Lecture Notes in Mathematics 1467

Wolfgang M. Schmidt

Diophantine Approximations and Diophantine Equations

Springer-Verlag

Wolfgang M. Schmidt

Diophantine Approximations and Diophantine Equations

Springer-Verlag
Berlin Heidelberg New York
London Paris Tokyo
Hong Kong Barcelona
Budapest

Author

Wolfgang M. Schmidt
Department of Mathematics, University of Colorado
Boulder, Colorado, 80309-0426, USA

Mathematics Subject Classification (1991): 11J68, 11J69, 11O57, 11D61

ISBN 3-540-54058-X Springer-Verlag Berlin Heidelberg New York
ISBN 0-387-54058-X Springer-Verlag New York Berlin Heidelberg

Typesetting: Camera ready by author
Printing and binding: Druckhaus Beltz, Hemsbach/Bergstr.
2146/3140-543210 - Printed on acid-free paper

Preface

The present notes are the outcome of lectures I gave at Columbia University in the fall of 1987, and at the University of Colorado 1988/1989. Although there is necessarily some overlap with my earlier Lecture Notes on Diophantine Approximation (Springer Lecture Notes 785, 1980), this overlap is small. In general, whereas in the earlier Notes I gave a systematic exposition with all the proofs, the present notes present a varirety of topics, and sometimes quote from the literature wihtout giving proofs. Nevertheless, I believe that the pace is again leisurely.

Chapter I contains a fairly thorough discussion of Siegel's Lemma and of heights. Chapter II is devoted to Roth's Theorem. Rather than Roth's Lemma, I use a generalization of Dyson's Lemma as given by Esnault and Viehweg. A proof of this generalized lemma is not given; it is beyond the scope of the present notes. An advantage of the lemma is that it leads to new bounds on the number of exceptional approximations in Roth's Theorem, as given recently by Bombieri and Van der Poorten. These bounds turn out to be best possible in some sense. Chapter III deals with the Thue equation. Among the recent developments are bounds by Bombieri and author on the number of solutions of such equations, and by Mueller and the author on the number of solutions of Thue equations with few nonzero coefficients, say s such coefficients (apart from the constant term). I give a proof of the former, but deal with the latter only up to $s = 3$, i.e., to trinomial Thue equations. Chapter IV is about S-unit equations and hyperelliptic equations. S-unit equations include equations such as $2^x + 3^y = 4^z$. I present Evertse's remarkable bounds for such equations. As for elliptic and hyperelliptic equations, I mention a few basic facts, often without proofs, and proceed to counting the number of solutions as in recent works of Evertse, and of Silverman, where the connection with the Mordell–Weil rank is explored. Chapter V is on certain diophantine equations in more than two variables. A tool here is my Subspace Theorem, of which I quote several versions, but without proofs. I study generalized S-unit equations, such as, e.g. $\pm a_1^{x_1} \pm a_2^{x_2} \pm \cdots \pm a_n^{x_n} = 0$ with given integers $a_i > 1$, as well as norm form equations. Recent advances permit to give explicit estimates on the number of solutions. The notes end with an Epilogue on the abc-conjecture of Oesterlé and Masser.

Hand written notes of my lectures were taken at Columbia University by Mr. Agboola, and at the University of Colorado by Ms. Deanna Caveny. The manuscript was typed by Ms. Andrea Hennessy and Ms. Elizabeth Stimmel. My thanks are due to them.

January 1991 Wolfgang M. Schmidt

Table of Contents

Table of Contents (cont.)

I. Siegel's Lemma and Heights

§1. Siegel's Lemma.

Consider a system of homogeneous linear equations

$$\begin{aligned} a_{11}x_1 + \cdots + a_{1n}x_n &= 0 \\ &\vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n &= 0 \end{aligned} \tag{1.1}$$

If $m < n$ and the coefficients lie in a field, then there is a nontrivial solution with components in the field. If $m < n$ and the coefficients lie in $\mathbb{Z}$ (the integers), then there is a nontrivial solution in integers. (Just take a solution with rational components and multiply by the common denominator.) It is reasonable to believe that if the coefficients are small integers, then there will also be a solution in small integers. This idea was used by A. Thue (1909) and formalized by Siegel in (1929; on p. 213 of his Collected Works).

LEMMA 1. *Suppose that in (1.1) the coefficients a_{ij} lie in $\mathbb{Z}$ and have $|a_{ij}| \leqq A$ $(1 \leqq i, j \leqq n)$ where A is natural. Then there is a nontrivial solution in $\mathbb{Z}$ with*

$$|x_i| < 1 + (nA)^{m/(n-m)} \qquad (i = 1, \dots, n).$$

Proof. We follow Siegel. Let H be an integer parameter to be specified later. Let C be the cube consisting of points

$$\underline{\underline{x}} = (x_1, \dots, x_n)$$

with

$$|x_i| \leqq H \qquad (i = 1, \dots, n).$$

There are $(2H+1)^n$ integer points in this cube, since there are $2H+1$ possibilities for each coordinate. Let T be the linear map $\mathbb{R}^n \to \mathbb{R}^m$ with

$$T\underline{\underline{x}} = (a_{11}x_1 + \cdots + a_{1n}x_n, \dots, a_{m1}x_1 + \cdots + a_{mn}x_n).$$

