3. The Geometry of Numbers

The, so-called, geometry of numbers, in its simplest form, is concerned with the following kind of thing. Suppose that S is a measurable subset of $\mathbb{R}^n$ with some reasonable properties as regarding its general shape. Then can we make deductions along the lines that if the (n-dimensional) measure is large enough, then S contains at least one point of $\mathbb{Z}^n$? Alternatively, we might insist that $\mathbf{0} \in S$ and ask for another point of $\mathbb{Z}^n$. If we can for reasonable S, then it is natural to consider also what happens when $\mathbb{Z}^n$ is replaced by its linear transformations. The set $\mathbb{Z}^n$ and its generalizations are usually thought of as "lattice points".

There is a simple lemma which makes a good starting point.

Lemma 3.1. . *Let $f(\mathbf{x})$ be a non-negative integrable function on $\mathbb{R}^n$ with*

$$\int_{\mathbb{R}^n} f(\mathbf{x})d\mathbf{x} < \infty.$$

Then

$$\int_{\mathbb{R}^n} f(\mathbf{x})d\mathbf{x} \leq \sup_{\mathbf{y}\in\mathbb{R}^n} \sum_{\mathbf{k}\in\mathbb{Z}^n} f(\mathbf{k}+\mathbf{y}).$$

(Here the integrals can be either Riemann or Lebesque)

Proof. We may suppose that the series on the right is uniformly bounded in y, since otherwise the conclusion is trivial. Now we have

$$\begin{aligned}\int_{\mathbb{R}^n} f(\mathbf{x})d\mathbf{x} &= \sum_{\mathbf{k}\in\mathbb{Z}^n} \int_{[0,1)^n} f(\mathbf{k}+\mathbf{x})d\mathbf{x} \\ &= \int_{[0,1)^n} \sum_{\mathbf{k}\in\mathbb{Z}^n} f(\mathbf{k}+\mathbf{x})d\mathbf{x}\end{aligned}$$

where the interchange is justified by dominated convergence since the series is uniformly bounded. The lemma follows at once because the supremum of the integrand must be at least as large as its average.

Theorem 3.2 (Blichfeldt's principle, 1914). *Let $\mathcal{S}$ be a measurable set in $\mathbb{R}^n$ with $\mu(\mathcal{S}) > 1$. Then there exist distinct $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{R}^n$ such that $\mathbf{x}_1, \mathbf{x}_2 \in \mathcal{S}$ and $\mathbf{x}_2 - \mathbf{x}_1 \in \mathbb{Z}^n$.*

Proof. Let f be the characteristic function of $\mathcal{S}$. By Lemma 3.1 there exists an $\mathbf{y}$ such that $\sum f(\mathbf{k}+\mathbf{y}) > 1$. For such a $\mathbf{y}$ there exist distinct $\mathbf{k}_1, \mathbf{k}_2 \in \mathbb{Z}^n$ such that $\mathbf{k}_1 + \mathbf{x}, \mathbf{k}_2 + \mathbf{x} \in \mathcal{S}$. Put $\mathbf{x}_j = \mathbf{k}_j + \mathbf{x}$. Then the theorem follows.

Now we can state and prove the first main theorem of the geometry of numbers.

Theorem 3.3 (Minkowski's convex body theorem). *Let $\mathcal{C}$ be a convex body in $\mathbb{R}^n$ which is symmetric about $\mathbf{0}$. If $\mathrm{vol}(\mathcal{C}) > 2^n$, then there is a $\mathbf{k} \in \mathbb{Z}^n$, $\mathbf{k} \neq \mathbf{0}$, such that $\mathbf{k} \in \mathcal{C}$.*

Proof. Let $\mathcal{S} = \frac{1}{2}\mathcal{C}$, i.e., $\mathcal{S} = \{\mathbf{x} : 2\mathbf{x} \in \mathcal{C}\}$. Then $\mathrm{vol}(\mathcal{S}) > 1$. By Blichfeldt's principle there exist $\mathbf{x}_1, \mathbf{x}_2 \in \mathcal{S}$ such that $\mathbf{x}_1 \neq \mathbf{x}_2$, $\mathbf{x}_1 - \mathbf{x}_2 \in \mathbb{Z}^n$. Since $\mathcal{S}$ is symmetric about $\mathbf{0}$ we also have $-\mathbf{x}_2 \in \mathcal{S}$. Since $\mathcal{S}$ is convex, the line joining $\mathbf{x}_1$ to $-\mathbf{x}_2$ lies in $\mathcal{S}$. That is, $\lambda\mathbf{x}_1 - (1-\lambda)\mathbf{x}_2 \in \mathcal{S}$ whenever $0 \leq \lambda \leq 1$. In particular, the midpoint (given by $\lambda = \frac{1}{2}$) lies in $\mathcal{S}$. That is $\frac{1}{2}\mathbf{x}_1 - \frac{1}{2}\mathbf{x}_2 \in \mathcal{S}$, which is to say $\mathbf{x}_1 - \mathbf{x}_2 \in 2\mathcal{S} = \mathcal{C}$. Thus $\mathbf{k} = \mathbf{x}_1 - \mathbf{x}_2$ has the desired properties.

Observe that the condition $\mathrm{vol}(\mathcal{C}) > 2^n$ cannot be weakened since the the hypersquare $\{\mathbf{x} \in \mathbb{R}^n : |x_j| < 1\}$ has volume 2^n but only shares $\mathbf{0}$ with $\mathbb{Z}^n$.

Let A be a $n \times n$ matrix with real elements and column vectors $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$, and suppose that these n vectors are linearly independent. Then every point in $\mathbb{R}^n$ is uniquely of the form

$$A\mathbf{x} = x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_n\mathbf{a}_n,$$

with $x_j \in \mathbb{R}$. The points $A\mathbf{k} = k_1\mathbf{a}_1 + k_2\mathbf{a}_2 + \cdots + k_n\mathbf{a}_n$ for which the k_j are integral consitute a *lattice*. The vectors $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$ are a *basis* for the lattice. Taking A to be the identity matrix we see that $\mathbb{Z}^n$ is a lattice, called the *lattice of integral points*, or *integer lattice*. It is useful to determine when one lattice is a sublattice (i.e. a subset) of another.

