbb{R}^n let ρ be the radius of the sphere and define Δ , the packing density, in the following way. For any real number x let S_x be a sphere of radius x in \mathbb{R}^n . We put

$$\Delta = \overline{\lim}_{R \rightarrow \infty} \left(\frac{\text{the number of spheres in the collection of radius } \rho \text{ inside } S_R^{(0)} \cdot \text{volume}(S_\rho)}{\text{volume}(S_R^{(0)})} \right).$$

Δ measures the “proportion” of \mathbb{R}^n covered by the spheres in the sphere packing.

We now define Δ_n for $n = 1, 2, \dots$ by

$$\Delta_n = \sup_{\text{sphere packing in } \mathbb{R}^n} \Delta;$$

here the sup is taken over all sphere packings in \mathbb{R}^n . Similarly

$$\Delta_n(L) = \sup_{\text{lattice packing in } \mathbb{R}^n} \Delta;$$

here the sup is taken over all sphere packings in \mathbb{R}^n which are lattice packings. Notice that if L is a lattice then the largest radius ρ_0 of spheres in a sphere packing associated with the lattice is $\frac{1}{2}$ 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 L then

$$\Delta(L) = \frac{\text{volume } S_{\rho_0}}{\text{volume fundamental region of } L} = \frac{\text{volume } S_{\rho_0}}{d(L)}.$$

Certainly $\Delta_n \geq \Delta_n(L)$ for $n = 1, 2, \dots$. In fact $\Delta_1 = \Delta_1(L)$, $\Delta_2 = \Delta_2(L)$. For $n = 2$ the hexagonal lattice yields Δ_2 . We have

$$\Delta_2 = \Delta_2(L) = \frac{\pi(\frac{1}{2}\sqrt{4/3})^2}{1} = \frac{\pi}{\sqrt{12}} = 0.9069\dots$$

Let us compute the packing density Δ of D_4 . Since the minimal non-zero distance between two lattice points in $D_4(\sqrt{2})$ is $\sqrt{2}$ we may take $\rho_0 = \frac{1}{2}\sqrt{2} = \frac{1}{\sqrt{2}}$ and we have

$$\Delta(D_4) = \frac{\frac{\pi^2}{2} \left(\frac{1}{\sqrt{2}}\right)^4}{2} = \frac{\pi^2}{16} = 0.6169\dots$$

This is the largest lattice packing density known in \mathbb{R}^4 .

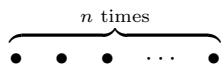
It was proved by Korkine and Zolotareff in 1872 that

$$\Delta_4(L) = \Delta_4(D_4).$$

⁹⁾Let $S_R^{(0)}$ be the sphere of radius R centred at the origin.

PMATH 944 Lecture 16: November 5, 2009



Let us consider the lattice of integer points in \mathbb{R}^n denoted by A_0 . The diagram is:

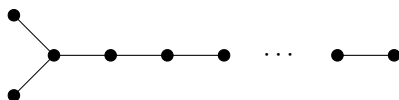


We have $d(A_0) = 1$. The vectors of minimal non-zero length in A_0 are $(\pm 1, 0, \dots, 0), \dots, (0, \dots, 0, \pm 1)$ and they are of length 1. Thus the lattice packing associated with A_0 consists of spheres of radius $\frac{1}{2}$ around each integer point. Thus the packing density is

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

In \mathbb{R}^2 , it is $\frac{\pi}{4} = 0.785\dots$, in \mathbb{R}^3 it is $\frac{\pi}{6} = 0.529\dots$, in \mathbb{R}^4 , $\frac{\pi^2}{32} = 0.308\dots$. The kissing number associated with A_0 is $2n$.

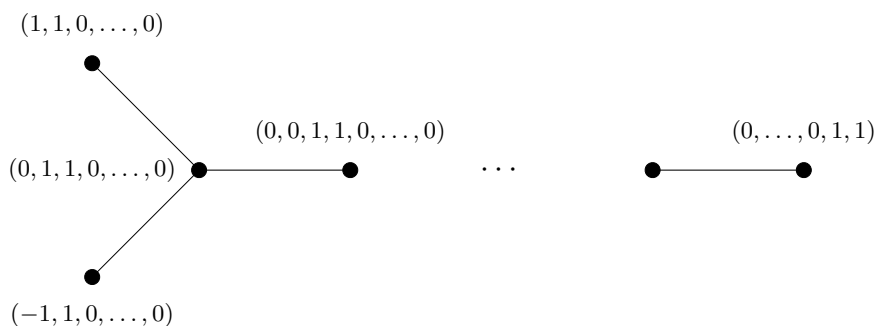
The lattice A_3 associated with  may also be associated with  which we call D_3 . For $n = 3, 4, \dots$ we denote by D_n the lattice associated with



We can represent D_n as the sublattice of A_0 given by

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

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

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

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 $4 \cdot \binom{n}{2} = 2n(n-1)$. We have a central sphere around $(0, \dots, 0)$ of radius $\frac{1}{2}\sqrt{2} = \frac{1}{\sqrt{2}}$ and it is touched by the $2n(n-1)$ non-overlapping spheres of radius $\frac{1}{\sqrt{2}}$ around $(\pm 1, \pm 1, 0, \dots, 0), \dots$,