Writing $T\underline{\underline{x}} = \underline{\underline{y}} = (y_1, \dots, y_m)$, we observe that the image of C lies in the cube

$$C' : \quad |y_j| \leqq nAH \quad (j = 1, \dots, m)$$

of $\mathbb{R}^m$. The number of integer points in C' is $(2nAH+1)^n$. Suppose now that

$$(2nAH+1)^m < (2H+1)^n. \tag{1.2}$$

Then T restricted to the integers in the original cube will not be one-to-one. So there exist $\underline{\underline{x}}', \underline{\underline{x}}''$ in C, with $\underline{\underline{x}}' \neq \underline{\underline{x}}''$, such that $T\underline{\underline{x}}' = T\underline{\underline{x}}''$. Put $\underline{\underline{x}} = \underline{\underline{x}}' - \underline{\underline{x}}''$. Then $T\underline{\underline{x}} = \underline{\underline{0}}$. So $\underline{\underline{x}}$ is a solution to the system with integer coordinates. Note that $|x_i| \leqq 2H$ $(i =$

$1, \ldots, n)$, because $|x_i'|, |x_i''| \leqq H$ $(i = 1, \ldots, n)$, so $|x_i| = |x_i' - x_i''| \leqq |x_i'| + |x_i''| \leqq 2H$. Choose H to be the natural number satisfying

$$(nA)^{m/(n-m)} \leqq 2H \leqq (nA)^{m/(n-m)} + 1.$$

Then

$$\begin{aligned}(2H+1)^n &= (2H+1)^m (2H+1)^{n-m} \\ &\leqq (2H+1)^m (nA)^m \\ &> (2nAH+1)^m.\end{aligned}$$

So there exists an $\underline{\underline{x}}$ satisfying $|x_i| \leqq 2H < 1 + (nA)^{m/(n-m)}$.

So the proof of Siegel's Lemma uses the box principle.

Can the exponent be improved? The answer is "no".

Siegel's Lemma is almost best possible. Put $k = n - m$, and for large P pick distinct primes p_{ij} $(i \leqq i \leqq k, 1 \leqq j \leqq m)$ with $P\eta < p_{ij} < P$, where $0 < \eta < 1$ is given. Put

$$P_j = p_{1j} p_{2j} \cdots p_{kj} \quad (1 \leqq j \leqq m), \qquad P_{ij} = P_j / p_{ij} \quad (1 \leqq i \leqq k,\ 1 \leqq j \leqq m),$$

$$Q_i = p_{i1} p_{i2} \cdots p_{im} \quad (1 \leqq i \leqq k), \qquad Q_{ij} = Q_i / p_{ij} \quad (1 \leqq i \leqq k,\ 1 \leqq j \leqq m).$$

Consider the system of m equations in n variables:

$$\begin{array}{llllll} P_{11}x_1 + \cdots + P_{k1}x_k & - P_1 y_1 & & & & = 0 \\ P_{12}x_1 + \cdots + P_{k2}x_k & & -P_2 y_2 & & & = 0 \\ \vdots & & & & & \\ P_{1m}x_1 + \cdots + P_{km}x_k & & & & -P_m y_m & = 0. \end{array}$$

The maximum modulus A of its coefficients has $A \leqq P^k$. The following are solution vectors:

$$\begin{aligned} \underline{\underline{z}}_1 &= (Q_1, 0, \ldots, 0, Q_{11}, Q_{12}, \ldots, Q_{1m}) \\ &\ \ \vdots \\ \underline{\underline{z}}_k &= (0, 0, \ldots, Q_k, Q_{k1}, Q_{k2}, \ldots, Q_{km}). \end{aligned}$$

It is clear that every solution of our system of equations is a linear combination $c_1 \underline{\underline{z}}_1 + \cdots + c_k \underline{\underline{z}}_k$. For integer solutions, c_i is necessarily a rational number whose denominator is Q_i, and then every component of $c_i \underline{\underline{z}}_i$ has denominator Q_i. Moreover, if, say, c_1 is not integral, then (since $Q_{11}, Q_{12}, \ldots, Q_{1m}$ are coprime), $c_1 \underline{\underline{z}}_1$ is not integral and has some component whose denominator is a prime p_{1j}. But since $c_2 \underline{\underline{z}}_2, \ldots, c_k \underline{\underline{z}}_k$ don't have p_{1j} in the denominator, $c_1 \underline{\underline{z}}_1 + c_2 \underline{\underline{z}}_2 + \cdots + c_k \underline{\underline{z}}_k$ cannot be integral—a contradiction. Therefore the integer solutions are $\underline{\underline{z}} = c_1 \underline{\underline{z}}_1 + \cdots + c_k \underline{\underline{z}}_k$ with $c_1, \ldots, c_k$ in $\mathbb{Z}$. When $\underline{\underline{z}} \neq \underline{\underline{0}}$, say $c_1 \neq 0$, the first component x_1 of $\underline{\underline{z}}$ has

$$|x_1| \geqq Q_1 > (\eta P)^m = \eta^m P^m > \eta^m A^{m/k} = \eta^m A^{m/(n-m)}.$$

Therefore every integer solution $(x_1, \ldots, x_k,\ y_1, \ldots, y_m) \neq (0, \ldots, 0)$ has

$$\max(|x_1|, \ldots, |x_k|) > \eta^m A^{m/(n-m)}.$$

Here η may be taken arbitrarily close to 1.