Theorem 3.4. *Let A and B be nonsingular $n \times n$ matrices, and put $\Lambda_1 = A\mathbb{Z}^n$, $\Lambda_2 = B\mathbb{Z}^n$. Then $\Lambda_2 \subset \Lambda_1$ if and only if B is of the form $B = AK$ where K is an $n \times n$ matrix with integral elements.*

Proof. Put $K = A^{-1}B$ and suppose that K has integral elements. If $\mathbf{x} \in \mathbb{Z}^n$, then also $K\mathbf{x} \in \mathbb{Z}^n$. That is, $K\mathbb{Z}^n \subset \mathbb{Z}^n$, so that $B\mathbb{Z}^n = (AK)\mathbb{Z}^n = A(K\mathbb{Z}^n) \subset A\mathbb{Z}^n$.

Suppose conversely that $\Lambda_2 \subset \Lambda_1$. Then each column $\mathbf{b}_j$ of B equals $= B\mathbf{e}_j$ where $\mathbf{e}_j$ is the jth column of the identity matrix and so is in Λ_2. Therefore it can be written as

$$\mathbf{b}_j = k_{1j}\mathbf{a}_1 + k_{2j}\mathbf{a}_2 + \cdots + k_{nj}\mathbf{a}_n$$

where the k_{ij} are integers and $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$ are the columns of A. Therefore $B = AK$.

There is a special subclass of matrices with integral elements, those which are invertible and whose inverse also has integral elements. It is obvious, by considering the adjoint, for example, that to be sure that the inverse also has integral matrices we should restrict our attention to those for which the determinant is ± 1. Such a matrix is called *unimodular* and can be characterised as follows.

Theorem 3.5. *Let U be a $n \times n$ matrix with integral elements. The following assertions are equivalent.*
(i) U is unimodular.
(ii) U is a product of elementary row matrices, each one unimodular.
(iii) The inverse matrix U^{-1} exists and has integral elements.

We remark that the set of all $n \times n$ unimodular matrices forms a group. To see this observe that if U and V are unimodular, then $\det(UV) = \det(U)\det(V) = \pm 1$ and that by (iii) U^{-1} is unimodular.

Proof. Let A be an $n \times n$ matrix with integral elements. We apply row operations to A in the maner of Gaussian elimination, but using only *integer* arithmetic, until we reach a triangular matrix T. Expressed in matrix form we have $E_k E_{k-1} \cdots E_1 A = T$ where the E_i are elementary unimodular matrices. Thus $\det(A) = \det(T)$, so that A is unimodular if and only if T is. But T is unimodular if and only if all its diagonal elements are ± 1, in which case further elementary row operations reduce T to the identity matrix. Thus (i) and (ii) are equivalent.

We now advert to (iii). Note that if (i), and so (ii), hold then U is a product of elementary matrices $E_1 E_2 \cdots E_r$ and so $E_r^{-1} \cdots E_2^{-1} E^{-1}$ exists, has integral coefficients and is the inverse of U.

On the other hand, if U^{-1} exists and has integral elements, then $\det(U^{-1})$ is an integer, and hence $\det(U)$ and $\det(U^{-1})$ are both integers whose product is $\det(UU^{-1}) = 1$ and so U is unimdoular.

Let U denote a unimodular matrix. Obviously $U\mathbb{Z}^n \subset \mathbb{Z}^n$. Moreover for any point $\mathbf{z}$ of $\mathbb{Z}^n$, let $\mathbf{w} = U^{-1}\mathbf{z}$. Since the inverse also has inegral entries we have $\mathbf{w} \in \mathbb{Z}^n$ and so $U\mathbf{w} = \mathbf{z}$. Thus $\mathbb{Z}^n \subset U\mathbb{Z}^n$. Hence $U\mathbb{Z}^n = \mathbb{Z}^n$.

Theorem 3.6. *Put $\Lambda_1 = A\mathbb{Z}^n$, $\Lambda_2 = B\mathbb{Z}^n$ where A and B are non-singular. Then $\Lambda_1 = \Lambda_2$ if and only if $A^{-1}B$ is unimodular.*

At once from the theorem, $U\mathbb{Z}^n = \mathbb{Z}^n$ if and only if U is unimodular. Whenever U is unimodular, the columns of U form a basis for $\mathbb{Z}^n$, and vice versa.

Proof. If $A^{-1}B$ is unimodular, then we have $A^{-1}B\mathbb{Z}^n = \mathbb{Z}^n$ and so $\Lambda_1 = \Lambda_2$. On the other hand, if $\Lambda_1 = \Lambda_2$, then by Theorem 3.4, both $A^{-1}B$ and $B^{-1}A$ have integer entries and so have integer determinants. But the product of their determinants is 1 so they are both unimodular.

Let A be a non-singular $n \times n$ matrix, so that $\Lambda = A\mathbb{Z}^n$ is a lattice Λ. The *determinant* of Λ, $d(\Lambda)$, is $|\det(A)|$. Since the basic vectors $\mathbf{e}_k$ are in $\mathbb{Z}^n$, the column vectors of A are in Λ. Moreover, for each $\boldsymbol{\lambda} \in \Lambda$ there is a $\mathbf{z} \in \mathbb{Z}^n$ such that

$$\boldsymbol{\lambda} = z_1\mathbf{a}_1 + \cdots + z_n\mathbf{a}_n$$

where the $\mathbf{a}_j$ are the columns of A. Thus we say that the columns of A form a basis for Λ. A lattice may have many bases, but nevertheless the determinant is

well defined, for if we write $\Lambda = B\mathbb{Z}^n$, then $B = AU$ with U unimodular and so $|\det(B)| = |\det(A)|$.

Let A be a nonsingular $n \times n$ matrix with real elements. The linear map $\mathbf{x} \mapsto A\mathbf{x}$ preserves lines. Hence $\mathcal{C}$ is convex if and only if $A\mathcal{C}$ is, and $\mathcal{S}$ is symmetric about $\mathbf{0}$ if and only if $A\mathcal{S}$ is. Moreover, if $\mathcal{S}$ is a measurable set, then so is $A\mathcal{S}$, and $\mathrm{vol}(A\mathcal{S}) = |\det(A)|\mathrm{vol}(\mathcal{S})$. Lemma 3.1 and Theorems 3.2 and 3.3 can be generalised by replacing $\mathbb{Z}^n$ by an arbitrary lattice Λ. This may be accomplished by repeating the proofs already given, or by systematically applying the linear transformation $\mathbf{x} \mapsto A\mathbf{x}$. In either case the results are as follows.

