Zeitschrift: L'Enseignement Mathématique
Herausgeber: Commission Internationale de l'Enseignement Mathématique
Band: 17 (1971)
Heft: 1: L'ENSEIGNEMENT MATHÉMATIQUE

Artikel: APPROXIMATION TO ALGEBRAIC NUMBERS
Autor: Schmidt, Wolfgang M.
DOI: https://doi.org/10.5169/seals-44578

Nutzungsbedingungen
Die ETH-Bibliothek ist die Anbieterin der digitalisierten Zeitschriften auf E-Periodica. Sie besitzt keine Urheberrechte an den Zeitschriften und ist nicht verantwortlich für deren Inhalte. Die Rechte liegen in der Regel bei den Herausgebern beziehungsweise den externen Rechteinhabern. Das Veröffentlichen von Bildern in Print- und Online-Publikationen sowie auf Social Media-Kanälen oder Webseiten ist nur mit vorheriger Genehmigung der Rechteinhaber erlaubt. Mehr erfahren

Conditions d'utilisation
L'ETH Library est le fournisseur des revues numérisées. Elle ne détient aucun droit d'auteur sur les revues et n'est pas responsable de leur contenu. En règle générale, les droits sont détenus par les éditeurs ou les détenteurs de droits externes. La reproduction d'images dans des publications imprimées ou en ligne ainsi que sur des canaux de médias sociaux ou des sites web n'est autorisée qu'avec l'accord préalable des détenteurs des droits. En savoir plus

Terms of use
The ETH Library is the provider of the digitised journals. It does not own any copyrights to the journals and is not responsible for their content. The rights usually lie with the publishers or the external rights holders. Publishing images in print and online publications, as well as on social media channels or websites, is only permitted with the prior consent of the rights holders. Find out more

Download PDF: 16.10.2025

ETH-Bibliothek Zürich, E-Periodica, https://www.e-periodica.ch

vide-leer-empty

APPROXIMATION TO ALGEBRAIC NUMBERS [1]

by Wolfgang M. SCHMIDT

Contents

0. Introduction 188

1. Approximation to real numbers by rationals 188

2. Approximation to real algebraic numbers by rationals. Roth's Theorem 193

3. An outline of the proof of Roth's Theorem 199

4. Some generalizations of Roth's Theorem 205

5. Effective methods. Baker's Theorem 211

6. Simultaneous approximation to real numbers by rationals . . 216

7. Simultaneous approximation to algebraic numbers by rationals 223

8. Tools from the Geometry of Numbers 229

9. Outline of the proof of the theorems on simultaneous approximation to algebraic numbers 233

10. Norm forms 241

11. Generalizations and open problems 246

References 248

---

[1] Survey lectures given by invitation at the Institute for Advanced Study on February 1, 4, 8, 11 and 15, 1971, under the sponsorship of the International Mathematical Union while the author was supported in part by Air Force Office of Scientific Research grant AF-AFOSR-69-1712.

0. Introduction

Our subject is part of the more general field of *diophantine approximation*, i.e. the study of rational approximation to real numbers. Books on diophantine approximation in general are due to Minkowski (1907) [1]), Koksma (1936), Cassels (1957), Niven (1963) and Lang (1966b), and a p-adic version is treated by Lutz (1955).

In the present survey we shall be concerned with the more special problem of rational approximation to real *algebraic* numbers. Contributions to this problem were first made by Liouville (1844), and deep theorems were proved among others by Thue (1908), Siegel (1921a), Roth (1955a) and Baker (1968b). We shall also discuss the more general questions of approximation to an algebraic number by algebraic numbers in a fixed number field or by algebraic numbers of fixed degree, and the question of simultaneous approximation to real algebraic numbers by rationals. As is well known, many results on approximation to algebraic numbers have applications to diophantine equations.

Of the books listed above, the one by Cassels (1957) has a chapter (ch. VI) on approximation to algebraic numbers. This subject also is the main topic in the book by Mahler (1961) and is the topic of chapter 6 of Le Veque (1955), of Kapitel 1 of Schneider (1957) and of chapter 6 of Lang (1962). Also see Lang (1971) and chapter 1 of Feldman and Shidlovskii (1967).

Until recently all the deep theorems on approximation to algebraic numbers were obtained by the method of Thue, Siegel and Roth, and accordingly most of the present survey is devoted to this method. In view of Baker's (1968b) results it is possible that the method of Gelfond and Baker will play an increasing role in the future. Rather than attempting to give a complete account of the literature, I tried to explain the main ideas in the proofs of the principal theorems.

1. Approximation to real numbers by rationals

1.1. This section is intended for the benefit of a reader who is not familiar with diophantine approximation, to provide a background for the

[1]) References are listed at the end. They are listed alphabetically by the name of the author, by the year, and finally by *a*, *b*, ... if there are several works by the same author in the same year.

more special problem of approximation to real algebraic numbers which will be discussed later.

THEOREM 1A (Dirichlet 1842). *Suppose α is a real number and Q is a real number greater than* 1. *Then there are integers p, q with*

$$1 \leqq q < Q \quad \textit{and} \quad |\alpha q - p| \leqq Q^{-1}. \tag{1.1}$$

Let us recall the well known proof. Every real number ξ may be written as

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

where $[\xi]$ is a rational integer, the *integer part* of ξ, and where $\{\xi\}$, the *fractional part* of ξ, satisfies $0 \leqq \{\xi\} < 1$. Assume now that Q is an integer. The $Q + 1$ numbers

$$0, 1, \{\alpha\}, \{2\alpha\}, \ldots, \{(Q-1)\alpha\}$$

lie in the unit-interval $0 \leqq \xi \leqq 1$; hence there are two of these numbers whose difference has absolute value at most Q^{-1}. Thus there are integers r_1, r_2, s_1, s_2 with $0 \leqq r_i \leqq Q - 1$ $(i=1, 2)$ and $r_1 \neq r_2$ such that

$$|(r_1\alpha - s_1) - (r_2\alpha - s_2)| \leqq Q^{-1}.$$

If, say, $r_1 > r_2$, then $p = s_1 - s_2$ and $q = r_1 - r_2$ satisfy (1.1). This proves Dirichlet's Theorem when Q is an integer.