Another approach is as follows. When $m = n - 1$, consider the system of equations

$$Ax_i - x_{i+1} = 0 \qquad (i = 1, \dots, n-1).$$

Every nontrivial solution, in fact every nontrivial complex solution, has $x_n/x_1 = A^{n-1}$. Thus if we set

$$q(\underline{\underline{x}}) = \max |x_i/x_j|,$$

with the maximum over i, j in $1 \leqq i,\ j \leqq n$ with $x_j \neq 0$, then $q(\underline{\underline{x}}) = A^{n-1} = A^{m/(n-m)}$. But then for integer solutions, $\max(|x_1|, \dots, |x_n|) \geqq A^{m/(n-m)}$.

Exercise 1a. Suppose now that $m = 1$. For large A, construct an equation

$$a_1x_1 + \cdots + a_nx_n = 0$$

with integral coefficients and $|a_i| \leqq A$ $(i = 1, \dots, n)$, such that every nontrivial solution $\underline{\underline{x}}$ with complex components has

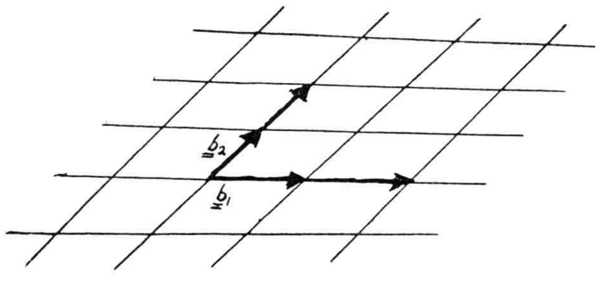
$$q(\underline{\underline{x}}) \geqq c(n)A^{1/(n-1)} = c_1(n,m)A^{m/(n-m)} > 0.$$

This approach can be carried out for general n, m. See Schmidt (1985).

§2. Geometry of Numbers.

The subject was founded by Minkowski (1896 & 1910). Other references are Cassels (1959), Gruber and Lekkerkerker (1987), and Schmidt (1980, Chapter IV).

A *lattice* Λ is a subgroup of $\mathbb{R}^n$ which is generated by n linearly independent vectors $\underline{\underline{b}}_1, \dots, \underline{\underline{b}}_n$ (linearly independent over $\mathbb{R}^n$). The elements of this lattice are $c_1\underline{\underline{b}}_1 + \cdots + c_n\underline{\underline{b}}_n$ with $c_i \in \mathbb{Z}$.



The set $\underline{\underline{b}}_1, \dots, \underline{\underline{b}}_n$ is called a *basis*. A basis is not uniquely determined. For example, $\underline{\underline{b}}_1, \underline{\underline{b}}_1 + \underline{\underline{b}}_2, \underline{\underline{b}}_3, \dots, \underline{\underline{b}}_n$ is another basis.

How unique is a basis? Suppose $\underline{\underline{b}}'_1, \dots, \underline{\underline{b}}'_n$ is another basis. Then

$$\underline{\underline{b}}'_i = \sum_{j=1}^{n} c_{ij}\underline{\underline{b}}_j \quad \text{and} \quad c_{ij} \in \mathbb{Z}$$

and

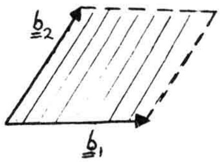
$$\underline{\underline{b}}_j = \sum_{i=1}^{n} c'_{ij}\underline{\underline{b}}_i \quad \text{and} \quad c'_{ji} \in \mathbb{Z}.$$

So the matrices (c_{ij}) and (c'_{ji}) are inverse to each other and $c_{ij}, c'_{ji} \in \mathbb{Z}$, so $\det(c_{ij}) = \det(c'_{ji}) = \pm 1$. Thus the matrix (c_{ij}) is *unimodular*, where by definition a unimodular matrix is a square matrix with integer entries and determinant 1 or -1.

LEMMA 2A. *A necessary and sufficient condition for a subset Λ of $\mathbb{R}^n$ to be a lattice is the following:*
(i) *Λ is a group under addition.*
(ii) *Λ contains n linearly independent vectors.*
(iii) *Λ is discrete.*

For a proof, see e.g., Schmidt (1980, Ch. IV, Theorem 8A).

Consider $\mathbb{R}^n$ with the Euclidean metric and Λ a lattice with $\underline{\underline{b}}_1, \dots, \underline{\underline{b}}_n$ as basis. Let Π be the set of linear combinations $\lambda_1\underline{\underline{b}}_1 + \cdots + \lambda_n\underline{\underline{b}}_n$ with $0 \leqq \lambda_i < 1$ $(i = 1, \dots, n)$. Then Π is called a *fundamental parallelepiped* of Λ.



The fundamental parallelepiped does depend on which basis is chosen. The volume of Π is given by $V(\Pi) = |\det(\underline{\underline{b}}_1, \dots, \underline{\underline{b}}_n)|$ where the right-hand side involves the matrix whose rows are respectively made up of the coordinates of $\underline{\underline{b}}_1, \dots, \underline{\underline{b}}_n$ with respect to an orthonormal bases of $\mathbb{R}^n$. This volume is independent of the chosen basis of the lattice, since different bases are connected by unimodular transformations. It is an invariant of the lattice.