$(0, \dots, 0, \pm 1, \pm 1)$. Put spheres of radius $\frac{1}{\sqrt{2}}$ around each lattice point to give a sphere packing. The sphere packing density

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

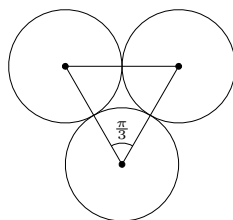
Note that

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

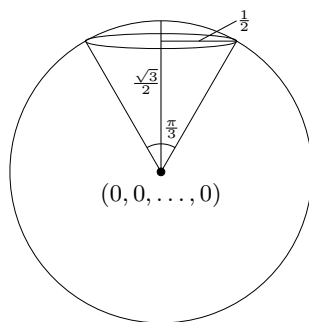
The sphere packing of D_3 corresponds to the cannonball packing. In 1831 Gauss proved that $\Delta_3(L) = \Delta(D_3)$, that is to say that D_3 gives the sphere packing associated with a lattice of maximal density.

Kepler conjectured that $\Delta_3 = \Delta_3(L) = \Delta(D_3)$, or equivalently that the most efficient packing of spheres in \mathbb{R}^3 is given by the cannonball packing. In 1958 Rogus proved $\Delta_3 \leq 0.7796$ and in 1983 Lindsay proved $\Delta_3 \leq 0.7784$. (In 1993 Hsiang claimed a proof that $\Delta_3 = \Delta(D_3)$ and his “proof” appeared in a 92 page paper in the International Journal of Mathematics.) Hales in 2005 in a 120 page paper in the Annals of Math gave a proof of Kepler’s conjecture. It depended on a massive amount of computation and this part of the argument is very hard to check.

Consider the kissing number problem in \mathbb{R}^3 . Three spheres touching in \mathbb{R}^3 :



The centres 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 $\frac{\pi}{3}$ from the origin.



The surface area of the shadow is $2\pi h$ where h is the height of the spherical cap. Here $h = 1 - \frac{\sqrt{3}}{2}$ 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 $\frac{4\pi}{(2-\sqrt{3})\pi} = 8 + 4\sqrt{3} < 15$. Thus $\tau_3 \leq 14$. The packing associated with D_3 gives $\tau_3 \geq 12$.

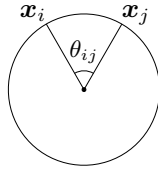
In fact $\tau_3 = 12$ as was first proved in 1953 by Schutte and van der Waerden. The following kissing numbers are known: $\tau_1 = 2$, $\tau_2 = 6$, $\tau_3 = 12$, $\tau_4 = 24$, $\tau_5 = 240$ and $\tau_{24} = 196,560$. How do we find such results?

The arguments depend on linear programming and the study of positive semidefinite functions on the sphere S^{n-1} in \mathbb{R}^n .

Let $\{\mathbf{x}_1, \dots, \mathbf{x}_m\}$ be points on S^{n-1} in \mathbb{R}^n . Thus $\mathbf{x}_i \cdot \mathbf{x}_i = 1$ for $i = 1, \dots, m$ and $\mathbf{x}_i \in \mathbb{R}^n$.

$$S^{n-1} = \{ (x_1, \dots, x_n) \in \mathbb{R}^n : x_1^2 + \dots + x_n^2 = 1 \}.$$

Let θ_{ij} be the distance between \mathbf{x}_i and \mathbf{x}_j on the surface of S^{n-1} , so the length of the geodesic between \mathbf{x}_i and \mathbf{x}_j .



It is just the angle in radians determined by the points.

Notice that for any real numbers t_1, \dots, t_m we have

$$\begin{aligned} \|t_1\mathbf{x}_1 + \dots + t_m\mathbf{x}_m\|^2 &= (t_1\mathbf{x}_1 + \dots + t_m\mathbf{x}_m, t_1\mathbf{x}_1 + \dots + t_m\mathbf{x}_m) \\ &= \sum_{i=1}^m \sum_{j=1}^m t_i t_j \cos(\theta_{ij}) \geq 0 \end{aligned}$$

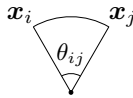
Equivalently the matrix

$$(\cos(\theta_{ij}))_{\substack{i=1, \dots, m \\ j=1, \dots, m}}$$

is positive semidefinite.

PMATH 944 Lecture 17: November 10, 2009

Let $\mathbf{x}_1, \dots, \mathbf{x}_m$ be on S^{n-1} in \mathbb{R}^n . Let θ_{ij} be the (angular) distance between \mathbf{x}_i and \mathbf{x}_j on S^{n-1} .



Notice that for any real numbers t_1, \dots, t_m we have

$$\begin{aligned} \|t_1\mathbf{x}_1 + \dots + t_m\mathbf{x}_m\|^2 &= (t_1\mathbf{x}_1 + \dots + t_m\mathbf{x}_m, t_1\mathbf{x}_1 + \dots + t_m\mathbf{x}_m) \\ &= \sum_{i=1}^m \sum_{j=1}^m t_i t_j \cos \theta_{ij} \geq 0. \end{aligned}$$

Equivalently

$$(t_1, \dots, t_m) (\cos \theta_{ij}) \begin{pmatrix} t_1 \\ \vdots \\ t_m \end{pmatrix} \geq 0.$$

Thus the matrix $(\cos \theta_{ij})_{\substack{i=1, \dots, m \\ j=1, \dots, m}}$ is positive semi-definite.