Now suppose that Q is *not* an integer. Since $Q' = [Q] + 1$ is an integer, there are integers p, q with $1 \leqq q < Q'$ and $|\alpha q - p| \leqq Q'^{-1}$, whence with $1 \leqq q < Q$ and $|\alpha q - p| < Q^{-1}$.

1.2. The greatest common factor of integers p, q will be denoted by (p, q). It is clear that in Dirichlet's Theorem one could stipulate that $(p, q) = 1$. The inequalities (1.1) in Dirichlet's Theorem yield

$$\left|\alpha - \frac{p}{q}\right| < \frac{1}{q^2}. \tag{1.2}$$

COROLLARY 1B. *Suppose that α is irrational. Then there exist infinitely many rationals p/q with $(p, q) = 1$ and with* (1.2).

For since α is irrational, the inequality $|\alpha q - p| \leqq Q^{-1}$ in (1.1) can for fixed integers p, q with $q \neq 0$ hold only for bounded values of Q, say for $Q \leqq Q_0(p, q)$. Hence as $Q \to \infty$, there will be infinitely many distinct

pairs of coprime integers p, q in Dirichlet's Theorem, giving rise to infinitely many rationals p/q with (1.2).

One should remark that this corollary does not hold for rationals α. For if $\alpha = a/b$ and if $p/q \neq a/b$, then

$$\left| \alpha - \frac{p}{q} \right| \geqq |bq|^{-1} > \frac{1}{q^2} \quad \textit{if} \quad |q| > |b| .$$

Corollary 1B can be strengthened:

THEOREM 1C (Hurwitz 1891).

(i) *For every irrational α there are infinitely many rationals p/q with $(p, q) = 1$ and with*

$$\left| \alpha - \frac{p}{q} \right| < \frac{1}{\sqrt{5}\, q^2} . \tag{1.3}$$

(ii) *This would no longer be true if $\sqrt{5}$ were replaced by a larger constant.*

The second statement can easily be proved: Suppose α is a real quadratic irrational and suppose there are infinitely many rationals p/q with

$$\left| \alpha - \frac{p}{q} \right| < \frac{1}{Aq^2} . \tag{1.4}$$

Let $P(x) = ax^2 + bx + c$ be a polynomial with rational integer coefficients and with root α; then $P(x) = a(x-\alpha)(x-\alpha')$ where α' is the conjugate of α. For every p/q with (1.4) we have

$$\frac{1}{q^2} \leqq \left| P\left(\frac{p}{q}\right) \right| = \left| \alpha - \frac{p}{q} \right| \left| a\left(\alpha' - \frac{p}{q}\right) \right| < \frac{1}{Aq^2} \left| a\left(\alpha' - \alpha + \alpha - \frac{p}{q}\right) \right|$$

$$< \frac{\sqrt{D}}{Aq^2} + \frac{|a|}{A^2 q^4} ,$$

where $D = b^2 - 4ac = a^2(\alpha-\alpha')^2$ is the discriminant of P. It follows that $A \leqq \sqrt{D}$. In the special case when $\alpha = \dfrac{\sqrt{5}-1}{2}$, $P(x) = x^2 + x - 1$, we have $D = 5$ whence $A \leqq \sqrt{5}$.

Note that $\dfrac{1}{\sqrt{5}}$ is a quadratic irrationality. It can be shown that the

numbers α for which this constant is best possible are certain numbers in the quadratic number field generated by $\sqrt{5}$.

One calls an irrational number α *badly approximable* if there is a $c = c(\alpha) > 0$ such that

$$\left| \alpha - \frac{p}{q} \right| > \frac{c(\alpha)}{q^2}$$

for every rational p/q. We have just seen that the quadratic irrationals are badly approximable.

1.3. Certain results on diophantine approximation are closely related to continued fractions. Continued fractions are discussed in the books mentioned at the beginning, and a fuller account of them is given in Perron (1954). The rational function

$$a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cdots + \cfrac{1}{a_n}}}$$

will be denoted by $[a_0, a_1, \ldots, a_n]$ and will be called a *continued fraction*. Every rational number r may be written $r = [a_0, a_1, \ldots, a_n]$ where $n \geqq 0$ and where a_0 is an integer and $a_1, \ldots, a_n$ are positive integers. For every irrational α there exist unique rational integers $a_0, a_1, a_2, \ldots$ such that $a_1, a_2, \ldots$ are positive and $\lim_{n\to\infty} [a_0, a_1, \ldots, a_n] = \alpha$. The integers $a_0, a_1, a_2, \ldots$ are called the *partial quotients* of the *continued fraction expansion* of α. Define coprime integers p_n, q_n with $q_n > 0$ by $p_n/q_n = [a_0, a_1, \ldots, a_n]$. The rationals p_n/q_n converge to α, and they are called the *convergents to* α. These convergents are important for diophantine approximation because it can easily be shown that

$$\frac{1}{(a_{n+1}+2)\, q_n^2} < \left| \alpha - \frac{p_n}{q_n} \right| < \frac{1}{a_{n+1} q_n^2} \tag{1.5}$$

for every $n \geqq 0$, which implies in particular that $\left|\alpha - \frac{p_n}{q_n}\right| < \frac{1}{q_n^2}$, and because it was shown by Legendre that if

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

then $\frac{p}{q}$ is necessarily a convergent to α. It follows that α is badly approximable precisely if the partial quotients of the continued fraction expansion of α are bounded. In particular, the real quadratic irrationals have bounded partial quotients. (In fact it is well known that these numbers have a " periodic " continued fraction expansion.)