Lemma 3.7. . *Let $f(\mathbf{x})$ be a non-negative integrable function on $\mathbb{R}^n$ with*

$$\int_{\mathbb{R}^n} f(\mathbf{x})d\mathbf{x} < \infty$$

and let Λ be a lattice in $\mathbb{R}^n$. Then there is a $\mathbf{y}$ such that

$$\sum_{\boldsymbol{\lambda}\in\Lambda} f(\boldsymbol{\lambda} + \mathbf{y}) \geq \frac{1}{d(\Lambda)} \int_{\mathbb{R}^n} f(\mathbf{x})d\mathbf{x}.$$

Theorem 3.8 (Blichfeldt's principle, 1914, II). *Let Λ be a lattice in $\mathbb{R}^n$, and let $\mathcal{S}$ be a measurable set in $\mathbb{R}^n$ with $\mu(\mathcal{S}) > d(\Lambda)$. Then there exist distinct $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{R}^n$ such that $\mathbf{x}_1, \mathbf{x}_2 \in \mathcal{S}$ and $\mathbf{x}_2 - \mathbf{x}_1 \in \Lambda$.*

For the conclusion here we sometimes write $\mathbf{x}_1 \equiv \mathbf{x}_2 \pmod{\Lambda}$.

Theorem 3.9 (Minkowski's convex body theorem, II). *Let Λ be a lattice in $\mathbb{R}^n$, and let $\mathcal{C}$ be a convex body in $\mathbb{R}^n$ which is symmetric about $\mathbf{0}$. If $\mathrm{vol}(\mathcal{C}) > 2^n d(\Lambda)$, then there is a $\boldsymbol{\lambda} \in \Lambda$, $\boldsymbol{\lambda} \neq \mathbf{0}$, such that $\boldsymbol{\lambda} \in \mathcal{C}$.*

There is a variant of this which can be stated as follows

Theorem 3.10 (Minkowski's convex body theorem, III). *Let Λ be a lattice in $\mathbb{R}^n$, and let $\mathcal{C}$ be a convex body in $\mathbb{R}^n$ which is symmetric about $\mathbf{0}$. If $\mathcal{C}$ is closed and $\mathrm{vol}(\mathcal{C}) \geq 2^n d(\Lambda)$, then there is a $\boldsymbol{\lambda} \in \Lambda$, $\boldsymbol{\lambda} \neq \mathbf{0}$, such that $\boldsymbol{\lambda} \in \mathcal{C}$.*

Proof. First suppose that $\mathcal{C}$ is unbounded. Since $\mathrm{vol}(\mathcal{C}) > 0$, it follows that $\mathcal{C}$ has non-empty interior, and hence $\mathrm{vol}(\mathcal{C}) = \infty$, when the result is immediate from the previous theorem.

Now suppose that $\mathcal{C}$ is bounded. Then $\mathcal{C}$ is compact. Let ε_k be a sequence of positive real numbers tending monotonically to 0. Put $\mathcal{C}_k = (1 + \varepsilon_k)\mathcal{C}$. Then $\mathrm{vol}(\mathcal{C}_k) = (1 + \varepsilon_k)\mathrm{vol}(\mathcal{C}) > 2^n d(\Lambda)$. Thus, by Theorem 3.9, there is a point $\boldsymbol{\lambda}_k \in \Lambda$, $\boldsymbol{\lambda}_k \neq \mathbf{0}$, $\boldsymbol{\lambda}_k \in \mathcal{C}_k$. Since all the points λ_k lie in the compact set $\mathcal{C}_1$ there is a subsequence which converges. Since the points are discrete, it follows that the corresponding limit $\boldsymbol{\lambda}$ point occurs in the sequence infinitely many times. Since $\boldsymbol{\lambda} \in \mathcal{C}_k$ for arbitrarily large k, we conclude from the compactness of $\mathcal{C}$ that $\boldsymbol{\lambda} \in \mathcal{C}$.

There is a simple application of Minkowski's Convex Body Theorem which has many uses.

Theorem 3.11 (Minkowski's linear forms theorem). *Let $A = [a_{ij}]$ be an $n \times n$ matrix with real elements. Let $c_1, c_2, \dots, c_n$ be positive real numbers such that $c_1 c_2 \cdots c_N > |\det(A)|$. Then there exist integers $x_1, \dots, x_n$, not all 0, such that*

$$\left| \sum_{j=1}^{n} a_{ij} x_j \right| < c_i \quad (1 \le i \le n).$$

Proof. Let $\mathcal{C} = \{\mathbf{x} \in \mathbb{R}^n : |\sum a_{ij} x_j| < c_j,\ 1 \le i \le n\}$. If $\det(A) = 0$, then $\mathcal{C}$ is unboundend and $\mathrm{vol}(\mathcal{C}) = \infty$. Otherwise $\mathrm{vol}(\mathcal{C}) = \frac{2^n c_1 c_2 \cdots c_n}{|\det(A)|}$. Now we wish to find solutions in the integer lattice. Thus the result follows from Minkowski's convex body theorem (Theorem 3.9) with $\Lambda = \mathbb{Z}^n$, so that $d(\Lambda) = 1$.

As in Theorem 3.10, we can weaken the hypothesis to $c_1 c_2 \cdots c_n \ge |\det(A)|$ provided that one of the inequalities is weakened, say the last one. This we can do by a limiting argument, based on replacing c_n, say, by $(1 + \varepsilon_k) c_n$, analoguous to that of the proof of Theorem 3.10.

Theorem 3.12 (Minkowski's linear forms theorem, II). *Let $A = [a_{ij}]$ be an $n \times n$ matrix with real elements. Let $c_1, c_2, \dots, c_n$ be positive real numbers such that $c_1 c_2 \cdots c_n \ge |\det(A)|$. Then there exist integers $x_1, \dots, x_n$, not all 0, such that*

$$\left| \sum_{j=1}^{n} a_{ij} x_j \right| < c_i \quad (1 \le i \le n-1)$$

and

$$\left| \sum_{j=1}^{n} a_{nj} x_j \right| \le c_n$$

We can apply Theorem 3.12 to give another proof of Dirichlet's Theorem, Theorem 1.1. We take $n = 2$, $c_1 = 1/(Q+1)$, $c_2 = Q+1$ and consider the linear forms q and $\alpha q - a$, so that $A = \begin{bmatrix} 1 & 0 \\ \alpha & -1 \end{bmatrix}$. Then there are integers a and q, not both 0, such that $|q| < Q+1$ and $|\alpha q - a| \le 1/(Q+1)$. But $q = 0$ implies that $a = 0$, so $q \ne 0$. Moreover, if $q < 0$, then $-a$, $-q$ is also a solution. This gives Theorem 1.1 again.