We define

$$\det \Lambda = V(\Pi).$$

Notice that when $\underline{\underline{b}}_i = (b_{i1}, \dots, b_{in})$, then

$$V^2 = \det \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & & & \\ b_{n1} & b_{n2} & \cdots & b_{nn} \end{pmatrix} \cdot \det \begin{pmatrix} b_{11} & b_{21} & \cdots & b_{n1} \\ b_{12} & b_{22} & \cdots & b_{n2} \\ \vdots & & & \\ b_{1n} & b_{2n} & \cdots & b_{nn} \end{pmatrix}$$

$$= \det \begin{pmatrix} \underline{\underline{b}}_1\underline{\underline{b}}_1 & \underline{\underline{b}}_1\underline{\underline{b}}_2 & \cdots & \underline{\underline{b}}_1\underline{\underline{b}}_n \\ \underline{\underline{b}}_2\underline{\underline{b}}_1 & \underline{\underline{b}}_2\underline{\underline{b}}_2 & \cdots & \underline{\underline{b}}_2\underline{\underline{b}}_n \\ \vdots & & & \\ \underline{\underline{b}}_n\underline{\underline{b}}_1 & \underline{\underline{b}}_n\underline{\underline{b}}_2 & \cdots & \underline{\underline{b}}_n\underline{\underline{b}}_n \end{pmatrix}, \tag{2.1}$$

where the inner product of vectors $\underline{\underline{x}}$, $\underline{\underline{y}}$ is denoted by $\underline{\underline{x}}\,\underline{\underline{y}}$.

Every $\underline{\underline{x}}$ in $\mathbb{R}^n$ may uniquely be written as $\underline{\underline{x}} = \underline{\underline{x}}' + \underline{\underline{x}}''$ where $\underline{\underline{x}}' \in \Pi$ and $\underline{\underline{x}}'' \in \Lambda$.

$$\underline{\underline{x}} = \sum_{i=1}^{n} \xi_i \underline{\underline{b}}_i = \underbrace{\sum_{i=1}^{n} \{\xi_i\} \underline{\underline{b}}_i}_{\in \Pi} + \underbrace{\sum_{i=1}^{n} [\xi_i] \underline{\underline{b}}_i}_{\in \Lambda} .$$

Here we used the notation that uniquely

$$\xi = [\xi] + \{\xi\}$$

where $[\xi]$ is an integer, called the *integer part* of ξ, and $\{\xi\}$ satisfies $0 \le \{\xi\} < 1$ and is called the *fractional part* of ξ.

$\mathbb{Z}^n$ is a lattice with basis $\underline{\underline{e}}_1, \dots, \underline{\underline{e}}_n$ where $\underline{\underline{e}}_i = (\overbrace{0, \dots, 0, 1}^{i}, 0, \dots, 0)$, $(i = 1, \dots, n)$, and with $\det \mathbb{Z}^n = 1$. If Λ is an arbitrary lattice with basis $\underline{\underline{b}}_1, \dots, \underline{\underline{b}}_n$, then there exists a linear transformation T such that $T\underline{\underline{e}}_i = \underline{\underline{b}}_i$ $(i = 1, \dots, n)$. So $T\mathbb{Z}^n = \Lambda$.

Is T unique? Suppose $T\mathbb{Z}^n = T'\mathbb{Z}^n$. Then $(T')^{-1}T\mathbb{Z}^n = \mathbb{Z}^n$, so $\det((T')^{-1}T) = \pm 1$ and $(T')^{-1}T$ is unimodular. Call it U. Then $T = T'U$. Observe that

$$\det \Lambda = |\det T|.$$

THEOREM 2B. (Minkowski's First Theorem on Convex Sets.) *Let $B \subseteq \mathbb{R}^n$ be a convex set which is symmetric about the origin (i.e., $\underline{\underline{x}} \in B$ if and only if $-\underline{\underline{x}} \in B$) of volume*

$$V(B) > 2^n \det \Lambda \tag{2.2}$$

where Λ is a lattice. Then B contains a non-zero lattice point.

Comments. The volume here is the Jordan volume, i.e., the Riemann integral over the characteristic function of the set. Every bounded convex set has such a volume. Let $\underline{\underline{g}} \in B$ be a non-zero lattice point in B. Then $-\underline{\underline{g}} \ne \underline{\underline{0}}$ and $-\underline{\underline{g}} \in B$ by symmetry, so B contains actually at least two non-zero lattice points.