1.4. We have seen that for certain irrationals α the number $\frac{1}{\sqrt{5}}$ in (1.3) cannot be replaced by a smaller factor. But for most irrationals α the inequality (1.3) can be improved:

THEOREM 1D (Khintchine 1926b). *Suppose $\psi(q)$ is a positive, non-increasing function defined for $q = 1, 2, \ldots$. Consider the inequality*

$$\left|\alpha - \frac{p}{q}\right| < \frac{\psi(q)}{q} \tag{1.6}$$

and the sum

$$\sum_{q=1}^{\infty} \psi(q). \tag{1.7}$$

If the sum is convergent, then (1.6) *has only finitely many solutions in rationals p/q with $q > 0$ for almost all α (in the sense of Lebesgue measure). If the sum is divergent, then* (1.6) *has infinitely many solutions for almost all α.*

In particular, for every $\delta > 0$, the inequality

$$\left|\alpha - \frac{p}{q}\right| < \frac{1}{q^{2+\delta}} \tag{1.8}$$

has only finitely many solutions for almost all α, but

$$(1.9) \qquad \left|\alpha - \frac{p}{q}\right| < \frac{1}{q^2 \log q}$$

has infinitely many solutions for almost all α.

Given a number like $\sqrt{2}$, e, π or $\sqrt[3]{2}$, it is of interest to know whether it behaves like almost every number. Quadratic irrationals are badly approximable and hence behave like almost every number with respect to (1.8) but not with respect to (1.9). From the known continued fraction expansion of e it is easy to deduce that neither of the inequalities (1.8), (1.9) has infinitely many solutions if $\alpha = e$. Mahler (1953) showed that $\left|\pi - \frac{p}{q}\right| < q^{-42}$ has only finitely many solutions, and Wirsing (unpublished) could reduce 42 to 21. The behavior of $\sqrt[3]{2}$ and of real algebraic numbers in general will be discussed in the next section.

2. Approximation to algebraic numbers by rationals. Roth's Theorem

2.1. Theorem 2A (Liouville 1844). *Suppose α is a real algebraic number of degree d. Then there is a constant $c(\alpha)$* [1] > 0 *such that*

$$\left|\alpha - \frac{p}{q}\right| > \frac{c(\alpha)}{q^d}$$

for every rational $\frac{p}{q}$ distinct from α.

This theorem was used by Liouville to construct transcendental numbers. For example, put $\alpha = \sum_{v=1}^{\infty} 2^{-v!}$, $q(k) = 2^{k!}$, $p(k) = 2^{k!} \sum_{v=1}^{k} 2^{-v!}$. Then

$$\left|\alpha - \frac{p(k)}{q(k)}\right| = \sum_{v=k+1}^{\infty} 2^{-v!} < 2 \cdot 2^{-(k+1)!} = 2(q(k))^{-k-1}.$$

Hence for any d and any constant $c > 0$ one has

$$\left|\alpha - \frac{p(k)}{q(k)}\right| < \frac{c}{(q(k))^d}$$

[1]) The constants $c, c_1, c_2, \ldots$ of different subsections are independent.

if k is large. By Liouville's Theorem, α cannot be algebraic of any degree d, and hence α is transcendental.

For the sake of later refinements we shall break the extremely simple proof of Liouville's Theorem into three parts (a), (b) and (c).

(a) Let $P(x)$ be the *defining polynomial* of α, i.e. the polynomial of degree d with root α which has coprime integer coefficients and a positive leading coefficient.

(b) Taylor's formula yields

$$\left| P\left(\frac{p}{q}\right) \right| = \left| \sum_{i=1}^{d} \left(\frac{p}{q} - \alpha\right)^i \frac{1}{i!} P^{(i)}(\alpha) \right| < \frac{1}{c(\alpha)} \left| \frac{p}{q} - \alpha \right|$$

if

$$\left| \frac{p}{q} - \alpha \right| \leqq 1 .$$

(c) $P\left(\frac{p}{q}\right) \neq 0$, whence $\left| P\left(\frac{p}{q}\right) \right| \geqq \frac{1}{q^d}$, and combining this with (b) we obtain Liouville's Theorem if $\left| \frac{p}{q} - \alpha \right| \leqq 1$. The Theorem is obvious if $\left| \frac{p}{q} - \alpha \right| > 1$.

2.2. Now suppose that α is a real algebraic number of degree d and consider the inequality

$$\left| \alpha - \frac{p}{q} \right| < \frac{1}{q^\mu}, \tag{2.1}$$

where $\frac{p}{q}$ is rational with a positive denominator q. By Liouville's Theorem this inequality has only finitely many solutions if $\mu > d$. Thue (1908, 1909) made the important discovery that this is still true under the weaker assumption that $\mu > (d/2) + 1$. Then Siegel (1921a) showed that it suffices to have $\mu > 2\sqrt{d}$. (Actually his result was slightly better, with a more complicated function in place of $2\sqrt{d}$.) These results of Thue and of Siegel will be referred to as Thue's Theorem and as Siegel's Theorem. Dyson (1947) improved $\mu > 2\sqrt{d}$ to $\mu > \sqrt{2d}$. (See also Gelfond (1952), ch. 1.) Finally

Roth (1955a) showed that (2.1) with $\mu > 2$ has only finitely many solutions. His result may be formulated as follows.

THEOREM 2B. *Suppose α is a real algebraic number. Then for every $\delta > 0$ there are only finitely many distinct rationals $\frac{p}{q}$ with $q > 0$ and with*

$$\left|\alpha - \frac{p}{q}\right| < \frac{1}{q^{2+\delta}}. \tag{2.2}$$

In view of Dirichlet's Theorem, the exponent 2 is best possible here. But it is conceivable that the factor q^δ could be replaced by a smaller factor. But nothing is known in this direction. The metrical result (Theorem 1D) suggests that

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

has only finitely many solutions for every positive δ. The first written account of this conjecture appears to be in Lang (1965a).

For real quadratic irrationals α we have $\left|\alpha - \frac{p}{q}\right| > \frac{c(\alpha)}{q^2}$ by Liouville's Theorem (or by the fact that α is badly approximable, as was shown below Theorem 1C). Hence for such numbers α, Liouville's Theorem is stronger than Roth's Theorem. It is not known whether there exists an algebraic number of degree $d \geqq 3$ whose partial quotients are bounded, or whether there exists such a number whose partial quotients are unbounded. In view of Theorem 1D it is likely that every algebraic number of degree $d \geqq 3$ has unbounded partial quotients. Some numerical evidence for this was given by von Neumann and Tuckerman (1955), Richtmyer et al. (1962), and Bruyno (1964).