In the alternative notation, the natural metric on $\mathbb{R}/\mathbb{Z}$, $\|\theta\| = \min_{n \in \mathbb{Z}} |\theta - n|$, this asserts that $\|q\alpha\| \le 1/(Q+1)$ for some q with $1 \le q \le Q$.

Minkowski's linear forms theorem provides a convenient way of establishing generalisations of Theorem 1.1 to simultaneous approximations. Let $\alpha_1, \dots, \alpha_m$ be real numbers and Q a positive integer. Then there is a q, $1 \le q \le Q$ such that $\|q\alpha_i\| < Q^{-1/m} (1 \le i \le m)$. This is a special case of the following general result.

Theorem 3.13. *Let A be an $m \times n$ matrix with real elements, and let Q be a real number, $Q \geq 1$. Then there exist integers $q_1, \ldots, q_n$, not all 0, such that*

$$\left\| \sum_{j=1}^{n} a_{ij} q_j \right\| < Q^{-n/m} \quad (1 \leq i \leq m)$$

and $|q_j| \leq Q$.

Proof. We apply Minkowski's linear forms theorem II, Theorem 3.12, to the $(m+n) \times (m+n)$ matrix $\begin{bmatrix} A & I_m \\ I_n & 0 \end{bmatrix}$, whose determinant is ± 1, and take $c_1 = \cdots = c_m = Q^{-n/m}$ and $c_{m+1} = \cdots = c_{m+n} = Q$. The corresponding linear forms are $\sum_{j=1}^{n} a_{ij} q_j + a_i$ $(1 \leq i \leq m)$ and q_j $(1 \leq j \leq n)$. If all the q_j were to be zero, then all the a_i would be zero. Thus at least one q_j is non-zero.

As in the case when $m = n = 1$, it is possible to show that for suitable configurations of algebraic numbers this theorem cannot, in general, be improved apart from constants. We now take a detour from the geometry of numbers to establish this.

Theorem 3.14. *For any positive integers m and n there exists a constant $c > 0$ and an $m \times n$ matrix A with real entries such that*

$$\max_{1 \leq i \leq m} \left\| \sum_{j=1}^{n} a_{ij} q_j \right\| \geq c \left(\max_{1 \leq j \leq n} |q_j| \right)^{-n/m}$$

whenever the q_j are integers not all zero.

Proof. Let $l = m+n$, so that $l > 1$, and let $\alpha_1, \alpha_2, \ldots, \alpha_l$ be a set of real conjugate algebraic integers. Such algebraic integers may be constructed in a variety of ways. For example, let p be a large prime and put $P(z) = \prod_{k=1}^{l} (z - kp) + p$. Then $P(z)$ is irreducible by Eisenstein's criterion. As $P(z)$ is monic with integer coefficients its roots form a set of conjugate algebraic numbers. Moreover

$$P\left(\frac{2l+1}{2}p\right) > 0, \;\; P\left(\frac{2l-1}{2}p\right) < 0, \;\; P\left(\frac{2l-3}{2}p\right) > 0, \ldots .$$

Hence each of the intervals

$$\left(\frac{1}{2}p, \frac{3}{2}p\right), \;\; \left(\frac{3}{2}p, \frac{5}{2}p\right), \;\; \ldots, \left(\frac{2l-1}{2}p, \frac{2l+1}{2}p\right)$$

contains at least one change of sign, and so at least one root of $P(z)$. Since $P(z)$ has precisely l roots in the complex plane, this accounts for them all, and so all the roots of $P(z)$ are real, with exactly one in each of the above intervals.

For $1 \le k \le l$ put

$$L_k(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^{m} \alpha_k^{i-1} y_i + \sum_{j=1}^{n} \alpha_k^{m+j-1} x_j.$$

No α_k satisfies a polynomial of degree $< l$ with integer coefficients. Thus $L_k(\mathbf{x}, \mathbf{y}) \neq 0$ when $\mathbf{x} \in \mathbb{Z}^n$ and $\mathbf{y} \in \mathbb{Z}^m$ and they are not both $\mathbf{0}$. The expression $\prod_k L_k$ is symmetric in the α_k, and so it is an integer whenever the x_i and y_j are all integers. Hence $|\prod_k L_k| \ge 1$ when $\mathbf{x} \in \mathbb{Z}^n$ and $\mathbf{y} \in \mathbb{Z}^m$ and they are not both zero vectors.

The algebraic numbers $\alpha_1, \dots, \alpha_m$ are distinct. Therefore the $m \times m$ van der Monde matrix M defined by $[M]_{ki} = \alpha_k^{i-1}$ is invertible. Let B denote the $m \times n$ matrix defined by $[B]_{kj} = -\alpha_k^{m+j-1}$. Now define the $m \times n$ matrix $A = a_{ij}$ by $A = M^{-1}B$. Then $MA\mathbf{x} = B\mathbf{x}$. In other words, if $A_i(\mathbf{x}) = \sum_{j=1}^{n} a_{ij} x_j$, then

$$\sum_{i=1}^{m} \alpha_k^{i-1} A_i(\mathbf{x}) = -\sum_{j=1}^{n} \alpha_k^{m+j-1} x_j \ \ (1 \le k \le m).$$

Substituting in the definition of L_k gives

$$L_k(\mathbf{x}), \mathbf{y}) = \sum_{i=1}^{m} \alpha_k^{i-1} (y_i - A_i(\mathbf{x})) \ \ (1 \le k \le m).$$

We cannot expect this relationship to persist for larger values of k. However, we may define

$$b_{kj} = \alpha_k^{m+j-1} + \sum_{i=1}^{m} \alpha_k^{i-1} a_{ij}$$

and $B_k(\mathbf{x}) = \sum_{j=1}^{n} b_{kj} x_j$. Then

$$L_k(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^{m} \alpha_k^{i-1} (y_i - A_i(\mathbf{x})) + B_k(\mathbf{x}) \ \ (m < k \le m+n).$$

Suppose that $q_1, \dots, q_n$ are integers not all 0. Put $Q = \max_{1 \le j \le n} |q_j|$, set $x_j = q_j$, and choose integral values of y_i so that $|y_i - A_i(\mathbf{q})| \le \frac{1}{2}$. Now take $\delta = \max_{1 \le i \le m} \|A_i(\mathbf{q})\|$, so that $\delta \le \frac{1}{2}$. Then $L_k(\mathbf{q}, \mathbf{y}) \ll \delta$ for $1 \le k \le m$. Similarly $B_k(\mathbf{q}) \ll Q$. Thus $L_k(\mathbf{q}, \mathbf{y}) \ll Q$ for $m < k \le m+n$. Therefore,

$$1 \le \prod |L_k(\mathbf{q}, \mathbf{y})| \ll \delta^m Q^n,$$

and the theorem follows from this.