In 1943 Schoenberg proved that the matrix $(G_k^{(n)}(\cos(\theta_{ij})))_{\substack{i=1, \dots, m \\ j=1, \dots, m}}$ is again positive semi-definite for any set of points $\mathbf{x}_1, \dots, \mathbf{x}_m$ on S^{n-1} where the $G_k^{(n)}$ s are Gegenbauer polynomials.

Schoenberg also proved that if $(f(\cos \theta_{ij}))_{\substack{i=1, \dots, m \\ j=1, \dots, m}}$ is positive semi-definite for all choices of $\mathbf{x}_1, \dots, \mathbf{x}_m$ in S^{n-1} then f can be expressed as a linear combination (perhaps infinite) with non-negative coefficients of Gegenbauer polynomials.

We may define polynomials $C_k^{(n)}(t)$ by the expansion

$$(1 - 2rt + r^2)^{(2-n)/2} = \sum_{k=0}^{\infty} r^k C_k^{(n)}(t) \quad \text{for } n = 3, 4, \dots$$

We then put

$$G_k^{(n)}(t) = \frac{C_k^{(n)}(t)}{C_k^{(n)}(1)}, \quad \text{so } G_k^{(n)}(1) = 1.$$

We may also define $G_k^{(n)}(t)$ for $n = 1, 2, \dots$ recursively by the rules

$$G_0^{(n)}(t) = 1, \quad G_1^{(n)}(t) = t \quad \text{and}$$

$$G_k^{(n)}(t) = \frac{(2k+n-4)tG_{k-1}^{(n)}(t) - (k-1)G_{k-2}^{(n)}(t)}{k+n-3}$$

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

Since $(G_k^{(n)}(\cos \theta_{ij}))_{\substack{i=1, \dots, m \\ j=1, \dots, m}}$ is positive semi-definite, if a_0, \dots, a_d are non-negative real numbers then

$$(a_0 G_0^{(n)} + \dots + a_d G_d^{(n)}(\cos \theta_{ij}))$$

is also positive semi-definite. We put

$$f(n, a_0, \dots, a_d)(t) = f(t) = a_0 G_0^{(n)}(t) + \dots + a_d G_d^{(n)}(t)$$

and we define $S_f(\mathbf{x}_1, \dots, \mathbf{x}_m)$ by

$$S_f(\mathbf{x}_1, \dots, \mathbf{x}_m) = \sum_{i=1}^m \sum_{j=1}^m f(\cos \theta_{ij})$$

$$= \sum_{k=0}^d a_k \sum_{i=1}^m \sum_{j=1}^m G_k^{(n)}(\cos(\theta_{ij})).$$

Thus, since a_0, \dots, a_d are non-negative and $\sum_{i=1}^m \sum_{j=1}^m G_k^{(n)}(\cos \theta_{ij}) \geq 0$ for $k = 0, \dots, d$ we see that

$$S_f(\mathbf{x}_1, \dots, \mathbf{x}_m) \geq a_0 \sum_{i=1}^m \sum_{j=1}^m G_0^{(n)}(\cos(\theta_{ij})) = a_0 m^2. \quad (10)$$

Let us suppose now that $\mathbf{x}_1, \dots, \mathbf{x}_m$ is a configuration of m points on S^{n-1} which correspond to the m points of contact by m spheres of radius 1 which surround S^{n-1} in a kissing configuration. Then $\theta_{ij} \geq \frac{\pi}{3}$ provided that $i \neq j$ hence $\cos(\theta_{ij}) \leq \frac{1}{2}$ for $i \neq j$.

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

$$S_f(\mathbf{x}_1, \dots, \mathbf{x}_m) \leq m f(1) \quad \text{and so by (10), if } a_0 > 0,$$

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

The strategy is now to choose a_0, \dots, a_d so that $f(t) \leq 0$ for $[-1, \frac{1}{2}]$ and such that a_0 is large and $f(1)$ is small. There are two amazing applications of this approach. They were found independently in 1979 by Odlyzko and Sloan and by Levenshtein and they treat the cases $n = 8$ and $n = 24$. 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

$$f(t) = \frac{320}{3}(t+1)(t+\frac{1}{2})^2 t^2 (t-\frac{1}{2}).$$

One can check that $f(t) \leq 0$ for $[-1, \frac{1}{2}]$. Thus $\tau_8 \leq \frac{320}{3} \cdot \frac{1}{3} \cdot \frac{3^{\frac{1}{2}}}{2^2} \cdot \frac{1}{2} = 240$.

