VIII
The geometry of numbers

It was shown by Hermite (1850) that if

$$f(x) = x^t A x$$

is a positive definite quadratic form in n real variables, then there exists a vector x with *integer* coordinates, not all zero, such that

$$f(x) \leq c_n (\det A)^{1/n},$$

where c_n is a positive constant depending only on n. Minkowski (1891) found a new and more geometric proof of Hermite's result, which gave a much smaller value for the constant c_n. Soon afterwards (1893) he noticed that his proof was valid not only for an n-dimensional ellipsoid $f(x) \leq$ const., but for any convex body which was symmetric about the origin. This led him to a large body of results, to which he gave the somewhat paradoxical name 'geometry of numbers'. It seems fair to say that Minkowski was the first to realize the importance of convexity for mathematics, and it was in his lattice point theorem that he first encountered it.

1 Minkowski's lattice point theorem

A set $C \subseteq \mathbb{R}^n$ is said to be *convex* if $x_1, x_2 \in C$ implies $\theta x_1 + (1-\theta)x_2 \in C$ for $0 < \theta < 1$. Geometrically, this means that whenever two points belong to the set the whole line segment joining them is also contained in the set.

The *indicator function* or 'characteristic function' of a set $S \subseteq \mathbb{R}^n$ is defined by $\chi(x) = 1$ or 0 according as $x \in S$ or $x \notin S$. If the indicator function is Lebesgue integrable, then the set S is said to have *volume*

$$\lambda(S) = \int_{\mathbb{R}^n} \chi(x)\, dx.$$

The indicator function of a convex set C is actually Riemann integrable. It is easily seen that if a convex set C is not contained in a hyperplane of $\mathbb{R}^n$, then its *interior* int C (see §4 of

Chapter I) is not empty. It follows that $\lambda(C) = 0$ if and only if C is contained in a hyperplane, and $0 < \lambda(C) < \infty$ if and only if C is bounded and is not contained in a hyperplane.

A set $S \subseteq \mathbb{R}^n$ is said to be *symmetric* (with respect to the origin) if $x \in S$ implies $-x \in S$. Evidently any (nonempty) symmetric convex set contains the origin.

A point $x = (\xi_1,...,\xi_n) \in \mathbb{R}^n$ whose coordinates $\xi_1,...,\xi_n$ are all integers will be called a *lattice point*. Thus the set of all lattice points in $\mathbb{R}^n$ is $\mathbb{Z}^n$.

These definitions are the ingredients for Minkowski's *lattice point theorem*:

THEOREM 1 *Let C be a symmetric convex set in $\mathbb{R}^n$. If $\lambda(C) > 2^n$, or if C is compact and $\lambda(C) = 2^n$, then C contains a nonzero point of $\mathbb{Z}^n$.*

The proof of Theorem 1 will be deferred to §3. Here we illustrate the utility of the result by giving several applications, all of which go back to Minkowski himself.

PROPOSITION 2 *If A is an $n \times n$ positive definite real symmetric matrix, then there exists a nonzero point $x \in \mathbb{Z}^n$ such that*

$$x^tAx \leq c_n(\det A)^{1/n},$$

where $c_n = (4/\pi)\{(n/2)!\}^{2/n}$.

Proof For any $\rho > 0$ the ellipsoid $x^tAx \leq \rho$ is a compact symmetric convex set. By putting $A = T^tT$, for some nonsingular matrix T, it may be seen that the volume of this set is $\kappa_n\rho^{n/2}(\det A)^{-1/2}$, where κ_n is the volume of the n-dimensional unit ball. It follows from Theorem 1 that the ellipsoid contains a nonzero lattice point if $\kappa_n\rho^{n/2}(\det A)^{-1/2} = 2^n$. But, as we will see in §4 of Chapter IX, $\kappa_n = \pi^{n/2}/(n/2)!$, where $x! = \Gamma(x + 1)$. This gives the value c_n for ρ. □

It follows from Stirling's formula (Chapter IX, §4) that $c_n \sim 2n/\pi e$ for $n \to \infty$. Hermite had proved Proposition 2 with $c_n = (4/3)^{(n-1)/2}$. Hermite's value is smaller than Minkowski's for $n \leq 8$, but much larger for large n.

As a second application of Theorem 1 we prove Minkowski's *linear forms theorem*:

PROPOSITION 3 *Let A be an $n \times n$ real matrix with determinant $\pm$ 1. Then there exists a nonzero point $x \in \mathbb{Z}^n$ such that $Ax = y = (\eta_k)$ satisfies*

$$|\eta_1| \leq 1, \quad |\eta_k| < 1 \textit{ for } 1 < k \leq n.$$

Proof For any positive integer m, let C_m be the set of all $x \in \mathbb{R}^n$ such that $Ax \in D_m$, where

$$D_m = \{y = (\eta_k) \in \mathbb{R}^n \colon |\eta_1| \leq 1 + 1/m, |\eta_k| < 1 \text{ for } 2 \leq k \leq n\}.$$

Then C_m is a symmetric convex set, since A is linear and D_m is symmetric and convex. Moreover $\lambda(C_m) = 2^n(1 + 1/m)$, since $\lambda(D_m) = 2^n(1 + 1/m)$ and A is volume-preserving. Therefore, by Theorem 1, C_m contains a lattice point $x_m \neq O$. Since $C_m \subset C_1$ for all $m > 1$ and the number of lattice points in C_1 is finite, there exist only finitely many distinct points x_m. Thus there exists a lattice point $x \neq O$ which belongs to C_m for infinitely many m. Evidently x has the required properties. □

The continued fraction algorithm enables one to find rational approximations to irrational numbers. The subject of *Diophantine approximation* is concerned with the more general problem of solving inequalities in integers. From Proposition 3 we can immediately obtain a result in this area due to Dirichlet (1842):

PROPOSITION 4 *Let $A = (\alpha_{jk})$ be an $n \times m$ real matrix and let $t > 1$ be real. Then there exist integers $q_1,...,q_m,p_1,...,p_n$, with $0 < \max(|q_1|,...,|q_m|) < t^{n/m}$, such that*

$$|\textstyle\sum_{k=1}^m \alpha_{jk}q_k - p_j| \leq 1/t \quad (1 \leq j \leq n).$$

Proof Since the matrix

$$\begin{pmatrix} t^{-n/m}I_m & 0 \\ tA & tI_n \end{pmatrix}$$

has determinant 1, it follows from Proposition 3 that there exists a nonzero vector

$$x = \begin{pmatrix} q \\ -p \end{pmatrix} \in \mathbb{Z}^{n+m}$$

such that

$$|q_k| < t^{n/m} \quad (k = 1,...,m),$$

$$|\textstyle\sum_{k=1}^m \alpha_{jk}q_k - p_j| \leq 1/t \quad (j = 1,...,n).$$

Since $q = O$ would imply $|p_j| < 1$ for all j and hence $p = O$, which contradicts $x \neq O$, we must have $\max_k |q_k| > 0$. □

COROLLARY 5 *Let $A = (\alpha_{jk})$ be an $n \times m$ real matrix such that $Az \notin \mathbb{Z}^n$ for any nonzero vector $z \in \mathbb{Z}^m$. Then there exist infinitely many $(m+n)$-tuples $q_1,...,q_m,p_1,...,p_n$ of integers with greatest common divisor 1 and with arbitrarily large values of*

$$\|q\| = \max(|q_1|,...,|q_m|)$$

such that

$$|\textstyle\sum_{k=1}^m \alpha_{jk}q_k - p_j| < \|q\|^{-m/n} \quad (1 \leq j \leq n).$$

Proof Let $q_1,...,q_m,p_1,...,p_n$ be integers satisfying the conclusions of Proposition 4 for some $t > 1$. Evidently we may assume that $q_1,...,q_m,p_1,...,p_n$ have no common divisor greater than 1. For given $q_1,...,q_m$, let δ_j be the distance of $\sum_{k=1}^{m} \alpha_{jk}q_k$ from the nearest integer and put $\delta = \max \delta_j$ $(1 \leq j \leq n)$. By hypothesis $0 < \delta < 1$, and by construction

$$\delta \leq 1/t < \|q\|^{-m/n}.$$

Choosing some $t' > 2/\delta$, we find a new set of integers $q_1',...,q_m',p_1',...,p_n'$ satisfying the same requirements with t replaced by t', and hence with $\delta' \leq 1/t' < \delta/2$. Proceeding in this way, we obtain a sequence of $(m + n)$-tuples of integers $q_1^{(\nu)},...,q_m^{(\nu)},p_1^{(\nu)},...,p_n^{(\nu)}$ for which $\delta^{(\nu)} \rightarrow 0$ and hence $\|q^{(\nu)}\| \rightarrow \infty$, since we cannot have $q^{(\nu)} = q$ for infinitely many ν. □

The hypothesis of the corollary is certainly satisfied if $1,\alpha_{j1},...,\alpha_{jm}$ are linearly independent over the field $\mathbb{Q}$ of rational numbers for some $j \in \{1,...,n\}$.

Minkowski also used his lattice point theorem to give the first proof that the discriminant of any algebraic number field, other than $\mathbb{Q}$, has absolute value greater than 1. The proof is given in most books on algebraic number theory.

2 Lattices

In the previous section we defined the set of lattice points to be $\mathbb{Z}^n$. However, this definition is tied to a particular coordinate system in $\mathbb{R}^n$. It is useful to consider lattices from a more intrinsic point of view. The key property is 'discreteness'.

With vector addition as the group operation, $\mathbb{R}^n$ is an abelian group. A subgroup Λ is said to be *discrete* if there exists a ball with centre O which contains no other point of Λ. (More generally, a subgroup H of a topological group G is said to be discrete if there exists an open set $U \subseteq G$ such that $H \cap U = \{e\}$, where e is the identity element of G.)

If Λ is a discrete subgroup of $\mathbb{R}^n$, then any bounded subset of $\mathbb{R}^n$ contains at most finitely many points of Λ since, if there were infinitely many, they would have an accumulation point and their differences would accumulate at O. In particular, Λ is a closed subset of $\mathbb{R}^n$.

PROPOSITION 6 *If $x_1,...,x_m$ are linearly independent vectors in $\mathbb{R}^n$, then the set*

$$\Lambda = \{\zeta_1 x_1 + ... + \zeta_m x_m : \zeta_1,...,\zeta_m \in \mathbb{Z}\}$$

is a discrete subgroup of $\mathbb{R}^n$.

Proof It is clear that Λ is a subgroup of $\mathbb{R}^n$, since $x,y \in \Lambda$ implies $x - y \in \Lambda$. If Λ is not discrete, then there exist $y^{(v)} \in \Lambda$ with $|y^{(1)}| > |y^{(2)}| > \ldots$ and $|y^{(v)}| \to 0$ as $v \to \infty$. Let V be the vector subspace of $\mathbb{R}^n$ with basis $x_1,\ldots,x_m$ and for any vector

$$x = \alpha_1 x_1 + \ldots + \alpha_m x_m,$$

where $\alpha_k \in \mathbb{R}$ $(1 \leq k \leq m)$, put

$$\|x\| = \max(|\alpha_1|,\ldots,|\alpha_m|).$$

This defines a norm on V. We have

$$y^{(v)} = \zeta_1^{(v)} x_1 + \ldots + \zeta_m^{(v)} x_m,$$

where $\zeta_k^{(v)} \in \mathbb{Z}$ $(1 \leq k \leq m)$. Since any two norms on a finite-dimensional vector space are equivalent (Lemma VI.7), it follows that $\zeta_k^{(v)} \to 0$ as $v \to \infty$ $(1 \leq k \leq m)$. Since $\zeta_k^{(v)}$ is an integer, this is only possible if $y^{(v)} = O$ for all large v, which is a contradiction. □

The converse of Proposition 6 is also valid. In fact we will prove a sharper result:

PROPOSITION 7 *If Λ is a discrete subgroup of $\mathbb{R}^n$, then there exist linearly independent vectors $x_1,\ldots,x_m$ in $\mathbb{R}^n$ such that*

$$\Lambda = \{\zeta_1 x_1 + \ldots + \zeta_m x_m : \zeta_1,\ldots,\zeta_m \in \mathbb{Z}\}.$$

Furthermore, if $y_1,\ldots,y_m$ is any maximal set of linearly independent vectors in Λ, we can choose $x_1,\ldots,x_m$ so that

$$\Lambda \cap \langle y_1,\ldots,y_k \rangle = \{\zeta_1 x_1 + \ldots + \zeta_k x_k : \zeta_1,\ldots,\zeta_k \in \mathbb{Z}\} \quad (1 \leq k \leq m),$$

where $\langle Y \rangle$ denotes the vector subspace generated by the set Y.

Proof Let S_1 denote the set of all $\alpha_1 > 0$ such that $\alpha_1 y_1 \in \Lambda$ and let μ_1 be the infimum of all $\alpha_1 \in S_1$. We are going to show that $\mu_1 \in S_1$. If this is not the case there exist $\alpha_1^{(v)} \in S_1$ with $\alpha_1^{(1)} > \alpha_1^{(2)} > \ldots$ and $\alpha_1^{(v)} \to \mu_1$ as $v \to \infty$. Since the ball $|x| \leq (1 + \mu_1)|y_1|$ contains only finitely many points of Λ, this is a contradiction.

Any $\alpha_1 \in S_1$ can be written in the form $\alpha_1 = p\mu_1 + \theta$, where p is a positive integer and $0 \leq \theta < \mu_1$. Since $\theta > 0$ would imply $\theta \in S_1$, contrary to the definition of μ_1, we must have $\theta = 0$. Hence if we put $x_1 = \mu_1 y_1$, then

$$\Lambda \cap \langle y_1 \rangle = \{\zeta_1 x_1 : \zeta_1 \in \mathbb{Z}\}.$$

Assume that, for some positive integer k $(1 \leq k < m)$, we have found vectors $x_1,\ldots,x_k \in \Lambda$ such that

$$\Lambda \cap \langle y_1,\ldots,y_k\rangle = \{\zeta_1 x_1 + \ldots + \zeta_k x_k \colon \zeta_1,\ldots,\zeta_k \in \mathbb{Z}\}.$$

We will prove the proposition by showing that this assumption continues to hold when k is replaced by $k + 1$.