We now return to the geometry of numbers and consider further applications of Minkowski's convex body theorem. We have already seen in an exercise that Dirichlet's theorem on diophantine approximation furnishes a short proof of the following famous theorem of Fermat, and it is no surprise that the geometry of numbers provides another.

Theoerem 3.15 (Fermat). *Suppose that $p \equiv 1 \pmod 4$. Then there exist integers λ_1 and λ_2 such that $\lambda_1^2 + \lambda_2^2 = p$.*

Proof. Since $p \equiv 1 \pmod 4$, the congruence $z^2 \equiv -1 \pmod p$ is soluble. This follows from Wilson's theorem by taking $z = \left(\frac{p-1}{2}\right)!$ or from Euler's criterion. Let $A = \begin{bmatrix} p & z \\ 0 & 1 \end{bmatrix}$ and $\Lambda = A\mathbb{Z}^2$. If $\boldsymbol{\lambda} \in \Lambda$, say $\lambda_1 = px_1 + zx_2$, $\lambda_2 = x_2$ where $x_i \in \mathbb{Z}$, then $\lambda_1^2+\lambda_2^2 = (px_1+zx_2)^2+x_2^2 \equiv (a^2+1)x_2^2 \equiv 0 \pmod p$. Let $\mathcal{C} = \{\mathbf{x} : |\mathbf{x}| < \sqrt{2p}\}$. Then $\mathcal{C}$ is a disc, center the origin, of area $\pi(\sqrt{2p})^2 = 2\pi p$. This is greater than $4p = 4d(\Lambda)$. Hence, by Theorem 3.9, there is a $\boldsymbol{\lambda} \in \Lambda$ differing from $\mathbf{0}$ such that $\boldsymbol{\lambda} \in \mathcal{C}$. Thus $0 < \lambda_1^2 + \lambda_2^2 < 2p$ and $\lambda_1^2 + \lambda_2^2 \equiv 0 \pmod p$.

As usual, a method which gives the two square theorem can also be adapted to give the four square theorem.

Theorem 3.16 (Lagrange). *Every positive integer can be expressed as a sum of four squares of integers.*

As in the proof of the two square theorem we require some appropriate local information, which we summarise as a lemma.

Lemma 3.17. *For any prime p there exist integers r and s such that $1+r^2+s^2 \equiv 0 \pmod p$.*

Proof of Lemma 3.17. The case $p = 2$ is trivial, so we may suppose that p is odd. Let $\mathcal{R} = \{1 + r^2 \pmod p\}$, $\mathcal{S} = \{-s^2 \pmod p\}$. Then $\mathrm{card}(\mathcal{R}) = \mathrm{card}(\mathcal{S}) = \frac{p+1}{2}$. By the pigeon hole principle, $\mathcal{R}$ and $\mathcal{S}$ will have a common element.

Alternative proofs of the above lemma may be obtained by showing that

$$\sum_{n=1}^{p} \left(\frac{n(n+1)}{p}\right) = -1,$$

or by use of the Chevally-Warning theorem.

Proof of Theorem 3.16. In view of Euler's identity

$$\begin{gathered}(x_1^2 + x_2^2 + x_3^2 + x_4^2)(y_1^2 + y_2^2 + y_3^2 + y_4^2) = \\ (x_1y_1 + x_2y_2 + x_3y_3 + x_4y_4)^2+(x_1y_2 - x_2y_1 + x_3y_4 - x_4y_3)^2+ \\ (x_1y_3 - x_2y_4 - x_3y_1 + x_4y_2)^2 + (x_1y_4 + x_2y_3 - x_3y_2 - x_4y_1)^2,\end{gathered}$$

we see that it suffices to represent prime numbers. Let p be prime and r and s be as in the lemma and define

$$A = \begin{bmatrix} p & 0 & r & s \\ 0 & p & s & -r \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Now we consider the lattice $\Lambda = A\mathbb{Z}^4$. We have $d(\Lambda) = p^2$. Suppose that $\mathbf{t} \in \mathbb{Z}^4$, so that $\mathbf{x} = A\mathbf{t} \in \Lambda$. Then

$$\begin{aligned} x_1^2 + x_2^2 + x_3^2 + x_4^2 &= (pt_1 + rt_3 + st_4)^2 + (pt_2 + st_3 - rt_4)^2 + t_3^2 + t_4^2 \\ &\equiv (1 + r^2 + s^2)(t_3^2 + t_4^2) \\ &\equiv 0 \pmod{p}. \end{aligned}$$

Let $\mathcal{C} = \{\mathbf{x} \in \mathbb{R}^4 : |\mathbf{x}| < \sqrt{2p}\}$. Then $\mathrm{vol}(\mathcal{C}) = \frac{\pi^2}{2}\left(\sqrt{2p}\right)^4 = 2\pi^2 p^2 > 2^4 d(\Lambda)$. Thus, by Minkowski's convex body theorem II, Theorem 3.9, $\mathcal{C}$ contains a point $\mathbf{x} \in \Lambda$ with $\mathbf{x} \neq \mathbf{0}$. Moreover $x_1^2+x_2^2+x_3^2+x_4^2 \equiv 0 \pmod{p}$ and $0 < x_1^2+x_2^2+x_3^2+x_4^2 < 2p$, so that $x_1^2 + x_2^2 + x_3^2 + x_4^2 = p$, as required.

We now proceed to explore the relationships between a lattice and its sublattices. We begin with a theorem on the factorization of matrices.