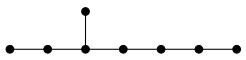
For $n = 24$ one can find a non-negative linear combination of the $G_k^{(24)}$ s to give $f(t)$ where

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

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

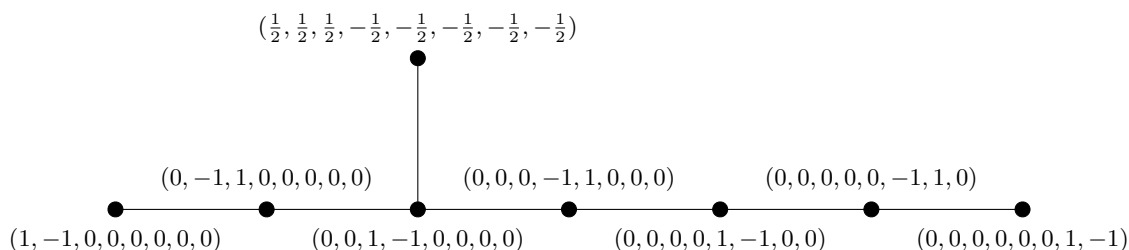
$$\tau_{24} \leq 196,560.$$

We'll show that the E_8 lattice has kissing number 240 and the Leech lattice has kissing number 196,560. Thus τ_8 and τ_{24} are determined. In general things don't go quite so smoothly. This approach gives $\tau_3 \leq 13$ and $\tau_4 \leq 25$, yet we know $\tau_3 = 12$ and $\tau_4 = 24$. The choice of a_0, \dots, a_d is made after running linear programming packages.

Let us now return to lattices. Recall E_8 has diagram . With each vector normalized to have length $\sqrt{2}$ we have that the matrix B of inner products is

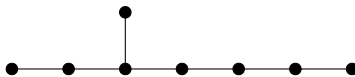
$$B = B(E_8(\sqrt{2})) = \begin{pmatrix} 2 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 2 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 2 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 2 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 2 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 2 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 2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

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



PMATH 944 Lecture 18: November 12, 2009

$E_8(\sqrt{2})$:



$\det(B(E_8(\sqrt{2}))) = 1$

All of the generating vectors have length $\sqrt{2}$. Further $\sqrt{2}$ is the minimal distance between distinct points in $E_8(\sqrt{2})$. Thus we may put a sphere of radius $\frac{\sqrt{2}}{2} = \frac{1}{\sqrt{2}}$ around each vector in the lattice. This will give us a sphere packing of \mathbb{R}^8 which is a lattice packing. Also the number of vectors in $E_8(\sqrt{2})$ of length $\sqrt{2}$ will be the kissing number of the lattice.

Notice that $2 \cdot (\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, -\frac{1}{2}, -\frac{1}{2}, -\frac{1}{2}, -\frac{1}{2}, -\frac{1}{2}) = (1, 1, 1, -1, -1, -1, -1, -1)$ is in the lattice. To fix ideas, how would we realize $(1, 1, 0, 0, 0, 0, 0, 0)$ in the lattice? Note that it suffices to realize $(1, 1, 1, -1, -1, -1, -1, -1) - (1, 1, 0, 0, 0, 0, 0, 0) - (0, 0, 1, -1, 0, 0, 0, 0) = (0, 0, 0, 0, -1, -1, -1, -1)$ or equivalently $(0, 0, 0, 0, 1, 1, 1, 1)$. But note¹⁰

$$\begin{aligned} (0, 0, 0, 0, 1, 1, 1, 1) + (1, 1, 1, -1, -1, -1, -1, -1) &= (1, 1, 1, -1, 0, 0, 0, 0) \\ (-1, -1, -1, 1, 1, 1, 1, 1) + 2(1, -1, 0, \dots, 0) - 4(0, -1, 1, 0, \dots, 0) &= (1, 1, -5, 1, 1, 1, 1, 1) \\ &= (1, 1, 0, 0, 0, 0, 0, 0) + (0, 0, -1, 1, 0, \dots, 0) + (0, 0, -1, 0, 1, 0, 0, 0) + \dots + (0, 0, -1, 0, \dots, 0, 1) \end{aligned}$$

Remark: Note that 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 the coordinates of the vector $(1, 1, 1, -1, -1, -1, -1, -1)$ is -2 hence we can realize all vectors whose sum is congruent to $0 \pmod{2}$.

¹⁰This bit caused some trouble; see the following remark instead.

Next note that $(\pm 1, \pm 1, 0, \dots, 0)$ is in the lattice and in fact so is any vector which has two coordinates from $\{1, -1\}$ and the others 0. This gives us $4 \cdot \binom{8}{2} = 112$ vectors of length $\sqrt{2}$. These vectors together with $(\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, -\frac{1}{2}, -\frac{1}{2}, -\frac{1}{2}, -\frac{1}{2}, -\frac{1}{2})$ allow us to show that the vectors $(\epsilon_1 \frac{1}{2}, \epsilon_2 \frac{1}{2}, \dots, \epsilon_8 \frac{1}{2})$ are in the lattice where ϵ_i is in $\{1, -1\}$ and $\prod_{i=1}^8 \epsilon_i = -1$. There are $2^7 = 128$ of these vectors of length $\sqrt{2}$. Thus we have found $112 + 128 = 240$ such vectors.

Notice that 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 one, then again the length exceeds $\sqrt{2}$.

Therefore $\tau_8(E_8(\sqrt{2})) = 240$ and since $\tau_8 \leq 240$ we conclude that $\tau_8 = 240$.

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

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

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

There are 240 vectors \mathbf{x} in $E_8(\sqrt{2})$ for which $\mathbf{x} \cdot \mathbf{x} = 2$. The next smallest norm in the lattice is 4 and there are 2,160 vectors \mathbf{x} in $E_8(\sqrt{2})$ for which $\mathbf{x} \cdot \mathbf{x} = 4$.