2.3. Let us see how much Roth's Theorem tells us about the partial quotients. In view of (1.5) it shows that

$$a_{n+1} < q_n^\delta$$

for every $\delta > 0$ and for $n > n_0(\alpha, \delta)$. Now it is well known that $q_0 = 1$, $q_1 = a_1 q_0 + 1$ and that $q_n = a_n q_{n-1} + q_{n-2}$ for $n \geqq 2$, hence that $q_n \leqq (a_1+1) \dots (a_n+1)$, and we obtain

$$a_{n+1} < ((a_1+1)\dots(a_n+1))^{\delta} \tag{2.3}$$

for every $\delta > 0$ and for $n > n_1(\alpha, \delta)$. On the other hand one has

$$q_n > (a_n a_{n-1}+1)\, q_{n-2} \geqq ((a_n+1)(a_{n-1}+1))^{1/2}\, q_{n-2} > \dots$$
$$> ((a_n+1)\dots(a_1+1))^{1/2},$$

and this shows that the truth of (2.3) for every $\delta > 0$ and sufficiently large n is equivalent to Roth's Theorem. Davenport and Roth (1955) proved that for real algebraic irrationals α one has

$$\log\log q_n < \frac{c_1(\alpha)\, n}{\sqrt{\log n}}.$$

Further results on the continued fraction expansion of algebraic numbers were given by Baker (1962).

2.4. Cugiani (1959) could show that if $\frac{p(1)}{q(1)}, \frac{p(2)}{q(2)}, \dots$ are solutions of

$$\left|\alpha - \frac{p}{q}\right| < q^{-2-20\,(\log\log\log q)^{-1/2}}$$

with $0 < q(1) < q(2) < \dots$, then

$$\limsup \frac{\log q(k+1)}{\log q(k)} = \infty. \tag{2.4}$$

Before Roth's Theorem was known, Schneider (1936) had shown that if $\frac{p(1)}{q(1)}, \frac{p(2)}{q(2)}, \dots$ with $0 < q(1) < q(2) < \dots$ are solutions of (2.2), then (2.4) holds. Schneider's Theorem in turn is a sharpening of a similar result of Siegel (1921b). Roth's Theorem enables one to prove the transcendency of a wider class of numbers than Liouville's Theorem, e.g. of $\alpha = \sum_{\nu=1}^{\infty} 2^{-3^{\nu}}$, but actually this can also be done with the earlier theorem of Schneider just mentioned.

2.5. Davenport and Roth (1955) have determined an explicit upper bound $B = B(\alpha, \delta)$ for the number of solutions of (2.2). However, at present one cannot give an upper bound $B^* = B^*(\alpha, \delta)$ for the denominators q of the solutions p/q of (2.2). Hence Roth's Theorem is " non-

effective". It is easy to see that Liouville's Theorem is effective, but the theorems of Thue, Siegel and Dyson are also non-effective. Some further remarks on this question will be made in §3.6.

But effective bounds for weaker inequalities than (2.2) were given by Baker and will be discussed in §5.

2.6. Now suppose that $F(x, y)$ is a not identically vanishing binary form of degree $d \geqq 3$ with rational coefficients which has no multiple factors of positive degree. Such a binary form can be factored as

$$F(x, y) = L_1(x, y) \dots L_d(x, y)$$

where $L_i(x, y) = \gamma_i x + \delta_i y$ $(i = 1, \dots, d)$ are linear forms whose coefficients are real or complex algebraic numbers. Since F has no multiple factors, any two linear forms L_i, L_j with $i \neq j$ are linearly independent.

Let (x, y) be an integer point with $F(x, y) \neq 0$. By rearranging the factors $L_1, \dots, L_d$ we may assume that

$$0 < |L_1(x, y)| \leqq \dots \leqq |L_d(x, y)|.$$

Now if $\gamma_1 = 0$ or if δ_1/γ_1 is rational, then it is clear that $|L_1(x, y)| \geqq c_1$ with a positive c_1 independent of (x, y). If $\gamma_1 \neq 0$ and $y = 0$, then $|L_1(x, y)| = |\gamma_1|(|x| + |y|)$. Finally if $\gamma_1 \neq 0$, δ_1/γ_1 is irrational and $y \neq 0$, then $L_1(x, y) = \gamma_1 y \left(\frac{x}{y} - \alpha\right)$ with $\alpha = -\delta_1/\gamma_1$, and for every $\delta > 0$ one has $|L_1(x, y)| \geqq c_2(\delta)|y|^{1-2-\delta} \geqq c_2(\delta)(|x| + |y|)^{-1-\delta}$ by Roth's Theorem. (Roth's Theorem is trivially true if α is complex.) Therefore it is true in general that $|L_1(x, y)| \geqq c_3(\delta)(|x| + |y|)^{-1-\delta}$ with $c_3(\delta) > 0$. On the other hand since L_1, L_2 are linearly independent, we have

$$|L_d(x, y)| \geqq \dots \geqq |L_2(x, y)| \geqq \tfrac{1}{2}(|L_1(x, y)| + |L_2(x, y)|)$$
$$\geqq c_4(|x| + |y|),$$

whence

$$|F(x, y)| \geqq c_3(\delta)\, c_4^{d-1}(|x| + |y|)^{-1-\delta+(d-1)} = c_5(\delta)(|x| + |y|)^{d-2-\delta}.$$

Thus the following holds.