Theorem 3.18. *Let A be an $m \times n$ matrix with integer elements. Then $A = UDV$ where U ia an $m \times m$ unimodular matrix, D is an $m \times n$ diagonal matrix (the only possibly non-zero entries in $D = [d_{ij}]$ are the d_{ii}) with non-negative integral elements, and V is an $n \times n$ unimodular matrix.*

Factorizations of this kind were investigated by H. J. S. Smith in the mid-nineteenth century. The above factorization is not unique, but by further reduction we may reach the *Smith normal form* D of A, which is unique. Here D is an $m \times n$ diagonal matrix with integer entries, as above, but in addition the first r diagonal elements $d_1, d_2, \ldots, d_r$ are positive, $d_1|d_2|\cdots|d_r$, and the remaining diagonal elements are 0.

Proof. Let δ_1 denote the greatest common divisor of the elements in the first column of A. Add integral multiples of one row to another, in the manner of the Euclidean algorithm, until only one element in the first column is non-zero. This element will be $\pm\delta_1$. Multiply the row in question by ± 1, and interchange the row in which δ_1 lies with the first row so that the first element in the first column is δ_1 and the remaining elements are 0. Let δ_2 denote the greatest common divisor of the elements in the first row. We have $\delta_2|\delta_1$. If $\delta_2 = \delta_1$, then add integral multiples of the first column to the other columns in such a manner that the remaining elements in the first row are 0. In this case we now have δ_2 in the first entry in the first row and all other entries in the first row and column are 0. If $\delta_2 < \delta_1$, then we apply column operations until the first row has first entry δ_2 and remaining gentries 0. Of course, after this series of operations the first column may have some non-zero entries below the first one. However, since $\delta_2 < \delta_1$ in that case, we can repeat this process until all the entries in the first row and first column, except possibly the first one, are non-zero. If this entry is negative we multiply the first row by -1. We now repeat the above process on the second column and row, and so on. Eventaully a diagonal matrix with non-negative diagonal entries is reached. Each of the row and column operations can be obtained by multiplying A be elementary unimodular matrices, $m \times m$ and premultiplied for the row operations, $n \times n$ and post multiplied

for the column operations. This $U'AV' = D$ for suitable unimodular U' and V' and multiplying by their inverses gives the desired factorization.

We can now say something about the structure of sublattices.

Theorem 3.19. *Suppose that Λ_1 is a lattice in $\mathbb{R}^n$ and Λ_2 is a sublattice of Λ_1. Then there is a basis $\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_n$ for Λ_1 and positive integers $d_1, d_2, \dots, d_n$ such that $d_1\mathbf{f}_1, d_2\mathbf{f}_2, \dots, d_n\mathbf{f}_n$ is a basis for Λ_2.*

Proof. Write $\Lambda_1 = A_1\mathbb{Z}^n$ and $\Lambda_2 = A_2\mathbb{Z}^n$. Then, by Theorem 3.4, there is an integral $n\times n$ matrix K such that $A_2 = A_1K$. By Theorem 3.18 we have $K = UDV$. By the definition of a lattice we know that $\det(A_2) \neq 0$. Hence $\det(K) \neq 0$, and consequently $\det(D) \neq 0$. Thus the diagonal entries $d_1, d_2, \dots, d_n$ of D are positive. Let $F = A_1U$. By Theorem 3.6, $\Lambda_1 = F\mathbb{Z}^n$, and $\Lambda_2 = (FD)\mathbb{Z}^n$. Let $\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_n$ be the columns of F. Then the $\mathbf{f}_j$ form a basis for Λ_1, and the columns of FD, namely the vectors $d_j\mathbf{f}_j$, form a basis for Λ_2.

Actually what we have here is that Λ_1 is a finitely generated additive group and Λ_2 is a subgroup. By counting the number of members of Λ_1 and λ_2 in a large box one can see that the index of Λ_2 in Λ_1 should be $d(\Lambda_2)/d(\Lambda_1)$, and the theorem above provides a proof of this. Moreover it also tells us that $\Lambda_1/\Lambda_2 \cong \mathbb{Z}_{d_1} \oplus \mathbb{Z}_{d_2} \oplus \cdots \oplus \mathbb{Z}_{d_n}$. In a similar manner it can be shown that if G is a finitely generated (multiplicative) group and H is a subgroup of G, then there are generators $g_1, g_2, \dots, g_n$ of G and positive integers $d_1, d_2, \dots, d_n$ such that $g_1^{d_1}, g_2^{d_2}, \dots, g_n^{d_n}$ generate H.

For us, the most useful aspect of Theorem 3.19 is that it enables us to describe the complete solution set for a system of simultaneous linear Diophantine equations. Let such a system be $A\mathbf{x} = \mathbf{b}$. Then, by Theorem 18 we have $A = UDV$. Put $V\mathbf{x} = \mathbf{y}$. By Theorem 3.6 we know that $\mathbf{x} \in \mathbb{Z}^n$ if and only if $\mathbf{y} \in \mathbb{Z}^n$. Thus $UD\mathbf{y} = \mathbf{b}$ is equivalent to the original system in the sense that there is a simple one-to-one correspondence between the solution sets. Moreover $D\mathbf{y} = U^{-1}\mathbf{b} = \mathbf{b}'$, say. Suppose that the first r elements of D, $d_1, d_2, \dots, d_r$, are positive and the remaining ones are zero. Then the system $D\mathbf{y} = \mathbf{b}'$ has integral solutions if and only if

(a) $d_i | b'_i$ for $1 \leq i \leq r$,
(b) $b'_i = 0$ for $r < i \leq n$.

When these conditions are met we have $y_i = b'_i/d_i$ for $1 \leq i \leq r$, whilst $y_{r+1}, \dots, y_n$ are free variables. Returning to the original system we see that the solutions are $\mathbf{x} = V^{-1}\mathbf{y}$ where $\mathbf{y}$ is as just described. Thus the solution set is either empty or given precisely in terms of $n - r$ integral parameters.