Proof. (Mordell, 1934). $V(B)/\det \Lambda$ is invariant under non-singular linear maps. Therefore, after applying a linear transformation, we may assume that $\Lambda = \mathbb{Z}^n$. Then the theorem reduces to: *If $V(B) > 2^n$, then B contains a non-zero integer point.* Let B_m be the set of rational points in B with common denominator m. Then

$$\frac{|B_m|}{m^n} \longrightarrow V(B) \quad \text{as} \quad m \to \infty$$

where $|\ |$ denotes the cardinality. For m sufficiently large, $|B_m|/m^n > 2^n$ and thus $|B_m| > (2m)^n$. So there are two points $\underline{\underline{a}} = (a_1/m, \dots, a_n/m)$, $\underline{\underline{b}} = (b_1/m, \dots, b_n/m)$ in B_m with

$$a_i \equiv b_i \pmod{2m} \qquad (i = 1, \dots, n).$$

Then

$$\frac{1}{2}(\underline{\underline{a}} - \underline{\underline{b}}) \in B$$

since the midpoint of $\underline{\underline{a}}$ and $-\underline{\underline{b}}$ is $\frac{1}{2}(\underline{\underline{a}}-\underline{\underline{b}}) \in B$. Let $\underline{\underline{g}} = \frac{1}{2}(\underline{\underline{a}}-\underline{\underline{b}})$. Clearly $\underline{\underline{g}}$ is a non-zero integer point.

Exercise 2a. If B is symmetric, convex, and $V(B) > 2^n k \det \Lambda$ where $k \in \mathbb{N}$, then B contains at least k pairs of non-zero lattice points.

A *convex body* is a compact, convex set containing $\underline{\underline{0}}$ as an interior point. Such a body clearly has $0 < V(B) < \infty$.

Remark 2C. If B is a symmetric, convex body and $V(B) \geqq 2^n \det \Lambda$, then B contains a non-zero lattice point. It is easy to show that 2C follows from 2B, and vice versa.

Remark 2D. Theorem 2B is best possible. Take $\Lambda = \mathbb{Z}^n$, B the cube with $|x_i| < 1$, $(i = 1, \dots, n)$. Then $V(B) = 2^n = 2^n \det \Lambda$ and there are no non-zero integer points in B.

Minkowski defines *successive minima* as follows: Given B, Λ where B is symmetric, convex, bounded, and with $\underline{\underline{0}}$ in its interior, let $\lambda_1 = \inf\{\lambda : \lambda B$ contains a non-zero lattice point$\}$. * More generally, for $1 \leqq j \leqq n$, $\lambda_j = \inf\{\lambda : \lambda B$ contains j linearly independent lattice points$\}$. Then

$$0 < \lambda_1 \leqq \lambda_2 \leqq \lambda_3 \leqq \cdots \leqq \lambda_n < \infty.$$

Here $\lambda_1 > 0$ since B is bounded and $\lambda_n < \infty$ since $\underline{\underline{0}}$ is an interior point.

THEOREM 2E. (Minkowski's Second Theorem on Convex Bodies.)

$$\frac{2^n}{n!} \det \Lambda \leqq \lambda_1 \cdots \lambda_n V(B) \leqq 2^n \det \Lambda. \tag{2.3}$$

Example. Let $n = 2$, $\Lambda = \mathbb{Z}^2$ and B the rectangle $|x_1| \leqq k$, $|x_2| \leqq 1/k$ where $k \geqq 1$. Then $\lambda_1 = 1/k$, $\lambda_2 = k$ and $\lambda_1\lambda_2 V(B) = V(B) = 4 = 2^2 \det \Lambda$.

A proof will not be given here. There is no really simple proof of the upper bound in (2.3). A weaker bound is proved in Schmidt (1980, p. 88).

Remark 2F. The constants $2^n/n!$ and 2^n are best possible. Let $\Lambda = \mathbb{Z}^n$ and B the cube $|x_i| \leqq 1$. Then $V(B) = 2^n$. Now $\underline{\underline{e}}_1, \dots, \underline{\underline{e}}_n$ are on the boundary, so $\lambda_1 = \cdots = \lambda_n = 1$. We have $\lambda_1 \cdots \lambda_n V(B) = 2^n$ and $2^n \det \Lambda = 2^n$. So we get equality on the right-hand side of (2.3). Next, let $\Lambda = \mathbb{Z}^n$ and B the "octahedron" $|x_1| + \cdots + |x_n| \leqq 1$. It may be seen that $V(B) = 2^n/n!$. We have again $\lambda_1 = \lambda_2 = \cdots = \lambda_n = 1$. Thus $(2^n/n!) \det \Lambda = 2^n/n!$ and $\lambda_1 \cdots \lambda_n V(B) = 2^n/n!$. We have equality on the left-hand side of (2.3).

Note that

$$\lambda_1^n V(B) \leqq 2^n \det \Lambda \tag{2.4}$$

---

* λB is the set of points $\lambda \underline{\underline{x}}$ with $\underline{\underline{x}} \in B$.

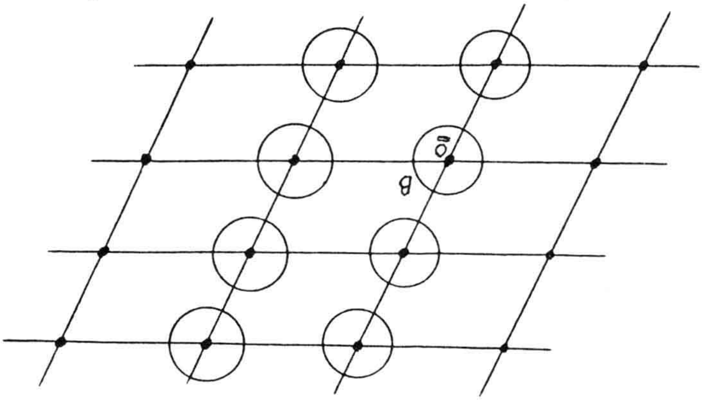
since $\lambda_1 \leqq \lambda_2 \leqq \cdots \leqq \lambda_n$. Suppose now that $V(B) > 2^n \det \Lambda$. Then $\lambda_1 < 1$ so that $B = 1B \supset \lambda_1 B$ contains a non-zero lattice point. Therefore (2.4) implies $2B$. It is easily seen that (2.4) and $2B$ are equivalent.