THEOREM 2C. *Suppose $F(x, y)$ is a binary form of degree $d \geqq 3$ with rational coefficients and without multiple factors. Then for any $\nu < d - 2$ there are only finitely many integer points (x, y) with*

$$0 < |F(x, y)| < (|x| + |y|)^{\nu}.$$

COROLLARY 2D. *Suppose $F(x, y)$ is a binary form as in Theorem 2C and suppose $F(x, y)$ has no rational linear factor. Let $G(x, y)$ be a polynomial of total degree $v < d - 2$. Then there are only finitely many integer points (x, y) with*

$$F(x, y) = G(x, y). \tag{2.5}$$

Namely, such a form F has $0 < |F(x, y)|$ for any non-zero integer point (x, y). We deduced Theorem 2C and its corollary from Roth's Theorem. If instead we would have used Thue's or Siegel's Theorem, then we would have had to replace the condition $v < d - 2$ by the stronger condition $v < (d/2) - 1$ or $v < d - 2\sqrt{d}$, respectively. In particular, Thue's Theorem suffices to deal with the equation

$$F(x, y) = m \tag{2.6}$$

where m is a constant, which is often called " Thue's equation ".

Using his (1921a) result, Siegel (1929) could classify all algebraic curves defined over the rationals on which there are infinitely many integer points. In particular these curves must be of genus zero.

Schinzel (1968) used this result of Siegel to prove a theorem which implies that in Corollary 2D the assumption that $v < d - 2$ may be replaced by the weaker assumption that $v < d$. Roth's Theorem is not required to obtain this sharper version of Corollary 2D.

2.7. Mahler (1933b), (1933c) gave upper bounds for the number of solutions of Thue's equation (2.6). Davenport and Roth (1955) derived upper bounds for the number of solutions of the equation (2.5) of Corollary 2D. Siegel (1970) showed that there is an explicit such bound for Thue's equation (2.6) which depends only on m and the degree d if $F(x, y) = (\alpha x + \beta y)^d + (\gamma x + \delta y)^d$ with $\alpha\delta - \beta\gamma \neq 0$. (In particular, every form $F(x, y)$ of degree $d = 3$ may be written in this way.) Perhaps Siegel's conclusion is true for arbitrary forms $F(x, y)$.

Mahler (1933c) gave an asymptotic formula

$$N(m) \approx c_1(F)\, m^{2/d}$$

for the number $N(m)$ of solutions of $|F(x, y)| \leqq m$ in coprime integers x, y. Now let $N'(m)$ be the number of integers n with $|n| \leqq m$ which may be represented at least once as $n = F(x, y)$ with coprime x, y. Hooley (see (1967) and the references given there) developed powerful analytic methods to show that

$$N'(m) \approx c_2(F)\, m^{2/d}$$

if either $d = 3$ and the discriminant of F is not of some rather special type, or if $F(x, y) = x^d + y^d$ for some $d \geqq 3$. To generalize these results to arbitrary forms $F(x, y)$ appears to be extremely difficult.

The methods of Thue, Siegel and Roth do not enable one to find bounds for the size $|x| + |y|$ of solutions of Thue's equation, and hence they provide no method to find all the solutions of such an equation. Therefore these methods are called " non-effective ". Effective results will be discussed in §5.

3. An outline of the proof of Roth's Theorem

3.1. We shall follow Cassel's rearrangement (Cassels (1957), ch. VI) of Roth's proof. It is easy to see that we may restrict ourselves to the case when α is an algebraic *integer* of degree $d > 1$.

Suppose we tried to modify the proof of Liouville's Theorem as follows. In step (a) we pick a polynomial $P(x)$ with rational integer coefficients which has a root at α of order i and which has degree r. Next, in step (b) we suppose that

$$\left|\alpha - \frac{p}{q}\right| < \frac{1}{q^{\mu}}, \tag{3.1}$$

and Taylor's expansion

$$P\left(\frac{p}{q}\right) = \sum_{j=i}^{r} \left(\frac{p}{q} - \alpha\right)^{j} \frac{1}{j!} P^{(j)}(\alpha)$$

yields $\left|P\left(\frac{p}{q}\right)\right| \leqq cq^{-\mu i}$. Finally (c) we have $P\left(\frac{p}{q}\right) \neq 0$ whence $\left|P\left(\frac{p}{q}\right)\right| \geqq q^{-r}$ for all but finitely many rationals $\frac{p}{q}$. Hence if (3.1) has infinitely many solutions, then $\mu i \leqq r$ or

$$\mu \leqq \left(\frac{i}{r}\right)^{-1}.$$

Hence one should try to make $\frac{i}{r}$ as large as possible. But it is clear that

always $\frac{i}{r} \leqq \frac{1}{d}$, and that $\frac{i}{r} = \frac{1}{d}$ if $P(x)$ is a power of the defining polynomial of α. Hence this method only gives $\mu \leqq d$, i.e. nothing better than Liouville's result.

3.2. In order to improve on this estimate, Thue and Siegel use a polynomial $P(x_1, x_2)$ in two variables, and Schneider (1936) and Roth use a polynomial $P(x_1, \ldots, x_m)$ in many variables. It is necessary to define the order of vanishing of $P(x_1, \ldots, x_m)$ at a given point $(\xi_1, \ldots, \xi_m)$. The simplest definition would be to take the smallest value of $i_1 + \ldots + i_m$ for which the mixed partial derivative

$$P^{(i_1, \ldots, i_m)}(\xi_1, \ldots, \xi_m) \neq 0 . \tag{3.2}$$

But it is necessary to study polynomials $P(x_1, \ldots, x_m)$ which have rather different degrees in $x_1, \ldots, x_m$, and hence it will be better to attach different weights to the integers $i_1, \ldots, i_m$ in (3.2). Thus Roth defines the *index of* P *at* $(\xi_1, \ldots, \xi_m)$ with respect to a given m-tuple of positive integers $(r_1, \ldots, r_m)$ as the least value of

$$\frac{i_1}{r_1} + \ldots + \frac{i_m}{r_m}$$

for which (3.2) holds, if $P \not\equiv 0$, and as $+\infty$ if $P \equiv 0$.