Any $x \in \Lambda \cap \langle y_1,\ldots,y_{k+1}\rangle$ has the form

$$x = \alpha_1 x_1 + \ldots + \alpha_k x_k + \alpha_{k+1} y_{k+1},$$

where $\alpha_1,\ldots,\alpha_{k+1} \in \mathbb{R}$. Let S_{k+1} denote the set of all $\alpha_{k+1} > 0$ which arise in such representations and let μ_{k+1} be the infimum of all $\alpha_{k+1} \in S_{k+1}$. We will show that $\mu_{k+1} \in S_{k+1}$. If $\mu_{k+1} \notin S_{k+1}$, there exist $\alpha_{k+1}{}^{(\nu)} \in S_{k+1}$ with $\alpha_{k+1}{}^{(1)} > \alpha_{k+1}{}^{(2)} > \ldots$ and $\alpha_{k+1}{}^{(\nu)} \to \mu_{k+1}$ as $\nu \to \infty$. Then Λ contains a point

$$x^{(\nu)} = \alpha_1{}^{(\nu)} x_1 + \ldots + \alpha_k{}^{(\nu)} x_k + \alpha_{k+1}{}^{(\nu)} y_{k+1},$$

where $\alpha_j{}^{(\nu)} \in \mathbb{R}$ $(1 \leq j \leq k)$. In fact, by subtracting an integral linear combination of $x_1,\ldots,x_k$ we may assume that $0 \leq \alpha_j{}^{(\nu)} < 1$ $(1 \leq j \leq k)$. Since only finitely many points of Λ are contained in the ball $|x| \leq |x_1| + \ldots + |x_k| + (1+\mu_{k+1})|y_{k+1}|$, this is a contradiction.

Hence $\mu_{k+1} > 0$ and Λ contains a vector

$$x_{k+1} = \alpha_1 x_1 + \ldots + \alpha_k x_k + \mu_{k+1} y_{k+1}.$$

As for S_1, it may be seen that S_{k+1} consists of all positive integer multiples of μ_{k+1}. Hence any $x \in \Lambda \cap \langle y_1,\ldots,y_{k+1}\rangle$ has the form

$$x = \zeta_1 x_1 + \ldots + \zeta_k x_k + \zeta_{k+1} x_{k+1},$$

where $\zeta_1,\ldots,\zeta_k \in \mathbb{R}$ and $\zeta_{k+1} \in \mathbb{Z}$. Since

$$x - \zeta_{k+1} x_{k+1} \in \Lambda \cap \langle y_1,\ldots,y_k\rangle,$$

we must actually have $\zeta_1,\ldots,\zeta_k \in \mathbb{Z}$. □

By being more specific in the proof of Proposition 7 it may be shown that there is a *unique* choice of $x_1,\ldots,x_m$ such that

$$\begin{aligned} y_1 &= p_{11} x_1 \\ y_2 &= p_{21} x_1 + p_{22} x_2 \\ &\cdots\cdots \\ y_m &= p_{m1} x_1 + p_{m2} x_2 + \ldots + p_{mm} x_m, \end{aligned}$$

where $p_{ij} \in \mathbb{Z}$, $p_{ii} > 0$, and $0 \le p_{ij} < p_{ii}$ if $j < i$ (*Hermite's normal form*).

It is easily seen that in Proposition 7 we can choose $x_i = y_i$ $(1 \le i \le m)$ if and only if, for any $x \in \Lambda$ and any positive integer h, x is an integral linear combination of $y_1,...,y_m$ whenever hx is.

By combining Propositions 6 and 7 we obtain

PROPOSITION 8 *For a set* $\Lambda \subseteq \mathbb{R}^n$ *the following two conditions are equivalent*:

(i) Λ *is a discrete subgroup of* $\mathbb{R}^n$ *and there exists* $R > 0$ *such that, for each* $y \in \mathbb{R}^n$, *there is some* $x \in \Lambda$ *with* $|y - x| < R$;
(ii) *there exist* n *linearly independent vectors* $x_1,...,x_n$ *in* $\mathbb{R}^n$ *such that*

$$\Lambda = \{\zeta_1 x_1 + ... + \zeta_n x_n : \zeta_1,...,\zeta_n \in \mathbb{Z}\}.$$

Proof If (i) holds, then in the statement of Proposition 7 we must have $m = n$, i.e. (ii) holds. On the other hand, if (ii) holds then Λ is a discrete subgroup of $\mathbb{R}^n$, by Proposition 6. Moreover, for any $y \in \mathbb{R}^n$ we can choose $x \in \Lambda$ so that

$$y - x = \theta_1 x_1 + ... + \theta_n x_n,$$

where $0 \le \theta_j < 1$ $(j = 1,...,n)$, and hence

$$|y - x| < |x_1| + ... + |x_n|. \quad \Box$$

A set $\Lambda \subseteq \mathbb{R}^n$ satisfying either of the two equivalent conditions of Proposition 8 will be called a *lattice* and any element of Λ a *lattice point*. The vectors $x_1,...,x_n$ in (ii) will be said to be a *basis* for the lattice.

A lattice is sometimes defined to be any discrete subgroup of $\mathbb{R}^n$, and what we have called a lattice is then called a 'nondegenerate' lattice. Our definition is chosen simply to avoid repetition of the word 'nondegenerate'. We may occasionally use the more general definition and, with this warning, believe it will be clear from the context when this occurs.

The basis of a lattice is not uniquely determined. In fact $y_1,...,y_n$ is also a basis if

$$y_j = \sum_{k=1}^{n} \alpha_{jk} x_k \quad (j = 1,...,n),$$

where $A = (\alpha_{jk})$ is an $n \times n$ matrix of integers such that $\det A = \pm 1$, since A^{-1} is then also a matrix of integers. Moreover, every basis $y_1,...,y_n$ is obtained in this way. For if

$$y_j = \sum_{k=1}^{n} \alpha_{jk} x_k, \; x_i = \sum_{j=1}^{n} \beta_{ij} y_j, \quad (i,j = 1,...,n),$$

where $A = (\alpha_{jk})$ and $B = (\beta_{ij})$ are $n \times n$ matrices of integers, then $BA = I$ and hence $(\det B)(\det A) = 1$. Since $\det A$ and $\det B$ are integers, it follows that $\det A = \pm 1$.

Let $x_1,\ldots,x_n$ be a basis for a lattice $\Lambda \subseteq \mathbb{R}^n$. If

$$x_k = \sum_{j=1}^{n} \gamma_{jk} e_j \quad (k = 1,\ldots,n),$$

where $e_1,\ldots,e_n$ is the canonical basis for $\mathbb{R}^n$ then, in terms of the nonsingular matrix $T = (\gamma_{jk})$, the lattice Λ is just the set of all vectors Tz with $z \in \mathbb{Z}^n$. The absolute value of the determinant of the matrix T does not depend on the choice of basis. For if $x_1',\ldots,x_n'$ is any other basis, then

$$x_i' = \sum_{j=1}^{n} \alpha_{ij} x_j \quad (i = 1,\ldots,n),$$

where $A = (\alpha_{ij})$ is an $n \times n$ matrix of integers with $\det A = \pm 1$. Thus

$$x_k' = \sum_{j=1}^{n} \gamma_{jk}' e_j \quad (k = 1,\ldots,n),$$

where $T' = (\gamma_{jk}')$ satisfies $T' = TA^t$ and hence

$$|\det T'| = |\det T|.$$

The uniquely determined quantity $|\det T|$ will be called the *determinant* of the lattice Λ and denoted by $\mathrm{d}(\Lambda)$. (Some authors, e.g. Conway and Sloane [14], call $|\det T|^2$ the determinant of Λ, but others prefer to call this the *discriminant* of Λ.)

The determinant $\mathrm{d}(\Lambda)$ has a simple geometrical interpretation. In fact it is the volume of the parallelotope Π, consisting of all points $y \in \mathbb{R}^n$ such that

$$y = \theta_1 x_1 + \ldots + \theta_n x_n,$$

where $0 \le \theta_k \le 1$ $(k = 1,\ldots,n)$. The interior of Π is a *fundamental domain* for the subgroup Λ, since

$$\mathbb{R}^n = \bigcup_{x \in \Lambda} (\Pi + x),$$

$$\operatorname{int}(\Pi + x) \cap \operatorname{int}(\Pi + x') = \emptyset \quad \text{if } x, x' \in \Lambda \text{ and } x \neq x'.$$

For any lattice $\Lambda \subseteq \mathbb{R}^n$, the set Λ^* of all vectors $y \in \mathbb{R}^n$ such that $y^t x \in \mathbb{Z}$ for every $x \in \Lambda$ is again a lattice, the *dual* (or 'polar' or 'reciprocal') of Λ. In fact,

$$\text{if } \Lambda = \{Tz : z \in \mathbb{Z}^n\}, \text{ then } \Lambda^* = \{(T^t)^{-1} w : w \in \mathbb{Z}^n\}.$$

Hence Λ is the dual of Λ^* and $\mathrm{d}(\Lambda)\mathrm{d}(\Lambda^*) = 1$. A lattice Λ is *self-dual* if $\Lambda^* = \Lambda$.

3 Proof of the lattice point theorem, and some generalizations

In this section we take up the proof of Minkowski's lattice point theorem. The proof will be based on a very general result, due to Blichfeldt (1914), which is not restricted to convex sets.

PROPOSITION 9 *Let S be a Lebesgue measurable subset of $\mathbb{R}^n$, Λ a lattice in $\mathbb{R}^n$ with determinant* $\mathrm{d}(\Lambda)$ *and m a positive integer.*

If $\lambda(S) > m\,\mathrm{d}(\Lambda)$, or if S is compact and $\lambda(S) = m\,\mathrm{d}(\Lambda)$, then there exist $m + 1$ distinct points $x_1,\dots,x_{m+1}$ of S such that the differences $x_j - x_k$ $(1 \le j,k \le m + 1)$ all lie in Λ.

Proof Let $b_1,\dots,b_n$ be a basis for Λ and let P be the half-open parallelotope consisting of all points $x = \theta_1 b_1 + \dots + \theta_n b_n$, where $0 \le \theta_i < 1$ $(i = 1,\dots,n)$. Then $\lambda(P) = \mathrm{d}(\Lambda)$ and

$$\mathbb{R}^n = \bigcup_{z\in\Lambda}(P + z), \quad (P + z) \cap (P + z') = \emptyset \quad \text{if } z \ne z'.$$

Suppose first that $\lambda(S) > m\,\mathrm{d}(\Lambda)$. If we put

$$S_z = S \cap (P + z), \quad T_z = S_z - z,$$

then $T_z \subseteq P$, $\lambda(T_z) = \lambda(S_z)$ and

$$\lambda(S) = \sum_{z\in\Lambda}\lambda(S_z).$$

Hence

$$\sum_{z\in\Lambda}\lambda(T_z) = \lambda(S) > m\,\mathrm{d}(\Lambda) = m\,\lambda(P).$$

Since $T_z \subseteq P$ for every z, it follows that some point $y \in P$ is contained in at least $m + 1$ sets T_z. (In fact this must hold for all y in a subset of P of positive measure.) Thus there exist $m + 1$ distinct points $z_1,\dots,z_{m+1}$ of Λ and points $x_1,\dots,x_{m+1}$ of S such that $y = x_j - z_j$ $(j = 1,..,m+1)$. Then $x_1,\dots,x_{m+1}$ are distinct and

$$x_j - x_k = z_j - z_k \in \Lambda \quad (1 \le j,k \le m + 1).$$

Suppose next that S is compact and $\lambda(S) = m\,\mathrm{d}(\Lambda)$. Let $\{\varepsilon_\nu\}$ be a decreasing sequence of positive numbers such that $\varepsilon_\nu \to 0$ as $\nu \to \infty$, and let S_ν denote the set of all points of $\mathbb{R}^n$ distant at most ε_ν from S. Then S_ν is compact, $\lambda(S_\nu) > \lambda(S)$ and

$$S_1 \supset S_2 \supset \dots, \quad S = \bigcap_\nu S_\nu .$$

By what we have already proved, there exist $m + 1$ distinct points $x_1^{(\nu)},\dots,x_{m+1}^{(\nu)}$ of S_ν such that $x_j^{(\nu)} - x_k^{(\nu)} \in \Lambda$ for all j,k. Since $S_\nu \subseteq S_1$ and S_1 is compact, by restricting attention to a subsequence we may assume that $x_j^{(\nu)} \to x_j$ as $\nu \to \infty$ $(j = 1,\ldots,m+1)$. Then $x_j \in S$ and $x_j^{(\nu)} - x_k^{(\nu)} \to x_j - x_k$. Since $x_j^{(\nu)} - x_k^{(\nu)} \in \Lambda$, this is only possible if $x_j - x_k = x_j^{(\nu)} - x_k^{(\nu)}$ for all large ν. Hence $x_1,\ldots,x_{m+1}$ are distinct. ◻

Siegel (1935) has given an analytic formula which underlies Proposition 9 and enables it to be generalized. Although we will make no use of it, this formula will now be established. For notational simplicity we restrict attention to the (self-dual) lattice $\Lambda = \mathbb{Z}^n$.

PROPOSITION 10 *If ψ: $\mathbb{R}^n \to \mathbb{C}$ is a bounded measurable function which vanishes outside some compact set, then*

$$\int_{\mathbb{R}^n} \psi(x)\overline{\phi(x)}\, dx = \sum_{w\in\mathbb{Z}^n} \left|\int_{\mathbb{R}^n} \psi(x)e^{-2\pi i w^t x}\, dx\right|^2,$$

where

$$\phi(x) = \sum_{z\in\mathbb{Z}^n} \psi(x+z).$$

Proof Since ψ vanishes outside a compact set, there exists a finite set $T \subseteq \mathbb{Z}^n$ such that $\psi(x+z) = 0$ for all $x \in \mathbb{R}^n$ if $z \in \mathbb{Z}^n \backslash T$. Thus the sum defining $\phi(x)$ has only finitely many nonzero terms and ϕ also is a bounded measurable function which vanishes outside some compact set.