We may use the above to establish an algorithm to obtain the complete solution set. We start from the following $(m+n)\times(n+1)$ array (essentially a non-rectangular array but we have included an irrelevant null block for notational convenience)

$$\begin{array}{cc} & \begin{array}{cc} n & 1 \end{array} \\ \begin{array}{c} m \\ n \end{array} & \begin{bmatrix} A & \mathbf{b} \\ I_n & \mathbf{0} \end{bmatrix} \end{array}.$$

We perform row operations on the first m rows, and column operations on the first n columns, as described in the proof of Theorem 3.18, and equivalent to premultiplication by U^{-1} and post multiplication by V^{-1}. The resulting array is

$$\begin{array}{cc} & \begin{array}{cc} n & 1 \end{array} \\ \begin{array}{c} m \\ n \end{array} & \begin{bmatrix} D & \mathbf{b}' \\ V^{-1} & \mathbf{0} \end{bmatrix} \end{array}.$$

The desired solution is readily obtained from this.

We now return to further study of Minkowski's convex body theorem. Let $\mathcal{C} = \{\mathbf{x} \in \mathbb{R}^2 : |\mathbf{x}| < 1\}$. The disk $\mathcal{C}$ has volume(area) π. Thus by Minkowski's theorem we know that when Λ is a lattice with $d(\Lambda) < \frac{\pi}{4}$, it contains a non-zero point of $\mathcal{C}$. This is useful, and we have already made applied this underlying principle in the proof of the two square theorem (Theorem 3.15), but it is not sharp. It is known that $\frac{\pi}{4}$ can be replaced by the larger number $\frac{\sqrt{3}}{2}$. This is best possible since the lattice generated by $(1, 0)$ and $(\frac{1}{2}, \frac{\sqrt{3}}{2})$ contains no non-zero point of $\mathcal{C}$.

In 1845, Ch. Hermite wrote four letters to Jacobi. In the first of these, Hermite proved that when $f(\mathbf{x}) = \sum_{i,j} a_{ij}x_ix_j = \mathbf{x}^T A\mathbf{x}$ (with A symmetric) is a positive definite quadratic form with discriminant $D = \det(A)$, then there exists an $\mathbf{x} \in \mathbb{Z}^n$ with $\mathbf{x} \neq \mathbf{0}$ and $f(\mathbf{x}) \leq \left(\frac{4}{3}\right)^{\frac{n-1}{2}} D^{\frac{1}{n}}$. Observe that, when $n = 2$, the discriminant here differs from the discriminant d above, as $D = -d/4$.

Hermite's proof is by induction on n. The case $n = 1$ is trivial. For the present we prove only the first non-trivial case, namely when $n = 2$.

Theorem 3.20. *Suppose that $f(x_1, x_2) = ax_1^2 + bx_1x_2 + cx_2^2$ is positive definite. Then there exist integers x_1 and x_2 such that $0 < f(x_1, x_2) \leq \sqrt{-d/3}$.*

This is best possible in view of the example $\sqrt{-d/3}(x_1^2 + x_1x_2 + x_2^2)$.

Proof. Choose λ large enough that the elliptical region $f(x_1, x_2) \leq \lambda$ contains at least one non-zero lattice point. By evaluating f at these points we find an integral point (u_1, u_2) at which f is minimal. Let g denote the greatest common divisor of u_1 and u_2. If $g > 1$, then $f(u_1/g, u_2/g) = f(u_1, u_2)/g^2 < f(u_1, u_2)$, contradicting the minimality. Thus $g = 1$. Put $u_{11} = u_1$, $u_{21} = u_2$, and choose integers u_{12}, u_{22} so that $U_1 = [\, u_{ij} \,]$ satisfies $\det U_1 = 1$. Set $A = \begin{bmatrix} a & \frac{1}{2}b \\ \frac{1}{2}b & c \end{bmatrix}$, $A_1 = U_1^T A U_1$, $\mathbf{y} = U_1^{-1}\mathbf{x}$. Then

$$\begin{aligned} f(\mathbf{x}) &= \mathbf{x}^T A\mathbf{x} = \mathbf{x}^T (U_1^T)^{-1} U_1^T A U_1 U_1^{-1}\mathbf{x} \\ &= (U_1^{-1}\mathbf{x})^T (U_1^T A U_1)(U_1^{-1}\mathbf{x}) \\ &= \mathbf{y}^T A_1 \mathbf{y} \\ &= f_1(\mathbf{y}), \end{aligned}$$

say. Since $\mathbf{x} \in \mathbb{Z}^2$ if and only if $\mathbf{y} \in \mathbb{Z}^2$, we find that

$$\{f(\mathbf{x}) : \mathbf{x} \in \mathbb{Z}^2\} = \{f_1(\mathbf{y}) : \mathbf{y} \in \mathbb{Z}^2\}.$$

By multipliying out the matrices defining A_1, we obtain $A_1 = \begin{bmatrix} a_1 & \frac{1}{2}b_1 \\ \frac{1}{2}b_1 & c_1 \end{bmatrix}$ where

$$\begin{aligned} a_1 &= au_{11}^2 + bu_{11}u_{21} + cu_{21}^2 = f(u_{11}, u_{21}), \\ b_1 &= 2au_{11}u_{12} + b(u_{11}u_{22} + u_{21}u_{12}) + 2cu_{21}u_{22}, \\ c_1 &= au_{12}^2 + bu_{12}u_{22} + cu_{22}^2 = f(u_{12}, u_{22}). \end{aligned}$$

The advantage of f_1 over f is that the minimal value is $f(u_{11}, u_{21}) = a_1 = f_1(1, 0)$.

We now make a second change of variable. We take $U_2 = \begin{bmatrix} 1 & m \\ 0 & 1 \end{bmatrix}$ and let $A_2 = U_2^T A_1 U_2$ and $f_2(\mathbf{z}) = \mathbf{z}^T A_2 \mathbf{z}$ where $\mathbf{z} = U_2^{-1}\mathbf{y}$. Then

$$\begin{aligned} a_2 &= a_1, \\ b_2 &= b_1 + 2a_1 m, \\ c_2 &= a_1 m^2 + b_1 m + c_1. \end{aligned}$$

Choose m so that $-a_1 < b_2 \leq a_1$. Since $c_2 = f_1(m, 1)$, and since a_1 is the minimal non-zero value of $f_1(\mathbf{y})$ at integral points, we see that $c_2 \geq a_1 = a_2$. Thus the coefficients of f_2 satisfy the inequality $-a_2 < b_2 \leq a_2 \leq c_2$. We also have $\det A = \det A_1 = \det A_2$ and so the discriminant of f_2 is the same as that of f. Thus $d = b_2^2 - 4a_2c_2$. Moreover

$$\begin{aligned} -d &= 4a_2c_2 - b_2^2 \\ &\geq 4a_2^2 - b_2^2 \quad (\text{since } c_2 \geq a_2) \\ &\geq 4a_2^2 - a_2^2 \quad (\text{since } |b_2| \leq a_2) \\ &= 3a_2^2. \end{aligned}$$

Thus $a_2 \leq \sqrt{-d/3}$, and the proof is complete.