These are of the form

$$(\pm 2, 0, 0, \dots, 0), (0, \pm 2, 0, \dots, 0), \dots, (0, \dots, 0, \pm 2).$$

Also $(\pm 1, \pm 1, \pm 1, \pm 1, 0, 0, 0, 0)$ where $\pm 1, \pm 1, \pm 1, \pm 1$ is put in any 4 coordinates and

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

and all permutations of the coordinates are allowed. There are 6,720 elements of norm 6, 17,520 of norm 8, and 30,240 of norm 10. In fact for each positive integer m the number $N(m)$ of \mathbf{x} in $E_8(\sqrt{2})$ for which $\mathbf{x} \cdot \mathbf{x} = 2m$ is given by

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

How do we get such a result?

Let Λ be a lattice in \mathbb{R}^n with $\mathbf{x} \cdot \mathbf{y} \in \mathbb{Z}$ for any \mathbf{x}, \mathbf{y} in Λ . Suppose $\mathbf{b}_1, \dots, \mathbf{b}_n$ is a basis for Λ and, as before, put $B = ((\mathbf{b}_i, \mathbf{b}_j))_{\substack{i=1, \dots, n \\ j=1, \dots, n}}$. Then for any $\mathbf{x} \in \Lambda$ there exist integers k_1, \dots, k_n such that

$$\mathbf{x} = k_1 \mathbf{b}_1 + \dots + k_n \mathbf{b}_n.$$

Then

$$\begin{aligned} \mathbf{x} \cdot \mathbf{x} &= k_1^2 (\mathbf{b}_1, \mathbf{b}_1) + \dots + k_n^2 (\mathbf{b}_n, \mathbf{b}_n) \\ &= 2 \sum_{\substack{i, j \\ i < j}} k_i k_j (\mathbf{b}_i, \mathbf{b}_j) \end{aligned}$$

and so this is a quadratic form in (k_1, \dots, k_n) . We have

$$(k_1, \dots, k_n) B \begin{pmatrix} k_1 \\ \vdots \\ k_n \end{pmatrix}.$$

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

$$\theta_\Lambda(z) = \sum_{\mathbf{x} \in \Lambda} q^{(\mathbf{x} \cdot \mathbf{x})/2} = \sum_{\mathbf{x} \in \Lambda} e^{(\mathbf{x} \cdot \mathbf{x})\pi iz}.$$

If B has integer entries and determinant 1 and $\mathbf{x} \cdot \mathbf{x} \equiv 0 \pmod{2}$ for all $\mathbf{x} \in \Lambda$ then it can be proved that $\theta_\Lambda(z)$ is a modular form of weight $\frac{n}{2}$. What is the significance of this claim?

PMATH 944 Lecture 19: November 17, 2009

TALKS Tue–Fri Dec 1–4.

Recall $q = e^{2\pi iz}$. We define the theta function of a lattice Λ in \mathbb{R}^n by

$$\theta_\Lambda(z) = \sum_{\mathbf{x} \in \Lambda} q^{(\mathbf{x} \cdot \mathbf{x})/2}.$$

If B has integer entries, determinant 1 and $\mathbf{x} \cdot \mathbf{x} \equiv 0 \pmod{2}$ for all $\mathbf{x} \in \Lambda$ then θ_Λ is a modular form of weight $n/2$.

Let $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \text{SL}(2, \mathbb{Z})$ so that a, b, c and d are integers with $ad - bc = 1$. $\text{SL}(2, \mathbb{Z})$ is a group which acts on the upper half plane $H = \{z \in \mathbb{C} : \text{Im}(z) > 0\}$ by, for each $g = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \text{SL}(2, \mathbb{Z})$ we put $gz = \frac{az+b}{cz+d}$. Let k be an integer. We say that a meromorphic function $f: H \rightarrow \mathbb{C}$ is said to be weakly modular of weight $2k$ if

$$f(z) = (cz + d)^{-2k} f\left(\frac{az + b}{cz + d}\right), \quad \text{for all } \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \text{SL}(2, \mathbb{Z}).$$

Note that if $g = \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ then $gz = z + 1$ and so if f is weakly modular of weight $2k$, $f(z + 1) = f(z)$ and so f can be expressed in terms of $q = e^{2\pi iz}$. In particular f determines a function $\tilde{f}(q)$ where

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

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

The space of modular forms of weight $2k$ ($k \geq 0$), forms a vector space M_{2k} over \mathbb{C} of dimension d_k where

$$d_k = \begin{cases} \left\lfloor \frac{k}{6} \right\rfloor, & k \equiv 1 \pmod{6}, k \geq 0 \\ \left\lfloor \frac{k}{6} \right\rfloor + 1, & k \not\equiv 1 \pmod{6}, k \geq 0 \end{cases}$$

Here $\lfloor x \rfloor$ denotes the greatest integer less than or equal to x .