If we write

$$x = (\xi_1,\ldots,\xi_n), \quad z = (\zeta_1,\ldots,\zeta_n),$$

then the sum defining $\phi(x)$ is unaltered by the substitution $\zeta_j \to \zeta_j + 1$ and hence ϕ has period 1 in each of the variables ξ_j $(j = 1,\ldots,n)$. Let Π denote the fundamental parallelotope

$$\Pi = \{x = (\xi_1,\ldots,\xi_n) \in \mathbb{R}^n\colon\ 0 \le \xi_j \le 1 \text{ for } j = 1,\ldots,n\}.$$

Since the functions $e^{2\pi i w^t x}$ $(w \in \mathbb{Z}^n)$ are an orthogonal basis for $L^2(\Pi)$, Parseval's equality (Chapter I, §10) holds:

$$\int_\Pi |\phi(x)|^2\, dx = \sum_{w\in\mathbb{Z}^n} |c_w|^2,$$

where

$$c_w = \int_\Pi \phi(x)e^{-2\pi i w^t x}\, dx.$$

But

$$\begin{aligned} c_w &= \int_\Pi \sum_{z\in\mathbb{Z}^n} \psi(x+z)e^{-2\pi i w^t x}\, dx \\ &= \int_\Pi \sum_{z\in\mathbb{Z}^n} \psi(x+z)e^{-2\pi i w^t (x+z)}\, dx\,, \end{aligned}$$

since $e^{2k\pi i} = 1$ for any integer k. Hence

$$c_w = \int_{\mathbb{R}^n} \psi(y)e^{-2\pi i w^t y}\, dy.$$

On the other hand,

$$\begin{aligned}\int_\Pi |\phi(x)|^2\, dx &= \int_\Pi \sum_{z',z''\in\mathbb{Z}^n} \psi(x+z')\overline{\psi(x+z'')}\, dx \\ &= \int_\Pi \sum_{z,z'\in\mathbb{Z}^n} \psi(x+z')\overline{\psi(x+z'+z)}\, dx \\ &= \int_{\mathbb{R}^n} \sum_{z\in\mathbb{Z}^n} \psi(y)\overline{\psi(y+z)}\, dy = \int_{\mathbb{R}^n} \psi(y)\overline{\phi(y)}\, dy.\end{aligned}$$

Substituting these expressions in Parseval's equality, we obtain the result. □

Suppose, in particular, that ψ takes only real nonnegative values. Then so also does ϕ and

$$\int_{\mathbb{R}^n} \psi(x)\phi(x)\, dx \leq \sup_{x\in\mathbb{R}^n} \phi(x) \int_{\mathbb{R}^n} \psi(x)\, dx.$$

On the other hand, omitting all terms with $w \neq 0$ we obtain

$$\sum_{w\in\mathbb{Z}^n} \left|\int_{\mathbb{R}^n} \psi(x)e^{-2\pi i w^t x}\, dx\right|^2 \geq \left(\int_{\mathbb{R}^n} \psi(x)\, dx\right)^2.$$

Hence, by Proposition 10,

$$\sup_{x\in\mathbb{R}^n} \phi(x) \geq \int_{\mathbb{R}^n} \psi(x)\, dx.$$

For example, let $S \subseteq \mathbb{R}^n$ be a measurable set with $\lambda(S) > m$. Then there exists a *bounded* measurable set $S' \subseteq S$ with $\lambda(S') > m$. If we take ψ to be the indicator function of S', then

$$\int_{\mathbb{R}^n} \psi(x)\, dx = \lambda(S') > m$$

and we conclude that there exists $y \in \mathbb{R}^n$ such that

$$\sum_{z\in\mathbb{Z}^n} \psi(y+z) = \phi(y) > m.$$

Since the only possible values of the summands on the left are 0 and 1, it follows that there exist $m+1$ distinct points $z_1,\ldots,z_{m+1} \in \mathbb{Z}^n = \Lambda$ such that $y+z_1,\ldots,y+z_{m+1} \in S$. The proof of Proposition 9 can now be completed in the same way as before.

Let $\{K_\alpha\}$ be a family of subsets of $\mathbb{R}^n$, where each K_α is the *closure* of a nonempty open set G_α, i.e. K_α is the intersection of all closed sets containing G_α. The family $\{K_\alpha\}$ is said to be a *packing* of $\mathbb{R}^n$ if $\alpha \neq \alpha'$ implies $G_\alpha \cap G_{\alpha'} = \emptyset$ and is said to be a *covering* of $\mathbb{R}^n$ if $\mathbb{R}^n = \bigcup_\alpha K_\alpha$. It is said to be a *tiling* of $\mathbb{R}^n$ if it is both a packing and a covering.

For example, if Π is a fundamental parallelotope of a lattice Λ, then the family $\{\Pi + a\colon a \in \Lambda\}$ is a tiling of $\mathbb{R}^n$. More generally, if G is a nonempty open subset of $\mathbb{R}^n$ with closure K, we may ask whether the family $\{K + a\colon a \in \Lambda\}$ of all Λ-translates of K is either a packing or a covering of $\mathbb{R}^n$. Some necessary conditions may be derived with the aid of Proposition 9:

PROPOSITION 11 *Let K be the closure of a bounded nonempty open set $G \subseteq \mathbb{R}^n$ and let Λ be a lattice in $\mathbb{R}^n$.*

If the Λ-translates of K are a covering of $\mathbb{R}^n$ then $\lambda(K) \geq \mathrm{d}(\Lambda)$, with strict inequality if they are not also a packing.

If the Λ-translates of K are a packing of $\mathbb{R}^n$ then $\lambda(K) \leq \mathrm{d}(\Lambda)$, with strict inequality if they are not also a covering.

Proof Suppose first that the Λ-translates of K cover $\mathbb{R}^n$. Then every point of a fundamental parallelotope Π of Λ has the form $x - a$, where $x \in K$ and $a \in \Lambda$. Hence

$$\begin{aligned}\lambda(K) &= \sum_{a\in\Lambda} \lambda(K \cap (\Pi + a)) \\ &= \sum_{a\in\Lambda} \lambda((K-a) \cap \Pi) \geq \lambda(\Pi) = \mathrm{d}(\Lambda).\end{aligned}$$

Suppose, in addition, that the Λ-translates of K are not a packing of $\mathbb{R}^n$. Then there exist distinct points x_1, x_2 in the interior G of K such that $a = x_1 - x_2 \in \Lambda$. Let

$$B_\varepsilon = \{x \in \mathbb{R}^n : |x| \leq \varepsilon\}.$$

We can choose $\varepsilon > 0$ so small that the balls $B_\varepsilon + x_1$ and $B_\varepsilon + x_2$ are disjoint and contained in G. Then $G' = G \setminus (B_\varepsilon + x_1)$ is a bounded nonempty open set with closure $K' = K \setminus (\operatorname{int} B_\varepsilon + x_1)$. Since

$$B_\varepsilon + x_1 = B_\varepsilon + x_2 + a \subseteq K' + a,$$

the Λ-translates of K' contain K and therefore also cover $\mathbb{R}^n$. Hence, by what we have already proved, $\lambda(K') \geq \mathrm{d}(\Lambda)$. Since $\lambda(K) > \lambda(K')$, it follows that $\lambda(K) > \mathrm{d}(\Lambda)$.

Suppose now that the Λ-translates of K are a packing of $\mathbb{R}^n$. Then Λ does not contain the difference of two distinct points in the interior G of K, since $G + a$ and $G + b$ are disjoint if a,b are distinct points of Λ. It follows from Proposition 9 that

$$\lambda(K) = \lambda(G) \leq \mathrm{d}(\Lambda).$$

Suppose, in addition, that the Λ-translates of K do not cover $\mathbb{R}^n$. Thus there exists a point $y \in \mathbb{R}^n$ which is not in any Λ-translate of K. We will show that we can choose $\varepsilon > 0$ so small that y is not in any Λ-translate of $K + B_\varepsilon$.

If this is not the case then, for any positive integer ν, there exists $a_\nu \in \Lambda$ such that

$$y \in K + B_{1/\nu} + a_\nu.$$

Evidently the sequence a_ν is bounded and hence there exists $a \in \Lambda$ such that $a_\nu = a$ for infinitely many ν. But then $y \in K + a$, which is contrary to hypothesis.

We may in addition assume ε chosen so small that $|x| > 2\varepsilon$ for every nonzero $x \in \Lambda$. Then the set $S = G \cup (B_\varepsilon + y)$ has the property that Λ does not contain the difference of any two distinct points of S. Hence, by Proposition 9, $\lambda(S) \leq \mathrm{d}(\Lambda)$. Since

$$\lambda(K) = \lambda(G) < \lambda(S),$$

it follows that $\lambda(K) < \mathrm{d}(\Lambda)$. □

We next apply Proposition 9 to convex sets. Minkowski's lattice point theorem (Theorem 1) is the special case $m = 1$ (and $\Lambda = \mathbb{Z}^n$) of the following generalization, due to van der Corput (1936):

PROPOSITION 12 *Let C be a symmetric convex subset of $\mathbb{R}^n$, Λ a lattice in $\mathbb{R}^n$ with determinant* $\mathrm{d}(\Lambda)$, *and m a positive integer.*

If $\lambda(C) > 2^n m\, \mathrm{d}(\Lambda)$, or if C is compact and $\lambda(C) = 2^n m\, \mathrm{d}(\Lambda)$, then there exist $2m$ distinct nonzero points $\pm y_1, \ldots, \pm y_m$ of Λ such that

$$\begin{aligned} y_j &\in C \quad (1 \leq j \leq m), \\ y_j - y_k &\in C \quad (1 \leq j,k \leq m). \end{aligned}$$

Proof The set $S = \{x/2 : x \in C\}$ has measure $\lambda(S) = \lambda(C)/2^n$. Hence, by Proposition 9, there exist $m + 1$ distinct points $x_1, \ldots, x_{m+1} \in C$ such that $(x_j - x_k)/2 \in \Lambda$ for all j,k.

The vectors of $\mathbb{R}^n$ may be totally ordered by writing $x > x'$ if the first nonzero coordinate of $x - x'$ is positive. We assume the points $x_1, \ldots, x_{m+1} \in C$ numbered so that

$$x_1 > x_2 > \ldots > x_{m+1}.$$

Put

$$y_j = (x_j - x_{m+1})/2 \quad (j = 1, \ldots, m).$$

Then, by construction, $y_j \in \Lambda$ $(j = 1, \ldots, m)$. Moreover $y_j \in C$, since $x_1, \ldots, x_{m+1} \in C$ and C is symmetric, and similarly $y_j - y_k = (x_j - x_k)/2 \in C$. Finally, since

$$y_1 > y_2 > \ldots > y_m > O,$$

we have $y_j \neq O$ and $y_j \neq \pm y_k$ if $j \neq k$. □

The conclusion of Proposition 12 need no longer hold if C is not compact and $\lambda(C) = 2^n m\, \mathrm{d}(\Lambda)$. For example, take $\Lambda = \mathbb{Z}^n$ and let C be the symmetric convex set

$$C = \{x = (\xi_1,...,\xi_n) \in \mathbb{R}^n: |\xi_1| < m, |\xi_j| < 1 \text{ for } 2 \le j \le n\}.$$

Then $d(\Lambda) = 1$ and $\lambda(C) = 2^n m$. However, the only nonzero points of Λ in C are the $2(m-1)$ points $(\pm k,0,...,0)$ $(1 \le k \le m-1)$.

To provide a broader view of the geometry of numbers we now mention without proof some further results. A different generalization of Minkowski's lattice point theorem was already proved by Minkowski himself. Let Λ be a lattice in $\mathbb{R}^n$ and let K be a compact symmetric convex subset of $\mathbb{R}^n$ with nonempty interior. Then ρK contains no nonzero point of Λ for small $\rho > 0$ and contains n linearly independent points of Λ for large $\rho > 0$. Let μ_i be the infimum of all $\rho > 0$ such that ρK contains at least i linearly independent points of Λ $(i = 1,...,n)$. The *successive minima* $\mu_i = \mu_i(K,\Lambda)$ evidently satisfy the inequalities

$$0 < \mu_1 \le \mu_2 \le ... \le \mu_n < \infty.$$

Minkowski's lattice point theorem says that

$$\mu_1{}^n \lambda(K) \le 2^n d(\Lambda).$$

Minkowski's *theorem on successive minima* strengthens this to

$$2^n d(\Lambda)/n! \le \mu_1\mu_2 \cdots \mu_n \lambda(K) \le 2^n d(\Lambda).$$

The lower bound is quite easy to prove, but the upper bound is deeper-lying – notwithstanding simplifications of Minkowski's original proof. If $\Lambda = \mathbb{Z}^n$, then equality holds in the lower bound for the *cross-polytope* $K = \{(\xi_1,...,\xi_n) \in \mathbb{R}^n: \sum_{i=1}^n |\xi_i| \le 1\}$ and in the upper bound for the *cube* $K = \{(\xi_1,...,\xi_n) \in \mathbb{R}^n: |\xi_i| \le 1 \text{ for all } i\}$.

If K is a compact symmetric convex subset of $\mathbb{R}^n$ with nonempty interior, its *critical determinant* $\Delta(K)$ is defined to be the infimum of the determinant $d(\Lambda)$ for all lattices Λ with no nonzero point in the interior of K. A lattice Λ for which $d(\Lambda) = \Delta(K)$ is called a *critical lattice* for K. It will be shown in §6 that a critical lattice always exists.

It follows from Proposition 12 that $\Delta(K) \ge 2^{-n}\lambda(K)$. A conjectured sharpening of Minkowski's theorem on successive minima, which has been proved by Minkowski (1896) himself for $n = 2$ and for n-dimensional ellipsoids, and by Woods (1956) for $n = 3$, claims that

$$\mu_1\mu_2 \cdots \mu_n \Delta(K) \le d(\Lambda).$$

The successive minima of a convex body are connected with those of its dual body. If K is a compact symmetric convex subset of $\mathbb{R}^n$ with nonempty interior, then its *dual*

$$K^* = \{y \in \mathbb{R}^n: y^t x \le 1 \text{ for all } x \in K\}$$

has the same properties, and K is the dual of K^*. Mahler (1939) showed that the successive minima of the dual body K^* with respect to the dual lattice Λ^* are related to the successive minima of K with respect to Λ by the inequalities

$$1 \leq \mu_i(K,\Lambda)\mu_{n-i+1}(K^*,\Lambda^*) \quad (i = 1,...,n),$$

and hence, by applying Minkowski's theorem on successive minima also to K^* and Λ^*, he obtained inequalities in the opposite direction:

$$\mu_i(K,\Lambda)\mu_{n-i+1}(K^*,\Lambda^*) \leq 4^n/\lambda(K)\lambda(K^*) \quad (i = 1,...,n).$$

By further proving that $\lambda(K)\lambda(K^*) \geq 4^n(n!)^{-2}$, he deduced that

$$\mu_i(K,\Lambda)\mu_{n-i+1}(K^*,\Lambda^*) \leq (n!)^2 \quad (i = 1,...,n).$$

Dramatic improvements of these bounds have recently been obtained. Banaszczyk (1996), using techniques from harmonic analysis, has shown that there is a numerical constant $C > 0$ such that, for all $n \geq 1$ and all $i \in \{1,...,n)$,

$$\mu_i(K,\Lambda)\mu_{n-i+1}(K^*,\Lambda^*) \leq Cn(1 + \log n).$$

He had shown already (1993) that if $K = B_1$ is the n-dimensional closed unit ball, which is self-dual, then for all $n \geq 1$ and all $i \in \{1,...,n)$,

$$\mu_i(B_1,\Lambda)\mu_{n-i+1}(B_1,\Lambda^*) \leq n.$$

This result is close to being best possible, since there exists a numerical constant $C' > 0$ and self-dual lattices $\Lambda_n \subseteq \mathbb{R}^n$ such that

$$\mu_1(B_1,\Lambda_n)\mu_n(B_1,\Lambda_n) \geq \mu_1(B_1,\Lambda_n)^2 \geq C'n.$$

Two other applications of Minkowski's theorem on successive minima will be mentioned here. The first is a sharp form, due to Bombieri and Vaaler (1983), of 'Siegel's lemma'. In his investigations on transcendental numbers Siegel (1929) used Dirichlet's pigeonhole principle to prove that if $A = (\alpha_{jk})$ is an $m \times n$ matrix of integers, where $m < n$, such that $|\alpha_{jk}| \leq \beta$ for all j,k, then the system of homogeneous linear equations

$$Ax = 0$$

has a solution $x = (\xi_k)$ in integers, not all 0, such that $|\xi_k| \leq 1 + (n\beta)^{m/(n-m)}$ for all k. Bombieri and Vaaler show that, if A has rank m and if $g > 0$ is the greatest common divisor of all $m \times m$

subdeterminants of A, then there are $n - m$ linearly independent integral solutions $x_j = (\xi_{jk})$ $(j = 1,\ldots,n - m)$ such that

$$\Pi_{j=1}^{n-m} \|x_j\| \leq [\det(AA^t)]^{1/2}/g,$$

where $\|x_j\| = \max_k |\xi_{jk}|$.