3.3. The steps (a), (b), (c) in the proof of Liouville's Theorem are now replaced by new steps (a), (b), (c).

(a) LEMMA 3A. *Suppose α is an algebraic integer of degree $d > 1$. Suppose $\varepsilon > 0$ and m is an integer with*

$$m > 8d^2 \varepsilon^{-2} . \tag{3.3}$$

Let $r_1, \ldots, r_m$ be positive integers. Then there is a polynomial $P(x_1, \ldots, x_m) \not\equiv 0$ with rational integer coefficients such that

(i) *P has degree at most r_h in x_h $(h = 1, \ldots, m)$.*

(ii) *P has index at least $\frac{m}{2}(1-\varepsilon)$ at $(\alpha, \ldots, \alpha)$ with respect to $(r_1, \ldots, r_m)$.*

(iii) *$H(P) \leqq B^{r_1 + \ldots + r_m}$ where $B = B(\alpha)$.*

Here $H(P)$ is the *height* of P, i.e. the maximum of the absolute values

of its coefficients. By virtue of (ii) the average of $\frac{i_h}{r_h}$ $(h=1, \ldots, m)$ when $p^{(i_1,\ldots,i_m)}(\alpha, \ldots, \alpha) \neq 0$ is at least $\frac{1}{2}(1-\varepsilon)$, which is rather better than $\frac{i}{r} = \frac{1}{d}$ we had in the proof of Liouville's Theorem.

To prove the lemma we put

$$P(x_1, \ldots, x_m) = \sum_{j_1=0}^{r_1} \cdots \sum_{j_m=0}^{r_m} C(j_1, \ldots, j_m)\, x_1^{j_1} \ldots x_m^{j_m}.$$

The $N = (r_1+1) \ldots (r_m+1)$ coefficients $C(j_1, \ldots, j_m)$ are unknown integers we have to determine such that (ii) and (iii) hold. The condition (ii) means that

$$p^{(i_1,\ldots,i_m)}(\alpha, \ldots, \alpha) = 0 \tag{3.4}$$

whenever

$$\frac{i_1}{r_1} + \ldots + \frac{i_m}{r_m} < \frac{m}{2}(1-\varepsilon). \tag{3.5}$$

Since (3.4) is always true if $i_h > r_h$ for some h, the number of non-trivial equations (3.4) is the number of points $\left(\frac{i_1}{r_1}, \ldots, \frac{i_m}{r_m}\right)$ in the unit cube $(0 \leqq \xi_1 \leqq 1, \ldots, 0 \leqq \xi_m \leqq 1)$ with (3.5). One can show that

$$\frac{\text{The number of points in the cube with (3.5)}}{N = \text{the total number of points in the cube}} \to 0 \tag{3.6}$$

as $m \to \infty$, independently of $r_1, \ldots, r_m$. This is just the law of large numbers in probability theory, since the "independent variables" $i_1/r_1, \ldots, i_m/r_m$ each have expectation value $\frac{1}{2}$. In fact an appeal to probability theory is not necessary and a simple combinatorial argument shows that the left hand side of (3.6) is at most $2^{1/2}m^{-1/2}\varepsilon^{-1}$, and hence by (3.3) is at most $1/(2d)$. Thus the number of non-trivial conditions (3.4) is at most $N/(2d)$. Each condition (3.4) is a homogeneous linear equation in the unknowns $C(j_1, \ldots, j_m)$ with coefficients in the field $\mathbf{Q}(\alpha)$. (I.e. the field obtained by adjoining α to the field $\mathbf{Q}$ of rationals). Hence each condition follows from d linear homogeneous equations whose coefficients are rational integers. Hence altogether our unknown integers $C(j_1, \ldots, j_m)$ have to satisfy

at most $N/2$ linear homogeneous equations with rational integer coefficients. But it is known that any system of linear homogeneous equations with rational integer coefficients where the number of equations is at most $\frac{1}{2}$ times the number of unknowns has a non-trivial solution in rational integers which are bounded in terms of the size of the coefficients. Carrying out all the estimates one sees that (iii) can be satisfied in addition to (ii).

3.4. We now turn to step

(b) It can be shown that if $\frac{p_1}{q_1}, \ldots, \frac{p_m}{p_m}$ are solutions of $\left|\alpha - \frac{p}{q}\right| < q^{-2-\delta}$ and if some further conditions are satisfied, then $P\left(\frac{p_1}{q_1}, \ldots, \frac{p_m}{q_m}\right) = 0$. In fact a little more is true:

LEMMA 3B. *Suppose* $\left|\alpha - \frac{p_h}{q_h}\right| < q_h^{-2-\delta}$ $(h = 1, 2, \ldots, m)$ *where* $0 < \delta < \frac{1}{12}$. *Suppose* $0 < \varepsilon < \delta/20$ *and suppose that* $q_h \geqq c_0(\alpha, \delta)$ $(h = 1, \ldots, m)$ *and*

$$(3.7) \qquad r_1 \log q_1 \leqq r_h \log q_h \leqq (1+\varepsilon) r_1 \log q_1 \qquad (h = 1, \ldots, m).$$

Now if all the conditions of Lemma 3A *are satisfied and if* P *is the polynomial of that lemma, then the index of* P *at* $\left(\frac{p_1}{q_1}, \ldots, \frac{p_m}{q_m}\right)$ *with respect to* $(r_1, \ldots, r_m)$ *is* $\geqq \varepsilon m$.