Exercise 2b. Suppose B is a symmetric, convex set, Λ a lattice, λ_1 the first minimum. Given $\mu > 0$, the number of lattice points in μB is $\leqq ((2\mu/\lambda_1)+1)^n$.

§3. Lattice Packings.

A good reference for general packing and covering problems is C. A. Rogers (1964).

Let $B \subset \mathbb{R}^n$ be compact and measurable. Given a lattice Λ, write $B + \Lambda = \{\underline{\underline{b}} + \underline{\underline{\ell}} : \underline{\underline{b}} \in B,\ \underline{\underline{\ell}} \in \Lambda\}$.



If the translates of B by vectors of Λ are disjoint (as in the picture), then we call $B + \Lambda$ a *lattice packing* of B. Having disjoint translates is equivalent to having unique representations of the vectors of $B + \Lambda$ as $\underline{\underline{b}} + \underline{\underline{\ell}}$.

The *density* of such a lattice packing is

$$\delta(B, \Lambda) = \lim_{r \to 0} \frac{V(\Lambda + B, r)}{V(r)} \tag{3.1}$$

where $V(r)$ is the volume of a ball of radius r, and $V(\Lambda + B, r)$ is the volume of the intersection of $\Lambda + B$ with the ball of radius r and center $\underline{\underline{0}}$.

Exercise 3a. Show that the limit (3.1) always exists under our hypotheses, and that (with Π a fundamental parallelepiped)

$$\delta(B, \Lambda) = V(\Pi \cap (\Lambda + B))/V(\Pi) = V(B)/\det \Lambda.$$

We define

$$\delta(B) = \sup_{\substack{\Lambda \\ \Lambda + B \text{ a lattice} \\ \text{packing}}} \delta(B, \Lambda).$$

This is invariant under nonsingular linear transformations T, since $\delta(TB, T\Lambda) = \delta(B, \Lambda)$.

Suppose B is convex and symmetric about $\underline{\underline{0}}$. Suppose B contains no non-zero lattice point.

Claim: Then $\Lambda + \frac{1}{2}B$ is a lattice packing.

Justification: Suppose $\underline{\ell}_1 + \frac{1}{2}\underline{b}_1 = \underline{\ell}_2 + \frac{1}{2}\underline{b}_2$ where $\underline{\ell}_1, \underline{\ell}_2 \in \Lambda$, $\underline{b}_1, \underline{b}_2 \in B$. Then by an argument used above, $\frac{1}{2}\underline{b}_1 - \frac{1}{2}\underline{b}_2$ is in $B \cap \Lambda$. But $\frac{1}{2}\underline{b}_1 - \frac{1}{2}\underline{b}_2 = \underline{\ell}_2 - \underline{\ell}_1$ and B contains no non-zero lattice points, so that $\underline{\ell}_2 - \underline{\ell}_1 = \underline{0}$, $\underline{\ell}_2 = \underline{\ell}_1$, hence $\underline{b}_1 = \underline{b}_2$. Therefore $\Lambda + \frac{1}{2}B$ is indeed a lattice packing.

$$\delta(\frac{1}{2}B, \Lambda) = \frac{V(\frac{1}{2}B)}{\det \Lambda} = \frac{V(B)}{2^n \det \Lambda} \leqq \delta\left(\frac{1}{2}B\right) = \delta(B) \leqq 1$$

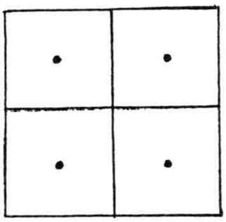
and therefore

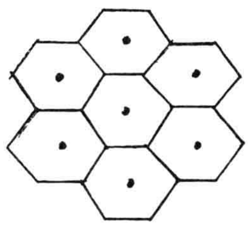
$$V(B) \leqq 2^n \delta(B) \det \Lambda.$$

THEOREM 3A. *If B is convex and symmetric about $\underline{0}$ and $V(B) > 2^n \delta(B) \det \Lambda$, then B contains a non-zero lattice point.*

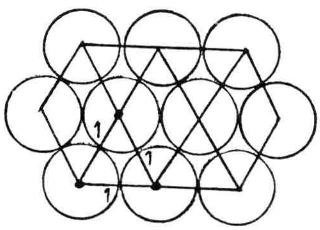
In particular, this happens when $V(B) > 2^n \det \Lambda$. So 3A is a strengthening of Minkowski's Theorem 2B.

Remark 3B. For B convex, symmetric about $\underline{0}$, one can show that $\delta(B) < 1$ except for certain polyhedra. E.g., the cube has density $\delta = 1$. So do regular hexagons in the plane.





Let $\delta_n = \delta(B)$ where B is a ball in $\mathbb{R}^n$. Consider the following picture.