The second application, due to Gillet and Soulé (1991), may be regarded as an arithmetic analogue of the Riemann–Roch theorem for function fields. Again let K be a compact symmetric convex subset of $\mathbb{R}^n$ with nonempty interior and let μ_i denote the infimum of all $\rho > 0$ such that ρK contains at least i linearly independent points of $\mathbb{Z}^n$ $(i = 1,\ldots,n)$. If $M(K)$ is the number of points of $\mathbb{Z}^n$ in K, and if h is the maximum number of linearly independent points of $\mathbb{Z}^n$ in the interior of K, then Gillet and Soulé show that $\mu_1 \cdots \mu_h/M(K)$ is bounded above and below by positive constants, which depend on n but not on K.

A number of results in this section have dealt with compact symmetric convex sets with nonempty interior. Since such sets may appear rather special, it should be pointed out that they arise very naturally in connection with normed vector spaces.

The vector space $\mathbb{R}^n$ is said to be *normed* if with each $x \in \mathbb{R}^n$ there is associated a real number $|x|$ with the properties

(i) $|x| \geq 0$, with equality if and only if $x = O$,
(ii) $|x + y| \leq |x| + |y|$ for all $x,y \in \mathbb{R}^n$,
(iii) $|\alpha x| = |\alpha|\,|x|$ for all $x \in \mathbb{R}^n$ and all $\alpha \in \mathbb{R}$.

Let K denote the set of all $x \in \mathbb{R}^n$ such that $|x| \leq 1$. Then K is bounded, since all norms on a finite-dimensional vector space are equivalent. In fact K is compact, since it follows from (ii) that K is closed. Moreover K is convex and symmetric, by (ii) and (iii). Furthermore, by (i) and (iii), $x/|x| \in K$ for each nonzero $x \in \mathbb{R}^n$. Hence the interior of K is nonempty and is actually the set of all $x \in \mathbb{R}^n$ such that $|x| < 1$.

Conversely, let K be a compact symmetric convex subset of $\mathbb{R}^n$ with nonempty interior. Then the origin is an interior point of K and for each nonzero $x \in \mathbb{R}^n$ there is a unique $\rho > 0$ such that ρx is on the boundary of K. If we put $|x| = \rho^{-1}$, and $|O| = 0$, then (i) obviously holds. Furthermore, since $|-x| = |x|$, it is easily seen that (iii) holds. Finally, if $y \in \mathbb{R}^n$ and $|y| = \sigma^{-1}$, then $\rho x, \sigma y \in K$ and hence, since K is convex,

$$\rho\sigma(\rho + \sigma)^{-1}(x + y) = \sigma(\rho + \sigma)^{-1}\rho x + \rho(\rho + \sigma)^{-1}\sigma y \in K.$$

Hence

$$|x + y| \leq (\rho + \sigma)/\rho\sigma = |x| + |y|.$$

Thus $\mathbb{R}^n$ is a normed vector space and K the set of all $x \in \mathbb{R}^n$ such that $|x| \leq 1$.

4 Voronoi cells

Throughout this section we suppose $\mathbb{R}^n$ equipped with the *Euclidean metric*:

$$d(y,z) = \|y - z\|,$$

where $\|x\| = (x^t x)^{1/2}$. We call $\|x\|^2 = x^t x$ the *square-norm* of x and we denote the scalar product $y^t z$ by (y,z).

Fix some point $x_0 \in \mathbb{R}^n$. For any point $x \neq x_0$, the set of all points which are equidistant from x_0 and x is the hyperplane H_x which passes through the midpoint of the segment joining x_0 and x and is orthogonal to this segment. Analytically, H_x is the set of all $y \in \mathbb{R}^n$ such that

$$(x - x_0,y) = (x - x_0,x + x_0)/2,$$

which simplifies to

$$2(x - x_0,y) = \|x\|^2 - \|x_0\|^2.$$

The set of all points which are closer to x_0 than to x is the open half-space G_x consisting of all points $y \in \mathbb{R}^n$ such that

$$2(x - x_0,y) < \|x\|^2 - \|x_0\|^2.$$

The closed half-space $\overline{G_x} = H_x \cup G_x$ is the set of all points at least as close to x_0 as to x.

Let X be a subset of $\mathbb{R}^n$ containing more than one point which is *discrete*, i.e. for each $y \in \mathbb{R}^n$ there exists an open set containing y which contains at most one point of X. It follows that each bounded subset of $\mathbb{R}^n$ contains only finitely many points of X since, if there were infinitely many, they would have an accumulation point. Hence for each $y \in \mathbb{R}^n$ there exists an $x_0 \in X$ whose distance from y is minimal:

$$d(x_0,y) \leq d(x,y) \quad \text{for every } x \in X. \tag{1}$$

For each $x_0 \in X$ we define its *Voronoi cell* $V(x_0)$ to be the set of all $y \in \mathbb{R}^n$ for which (1) holds. Voronoi cells are also called 'Dirichlet domains', since they were used by Dirichlet (1850) in $\mathbb{R}^2$ before Voronoi (1908) used them in $\mathbb{R}^n$.

If we choose $r > 0$ so that the open ball

$$\beta_r(x_0): = \{y \in \mathbb{R}^n: d(x_0,y) < r\}$$

contains no point of X except x_0, then $\beta_{r/2}(x_0) \subseteq V(x_0)$. Thus x_0 is an interior point of $V(x_0)$.

Since

$$\overline{G}_x = \{y \in \mathbb{R}^n: \mathrm{d}(x_0,y) \le \mathrm{d}(x,y)\},$$

we have $V(x_0) \subseteq \overline{G}_x$ and actually

$$V(x_0) = \bigcap_{x\in X\backslash x_0} \overline{G}_x. \qquad (2)$$

It follows at once from (2) that $V(x_0)$ is closed and convex. Hence $V(x_0)$ is the closure of its nonempty interior.

According to the definitions of §3, the Voronoi cells form a tiling of $\mathbb{R}^n$, since

$$\mathbb{R}^n = \bigcup_{x\in X} V(x),$$

$$\operatorname{int} V(x) \cap \operatorname{int} V(x') = \emptyset \quad \text{if } x,x' \in X \text{ and } x \neq x'.$$

A subset A of a convex set C is said to be a *face* of C if A is convex and, for any $c,c' \in C$, $(c,c') \cap A \neq \emptyset$ implies $c,c' \in A$. The tiling by Voronoi cells has the additional property that $V(x) \cap V(x')$ is a face of both $V(x)$ and $V(x')$ if $x,x' \in X$ and $x \neq x'$. We will prove this by showing that if y_1,y_2 are distinct points of $V(x)$ and if $z \in (y_1,y_2) \cap V(x')$, then $y_1 \in V(x')$.

Since $z \in V(x) \cap V(x')$, we have $\mathrm{d}(x,z) = \mathrm{d}(x',z)$. Thus z lies on the hyperplane H which passes through the midpoint of the segment joining x and x' and is orthogonal to this segment. If $y_1 \notin V(x')$, then $\mathrm{d}(x,y_1) < \mathrm{d}(x',y_1)$. Thus y_1 lies in the open half-space G associated with the hyperplane H which contains x. But then y_2 lies in the open half-space G' which contains x', i.e. $\mathrm{d}(x',y_2) < \mathrm{d}(x,y_2)$, which contradicts $y_2 \in V(x)$.

We now assume that the set X is not only discrete, but also *relatively dense*, i.e.

(†) there exists $R > 0$ such that, for each $y \in \mathbb{R}^n$, there is some $x \in X$ with $\mathrm{d}(x,y) < R$.

It follows at once that $V(x_0) \subseteq \beta_R(x_0)$. Thus $V(x_0)$ is bounded and, since it is closed, even compact. The ball $\beta_{2R}(x_0)$ contains only finitely many points $x_1,\ldots,x_m$ of X apart from x_0. We are going to show that

$$V(x_0) = \bigcap_{i=1}^{m} \overline{G}_{x_i}. \qquad (3)$$

By (2) we need only show that if $y \in \bigcap_{i=1}^{m} \overline{G}_{x_i}$, then $y \in \overline{G}_x$ for every $x \in X$.

Assume that $\mathrm{d}(x_0,y) \ge R$ and choose z on the segment joining x_0 and y so that $\mathrm{d}(x_0,z) = R$. For some $x \in X$ we have $\mathrm{d}(x,z) < R$ and hence $0 < \mathrm{d}(x_0,x) < 2R$. Consequently $x = x_i$ for some $i \in \{1,\ldots,m\}$. Since $\mathrm{d}(x_i,z) < R = \mathrm{d}(x_0,z)$, we have $z \notin \overline{G}_{x_i}$. But this is a contradiction, since $x_0,y \in \overline{G}_{x_i}$ and z is on the segment joining them.

We conclude that $\mathrm{d}(x_0,y) < R$. If $x \in X$ and $x \neq x_0,x_1,\ldots,x_m$, then

$$\begin{aligned} d(x,y) &\geq d(x_0,x) - d(x_0,y) \\ &\geq 2R - R = R > d(x_0,y). \end{aligned}$$

Consequently $y \in \overline{G}_x$ for every $x \in X$.

It follows from (3) that $V(x_0)$ is a polyhedron. Since $V(x_0)$ is bounded and has a nonempty interior, it is actually an *n-dimensional polytope*.

The faces of a polytope are an important part of its structure. An $(n-1)$-dimensional face of an n-dimensional polytope is said to be a *facet* and a 0-dimensional face is said to be a *vertex*. We now apply to $V(x_0)$ some properties common to all polytopes.

In the representation (3) it may be possible to omit some closed half-spaces $\overline{G}_{x_i}$ without affecting the validity of the representation. By omitting as many half-spaces as possible we obtain an *irredundant representation*, which by suitable choice of notation we may take to be

$$V(x_0) = \bigcap_{i=1}^{l} \overline{G}_{x_i}$$

for some $l \leq m$. The intersections $V(x_0) \cap H_{x_i}$ $(1 \leq i \leq l)$ are then the distinct facets of $V(x_0)$. Any nonempty proper face of $V(x_0)$ is contained in a facet and is the intersection of those facets which contain it. Furthermore, any nonempty face of $V(x_0)$ is the convex hull of those vertices of $V(x_0)$ which it contains.

It follows that for each x_i $(1 \leq i \leq l)$ there is a vertex v_i of $V(x_0)$ such that

$$d(x_0,v_i) = d(x_i,v_i).$$

For $d(x_0,v) \leq d(x_i,v)$ for every vertex v of $V(x_0)$. Assume that $d(x_0,v) < d(x_i,v)$ for every vertex v of $V(x_0)$. Then the open half-space G_{x_i} contains all vertices v and hence also their convex hull $V(x_0)$. But this is a contradiction, since $V(x_0) \cap H_{x_i}$ is a facet of $V(x_0)$.

To illustrate these results take $X = \mathbb{Z}^n$ and $x_0 = O$. Then the Voronoi cell $V(O)$ is the cube consisting of all points $y = (\eta_1,\ldots,\eta_n) \in \mathbb{R}^n$ with $|\eta_i| \leq 1/2$ $(i = 1,\ldots,n)$. It has the minimal number $2n$ of facets.

In fact any lattice Λ in $\mathbb{R}^n$ is discrete and has the property (†). *For a lattice* Λ *we can restrict attention to the Voronoi cell* $V(\Lambda)$: $= V(O)$, since an arbitrary Voronoi cell is obtained from it by a translation: $V(x_0) = V(O) + x_0$. The Voronoi cell of a lattice has additional properties. Since $x \in \Lambda$ implies $-x \in \Lambda$, $y \in V(\Lambda)$ implies $-y \in V(\Lambda)$. Also, if x_i is a lattice vector determining a facet of $V(\Lambda)$ and if $y \in V(\Lambda) \cap H_{x_i}$, then $\|y\| = \|y - x_i\|$. Since $x \in \Lambda$ implies $x_i - x \in \Lambda$, it follows that $y \in V(\Lambda) \cap H_{x_i}$ implies $x_i - y \in V(\Lambda) \cap H_{x_i}$. Thus *the Voronoi cell* $V(\Lambda)$ *and all its facets are centrosymmetric.*

In addition, any orthogonal transformation of $\mathbb{R}^n$ which maps onto itself the lattice Λ also maps onto itself the Voronoi cell $V(\Lambda)$. Furthermore the Voronoi cell $V(\Lambda)$ has volume $d(\Lambda)$, by Proposition 11, since the lattice translates of $V(\Lambda)$ form a tiling of $\mathbb{R}^n$.

We define a *facet vector* or 'relevant vector' of a lattice Λ to be a vector $x_i \in \Lambda$ such that $V(\Lambda) \cap H_{x_i}$ is a facet of the Voronoi cell $V(\Lambda)$. If $V(\Lambda)$ is contained in the closed ball $B_R = \{x \in \mathbb{R}^n : \|x\| \leq R\}$, then every facet vector x_i satisfies $\|x_i\| \leq 2R$. For, if $y \in V(\Lambda) \cap H_{x_i}$ then, by Schwarz's inequality (Chapter I, §4),

$$\|x_i\|^2 = 2(x_i,y) \leq 2\|x_i\| \, \|y\|.$$

The facet vectors were characterized by Voronoi (1908) in the following way:

PROPOSITION 13 *A nonzero vector $x \in \Lambda$ is a facet vector of the lattice $\Lambda \subseteq \mathbb{R}^n$ if and only if every vector $x' \in x + 2\Lambda$, except $\pm x$, satisfies $\|x'\| > \|x\|$.*

Proof Suppose first that $\|x\| < \|x'\|$ for all $x' \neq \pm x$ such that $(x' - x)/2 \in \Lambda$. If $z \in \Lambda$ and $x' = 2z - x$, then $(x' - x)/2 \in \Lambda$. Hence if $z \neq O,x$ then

$$\|x/2\| < \|z - x/2\|,$$

i.e. $x/2 \in G_z$. Since $\|x/2\| = \|x - x/2\|$, it follows that $x/2 \in V(\Lambda)$ and x is a facet vector.

Suppose next that there exists $x' \neq \pm x$ such that $w = (x' - x)/2 \in \Lambda$ and $\|x'\| \leq \|x\|$. Then also $z = (x' + x)/2 \in \Lambda$ and $z,w \neq O$. If $y \in \overline{G}_z \cap \overline{G}_{-w}$, then

$$2(z,y) \leq \|z\|^2, \quad -2(w,y) \leq \|w\|^2.$$

Hence, by the parallelogram law (Chapter I, §10),

$$\begin{aligned} 2(x,y) &= 2(z,y) - 2(w,y) \leq \|z\|^2 + \|w\|^2 \\ &= \|x\|^2/2 + \|x'\|^2/2 \leq \|x\|^2. \end{aligned}$$

That is, $y \in \overline{G}_x$. Thus $\overline{G}_x$ is not needed to define $V(\Lambda)$ and x is not a facet vector. □

Any lattice Λ contains a nonzero vector with minimal square-norm. Such a vector will be called a *minimal vector*. Its square-norm will be called the *minimum* of Λ and will be denoted by $m(\Lambda)$.