To prove this lemma we shall use Taylor's formula:

$$P\left(\frac{p_1}{q_1}, \ldots, \frac{p_m}{q_m}\right) = \sum_{i_1=0}^{r_1} \cdots \sum_{i_m=0}^{r_m} \left(\frac{p_1}{q_1} - \alpha\right)^{i_1} \cdots \left(\frac{p_m}{q_m} - \alpha\right)^{i_m} \frac{P^{(i_1, \ldots, i_m)}(\alpha, \ldots, \alpha)}{i_1! \ldots i_m!}.$$

By (ii) of Lemma 3A only terms with $\frac{i_1}{r_1} + \ldots + \frac{i_m}{r_m} \geqq \frac{m}{2}(1-\varepsilon)$ can be non-zero. For these terms we have

$$\left|\left(\frac{p_1}{q_1}-\alpha\right)^{i_1}\cdots\left(\frac{p_m}{q_m}-\alpha\right)^{i_m}\right| \leqq q_1^{-(2+\delta)i_1}\cdots q_m^{-(2+\delta)i_m}$$
$$= \left(q_1^{r_1(i_1/r_1)}\cdots q_m^{r_m(i_m/r_m)}\right)^{-2-\delta}$$
$$\leqq q_1^{-r_1(2+\delta)((i_1/r_1)+\ldots+(i_m/r_m))} \leqq q_1^{-r_1(2+\delta)\frac{1}{2}m(1-\varepsilon)}$$
$$< \left(q_1^{r_1}\cdots q_m^{r_m}\right)^{-\frac{1}{2}(2+\delta)(1-\varepsilon)/(1+\varepsilon)} < \left(q_1^{r_1}\cdots q_m^{r_m}\right)^{-(1+(\delta/4))}$$

by (3.7) and since $0 < \varepsilon < \delta/20$. Using this estimate as well as part (iii) of Lemma 3A it is not hard to show that

$$\left|P\left(\frac{p_1}{q_1},\ldots,\frac{p_m}{q_m}\right)\right| < \left(q_1^{r_1}\cdots q_m^{r_m}\right)^{-1}$$

if $q_h \geqq c_0(\alpha, \delta)$ $(h=1, \ldots, m)$, hence that

$$P\left(\frac{p_1}{q_1},\ldots,\frac{p_m}{q_m}\right) = 0\,.$$

Thus the index of P at $\left(\frac{p_1}{q_1},\ldots,\frac{p_m}{q_m}\right)$ is positive, and a slight extension of the argument shows that the index with respect to $(r_1, \ldots, r_m)$ is at least εm.

3.5. Finally we turn to step

(c) If one could show that $P\left(\frac{p_1}{q_1},\ldots,\frac{p_m}{q_m}\right) \neq 0$, then this would contradict Lemma 3B, and this contradiction would show that the inequality (2.2) has only finitely many solutions. But to show that $P\left(\frac{p_1}{q_1},\ldots,\frac{p_m}{q_m}\right) \neq 0$ was very easy for $m = 1$ and it is rather difficult when $m > 1$. To get a contradiction to Lemma 3B it will suffice to show that the index of P at $\left(\frac{p_1}{q_1},\ldots,\frac{p_m}{q_m}\right)$ with respect to $(r_1, \ldots, r_m)$ is less than εm. When $m = 2$ the situation is a little simpler than in the general case. Siegel (1921a) devised an algebraic argument to deal with this case, and Schneider (1936) devised a more general arithmetical argument. The latter argument was considerably sharpened by Roth. The following lemma of Roth is called Roth's Lemma.

LEMMA 3C. *Suppose* $0 < \varepsilon < 1/12$ *and let* m *be a positive integer. Put* $\omega = 24\cdot 2^{-m}(\varepsilon/12)^{2m-1}$. *Let* $r_1, \ldots, r_m$ *be positive integers with*

$$(3.8) \qquad \omega r_h \geqq r_{h+1} \qquad (h=1, \ldots, m-1).$$

Let $\frac{p_1}{q_1}, \ldots, \frac{p_m}{q_m}$ *be rationals in their lowest terms with positive denominators and with* $q_h^{\omega} \geqq 2^{3m}$ *and* $q_h^{r_h} \geqq q_1^{r_1}$. *Further let* $P(x_1, \ldots, x_m)$ *be a polynomial with rational integer coefficients, not identically zero, of degree* $\leqq r_h$ *in* x_h $(h=1, \ldots, m)$ *and with* $H(P) \leqq q_1^{\omega r_1}$. *Then the index of* P *at* $\left(\frac{p_1}{q_1}, \ldots, \frac{p_m}{q_m}\right)$ *with respect to* $(r_1, \ldots, r_m)$ *is* $\leqq \varepsilon$.

The proof of this lemma is ingenious and complicated and will not be given here. It uses " generalized Wronskians ", i.e. determinants whose entries are mixed partial derivatives of certain polynomials. Some condition like (3.8) is necessary, for otherwise if $m = 2$, say, the polynomial $P(x_1, x_2) = (x_1 - x_2)^r$ would have an index as large as 1 at every point (ξ, ξ).

The lemma is proved by induction on m. Only the case $m = 1$ is simple and will be proved here. Suppose $P(x)$ has a zero of order l at p/q. Then

$$P(x) = (qx - p)^l R(x)$$

where $R(x)$ has rational integer coefficients by Gauss' Lemma. We have

$$q^l \leqq H(P) \leqq q^{\omega r_1} = q^{\varepsilon r_1}$$

(since $\omega = \varepsilon$ when $m=1$), whence $l/r_1 \leqq \varepsilon$. But l/r_1 is the index of P at $\frac{p}{q}$ with respect to (r_1).

Now if there are infinitely many rationals $\frac{p}{q}$ with (2.2), then both Lemma 3B and Lemma 3C can be satisfied. (One picks $0 < \delta < 1/12$, then $0 < \varepsilon < \delta/20$, then m with (3.3), then rationals $\frac{p_1}{q_1}, \ldots, \frac{p_m}{q_m}$ with (2.2) and with rapidly increasing denominators, and finally one picks $r_1, \ldots, r_m$). These two lemmas together give the desired contradiction, and Roth's Theorem follows.

3.6. The reason why this proof is non-effective is that one needs m very good approximations to α rather than just one, in order to get the desired contradiction. For Thue's and Siegel's Theorems one needs two such approximations. In fact Davenport (1968) found a function $\kappa_0(d)$

defined for $d = 3, 4, \ldots$ with $\kappa_0(3) = 1 + \sqrt{3}$ and with $|\kappa_0(d) - \frac{1}{2}d| \leqq c_1$ such that for an algebraic number α of degree d and for $\kappa > \kappa_0(d)$ there is a computable constant $c_2 = c_2(\alpha, \kappa)$ such that the inequality $\left|\alpha - \frac{p}{q}\right| < q^{-\kappa}$ has at most one solution p/q with $(p, q) = 1$ and $q > c_2$. But earlier Schinzel (1967) had pointed out that this is true with $\kappa > 3\sqrt{d/2}$ in place of $\kappa > \kappa_0(d)$. If d is large, then $\kappa_0(d) > 3\sqrt{d/2}$, and hence in this case Schinzel's result is better than Davenport's. Earlier Siegel (1937) and Hyyrö (1964) had shown results of this kind for numbers α of the type $\alpha = \sqrt[d]{(a/b)}$, or rather for the corresponding Thue's equation $ax^d - by^d = m$.