The "triangle" lattice Λ has $\det \Lambda = \frac{1}{2}\sqrt{3}$. It is easy to guess that

$$\delta_2 = \frac{V(B)}{\det \Lambda} = \frac{(\frac{1}{2})^2 \pi}{\frac{1}{2}\sqrt{3}}.$$

This had already been proved implicitly by Gauss in his theory of positive definite binary quadratic forms. We know the values of $\delta_2, \delta_3, \dots, \delta_8$; see Cassels (1959, Appendix) and Exercise 3b below. The estimation of δ_n for large n remains among the central unsolved problems in the Geometry of Numbers. Blichfeldt (1929) proved that $\delta_n \leqq 2^{-n/2}(1+\frac{n}{2})$. Also, $\delta_n \geqq (\frac{1}{2}-\varepsilon)^n$ if $n > n_0(\varepsilon)$. See Cassels (1959, p. 249). More recently, G.A. Kabatjanskii and V.I. Levenšteĭn (1978) have shown that $\delta_n \leqq 2^{-0.599n(1-\epsilon)}$ for $n > n_0(\epsilon)$.

One may in a fairly obvious way define a general (not necessarily lattice) packing of a set B, and the maximum general packing density. For a disk in $\mathbb{R}^2$, Thue (1892) had shown that the maximum packing density is in fact achieved for a lattice–packing. It is not known whether a similar result holds for a ball in $\mathbb{R}^3$. It is generally believed that the densest packing density of a ball in $\mathbb{R}^n$ where n is sufficiently large is less than the smallest lattice packing density.

Now $V(B)\lambda_1^n \leqq 2^n \det \Lambda \delta(B)$, so that $\lambda_1 \leqq 2(\det \Lambda)^{1/n}(\delta(B)/V(B))^{1/n}$. For the unit ball B, $V(B) = V(n) = \pi^{n/2}/\Gamma(1+\frac{n}{2})$, so that by Stirling's formula we have the asymptotic relation

$$V(n)^{2/n} = V(B)^{2/n} \sim \frac{2\pi e}{n} \quad \text{as} \quad n \to \infty.$$

We define *Hermite's constant* γ_n to be least such that for any lattice Λ

$$\lambda_1 \leqq \gamma_n^{1/2}(\det \Lambda)^{1/n}$$

where $\lambda_1 = \lambda_1$ (unit ball). So

$$\gamma_n \leqq \frac{4\delta_n^{2/n}}{V(n)^{2/n}} \leqq \frac{2n}{\pi} \qquad \text{if} \quad n \geq 2.$$

We have $\overline{\lim}(\gamma_n/n) \leqq \frac{4}{2\pi e} = \frac{2}{\pi e}$, by using the trivial estimate $\delta_n \leqq 1$. If instead we use Blichfeldt's estimate, we obtain $\overline{\lim}(\gamma_n/n) \leqq \frac{1}{\pi e}$. The result of Kabatjanskii and Levenšteĭn quoted above yields $\overline{\lim}(\gamma_n/n) \leqq 2^{-0.197}(\pi e)^{-1}$.

Exercise 3b. Show that $\gamma_n = 4\delta_n^{2/n}/V(n)^{2/n}$.

Exercise 3c. Let $Q(\underline{\underline{X}}) = \sum_{i,j=1}^n a_{ij}X_iX_j$ be a positive definite quadratic form with real coefficients $a_{ij} = a_{ji}$. Then there exists a non-zero integer point $\underline{\underline{x}}$ with $Q(\underline{\underline{x}}) \leqq \gamma_n(\det Q)^{1/n}$. Moreover, γ_n is least with this property.

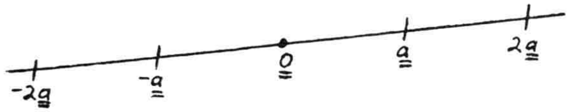
§4. Siegel's Lemma Again.

A *rational subspace* of $\mathbb{R}^n$ or $\mathbb{C}^n$ is a subspace spanned by vectors with rational coordinates. A rational subspace S^k of dimension k is spanned by k vectors $\underline{\underline{a}}_1, \dots, \underline{\underline{a}}_k \in \mathbb{Q}^n$. Such a space S^k may be defined by $n-k$ linear homogeneous equations with rational coefficients. The integer points in a subspace S^k form a set Λ which is a lattice of S^k by Lemma 2A.

The *height* of S^k is defined by

$$H(S^k) = \det \Lambda.$$

An integer point $\underline{a} = (a_1, \dots, a_n)$ is called *primitive* if $gcd(a_1, \dots, a_n) = 1$. Then $\underline{a}$ is "closest to the origin," i.e., there is no integer point on the line segment between $\underline{0}$ and $\underline{a}$. Either $\underline{a}$ or $-\underline{a}$ is a basis of the 1–dimensional lattice of integer points of the space S^1 spanned by $\underline{a}$. Thus $H(S^1) = |\underline{a}|$.



By the definition of Hermite's constant, there is on S^k an integer point $\underline{x} \neq \underline{0}$ with

$$|\underline{x}| \leqq \gamma_k^{1/2} H(S^k)^{1/k}.$$

LEMMA 4A. *Consider the unit cube C in $\mathbb{R}^n$, i.e., $|x_i| \leqq 1$ $(i = 1, \dots, n)$. Let S^k be a k–dimensional subspace. Then $C \cap S^k$ has k–dimensional volume $\geqq 2^k$.*

This result, due to J. Vaaler (1979), will not be proved here.