PROPOSITION 14 *If $\Lambda \subseteq \mathbb{R}^n$ is a lattice with minimum $m(\Lambda)$, then any nonzero vector in Λ with square-norm $< 2m(\Lambda)$ is a facet vector. In particular, any minimal vector is a facet vector.*

Proof Put $r = m(\Lambda)$ and let x be a nonzero vector in Λ with $\|x\|^2 < 2r$. If x is not a facet vector, there exists $y \neq \pm x$ with $(y - x)/2 \in \Lambda$ such that $\|y\| \leq \|x\|$. Since $(y \pm x)/2 \in \Lambda$, $\|x \pm y\|^2 \geq 4r$. Thus

$$4r \leq \|x\|^2 + \|y\|^2 \pm 2(x,y) < 4r \pm 2(x,y),$$

which is impossible. □

PROPOSITION 15 *For any lattice $\Lambda \subseteq \mathbb{R}^n$, the number of facets of its Voronoi cell $V(\Lambda)$ is at most* $2(2^n - 1)$.

Proof Let $x_1,...,x_n$ be a basis for Λ. Then any vector $x \in \Lambda$ has a unique representation $x = x' + x''$, where $x' \in 2\Lambda$ and

$$x'' = \alpha_1 x_1 + ... + \alpha_n x_n,$$

with $\alpha_j \in \{0,1\}$ for $j = 1,...,n$. Thus the number of cosets of 2Λ in Λ is 2^n. But, by Proposition 13, each coset contains at most one pair $\pm y$ of facet vectors. Since 2Λ itself does not contain any facet vectors, the total number of facet vectors is at most $2(2^n - 1)$. □

There exist lattices $\Lambda \subseteq \mathbb{R}^n$ for which the upper bound of Proposition 15 is attained, e.g. the lattice $\Lambda = \{Tz: z \in \mathbb{Z}^n\}$ with $T = I + \beta J$, where J is the $n\times n$ matrix every element of which is 1 and $\beta = \{(1 + n)^{1/2} - 1\}/n$.

PROPOSITION 16 *Every vector of a lattice $\Lambda \subseteq \mathbb{R}^n$ is an integral linear combination of facet vectors.*

Proof Let $b_1,...,b_m$ be the facet vectors of Λ and put

$$\Lambda' = \{x = \beta_1 b_1 + ... + \beta_m b_m: \beta_1,...,\beta_m \in \mathbb{Z}\}.$$

Evidently Λ' is a subgroup of $\mathbb{R}^n$ and actually a discrete subgroup, since $\Lambda' \subseteq \Lambda$. If Λ' were contained in a hyperplane of $\mathbb{R}^n$ any point on the line through the origin orthogonal to this hyperplane would belong to the Voronoi cell V of Λ, which is impossible because V is bounded. Hence Λ' contains n linearly independent vectors.

Thus Λ' is a sublattice of Λ. It follows that the Voronoi cell V of Λ is contained in the Voronoi cell V' of Λ'. But if $y \in V'$, then

$$\|y\| \leq \|b_i - y\|, \quad (i = 1,...,m)$$

and hence $y \in V$. Thus $V' = V$. Hence the Λ'-translates of V and the Λ-translates of V are both tilings of $\mathbb{R}^n$. Since $\Lambda' \subseteq \Lambda$, this is possible only if $\Lambda' = \Lambda$. □

Since every integral linear combination of facet vectors is in the lattice, Proposition 16 implies

COROLLARY 17 *Distinct lattices in $\mathbb{R}^n$ have distinct Voronoi cells.* □

Proposition 16 does not say that the lattice has a basis of facet vectors. It is known that every lattice in $\mathbb{R}^n$ has a basis of facet vectors if $n \leq 6$, but if $n > 6$ this is still an open question. It is known also that every lattice in $\mathbb{R}^n$ has a basis of minimal vectors when $n \leq 4$ but, when $n > 4$, there are lattices with no such basis. In fact a lattice may have no basis of minimal vectors, even though every lattice vector is an integral linear combination of minimal vectors.

Lattices and their Voronoi cells have long been used in crystallography. An n-dimensional *crystal* may be defined mathematically to be a subset of $\mathbb{R}^n$ of the form

$$F + \Lambda = \{x + y : x \in F, y \in \Lambda\},$$

where F is a finite set and Λ a lattice. Crystals may be studied by means of their symmetry groups.

An *isometry* of $\mathbb{R}^n$ is an invertible affine transformation which leaves unaltered the Euclidean distance between any two points. For example, any orthogonal transformation is an isometry and so is a translation by an arbitrary vector v. Any isometry is the composite of a translation and an orthogonal transformation. The *symmetry group* of a set $X \subseteq \mathbb{R}^n$ is the group of all isometries of $\mathbb{R}^n$ which map X to itself.

We define an n-dimensional *crystallographic group* to be a group G of isometries of $\mathbb{R}^n$ such that the vectors corresponding to translations in G form an n-dimensional lattice. It is not difficult to show that a subset of $\mathbb{R}^n$ is an n-dimensional crystal if and only if it is discrete and its symmetry group is an n-dimensional crystallographic group.

It was shown by Bieberbach (1911) that a group G of isometries of $\mathbb{R}^n$ is a crystallographic group if and only if it is discrete and has a compact fundamental domain D, i.e. the sets $\{g(D) : g \in G\}$ form a tiling of $\mathbb{R}^n$. He could then show that the translations in a crystallographic group form a torsion-free abelian normal subgroup of finite index. He showed later (1912) that two crystallographic groups G_1, G_2 are isomorphic if and only if there exists an invertible affine transformation A such that $G_2 = A^{-1}G_1A$. With the aid of results of Minkowski and Jordan it follows that, for a given dimension n, there are only finitely many non-isomorphic crystallographic groups. These results provided a positive answer to the first part of the 18th Paris problem of Hilbert (1900).

The structure of physical crystals is analysed by means of the corresponding 3-dimensional crystallographic groups. A stronger concept than isomorphism is useful for such applications.

Two crystallographic groups G_1,G_2 may be said to be *properly isomorphic* if there exists an orientation-preserving invertible affine transformation A such that $G_2 = A^{-1}G_1A$. An isomorphism class of crystallographic groups either coincides with a proper isomorphism class or splits into two distinct proper isomorphism classes.

Fedorov (1891) showed that there are 17 isomorphism classes of 2-dimensional crystallographic groups, none of which splits. Collating earlier work of Sohncke (1879), Schoenflies (1889) and himself, Fedorov (1892) also showed that there are 219 isomorphism classes of 3-dimensional crystallographic groups, 11 of which split. More recently, Brown *et al.* (1978) have shown that there are 4783 isomorphism classes of 4-dimensional crystallographic groups, 112 of which split.

5 Densest packings

The result of Hermite, mentioned at the beginning of the chapter, can be formulated in terms of lattices instead of quadratic forms. For any real non-singular matrix T, the matrix

$$A = T^tT$$

is a real positive definite symmetric matrix. Conversely, by a principal axes transformation or otherwise, it may be seen that any real positive definite symmetric matrix A may be represented in this way.

Let Λ be the lattice

$$\Lambda = \{y = Tx \in \mathbb{R}^n : x \in \mathbb{Z}^n\}$$

and put

$$\gamma(\Lambda) = m(\Lambda)/\mathrm{d}(\Lambda)^{2/n},$$

where $\mathrm{d}(\Lambda)$ is the determinant and $m(\Lambda)$ the minimum of Λ. Evidently $\gamma(\rho\Lambda) = \gamma(\Lambda)$ for any $\rho > 0$. Hermite's result that there exists a positive constant c_n, depending only on n, such that $0 < x^tAx \le c_n(\det A)^{1/n}$ for some $x \in \mathbb{Z}^n$ may be restated in the form

$$\gamma(\Lambda) \le c_n.$$

Hermite's constant γ_n is defined to be the least positive constant c_n such that this inequality holds for all $\Lambda \subseteq \mathbb{R}^n$.

It may be shown that $\gamma_n{}^n$ is a rational number for each n. It follows from Proposition 2 that $\overline{\lim}_{n\to\infty}\gamma_n/n \le 2/\pi e$. Minkowski (1905) showed also that

$$\underline{\lim}_{n\to\infty}\gamma_n/n \;\geq\; 1/2\pi e \;=\; 0.0585\ldots\,,$$

and it is possible that actually $\lim_{n\to\infty}\gamma_n/n = 1/2\pi e$. The significance of Hermite's constant derives from its connection with lattice packings of balls, as we now explain.

Let Λ be a lattice in $\mathbb{R}^n$ and K a subset of $\mathbb{R}^n$ which is the closure of a nonempty open set G. We say that Λ gives a *lattice packing* for K if the family of translates $K + x$ ($x \in \Lambda$) is a packing of $\mathbb{R}^n$, i.e. if for any two distinct points $x,y \in \Lambda$ the interiors $G + x$ and $G + y$ are disjoint. This is the same as saying that Λ does not contain the difference of any two distinct points of the interior of K, since $g + x = g' + y$ if and only if $g' - g = x - y$. If K is a compact symmetric convex set with nonempty interior G, it is the same as saying that the interior of the set $2K$ contains no nonzero point of Λ, since in this case $g,g' \in G$ implies $(g' - g)/2 \in G$ and $2g = g - (-g)$.

The *density* of the lattice packing, i.e. the fraction of the total space which is occupied by translates of K, is readily shown to be $\lambda(K)/\mathrm{d}(\Lambda)$. Hence the maximum density of any lattice packing for K is

$$\delta(K) \;=\; \lambda(K)/\Delta(2K) \;=\; 2^{-n}\lambda(K)/\Delta(K),$$

where $\Delta(K)$ is the critical determinant of K, as defined in §3. The use of the word 'maximum' is justified, since it will be shown in §6 that the infimum involved in the definition of critical determinant is attained.

We are interested in the special case of a closed ball: $K = B_\rho = \{x \in \mathbb{R}^n\colon \|x\| \leq \rho\}$. By what we have said, Λ gives a lattice packing for B_ρ if and only if the interior of $B_{2\rho}$ contains no nonzero point of Λ, i.e. if and only if $m(\Lambda)^{1/2} \geq 2\rho$. Hence

$$\delta(B_\rho) \;=\; \sup\,\{\lambda(B_\rho)/\mathrm{d}(\Lambda)\colon m(\Lambda)^{1/2} = 2\rho\}$$

$$=\; \kappa_n\rho^n \sup\,\{\mathrm{d}(\Lambda)^{-1}\colon m(\Lambda)^{1/2} = 2\rho\},$$

where $\kappa_n = \pi^{n/2}/(n/2)!$ again denotes the volume of the unit ball in $\mathbb{R}^n$. By homogeneity it follows that

$$\delta_n\colon = \delta(B_\rho) \;=\; 2^{-n}\kappa_n \sup_\Lambda \gamma(\Lambda)^{n/2},$$

where the supremum is now over all lattices $\Lambda \subseteq \mathbb{R}^n$, i.e. in terms of Hermite's constant γ_n,

$$\delta_n \;=\; 2^{-n}\kappa_n\gamma_n^{\,n/2}.$$

Thus γ_n, like δ_n, measures the densest lattice packing of balls. A lattice $\Lambda \subseteq \mathbb{R}^n$ for which $\gamma(\Lambda) = \gamma_n$, i.e. a critical lattice for a ball, will be called simply a *densest lattice*.

The densest lattice in $\mathbb{R}^n$ is known for each $n \leq 8$, and is uniquely determined apart from isometries and scalar multiples. In fact these densest lattices are all examples of indecomposable root lattices. These terms will now be defined.

A lattice Λ is said to be *decomposable* if there exist additive subgroups Λ_1,Λ_2 of Λ, each containing a nonzero vector, such that $(x_1,x_2) = 0$ for all $x_1 \in \Lambda_1$ and $x_2 \in \Lambda_2$, and every vector in Λ is the sum of a vector in Λ_1 and a vector in Λ_2. Since Λ_1 and Λ_2 are necessarily discrete, they are lattices in the wide sense (i.e. they are not full-dimensional). We say also that Λ is the *orthogonal sum* of the lattices Λ_1 and Λ_2. The orthogonal sum of any finite number of lattices is defined similarly. A lattice is *indecomposable* if it is not decomposable.

The following result was first proved by Eichler (1952).

PROPOSITION 18 *Any lattice Λ is an orthogonal sum of finitely many indecomposable lattices, which are uniquely determined apart from order.*

Proof (i) Define a vector $x \in \Lambda$ to be 'decomposable' if there exist nonzero $x_1,x_2 \in \Lambda$ such that $x = x_1 + x_2$ and $(x_1,x_2) = 0$. We show first that every nonzero $x \in \Lambda$ is a sum of finitely many indecomposable vectors.

By definition, x is either indecomposable or is the sum of two nonzero orthogonal vectors in Λ. Both these vectors have square-norm less than the square-norm of x, and for each of them the same alternative presents itself. Continuing in this way, we must eventually arrive at indecomposable vectors, since there are only finitely many vectors in Λ with square-norm less than that of x.

(ii) If Λ is the orthogonal sum of finitely many lattices L_ν then, by the definition of an orthogonal sum, every indecomposable vector of Λ lies in one of the sublattices L_ν. Hence if two indecomposable vectors are not orthogonal, they lie in the same sublattice L_ν.

(iii) Call two indecomposable vectors x,x' 'equivalent' if there exist indecomposable vectors $x = x_0,x_1,\ldots,x_{k-1},x_k = x'$ such that $(x_j,x_{j+1}) \neq 0$ for $0 \leq j < k$. Clearly 'equivalence' is indeed an equivalence relation and thus the set of all indecomposable vectors is partitioned into equivalence classes $\mathscr{C}_\mu$. Two vectors from different equivalence classes are orthogonal and, if Λ is an orthogonal sum of lattices L_ν as in (ii), then two vectors from the same equivalence class lie in the same sublattice L_ν.

(iv) Let Λ_μ be the subgroup of Λ generated by the vectors in the equivalence class $\mathscr{C}_\mu$. Then, by (i), Λ is generated by the sublattices Λ_μ. Since, by (iii), Λ_μ is orthogonal to $\Lambda_{\mu'}$ if $\mu \neq \mu'$, Λ is actually the orthogonal sum of the sublattices Λ_μ. If Λ is an orthogonal sum of lattices L_ν as in (ii), then each Λ_μ is contained in some L_ν. It follows that each Λ_μ is indecomposable and that these indecomposable sublattices are uniquely determined apart from order. □

Let Λ be a lattice in $\mathbb{R}^n$. If $\Lambda \subseteq \Lambda^*$, i.e. if $(x,y) \in \mathbb{Z}$ for all $x,y \in \Lambda$, then Λ is said to be *integral*. If (x,x) is an even integer for every $x \in \Lambda$, then Λ is said to be *even*. (It follows that an even lattice is also integral.) If Λ is even and every vector in Λ is an integral linear combination of vectors in Λ with square-norm 2, then Λ is said to be a *root lattice*.