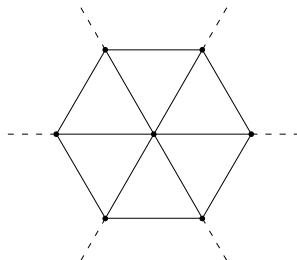
The lattice $\Lambda = E_8(\sqrt{2})$ determines $\theta_{E_8(\sqrt{2})}(z)$ which is a modular form of weight 4. Thus $\theta_{E_8(\sqrt{2})}(z)$ lies in M_4 a vector space of dimension 1 over \mathbb{C} . Now $E_2(z) = 1 + 240 \sum_{n=1}^{\infty} \sigma_3(n)q^n$ is in M_4 . We have

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

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

Associated to each lattice Λ in \mathbb{R}^n is $\text{Aut}(\Lambda)$, the group of symmetries of the lattice which fix the origin or equivalently the set of isometries (distance preserving maps) of \mathbb{R}^n which fix the origin and take the lattice to itself. For each lattice Λ in \mathbb{R}^n , $\text{Aut}(\Lambda)$ is a finite group. Each element of the group can be represented by an orthogonal matrix.

The automorphism group of the hexagonal lattice A_2



is generated by a rotation of $\frac{\pi}{3}$ and a reflection around the line determined by any non-zero vector. Thus it is isomorphic to the Dihedral group D_6 .

The automorphism group of E_8 ($E_8(\sqrt{2})$) 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.

We'll now construct an astonishing combinatorial object called the Leech lattice. It was found by Leech in 1965 and it was described by him in a paper in the Canadian Journal of Math in 1967. It is a lattice L in \mathbb{R}^{24} with determinant 1, the associated inner product matrix B has integer entries. The polar lattice L^* of L is L , in other words L is self-dual. Further if $\mathbf{x} \in L$ and $\mathbf{x} \neq \mathbf{0}$ then

$$\mathbf{x} \cdot \mathbf{x} \geq 4.$$

We'll now construct the Leech lattice L following Leech and Milnar.

Let \mathbb{F}_2^{24} be the 24 dimensional vector space over the field $\mathbb{F}_2 = \{0, 1\}$ of two elements.

Proposition 28: There exists a 12 dimensional subspace S of \mathbb{F}_2^{24} with the following property. For every non-zero vector $\mathbf{s} = (s_1, \dots, s_{24})$ in S the number of coordinates which are 1 is at least 8 and is congruent to 0 mod 4. Further $(1, 1, \dots, 1)$ is in S .

To prove this we'll realize S as the span of the rows of a 12×24 matrix over \mathbb{F}_2 which we shall construct. Let A denote a symmetric 11×11 matrix whose first row is

$$11101101000$$

and whose remaining rows are obtained by permuting the rows cyclically to the left so

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

One may check that each pair of rows have exactly three columns consisting of two 1s.

Next let B be the symmetric 12×12 matrix obtained by adjoining a first row of the form $(011 \dots 1)$ to A and completing the first column to be $(011 \dots 1)^{\text{tr}}$ also. Thus

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

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

$$B^2 = BB^{\text{tr}} = I.^{11)}$$

We now put

$$C = (I_{12} \mid B) \quad \text{a } 12 \times 24 \text{ matrix}$$

We claim that S is the subspace of \mathbb{F}_2^{24} generated by the rows of C .

¹¹⁾over \mathbb{F}_2

A generator matrix for the Leech lattice L , in terms of the standard basis. Notice the index of L as a sublattice of $\{\frac{1}{\sqrt{8}}\mathbf{e}_1, \dots, \frac{1}{\sqrt{8}}\mathbf{e}_{24}\}$ is 2^{36} .

Suppose that there is an element $\mathbf{x} \in L$ with $\mathbf{x} \cdot \mathbf{x} = 2$. Write

$$\mathbf{x} \cdot \mathbf{x} = t_1\mathbf{b}_1 + \dots + t_{24}\mathbf{b}_{24}^{12}), \quad t_i \in \mathbb{Z},$$

and then

$$\mathbf{x} \cdot \mathbf{x} = \frac{1}{8}(t_1^2 + \dots + t_{24}^2) = 2$$

hence

$$t_1^2 + \dots + t_{24}^2 = 16.$$

Notice that if the t_i are all odd then $t_1^2 + \dots + t_{24}^2 > 16$. Thus t_1, \dots, t_{24} are even. We have two possibilities, one of the t_i s is 4 and the others are 0 or four of the t_i s are 2 and the others are 0. But $t_1 + \dots + t_{24} \equiv 0 \pmod{8}$ which excludes the first possibility. The second possibility is excluded by Proposition 28 since the number of terms which are $\equiv 2 \pmod{4}$ is either 0 or at least 8. Therefore there is no $\mathbf{x} \in L$ for which $\mathbf{x} \cdot \mathbf{x} = 2$.

Thus $\mathbf{x} \cdot \mathbf{x} \geq 4$ for $\mathbf{x} \neq 0$ in L .

Next observe that for any $\mathbf{x}, \mathbf{y} \in L$

$$\mathbf{x} \cdot \mathbf{y} = \frac{1}{2}((\mathbf{x} + \mathbf{y}) \cdot (\mathbf{x} + \mathbf{y}) - \mathbf{x} \cdot \mathbf{x} - \mathbf{y} \cdot \mathbf{y})$$

and since $\mathbf{z} \cdot \mathbf{z} \equiv 0 \pmod{2}$, for all $\mathbf{z} \in L$ we see that $\mathbf{x} \cdot \mathbf{y} \in \mathbb{Z}$ for all $\mathbf{x}, \mathbf{y} \in L$.