Let S^k be a rational subspace, Λ the lattice of integer points associated with S^k, i.e., $\Lambda = \Lambda(S^k)$. Let $B = C \cap S^k$. By Minkowski's Theorem 2C, $\lambda_1^k V(B) \leqq 2^k \det \Lambda$. Now $V(B) \geqq 2^k$ so that

$$\lambda_1^k 2^k \leqq 2^k \det \Lambda,$$
$$\lambda_1^k \leqq \det \Lambda = H(S^k),$$
$$\lambda_1 \leqq H(S^k)^{1/k}.$$

Recall that $|\underline{x}|$ was the Euclidean norm. Let

$$\overline{|\underline{x}|} = \max(|x_1|, \dots, |x_n|)$$

be the maximum norm. Our results may be summarized in

LEMMA 4B. *Given a rational subspace S^k there is an integer point $\underline{x} \neq \underline{0}$ on S^k with*

$$|\underline{x}| \leqq \gamma_k^{1/2} H(S)^{1/k}.$$

Also, there is such a point $\underline{x}^1$ with

$$\overline{|\underline{x}^1|} \leqq H(S)^{1/k}.$$

When S^k is a rational subspace, then $(S^k)^\perp$ is a rational subspace of dimension $m = n - k$.

LEMMA 4C. $H(S^\perp) = H(S)$.

The proof is postponed to the next section (see Corollary 5J). To make the lemma correct for S^0 and $S^n = \mathbb{R}^n$, we set $H(S^0) = 1$.

Let us go back to a system of linear equations

$$\begin{aligned} a_{11}x_1 + \cdots + a_{1n}x_n &= 0 \\ &\vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n &= 0 \end{aligned}$$

where $0 < m < n$, $a_{ij} \in \mathbb{Z}$. If these equations are independent, they define a rational subspace S^k of dimension $k = n - m$. And $(S^k)^\perp$ is spanned by the row vectors $\underline{\underline{a}}_1, \ldots, \underline{\underline{a}}_m$ where $\underline{\underline{a}}_i = (a_{i1}, \ldots, a_{in})$. Then

$$H(S^k) = H(S^{k\perp}) \leqq |\underline{\underline{a}}_1| \cdots |\underline{\underline{a}}_m|.$$

The last inequality occurs because $\underline{\underline{a}}_1, \ldots, \underline{\underline{a}}_m$ can be written as linear combinations of basis vectors for the lattice, so that $\det|\underline{\underline{a}}_1, \ldots, \underline{\underline{a}}_m|$ is an integer multiple of the determinant of a basis.[‡] Thus

$$\det(\Lambda(S^{k\perp})) \leqq |\det(\underline{\underline{a}}_1, \ldots, \underline{\underline{a}}_m)| \leqq |\underline{\underline{a}}_1| \cdots |\underline{\underline{a}}_m|$$

by Hadamard's inequality, which is a consequence of Lemma 5E below.

LEMMA 4D. (Siegel's Lemma) *Given the system of equations above,*

(i) *there is a non-trivial integer solution $\underline{\underline{x}}$ with*

$$|\underline{\underline{x}}| \leqq \gamma_{n-m}^{1/2}(|\underline{\underline{a}}_1| \cdots |\underline{\underline{a}}_m|)^{1/(n-m)} \leqq \left(\frac{2}{\pi}n\right)^{1/2} (\sqrt{n}\, A)^{m/(n-m)}$$

if $|a_{ij}| \leqq A$ for every i, j.

(ii) *Also, there is a non-trivial integer solution $\underline{\underline{x}}^1$ with*

$$\overline{|\underline{\underline{x}}^1|} \leqq (|\underline{\underline{a}}_1| \cdots |\underline{\underline{a}}_m|)^{1/(n-m)} \leqq (\sqrt{n}\, A)^{m/(n-m)}.$$

In the first inequality we used that $\gamma_{n-m} < \frac{2}{\pi}(n-m) < \frac{2}{\pi}n$ if $n - m \geqq 2$, and $\gamma_{n-m} = 1 < \frac{2}{\pi}n$ if $n - m = 1$, so that $n \geqq 2$. It is clear that we do not have to restrict to the case when the m equations are independent.

Remark 4E. If Minkowski's Second Theorem (2E) is used, (ii) can be strengthened to get the following: *there are $n-m$ linearly independent solutions $\underline{\underline{x}}_1, \ldots, \underline{\underline{x}}_{n-m}$ of our system of equations such that*

$$\overline{|\underline{\underline{x}}_1|}\,\overline{|\underline{\underline{x}}_2|} \cdots \overline{|\underline{\underline{x}}_{n-m}|} \leqq |\underline{\underline{a}}_1| \cdots |\underline{\underline{a}}_m|.$$

The first assertion can be strengthened in the same way, but this is not so obvious.

§5. Grassman Algebra.

[‡]We think of $\underline{\underline{a}}_1, \ldots, \underline{\underline{a}}_m$ as vectors with m components in terms of an orthonormal coordinate system in $S^{k\perp}$