Thus in a root lattice the minimal vectors have square-norm 2. It may be shown by a long, but elementary, argument that any root lattice has a basis of minimal vectors such that every minimal vector is an integral linear combination of the basis vectors with coefficients which are all nonnegative or all nonpositive. Such a basis will be called a *simple* basis. The facet vectors of a root lattice are precisely the minimal vectors, and hence its Voronoi cell is the set of all $y \in \mathbb{R}^n$ such that $(y,x) \leq 1$ for every minimal vector x.

Any root lattice is an orthogonal sum of indecomposable root lattices. It was shown by Witt (1941) that the indecomposable root lattices can be completely enumerated; they are all listed in Table 1. We give also their minimal vectors in terms of the canonical basis $e_1,\ldots,e_n$ of $\mathbb{R}^n$.

$$\begin{aligned}
A_n &= \{x = (\xi_0,\xi_1,\ldots,\xi_n) \in \mathbb{Z}^{n+1}: \xi_0 + \xi_1 + \ldots + \xi_n = 0\} \quad (n \geq 1);\\
D_n &= \{x = (\xi_1,\ldots,\xi_n) \in \mathbb{Z}^n: \xi_1 + \ldots + \xi_n \text{ even}\} \quad (n \geq 3);\\
E_8 &= D_8 \cup D_8^\dagger, \text{ where } D_8^\dagger = (1/2,1/2,\ldots,1/2) + D_8;\\
E_7 &= \{x = (\xi_1,\ldots,\xi_8) \in E_8: \xi_7 = -\xi_8\};\\
E_6 &= \{x = (\xi_1,\ldots,\xi_8) \in E_8: \xi_6 = \xi_7 = -\xi_8\}.
\end{aligned}$$

Table 1: Indecomposable root lattices

The lattice A_n has $n(n+1)$ minimal vectors, namely the vectors $\pm(e_j - e_k)$ $(0 \leq j < k \leq n)$, and the vectors $e_0 - e_1, e_1 - e_2, \ldots, e_{n-1} - e_n$ form a simple basis. By calculating the determinant of B^tB, where B is the $(n+1)\times n$ matrix whose columns are the vectors of this simple basis, it may be seen that the determinant of the lattice A_n is $(n+1)^{1/2}$.

The lattice D_n has $2n(n-1)$ minimal vectors, namely the vectors $\pm e_j \pm e_k$ $(1 \leq j < k \leq n)$. The vectors $e_1 - e_2, e_2 - e_3, \ldots, e_{n-1} - e_n, e_{n-1} + e_n$ form a simple basis and hence the lattice D_n has determinant 2.

The lattice E_8 has 240 minimal vectors, namely the 112 vectors $\pm e_j \pm e_k$ $(1 \leq j < k \leq 8)$ and the 128 vectors $(\pm e_1 \pm \ldots \pm e_8)/2$ with an even number of minus signs. The vectors

$$v_1 = (e_1 - e_2 - \ldots - e_7 + e_8)/2,\ v_2 = e_1 + e_2,\ v_3 = e_2 - e_1,\ v_4 = e_3 - e_2,\ \ldots,\ v_8 = e_7 - e_6,$$

form a simple basis and hence the lattice has determinant 1.

The lattice E_7 has 126 minimal vectors, namely the 60 vectors $\pm e_j \pm e_k$ $(1 \leq j < k \leq 6)$, the vectors $\pm (e_7 - e_8)$ and the 64 vectors $\pm (\sum_{i=1}^{6} (\pm e_i) - e_7 + e_8)/2$ with an odd number of minus signs in the sum. The vectors $v_1,...,v_7$ form a simple basis and the lattice has determinant $\sqrt{2}$.

The lattice E_6 has 72 minimal vectors, namely the 40 vectors $\pm e_j \pm e_k$ $(1 \leq j < k \leq 5)$ and the 32 vectors $\pm (\sum_{i=1}^{5} (\pm e_i) - e_6 - e_7 + e_8)/2$ with an even number of minus signs in the sum. The vectors $v_1,...,v_6$ form a simple basis and the lattice has determinant $\sqrt{3}$.

We now return to lattice packings of balls. The densest lattices for $n \leq 8$ are given in Table 2. These lattices were shown to be densest by Lagrange (1773) for $n = 2$, Gauss (1831) for $n = 3$, Korkine and Zolotareff (1872,1877) for $n = 4,5$ and Blichfeldt (1925,1926,1934) for $n = 6,7,8$.

n	Λ	γ_n	δ_n
1	A_1	1	1
2	A_2	$(4/3)^{1/2} = 1.1547..$	$3^{1/2}\pi/6 = 0.9068..$
3	D_3	$2^{1/3} = 1.2599..$	$2^{1/2}\pi/6 = 0.7404..$
4	D_4	$2^{1/2} = 1.4142..$	$\pi^2/16 = 0.6168..$
5	D_5	$8^{1/5} = 1.5157..$	$2^{1/2}\pi^2/30 = 0.4652..$
6	E_6	$(64/3)^{1/6} = 1.6653..$	$3^{1/2}\pi^3/144 = 0.3729..$
7	E_7	$(64)^{1/7} = 1.8114..$	$\pi^3/105 = 0.2952..$
8	E_8	2	$\pi^4/384 = 0.2536..$

Table 2: Densest lattices in $\mathbb{R}^n$

Although the densest lattice in $\mathbb{R}^n$ is unknown for every $n > 8$, there are plausible candidates in some dimensions. In particular, a lattice discovered by Leech (1967) is believed to be densest in 24 dimensions. This lattice may be constructed in the following way. Let p be a prime such that $p \equiv 3 \bmod 4$ and let H_n be the Hadamard matrix of order $n = p + 1$ constructed by Paley's method (see Chapter V, §2). The columns of the matrix

$$T = (n/4 + 1)^{-1/2} \begin{pmatrix} (n/4+1)I_n & H_n - I_n \\ 0_n & I_n \end{pmatrix}$$

generate a lattice in $\mathbb{R}^{2n}$. For $p = 3$ we obtain the root lattice E_8 and for $p = 11$ the Leech lattice Λ_{24}.

Leech's lattice may be characterized as the unique even lattice Λ in $\mathbb{R}^{24}$ with $d(\Lambda) = 1$ and $m(\Lambda) > 2$. It was shown by Conway (1969) that, if G is the group of all orthogonal transformations of $\mathbb{R}^{24}$ which map the Leech lattice Λ_{24} onto itself, then the factor group $G/\{\pm I_{24}\}$ is a finite simple group, and two more finite simple groups are easily obtained as (stabilizer) subgroups. These are three of the 26 sporadic simple groups which were mentioned in §7 of Chapter V.

Leech's lattice has 196560 minimal vectors of square-norm 4. Thus the packing of unit balls associated with Λ_{24} is such that each ball touches 196560 other balls. It has been shown that 196560 is the maximal number of nonoverlapping unit balls in $\mathbb{R}^{24}$ which can touch another unit ball and that, up to isometry, there is only one possible arrangement.

Similarly, since E_8 has 240 minimal vectors of square-norm 2, the packing of balls of radius $2^{-1/2}$ associated with E_8 is such that each ball touches 240 other balls. It has been shown that 240 is the maximal number of nonoverlapping balls of fixed radius in $\mathbb{R}^8$ which can touch another ball of the same radius and that, up to isometry, there is only one possible arrangement.

In general, one may ask what is the *kissing number* of $\mathbb{R}^n$, i.e. the maximal number of nonoverlapping unit balls in $\mathbb{R}^n$ which can touch another unit ball? The question, for $n = 3$, first arose in 1694 in a discussion between Newton, who claimed that the answer was 12, and Gregory, who said 13. It was first shown by Hoppe (1874) that Newton was right, but the arrangement of the 12 balls in $\mathbb{R}^3$ is *not* unique up to isometry. One possibility is to take the centres of the 12 balls to be the vertices of a regular icosahedron, the centre of which is the centre of the unit ball they touch.

The kissing number of $\mathbb{R}^1$ is clearly 2. It is not difficult to show that the kissing number of $\mathbb{R}^2$ is 6 and that the centres of the six unit balls must be the vertices of a regular hexagon, the centre of which is the centre of the unit ball they touch. For $n > 3$ the kissing number of $\mathbb{R}^n$ is unknown, except for the two cases $n = 8$ and $n = 24$ already mentioned.

6 Mahler's compactness theorem

It is useful to study not only individual lattices, but also the family $\mathcal{L}_n$ of all lattices in $\mathbb{R}^n$. A sequence of lattices $\Lambda_k \in \mathcal{L}_n$ will be said to *converge* to a lattice $\Lambda \in \mathcal{L}_n$, in symbols $\Lambda_k \to \Lambda$, if there exist bases $b_{k1},\ldots,b_{kn}$ of Λ_k $(k = 1,2,\ldots)$ and a basis $b_1,\ldots,b_n$ of Λ such that

$$b_{kj} \to b_j \text{ as } k \to \infty \quad (j = 1,\ldots,n).$$

Evidently this implies that $d(\Lambda_k) \to d(\Lambda)$ as $k \to \infty$. Also, for any $x \in \Lambda$ there exist $x_k \in \Lambda_k$

such that $x_k \to x$ as $k \to \infty$. In fact if $x = \alpha_1 b_1 + \ldots + \alpha_n b_n$, where $\alpha_i \in \mathbb{Z}$ $(i = 1,\ldots,n)$, we can take $x_k = \alpha_1 b_{k1} + \ldots + \alpha_n b_{kn}$.

It is not obvious from the definition that the limit of a sequence of lattices is uniquely determined, but this follows at once from the next result.

PROPOSITION 19 *Let Λ be a lattice in $\mathbb{R}^n$ and let $\{\Lambda_k\}$ be a sequence of lattices in $\mathbb{R}^n$ such that $\Lambda_k \to \Lambda$ as $k \to \infty$. If $x_k \in \Lambda_k$ and $x_k \to x$ as $k \to \infty$, then $x \in \Lambda$.*

Proof With the above notation,

$$x = \alpha_1 b_1 + \ldots + \alpha_n b_n,$$

where $\alpha_i \in \mathbb{R}$ $(i = 1,\ldots,n)$, and similarly

$$x_k = \alpha_{k1} b_1 + \ldots + \alpha_{kn} b_n,$$

where $\alpha_{ki} \in \mathbb{R}$ and $\alpha_{ki} \to \alpha_i$ as $k \to \infty$ $(i = 1,\ldots,n)$.

The linear transformation T_k of $\mathbb{R}^n$ which maps b_i to b_{ki} $(i = 1,\ldots,n)$ can be written in the form

$$T_k = I - A_k,$$

where $A_k \to O$ as $k \to \infty$. It follows that

$$T_k^{-1} = (I - A_k)^{-1} = I + A_k + A_k^2 + \ldots = I + C_k,$$

where also $C_k \to O$ as $k \to \infty$. Hence

$$x_k = T_k^{-1}\,(\alpha_{k1} b_{k1} + \ldots + \alpha_{kn} b_{kn})$$

$$= (\alpha_{k1} + \eta_{k1}) b_{k1} + \ldots + (\alpha_{kn} + \eta_{kn}) b_{kn},$$

where $\eta_{ki} \to 0$ as $k \to \infty$ $(i = 1,\ldots,n)$. But $\alpha_{ki} + \eta_{ki} \in \mathbb{Z}$ for every k. Letting $k \to \infty$, we obtain $\alpha_i \in \mathbb{Z}$. That is, $x \in \Lambda$. □

It is natural to ask if the Voronoi cells of a convergent sequence of lattices also converge in some sense. The required notion of convergence is in fact older than the notion of convergence of lattices and applies to arbitrary compact subsets of $\mathbb{R}^n$.

The *Hausdorff distance* $\mathrm{h}(K,K')$ between two compact subsets K,K' of $\mathbb{R}^n$ is defined to be the infimum of all $\rho > 0$ such that every point of K is distant at most ρ from some point of K' and every point of K' is distant at most ρ from some point of K. We will show that this defines a metric, the *Hausdorff metric*, on the space of all compact subsets of $\mathbb{R}^n$.

Evidently

$$0 \leq \mathrm{h}(K,K') = \mathrm{h}(K',K) < \infty.$$

Moreover $\mathrm{h}(K,K') = 0$ implies $K = K'$. For if $x' \in K'$, there exist $x_k \in K$ such that $x_k \to x'$ and hence $x' \in K$, since K is closed. Thus $K' \subseteq K$, and similarly $K \subseteq K'$.

Finally we prove the triangle inequality

$$\mathrm{h}(K,K'') \leq \mathrm{h}(K,K') + \mathrm{h}(K',K'').$$

To simplify writing, put $\rho = \mathrm{h}(K,K')$ and $\rho' = \mathrm{h}(K',K'')$. For any $\varepsilon > 0$, if $x \in K$ there exist $x' \in K'$ such that $\|x - x'\| < \rho + \varepsilon$ and then $x'' \in K''$ such that $\|x' - x''\| < \rho' + \varepsilon$. Hence

$$\|x - x''\| < \rho + \rho' + 2\varepsilon.$$

Similarly, if $x'' \in K''$ there exists $x \in K$ for which the same inequality holds. But ε can be arbitrarily small.

The definition of Hausdorff distance can also be expressed in the form

$$\mathrm{h}(K,K') = \inf\{\rho \geq 0 : K \subseteq K' + B_\rho, K' \subseteq K + B_\rho\},$$

where $B_\rho = \{x \in \mathbb{R}^n : \|x\| \leq \rho\}$. A sequence K_j of compact subsets of $\mathbb{R}^n$ *converges* to a compact subset K of $\mathbb{R}^n$ if $\mathrm{h}(K_j,K) \to 0$ as $j \to \infty$.

It was shown by Hausdorff (1927) that any uniformly bounded sequence of compact subsets of $\mathbb{R}^n$ has a convergent subsequence. In particular, any uniformly bounded sequence of compact convex subsets of $\mathbb{R}^n$ has a subsequence which converges to a compact convex set. This special case of Hausdorff's result, which is all that we will later require, had already been established by Blaschke (1916) and is known as *Blaschke's selection principle.*

PROPOSITION 20 *Let $\{\Lambda_k\}$ be a sequence of lattices in $\mathbb{R}^n$ and let V_k be the Voronoi cell of Λ_k. If there exists a compact convex set V with nonempty interior such that $V_k \to V$ in the Hausdorff metric as $k \to \infty$, then V is the Voronoi cell of a lattice Λ and $\Lambda_k \to \Lambda$ as $k \to \infty$.*

Proof Since every Voronoi cell V_k is symmetric, so also is the limit V. Since V has nonempty interior, it follows that the origin is itself an interior point of V. Thus there exists $\delta > 0$ such that the ball $B_\delta = \{x \in \mathbb{R}^n : \|x\| \leq \delta\}$ is contained in V.