L is a sublattice of L_0 .

We can calculate the index of L in L_0 . It is $8 \cdot 4^{11} \cdot 2^{11} \cdot 1 = 2^3 \cdot 2^{22} \cdot 2^{11} \cdot 1 = 2^{36}$. But $d(L_0) = \left(\frac{1}{\sqrt{8}}\right)^{24} = \frac{1}{2^{(3/2) \cdot 24}} = \frac{1}{2^{36}}$ and so $d(L) = 1$.

Recall the notion of the polar lattice or dual lattice of a lattice Λ . Suppose $\mathbf{b}_1, \dots, \mathbf{b}_n$ is a basis for Λ . Define $\mathbf{b}_1^*, \dots, \mathbf{b}_n^*$ so that

$$\mathbf{b}_i^* \cdot \mathbf{b}_j = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}.$$

Then $\mathbf{b}_1^*, \dots, \mathbf{b}_n^*$ is a basis for the polar lattice Λ^* of Λ .

Recall:

Theorem 2: The polar lattice Λ^* of a lattice Λ in \mathbb{R}^n consists of all vectors \mathbf{v}^* in \mathbb{R}^n for which $\mathbf{v}^* \cdot \mathbf{v}$ is an integer for all \mathbf{v} in Λ . In addition $d(\Lambda) \cdot d(\Lambda^*) = 1$.

What is the polar lattice of the Leech lattice L ? Note that since $d(L) = 1$ we have $d(L^*) = 1$. Further $\mathbf{x} \cdot \mathbf{y}$ is an integer for all \mathbf{x}, \mathbf{y} in L . Thus L^* contains L and, since $d(L) = d(L^*)$ we see that $L^* = L$. Thus the Leech lattice is self-dual.

The theta series $\theta_L(z)$ associated with the Leech lattice is a modular form of weight $\frac{24}{2} = 12$. The vector space of modular forms of weight 12 has dimension 2 over \mathbb{C} . With $q = e^{2\pi iz}$ we have

$$\theta_L(z) = \sum_{n=0}^{\infty} N(n)q^n = 1 + 196,560q^4 + 16773120q^6 + \dots$$

In fact, for even positive integers m ,

$$N(m) = \frac{65,520}{691} \left(\sigma_{11} \left(\frac{m}{2} \right) - \tau \left(\frac{m}{2} \right) \right), \quad (16)$$

¹²⁾ $\mathbf{b}_i = \frac{1}{\sqrt{8}}\mathbf{e}_i$

where $\sigma_{11}(n) = \sum_{d|n} d^{11}$ and $\tau(n)$ is Ramanujan's tau function defined by

$$\begin{aligned}\Delta_{24}(z) &= q \prod_{m=1}^{\infty} (1 - q^m)^{24} = \sum_{m=0}^{\infty} \tau(m) q^m \\ &= q - 24q^2 + 252q^3 - 1472q^4 + \dots\end{aligned}$$

Thus $N(4) = \frac{65,520}{691} (2^{11} + 1 + 24) = 196,560$.

Since $N(2k)$ is an integer for $k = 1, 2, \dots$ we see that $\sigma_{11}(k) \equiv \tau(k) \pmod{691}$ for $k = 1, 2, \dots$

Another representation for $\theta_L(z)$ is

$$\theta_L(z) = (\theta_{E_8(\sqrt{2})}(z))^3 - 720\Delta_{24}(z).$$

Examining coefficients in the above representation yields (16).

Since $N(4) = 196,560$ we see that this is the kissing number of L . But by our earlier analysis we see that it is the kissing number of \mathbb{R}^{24} so $\tau_{24} = 196,560$.

What are the vectors of minimal length in L ? There are $2^7 \cdot 759 = 97,152$ of the form

$$\frac{1}{\sqrt{8}}(\underbrace{\pm 2, \dots, \pm 2}_8, 0, \dots, 0)$$

where any choice of sign and position is permitted provided that $t_1 + \dots + t_{24} \equiv 0 \pmod{8}$ and $\frac{1}{2}(t_1, \dots, t_{24})$ reduced mod 2 is in S . There are $2^{12} \cdot 24 = 98,304$ of the form $\frac{1}{\sqrt{8}}(\pm 3, \pm 1, \dots, \pm 1)$ where $t_1 + \dots + t_{24} \equiv 4 \pmod{8}$ and $\frac{1}{2}(1 + t_1, \dots, 1 + t_{24})$ reduced mod 2 is in S . Also there are $4 \cdot \binom{24}{2} = 1104$ of the form $(\pm 4, \pm 4, 0, \dots, 0)$ where any choice of sign and position is permitted.

Notice that the sphere packing density in \mathbb{R}^{24} associated with L is

$$\frac{\pi^{12}}{12!} = 0.001930 \dots$$

The automorphism group of L is Co_0 or the 0th Conway group and it is of order $2^{22} \cdot 3^9 \cdot 5^4 \cdot 7^2 \cdot 11 \cdot 13 \cdot 23$.