Two quadratic forms (a, b, c) and (a', b', c') are equivalent when there is a unimodular transformation U such that $\begin{bmatrix} a' & \frac{1}{2}b' \\ \frac{1}{2}b' & c' \end{bmatrix} = U^T \begin{bmatrix} a & \frac{1}{2}b \\ \frac{1}{2}b & c \end{bmatrix} U$. The transformation $U = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ maps the quadratic form (a, b, c) to $(c, -b, a)$. When $a = c$ we may apply this transformation to ensure that $0 \leq b \leq a$. A form for which $-a < b \leq a < c$ or for which $0 \leq b \leq a = c$ is called *reduced*. The argument of the previous theorem together with the transformation above shows that every positive definite binary quadratic form is equivalent to a reduced form, and it can be shown that reduced forms are unique.

There are two simple corollaries of Theorem 3.22. The first is easily established by showing that when $b^2 < 4ac$ and $a > 0$ the area inside $ax^2 + bxy + cy^2 = \lambda$ is $\frac{2\pi\lambda}{\sqrt{4ac-b^2}}$, and the second from the observation that $-\frac{1}{4}d = \det(A^T A) = d(\Lambda)^2$.

Corollary 3.21. *Suppose that $\mathcal{E} \subset \mathbb{R}^2$ is an ellipse with centre at $\mathbf{0}$ and area exceeding $\frac{2\pi}{\sqrt{3}}$, then $\mathcal{E}$ contains a non-zero point of $\mathbb{Z}^2$.*

Corollary 3.22. *Suppose that $\Lambda \subset \mathbb{R}^2$ is a lattice with $d(\Lambda) \leq \sqrt{3}/2$. Then Λ contains a non-zero point $\mathbf{x}$ such that $x_1^2 + x_2^2 \leq 1$.*

More generally we may examine situations analoguous to that in Corollary 3.22 in the following way. Let $\mathcal{S} \subset \mathbb{R}^n$. Then we call Λ admissible when $\mathcal{S} \cap \Lambda = \emptyset$ or $\{\mathbf{0}\}$, and we set

$$\Delta(\mathcal{S}) = \inf_{\Lambda \text{ admissible}} d(\Lambda).$$

This is the *critical determinant* or *lattice constant* of $\mathcal{S}$. One of the primary research interests of the geometry of numbers is to determine the values of $\Delta(\mathcal{S})$ for various important sets $\mathcal{S}$. Existing techniques have been quite successful in determining $\Delta(\mathcal{S})$ for sets $\mathcal{S}$ in $\mathbb{R}^2$, but in higher dimensions the lattice constant is known in only a few cases. Minkowski's convex body theorem assserts that if $\mathcal{C}$ is convex and symmetric about $\mathbf{0}$, then $\Delta(\mathcal{C}) \geq \mathrm{vol}(\mathcal{C})/2^n$. In the oppositie direction, the Minkowski-Hlawka theorem asserts that if $\mathcal{S} \in \mathbb{R}^n$ and if $\mathcal{S}$ has Jordan content, then $\Delta(\mathcal{S}) \leq \mathrm{vol}(\mathcal{S})$.

Many questions in simultaneous Diophantine approximation concerning the best constant could be settled if one had reliable methods for determining $\Delta(\mathcal{S})$. For example, let $F(\mathbf{x})$ be a continuous non-negative function on $\mathbb{R}^n$ which is homogeneous of degree 1, so that $F(c\mathbf{x}) = |c|F(\mathbf{x})$, and let $C(F)$ be the infimum of those constants C such that for any real numbers $\alpha_1, \alpha_2, \ldots, \alpha_n$ there exist infinitely many $(n+1)$-tuples $(q_0, q_1, q_2, \ldots, q_n)$ such that

$$F(q_0\alpha_1 - q_1, q_0\alpha_2 - q_2, \ldots, q_0\alpha_n - q_n) < \left(\frac{C}{q_0}\right)^{1/n}.$$

A function F with these properties is called a *distance function*. The region $\mathcal{S} = \{\mathbf{x} : F < 1\}$ is a star body, but it is not necessarily convex or symmetric about $\mathbf{0}$. Davenport ("On a theorem of Furtwängler", J. London Math. Soc. 30(1955), 186-195; Collected Works II, pp. 659-668. see also Cassels' Introduction to the Geometry of Numbers, pp. 165-174) showed that $C(F) = 1/\Delta(\mathcal{K})$ where $\mathcal{K}$ is the star body

$$\mathcal{K} = \{\mathbf{x} \in \mathbb{R}^{n+1} : F(x_1, \ldots, x_n)|x_{n+1}| \leq 1\},$$

provided that $F(\mathbf{x}) = 0$ if and only if $\mathbf{x} = \mathbf{0}$. That is, the star body $\mathcal{S}$ must be bounded, for if not, then $\mathbf{x}$ can be made arbitrarily large, and so taking x_i to be the largest coordinate and putting $y_j = x_j/x_i$ one finds that $F(\mathbf{y}) = F(\mathbf{x})/|x_i| < 1/|x_i|$, whence by compactness a zero $\mathbf{z}$ of F is obtained with $1 \leq |\mathbf{z}| \leq n$. There are many useful choices for F. One can take $F(\mathbf{x}) = \max_i |x_i|$, $F((\mathbf{x}) = \left(\sum_i |x_i|^2\right)^{1/2}$, or $F(\mathbf{x}) = \sum_i |x_i|$, but not $F(x_1, x_2) = |x_1x_2|^{1/2}$, which is relevant to a celebrated problem of Littlewood, namely whether or not $\liminf_{n\to\infty} n\|n\alpha\|\|n\beta\| = 0$ for all pairs of real numbers α and β. Cassels and Swinnerton-Dyer proved that this question is equivalent to whether or not $\Delta(\mathcal{S}) = \infty$ where

$$\mathcal{S} = \{\mathbf{x} \in \mathbb{R}^3 : |x_1x_2x_3| < 1 \text{ or } |x_1x_2(x_2 + x_3)| < 1\}.$$