It follows that $B_{\delta/2} \subseteq V_k$ for all large k. The quickest way to see this is to use *Rådström's cancellation law*, which says that if A,B,C are nonempty compact convex subsets of $\mathbb{R}^n$ such that $A + C \subseteq B + C$, then $A \subseteq B$. In the present case we have

$$B_{\delta/2} + B_{\delta/2} \subseteq B_\delta \subseteq V \subseteq V_k + B_{\delta/2} \text{ for } k \geq k_0,$$

and hence $B_{\delta/2} \subseteq V_k$ for $k \geq k_0$. Since also $V_k \subseteq V + B_{\delta/2}$ for all large k, there exists $R > 0$ such that $V_k \subseteq B_R$ for all k.

The lattice Λ_k has at most $2(2^n - 1)$ facet vectors, by Proposition 15. Hence, by restriction to a subsequence, we may assume that all Λ_k have the same number m of facet vectors. Let $x_{k1},...,x_{km}$ be the facet vectors of Λ_k and choose the notation so that $x_{k1},...,x_{kn}$ are linearly independent. Since they all lie in the ball B_{2R}, by restriction to a further subsequence we may assume that

$$x_{kj} \to x_j \text{ as } k \to \infty \quad (j = 1,...,m).$$

Evidently $\|x_j\| \geq \delta$ $(j = 1,...,m)$ since, for $k \geq k_0$, all nonzero $x \in \Lambda_k$ have $\|x\| \geq \delta$.

The set Λ of all integral linear combinations of $x_1,...,x_m$ is certainly an additive subgroup of $\mathbb{R}^n$. Moreover Λ is discrete. For suppose $y \in \Lambda$ and $\|y\| < \delta$. We have

$$y = \alpha_1 x_1 + ... + \alpha_m x_m,$$

where $\alpha_j \in \mathbb{Z}$ $(j = 1,...,m)$. If

$$y_k = \alpha_1 x_{k1} + ... + \alpha_m x_{km},$$

then $y_k \to y$ as $k \to \infty$ and hence $\|y_k\| < \delta$ for all large k. Since $y_k \in \Lambda_k$, it follows that $y_k = O$ for all large k and hence $y = O$.

Since the lattice Λ_k' with basis $x_{k1},...,x_{kn}$ is a sublattice of Λ_k, we have

$$d(\Lambda_k') \geq d(\Lambda_k) = \lambda(V_k) \geq \lambda(B_{\delta/2}).$$

Since $d(\Lambda_k') = |\det(x_{k1},...,x_{kn})|$, it follows that also

$$|\det(x_1,...,x_n)| \geq \lambda(B_{\delta/2}) > 0.$$

Thus the vectors $x_1,...,x_n$ are linearly independent. Hence Λ is a lattice.

Let $b_1,...,b_n$ be a basis of Λ. Then, by the definition of Λ,

$$b_i = \alpha_{i1} x_1 + ... + \alpha_{im} x_m,$$

where $\alpha_{ij} \in \mathbb{Z}$ $(1 \leq i \leq n,\ 1 \leq j \leq m)$. Put

$$b_{ki} = \alpha_{i1} x_{k1} + ... + \alpha_{im} x_{km}.$$

Then $b_{ki} \in \Lambda_k$ and $b_{ki} \to b_i$ as $k \to \infty$ $(i = 1,...,n)$. Hence, for all large k, the vectors $b_{k1},...,b_{kn}$ are linearly independent. We are going to show that $b_{k1},...,b_{kn}$ is a basis of Λ_k for all large k.

Since $b_1,...,b_n$ is a basis of Λ, we have

$$x_j = \gamma_{j1}b_1 + ... + \gamma_{jn}b_n,$$

where $\gamma_{ji} \in \mathbb{Z}$ $(1 \leq i \leq n,\ 1 \leq j \leq m)$. Hence, if

$$y_{kj} = \gamma_{j1}b_{k1} + ... + \gamma_{jn}b_{kn},$$

then $y_{kj} \in \Lambda_k$ and $y_{kj} \to x_j$ as $k \to \infty$ $(j = 1,...,m)$. Thus, for all large k,

$$\|y_{kj} - x_{kj}\| < \delta \ (j = 1,...,m).$$

Since $y_{kj} - x_{kj} \in \Lambda_k$, this implies that, for all large k, $y_{kj} = x_{kj}$ $(j = 1,...,m)$. Thus every facet vector of Λ_k is an integral linear combination of $b_{k1},...,b_{kn}$ and hence, by Proposition 16, every vector of Λ_k is an integral linear combination of $b_{k1},...,b_{kn}$. Since $b_{k1},...,b_{kn}$ are linearly independent, this shows that they are a basis of Λ_k.

Let W be the Voronoi cell of Λ. We wish to show that $V = W$. If $v \in V$, then there exist $v_k \in V_k$ such that $v_k \to v$. Assume $v \notin W$. Then $\|v\| > \|z - v\|$ for some $z \in \Lambda$, and so

$$\|v\| = \|z - v\| + \rho,$$

where $\rho > 0$. There exist $z_k \in \Lambda_k$ such that $z_k \to z$. Then, for all large k,

$$\|v\| > \|z_k - v\| + \rho/2$$

and hence, for all large k,

$$\|v_k\| > \|z_k - v_k\|.$$

But this contradicts $v_k \in V_k$.

This proves that $V \subseteq W$. On the other hand, V has volume

$$\begin{aligned}\lambda(V) &= \lim_{k\to\infty}\lambda(V_k) = \lim_{k\to\infty} d(\Lambda_k)\\ &= \lim_{k\to\infty} |\det (b_{k1},...,b_{kn})|\\ &= |\det (b_1,...,b_n)| = d(\Lambda) = \lambda(W).\end{aligned}$$

It follows that every interior point of W is in V, and hence $W = V$. Corollary 17 now shows that the same lattice Λ would have been obtained if we had restricted attention to some other subsequence of $\{\Lambda_k\}$.

Let $a_1,...,a_n$ be any basis of Λ. We are going to show that, for the sequence $\{\Lambda_k\}$ originally given, there exist $a_{ki} \in \Lambda_k$ such that

$$a_{ki} \to a_i \text{ as } k \to \infty \ \ (i = 1,...,n).$$

If this is not the case then, for some $i \in \{1,\ldots,n\}$ and some $\varepsilon > 0$, there exist infinitely many k such that

$$\|x - a_i\| > \varepsilon \text{ for all } x \in \Lambda_k.$$

From this subsequence we could as before pick a further subsequence $\Lambda_{k_\nu} \to \Lambda$. Then every $y \in \Lambda$ is the limit of a sequence $y_\nu \in \Lambda_{k_\nu}$. Taking $y = a_i$, we obtain a contradiction.

It only remains to show that $a_{k1},\ldots,a_{kn}$ is a basis of Λ_k for all large k. Since

$$\begin{aligned}\lim_{k\to\infty} |\det (a_{k1},\ldots,a_{kn})| &= |\det (a_1,\ldots,a_n)| \\ &= \mathrm{d}(\Lambda) = \lambda(V) = \lim_{k\to\infty} \lambda(V_k),\end{aligned}$$

for all large k we must have

$$0 < |\det (a_{k1},\ldots,a_{kn})| < 2\lambda(V_k).$$

But if $a_{k1},\ldots,a_{kn}$ were not a basis of Λ_k for all large k, then for infinitely many k we would have

$$|\det (a_{k1},\ldots,a_{kn})| \geq 2\mathrm{d}(\Lambda_k) = 2\lambda(V_k). \quad \square$$

Proposition 20 has the following counterpart:

PROPOSITION 21 *Let $\{\Lambda_k\}$ be a sequence of lattices in $\mathbb{R}^n$ and let V_k be the Voronoi cell of Λ_k. If there exists a lattice Λ such that $\Lambda_k \to \Lambda$ as $k \to \infty$, and if V is the Voronoi cell of Λ, then $V_k \to V$ in the Hausdorff metric as $k \to \infty$.*

Proof By hypothesis, there exists a basis $b_1,\ldots,b_n$ of Λ and a basis $b_{k1},\ldots,b_{kn}$ of each Λ_k such that $b_{kj} \to b_j$ as $k \to \infty$ $(j = 1,\ldots,n)$. Choose $R > 0$ so that the fundamental parallelotope of Λ is contained in the ball $B_R = \{x \in \mathbb{R}^n \colon \|x\| \leq R\}$. Then, for all $k \geq k_0$, the fundamental parallelotope of Λ_k is contained in the ball B_{2R}. It follows that, for all $k \geq k_0$, every point of $\mathbb{R}^n$ is distant at most $2R$ from some point of Λ_k and hence $V_k \subseteq B_{2R}$.

Consequently, by Blaschke's selection principle, the sequence $\{V_k\}$ has a subsequence $\{V_{k_\nu}\}$ which converges in the Hausdorff metric to a compact convex set W. Moreover,

$$\lambda(W) = \lim_{\nu\to\infty} \lambda(V_{k_\nu}) = \lim_{\nu\to\infty} \mathrm{d}(\Lambda_{k_\nu}) = \mathrm{d}(\Lambda) > 0.$$

Consequently, since W is convex, it has nonempty interior. It now follows from Proposition 20 that $W = V$.

Thus any convergent subsequence of $\{V_k\}$ has the same limit V. If the whole sequence $\{V_k\}$ did not converge to V, there would exist $\rho > 0$ and a subsequence $\{V_{k_\nu}\}$ such that

$$\mathrm{h}(V_{k_\nu}, V) \geq \rho \text{ for all } \nu.$$

By the Blaschke selection principle again, this subsequence would itself have a convergent subsequence. Since its limit must be V, this yields a contradiction. □

Suppose $\Lambda_k \in \mathcal{L}_n$ and $\Lambda_k \to \Lambda$ as $k \to \infty$. We will show that not only $d(\Lambda_k) \to d(\Lambda)$, but also $m(\Lambda_k) \to m(\Lambda)$ as $k \to \infty$. Since every $x \in \Lambda$ is the limit of a sequence $x_k \in \Lambda_k$, we must have $\overline{\lim}_{k\to\infty} m(\Lambda_k) \leq m(\Lambda)$. On the other hand, by Proposition 19, if $x_k \in \Lambda_k$ and $x_k \to x$, then $x \in \Lambda$. It follows that $\underline{\lim}_{k\to\infty} m(\Lambda_k) \geq m(\Lambda)$.

Suppose now that a subset $\mathcal{F}$ of $\mathcal{L}_n$ has the property that any infinite sequence Λ_k of lattices in $\mathcal{F}$ has a convergent subsequence. Then there exist positive constants ρ,σ such that

$$m(\Lambda) \geq \rho^2, \quad d(\Lambda) \leq \sigma \quad \text{for all } \Lambda \in \mathcal{F}.$$

For otherwise there would exist a sequence Λ_k of lattices in $\mathcal{F}$ such that either $m(\Lambda_k) \to 0$ or $d(\Lambda_k) \to \infty$, and this sequence could have no convergent subsequence.

We now prove the fundamental *compactness theorem* of Mahler (1946), which says that this necessary condition on $\mathcal{F}$ is also sufficient.

PROPOSITION 22 *If $\{\Lambda_k\}$ is a sequence of lattices in $\mathbb{R}^n$ such that*

$$m(\Lambda_k) \geq \rho^2, \quad d(\Lambda_k) \leq \sigma \quad \textit{for all } k,$$

where ρ,σ are positive constants, then the sequence $\{\Lambda_k\}$ has a convergent subsequence.

Proof Let V_k denote the Voronoi cell of Λ_k. We show first that the ball $B_{\rho/2} = \{x \in \mathbb{R}^n: \|x\| \leq \rho/2\}$ is contained in every Voronoi cell V_k. In fact if $\|x\| \leq \rho/2$ then, for every nonzero $y \in \Lambda_k$,

$$\|x - y\| \geq \|y\| - \|x\| \geq \rho - \rho/2 = \rho/2 \geq \|x\|,$$

and hence $x \in V_k$.

Let v_k be a point of V_k which is furthest from the origin. Then V_k contains the convex hull C_k of the set $v_k \cup B_{\rho/2}$. Since the volume of V_k is bounded above by σ, so also is the volume of C_k. But this implies that the sequence v_k is bounded. Thus there exists $R > 0$ such that the ball B_R contains every Voronoi cell V_k.

By Blaschke's selection principle, the sequence $\{V_k\}$ has a subsequence $\{V_{k_v}\}$ which converges in the Hausdorff metric to a compact convex set V. Since $B_{\rho/2} \subseteq V$, it follows from Proposition 20 that $\Lambda_{k_v} \to \Lambda$, where Λ is a lattice with Voronoi cell V. □

To illustrate the utility of Mahler's compactness theorem, we now show that, as stated in §3, any compact symmetric convex set K with nonempty interior has a critical lattice.

By the definition of the critical determinant $\Delta(K)$, there exists a sequence Λ_k of lattices with no nonzero points in the interior of K such that $d(\Lambda_k) \to \Delta(K)$ as $k \to \infty$. Since K contains a ball B_ρ with radius $\rho > 0$, we have $m(\Lambda_k) \geq \rho^2$ for all k. Hence, by Proposition 22, there is a subsequence Λ_{k_ν} which converges to a lattice Λ as $\nu \to \infty$. Since every point of Λ is a limit of points of Λ_{k_ν}, no nonzero point of Λ lies in the interior of K. Furthermore,

$$d(\Lambda) = \lim_{\nu\to\infty} d(\Lambda_{k_\nu}) = \Delta(K),$$

and hence Λ is a critical lattice for K.

7 Further remarks

The geometry of numbers is treated more extensively in Cassels [11], Erdös *et al.* [22] and Gruber and Lekkerkerker [27]. Minkowski's own account is available in [42]. Numerous references to the earlier literature are given in Keller [34]. Lagarias [36] gives an overview of lattice theory. For a simple proof that the indicator function of a convex set is Riemann integrable, see Szabo [57].

Diophantine approximation is studied in Cassels [12], Koksma [35] and Schmidt [50]. Minkowski's result that the discriminant of an algebraic number field other than $\mathbb{Q}$ has absolute value greater than 1 is proved in Narkiewicz [44], for example.

Minkowski's theorem on successive minima is proved in Bambah *et al.* [3]. For the results of Banaszczyk mentioned in §3, see [4] and [5]. Sharp forms of Siegel's lemma are proved not only in Bombieri and Vaaler [7], but also in Matveev [40]. The result of Gillet and Soulé appeared in [25]. Some interesting results and conjectures concerning the product $\lambda(K)\lambda(K^*)$ are described on pp. 425-427 of Schneider [51].