The automorphism group permutes the 196,560 non-zero vectors of minimal length transitively. Further the automorphism modulo its centre is a sporadic simple group of order $2^{21} \cdot 3^9 \cdot 5^4 \cdot 7^2 \cdot 11 \cdot 13 \cdot 23$. There are 26 sporadic simple groups.

PMATH 944 Lecture 22: November 26, 2009

Sporadic simple groups

A group is said to be *simple* if it has no proper normal subgroup. Why are simple groups important?

Every finite group has a composition series

$$G \triangleleft G_1 \triangleleft G_2 \triangleleft \dots \triangleleft G_n = \{1\}$$

where G_{i+1} is a normal subgroup of G_i for $i = 1, \dots, n-1$ and G_i/G_{i+1} is simple. Jordan and Holder proved that the set of groups G_i/G_{i+1} for $i = 1, \dots, n-1$ is uniquely determined or equivalently the sequence $(G_i/G_{i+1})_{i=1}^{n-1}$ is determined up to permutation. Thus the simple groups are the building blocks or "primes" of the set of finite groups.

If G is an abelian finite simple group then G is isomorphic to $\mathbb{Z}/p\mathbb{Z}$ for some prime p .

The alternating groups A_n are simple for $n \neq 4$. In fact there are several infinite families of simple groups which have been found. In addition there are 26 finite simple groups which do not fit in these families and they are known as the sporadic simple groups. The Classification Theorem tells us this is the complete list.

12 of the 26 sporadic simple groups arise as subquotients of M . Further the largest of the sporadic simple groups M is known as the *Monster* or the Fischer–Griess group or the Friendly Giant. M can be constructed from the Leech lattice, it was first discovered by Fischer and Griess in 1973 and formally constructed in 1980 by Griess. The order of M is

$$2^{46} \cdot 3^{20} \cdot 5^9 \cdot 7^6 \cdot 11^2 \cdot 13^3 \cdot 17 \cdot 19 \cdot 23 \cdot 29 \cdot 31 \cdot 41 \cdot 47 \cdot 59 \cdot 71$$

Monstrous Moonshine: Conway and Norton. (Borcherds)

A rich source of lattices is algebraic number theory. Let $K = \mathbb{Q}(\theta)$ with $[K : \mathbb{Q}] = n$. Let $\alpha \in K$ then the conjugates of α over \mathbb{Q} are $\alpha = \alpha_1, \alpha_2, \dots, \alpha_n$ where

$$\sigma_i(\alpha) = \alpha_i \quad \text{for } i = 1, \dots, n$$

and σ_i is one of the n isomorphisms of K into \mathbb{C} which fix \mathbb{Q} . We have the notion of the norm and trace of α given by

$$N_{K/\mathbb{Q}}(\alpha) = \alpha_1 \cdots \alpha_n$$

and

$$\text{Tr}_{K/\mathbb{Q}}(\alpha) = \alpha_1 + \cdots + \alpha_n.$$

The embeddings (isomorphic injection) of K into \mathbb{C} which fix \mathbb{Q} can be split into r embeddings into \mathbb{R} and $2s$ embeddings into \mathbb{C} which are not embeddings in \mathbb{R} . We may denote them by $\sigma_1, \dots, \sigma_r$ and $\sigma_{r+1}, \dots, \sigma_{r+2s}$ where $\sigma_{r+i} = \overline{\sigma_{r+s+i}}$ for $i = 1, \dots, s$. We introduce the map $\nu: K \rightarrow \mathbb{R}^n$ by

$$\nu(\alpha) = (\sigma_1(\alpha), \dots, \sigma_r(\alpha), \text{Re}(\sigma_{r+1}(\alpha)), \text{Im}(\sigma_{r+1}(\alpha)), \dots, \text{Re}(\sigma_{r+s}(\alpha)), \text{Im}(\sigma_{r+s}(\alpha))).$$

Let \mathcal{O}_K be the ring of algebraic integers of K . We can show that $\{\nu(\alpha) : \alpha \in \mathcal{O}_K\}$ forms a lattice in \mathbb{R}^n and if A is a non-zero ideal of \mathcal{O}_K then $\{\nu(\alpha) : \alpha \in A\}$ is a sublattice of this lattice.

Let us consider the totally real case where $r = n$. Then

$$\nu(\alpha) \cdot \nu(\alpha) = \alpha_1^2 + \cdots + \alpha_n^2 = \text{Tr}_{K/\mathbb{Q}}(\alpha^2).$$

Further

$$\frac{\alpha_1^2 + \cdots + \alpha_n^2}{n} \geq (\alpha_1^2 \cdots \alpha_n^2)^{1/n}$$

by the arithmetic–geometric mean inequality. If α is a non-zero algebraic integer then so is α^2 hence $\alpha_1^2 \cdots \alpha_n^2$, which is $N_{K/\mathbb{Q}}(\alpha^2) = (N_{K/\mathbb{Q}}(\alpha))^2$, is a positive integer. Therefore if α is a non-zero algebraic integer

$$\nu(\alpha) \cdot \nu(\alpha) \geq n$$

and this gives us a way to show that the minimal length of a non-zero vector in the lattice is large.

It is possible to realize many lattices in this way. For example the Leech lattice can be realized by considering $K = \mathbb{Q}(\zeta_{39})$.