An algorithm of Lovász, which first appeared in Lenstra, Lenstra and Lovász [38], produces in finitely many steps a basis for a lattice Λ in $\mathbb{R}^n$ which is 'reduced'. Although the first vector of a reduced basis is in general not a minimal vector, it has square-norm at most $2^{n-1}\, m(\Lambda)$. This suffices for many applications and the algorithm has been used to solve a number of apparently unrelated computational problems, such as factoring polynomials in $\mathbb{Q}[t]$, integer linear programming and simultaneous Diophantine approximation. There is an account of the basis reduction algorithm in Schrijver [52]. The algorithmic geometry of numbers is surveyed in Kannan [33].

Mahler [39] has established an analogue of the geometry of numbers for formal Laurent series with coefficients from an arbitrary field F, the roles of $\mathbb{Z}$,$\mathbb{Q}$ and $\mathbb{R}$ being taken by $F[t]$,

$F(t)$ and $F((t))$. In particular, Eichler [19] has shown that the Riemann–Roch theorem for algebraic functions may be derived by geometry of numbers arguments.

There is also a generalization of Minkowski's lattice point theorem to locally compact groups, with Haar measure taking the place of volume; see Chapter 2 (Lemma 1) of Weil [60].

Voronoi *diagrams* and their uses are surveyed in Aurenhammer [1]. Proofs of the basic properties of polytopes referred to in §4 may be found in Brøndsted [9] and Coppel [15]. Planar tilings are studied in detail in Grünbaum and Shephard [28].

Mathematical crystallography is treated in Schwarzenberger [53] and Engel [21]. For the physicist's point of view, see Burckhardt [10], Janssen [32] and Birman [6]. There is much theoretical information, in addition to tables, in [31].

For Bieberbach's theorems, see Vince [59], Charlap [13] and Milnor [41]. Various equivalent forms for the definitions of crystal and crystallographic group are given in Dolbilin *et al.* [17]. It is shown in Charlap [13] that crystallographic groups may be abstractly characterized as groups containing a finitely generated maximal abelian torsion-free subgroup of finite index. (An abelian group is *torsion-free* if only the identity element has finite order.) The fundamental group of a compact flat Riemannian manifold is a torsion-free crystallographic group and all torsion-free crystallographic groups may be obtained in this way. For these connections with differential geometry, see Wolf [61] and Charlap [13].

In more than 4 dimensions the complete enumeration of all crystallographic groups is no longer practicable. However, algorithms for deciding if two crystallographic groups are equivalent in some sense have been developed by Opgenorth *et al.* [45]. An interesting subset of all crystallographic groups consists of those generated by reflections in hyperplanes, since Stiefel (1941/2) showed that they are in 1-1 correspondence with the compact simply-connected semi-simple Lie groups. See the 'Note historique' in Bourbaki [8].

There has recently been considerable interest in tilings of $\mathbb{R}^n$ which, although not lattice tilings, consist of translates of finitely many n-dimensional polytopes. The first example, in $\mathbb{R}^2$, due to Penrose (1974), was explained more algebraically by de Bruijn (1981). A substantial generalization of de Bruijn's construction was given by Katz and Duneau (1986), who showed that many such 'quasiperiodic' tilings may be obtained by a method of cut and projection from ordinary lattices in a higher-dimensional space. The subject gained practical significance with the discovery by Shechtman *et al.* (1984) that the diffraction pattern of an alloy of aluminium and magnesium has icosahedral symmetry, which is impossible for a crystal. Many other 'quasicrystals' have since been found. The papers referred to are reproduced, with others, in Steinhardt and Ostlund [56]. The mathematical theory of quasicrystals is surveyed in Le *et al.* [37].

Skubenko [54] has given an upper bound for Hermite's constant γ_n. Somewhat sharper bounds are known, but they have the same asymptotic behaviour and the proofs are much more complicated. A lower bound for γ_n was obtained with a new method by Ball [2].

For the densest lattices in $\mathbb{R}^n$ ($n \leq 8$), see Ryshkov and Baranovskii [49]. The enumeration of all root lattices is carried out in Ebeling [18]. (A more general problem is treated in Chap. 3 of Humphreys [30] and in Chap. 6 of Bourbaki [8].) For the Voronoi cells of root lattices, see Chap. 21 of Conway and Sloane [14] and Moody and Patera [43]. For the *Dynkin diagrams* associated with root lattices, see also Reiten [47].

Rajan and Shende [46] characterize root lattices as those lattices for which every facet vector is a minimal vector, but their definition of root lattice is not that adopted here. Their argument shows that if every facet vector of a lattice is a minimal vector then, after scaling to make the minimal vectors have square-norm 2, it is a root lattice in our sense.

There is a fund of information about lattice packings of balls in Conway and Sloane [14]. See also Thompson [58] for the Leech lattice and Coxeter [16] for the kissing number problem.

We have restricted attention to lattice packings and, in particular, to lattice packings of balls. Lattice packings of other convex bodies are discussed in the books on geometry of numbers cited above. Non-lattice packings have also received much attention. The notion of density is not so intuitive in this case and it should be realized that the density is unaltered if finitely many sets are removed from the packing.

Packings and coverings are discussed in the texts of Rogers [48] and Fejes Tóth [23],[24]. For packings of balls, see also Zong [62]. Sloane [55] and Elkies [20] provide introductions to the connections between lattice packings of balls and coding theory.

The third part of Hilbert's 18th problem, which is surveyed in Milnor [41], deals with the densest lattice or non-lattice packing of balls in $\mathbb{R}^n$. It is known that, for $n = 2$, the densest lattice packing is also a densest packing. The original proof by Thue (1882/1910) was incomplete, but a complete proof was given by L. Fejes Tóth (1940). The famous *Kepler conjecture* asserts that, also for $n = 3$, the densest lattice packing is a densest packing. A computer-aided proof has recently been announced by Hales [29]. It is unknown if the same holds for any $n > 3$.

Propositions 20 and 21 are due to Groemer [26], and are of interest quite apart from the application to Mahler's compactness theorem. Other proofs of the latter are given in Cassels [11] and Gruber and Lekkerkerker [27]. Blaschke's selection principle and Rådström's cancellation law are proved in [15] and [51], for example.

8 Selected references

[1] F. Aurenhammer, Voronoi diagrams – a survey of a fundamental geometric data structure, *ACM Computing Surveys* **23** (1991), 345-405.

[2] K. Ball, A lower bound for the optimal density of lattice packings, *Internat. Math. Res. Notices* 1992, no. 10, 217-221.

[3] R.P. Bambah, A.C. Woods and H. Zassenhaus, Three proofs of Minkowski's second inequality in the geometry of numbers, *J. Austral. Math. Soc.* **5** (1965), 453-462.

[4] W. Banaszczyk, New bounds in some transference theorems in the geometry of numbers, *Math. Ann.* **296** (1993), 625-635.

[5] W. Banaszczyk, Inequalities for convex bodies and polar reciprocal lattices in $\mathbb{R}^n$. II. Application of K-convexity, *Discrete Comput. Geom.* **16** (1996), 305-311.

[6] J.L. Birman, *Theory of crystal space groups and lattice dynamics,* Springer-Verlag, Berlin, 1984. [Original edition in *Handbuch der Physik*, 1974]

[7] E. Bombieri and J. Vaaler, On Siegel's lemma, *Invent. Math.* **73** (1983), 11-32.

[8] N. Bourbaki, *Groupes et algèbres de Lie, Chapitres 4,5 et 6,* Masson, Paris, 1981.

[9] A. Brøndsted, *An introduction to convex polytopes*, Springer-Verlag, New York, 1983.

[10] J.J. Burckhardt, *Die Bewegungsgruppen der Kristallographie*, 2nd ed., Birkhäuser, Basel, 1966.

[11] J.W.S. Cassels, *An introduction to the geometry of numbers*, corrected reprint, Springer-Verlag, Berlin, 1997. [Original edition, 1959]

[12] J.W.S. Cassels, *An introduction to Diophantine approximation*, Cambridge University Press, 1957.

[13] L.S. Charlap, *Bieberbach groups and flat manifolds*, Springer-Verlag, New York, 1986.

[14] J.H. Conway and N.J.A. Sloane, *Sphere packings, lattices and groups*, 3rd ed., Springer-Verlag, New York, 1999.

[15] W.A. Coppel, *Foundations of convex geometry*, Cambridge University Press, 1998.

[16] H.S.M. Coxeter, An upper bound for the number of equal nonoverlapping spheres that can touch another of the same size, *Convexity* (ed. V. Klee), pp. 53-71, Proc. Symp. Pure Math. **7**, Amer. Math. Soc., Providence, Rhode Island, 1963.

[17] N.P. Dolbilin, J.C. Lagarias and M. Senechal, Multiregular point systems, *Discrete Comput. Geom.* **20** (1998), 477-498.

[18] W. Ebeling, *Lattices and codes*, Vieweg, Braunschweig, 1994.

[19] M. Eichler, Ein Satz über Linearformen in Polynombereichen, *Arch. Math.* **10** (1959), 81-84.

[20] N.D. Elkies, Lattices, linear codes, and invariants, *Notices Amer. Math. Soc.* **47** (2000), 1238-1245 and 1382-1391.

[21] P. Engel, Geometric crystallography, *Handbook of convex geometry* (ed. P.M. Gruber and J.M. Wills), Volume B, pp. 989-1041, North-Holland, Amsterdam, 1993. (The same volume contains several other useful survey articles relevant to this chapter.)

[22] P. Erdös, P.M. Gruber and J. Hammer, *Lattice points*, Longman, Harlow, Essex, 1989.

[23] L. Fejes Tóth, *Regular Figures*, Pergamon, Oxford, 1964.

[24] L. Fejes Tóth, *Lagerungen in der Ebene auf der Kugel und im Raum*, 2nd ed., Springer-Verlag, Berlin, 1972.

[25] H. Gillet and C. Soulé, On the number of lattice points in convex symmetric bodies and their duals, *Israel J. Math.* **74** (1991), 347-357.

[26] H. Groemer, Continuity properties of Voronoi domains, *Monatsh. Math.* **75** (1971), 423-431.

[27] P.M. Gruber and C.G. Lekkerkerker, *Geometry of numbers*, 2nd ed., North-Holland, Amsterdam, 1987.

[28] B. Grünbaum and G.C. Shephard, *Tilings and patterns*, Freeman, New York, 1987.

[29] T.C. Hales, Cannonballs and honeycombs, *Notices Amer. Math. Soc.* **47** (2000), 440-449.

[30] J.E. Humphreys, *Introduction to Lie algebras and representation theory*, Springer-Verlag, New York, 1972.

[31] *International tables for crystallography, Vols. A-C*, Kluwer, Dordrecht, 1983-1993.

[32] T. Janssen, *Crystallographic groups*, North-Holland, Amsterdam, 1973.

[33] R. Kannan, Algorithmic geometry of numbers, *Annual review of computer science* **2** (1987), 231-267.

[34] O.-H. Keller, *Geometrie der Zahlen*, Enzyklopädie der mathematischen Wissenschaften I-2, 27, Teubner, Leipzig, 1954.

[35] J.F. Koksma, *Diophantische Approximationen*, Springer-Verlag, Berlin, 1936. [Reprinted Chelsea, New York, 1950]

[36] J.C. Lagarias, Point lattices, *Handbook of Combinatorics* (ed. R. Graham, M. Grötschel and L. Lovász), Vol. I, pp. 919-966, Elsevier, Amsterdam, 1995.

[37] T.Q.T. Le, S.A. Piunikhin and V.A. Sadov, The geometry of quasicrystals, *Russian Math. Surveys* **48** (1993), no. 1, 37-100.

[38] A.K. Lenstra, H.W. Lenstra and L. Lovász, Factoring polynomials with rational coefficients, *Math. Ann.* **261** (1982), 515-534.

[39] K. Mahler, An analogue to Minkowski's geometry of numbers in a field of series, *Ann. of Math.* **42** (1941), 488-522.

[40] E.M. Matveev, On linear and multiplicative relations, *Math. USSR-Sb.* **78** (1994), 411-425.

[41] J. Milnor, Hilbert's Problem 18: On crystallographic groups, fundamental domains, and on sphere packing, *Mathematical developments arising from Hilbert problems* (ed. F.E. Browder), pp. 491-506, Proc. Symp. Pure Math. **28**, Part 2, Amer. Math. Soc., Providence, Rhode Island, 1976.

[42] H. Minkowski, *Geometrie der Zahlen*, Teubner, Leipzig, 1896. [Reprinted Chelsea, New York, 1953]

[43] R.V. Moody and J. Patera, Voronoi and Delaunay cells of root lattices: classification of their faces and facets by Coxeter–Dynkin diagrams, *J. Phys. A* **25** (1992), 5089-5134.

[44] W. Narkiewicz, *Elementary and analytic theory of algebraic numbers*, 2nd ed., Springer-Verlag, Berlin, 1990.

[45] J. Opgenorth, W. Plesken and T. Schulz, Crystallographic algorithms and tables, *Acta Cryst. A* **54** (1998), 517-531.

[46] D.S. Rajan and A.M. Shende, A characterization of root lattices, *Discrete Math.* **161** (1996), 309-314.

[47] I. Reiten, Dynkin diagrams and the representation theory of Lie algebras, *Notices Amer. Math. Soc.* **44** (1997), 546-556.

[48] C.A. Rogers, *Packing and covering*, Cambridge University Press, 1964.

[49] S.S. Ryshkov and E.P. Baranovskii, Classical methods in the theory of lattice packings, *Russian Math. Surveys* **34** (1979), no. 4, 1-68.

[50] W.M. Schmidt, *Diophantine approximation*, Lecture Notes in Mathematics **785**, Springer-Verlag, Berlin, 1980.

[51] R. Schneider, *Convex bodies: the Brunn–Minkowski theory*, Cambridge University Press, 1993.

[52] A. Schrijver, *Theory of linear and integer programming*, corrected reprint, Wiley, Chichester, 1989.

[53] R.L.E. Schwarzenberger, *N-dimensional crystallography*, Pitman, London, 1980.

[54] B.F. Skubenko, A remark on an upper bound on the Hermite constant for the densest lattice packings of spheres, *J. Soviet Math.* **18** (1982), 960-961.

[55] N.J.A. Sloane, The packing of spheres, *Scientific American* **250** (1984), 92-101.

[56] P.J. Steinhardt and S. Ostlund (ed.), *The physics of quasicrystals*, World Scientific, Singapore, 1987.

[57] L. Szabo, A simple proof for the Jordan measurability of convex sets, *Elem. Math.* **52** (1997), 84-86.

[58] T.M. Thompson, *From error-correcting codes through sphere packings to simple groups*, Carus Mathematical Monograph No. 21, Mathematical Association of America, 1983.

[59] A. Vince, Periodicity, quasiperiodicity and Bieberbach's theorem, *Amer. Math. Monthly* **104** (1997), 27-35.

[60] A. Weil, *Basic number theory*, 2nd ed., Springer-Verlag, Berlin, 1973.

[61] J.A. Wolf, *Spaces of constant curvature*, 3rd ed., Publish or Perish, Boston, Mass., 1974.

[62] C. Zong, *Sphere packings*, Springer-Verlag, New York, 1999.