4. Some generalizations of Roth's Theorem

4.1. In this section we shall discuss several generalizations of Roth's Theorem, but not the generalization to simultaneous approximation, which will be taken up in §7.

The *height* $H(\beta)$ of an algebraic number β is defined by $H(\beta) = H(P)$, where P is the defining polynomial of β. Roth (1955b) enunciated and LeVeque ((1955), vol. 2, ch. 4) proved

Theorem 4A. *Let α be algebraic, K an algebraic number field and $\delta > 0$. There are only finitely many elements β of K with*

$$|\alpha - \beta| < H(\beta)^{-2-\delta}. \tag{4.1}$$

Neither α nor K need to be real here. When K is the field $\mathbf{Q}$ of rationals, then Theorem 4A reduces to Roth's Theorem. Since every number field contains $\mathbf{Q}$ as a subfield, it follows from Dirichlet's Theorem that the number 2 in the exponent in (4.1) is best possible if α is real.

In fact if α and K are real, then the exponent is best possible in a somewhat less trivial sense: Suppose K is a real or complex number field of degree t and β in K has degree d. Then d is a divisor of t. Define the *field height* $H_K(\beta)$ of β by $H_K(\beta) = H(P^{t/d})$, where P is the defining polynomial of β. One can show (LeVeque (1955), vol. 2, ch. 4.2) that $c_1(K)\, H(\beta)^{t/d} \leqq H_K(\beta) \leqq c_2(K)\, H(\beta)^{t/d}$, and hence Theorem 4A remains true a fortiori if (4.1) is replaced by

$$|\alpha - \beta| < H_K(\beta)^{-2-\delta}. \tag{4.2}$$

One can show that if both α and K are real and if $\alpha \notin K$, then there are infinitely many β in K with

$$(4.3) \qquad |\alpha - \beta| < c_3(K)\, H_K(\beta)^{-2}.$$

Thus the exponent on the right hand side of (4.2) is best possible. Now if α is algebraic, then by Theorem 4A and since $H_K(\beta) \geqq c_1(K)\, H(\beta)^{t/d}$, the inequality (4.3) can hold for only finitely many elements β of K of degree $d < t$. Hence if α is a real algebraic number and if the field K is real, then there are infinitely many primitive elements β of K (i.e. elements of K of degree t) with (4.3), i.e. with $|\alpha - \beta| < c_3(K)\, H(\beta)^{-2}$. Hence Theorem 4A remains best possible if one restricts oneself to primitive elements β of K.

If α is a real or complex number which does not lie in a complex number field K, then there are infinitely many β in K with

$$(4.4) \qquad |\alpha - \beta| < c_4(K)\, H_K(\beta)^{-1}.$$

We shall see in §4.4 that the exponent in (4.2) may be improved to $-1-\delta$ in this case.

4.2. The field K of Theorem 4A can be enlarged to contain α, and hence there is no loss of generality in this theorem if one assumes that $\alpha \in K$. One could try to give a lower bound for $|\alpha - \beta|$ where both α, β vary in K. In fact I can show (unpublished) that there is a number $c_1(d, \delta)$ defined for $d = 1, 2, \ldots$ and for $\delta > 0$ such that in every number field K of degree d there are only finitely many pairs of elements α, β with

$$H(\beta) > H(\alpha)^{c_1} \quad \textit{and} \quad |\alpha - \beta| < H(\beta)^{-2-\delta}.$$

For example, if $0 < \delta < 1$, one may put $c_1 = e^{c_2}$ with $c_2 = d^{10^4\delta^{-2}}$. This implies the existence of a (non-effective) constant $c_3 = c_3(K, \delta)$ such that $|\alpha - \beta| > (H(\alpha)\, H(\beta))^{-2-\delta}$ if either $H(\beta) > H(\alpha)^{c_3}$ or if $H(\alpha) > H(\beta)^{c_3}$. It is conceivable that there is a $c_4 = c_4(K, \delta) > 0$ such that

$$|\alpha - \beta| > c_4 (H(\alpha)\, H(\beta))^{-2-\delta}$$

for any two distinct elements α, β of K.

S. Schanuel (oral communication) also has a version of Theorem 4A in which both α and β are allowed to vary. It should be remarked that the inequalities of this subsection, when both α and β lie in K, would become quite trivial if we had substituted field heights for heights. In fact it is easy to see that

$$|\alpha - \beta| > c_5(K)\left(H_K(\alpha)\, H_K(\beta)\right)^{-1}$$

if α, β are distinct elements of K.

4.3. A rather different question is that of approximation to an algebraic number α by algebraic numbers β of fixed degree d. Siegel (1921a) already had given some estimates, and using the method of Roth, Ramachandra (1966) had improved these estimates. Wirsing was the first to prove (but published only in (1971)) a result in which the exponent depends on d only, namely the following.

THEOREM 4B. *Suppose α is a real or complex algebraic number and suppose $d \geqq 1$, $\delta > 0$. There are only finitely many (real or complex) algebraic numbers β of degree d with*

$$|\alpha - \beta| < H(\beta)^{-2d-\delta} \tag{4.5}$$

Wirsing's Theorem becomes Roth's Theorem when $d = 1$. As we shall see in §7.5, the exponent $-2d - \delta$ in (4.5) may be replaced by $-d - 1 - \delta$. Nevertheless we shall now discuss the interesting idea underlying Wirsing's proof.

If one attempts to generalize Roth's method to prove Theorem 4B, a difficulty arises in part (b). One has to show that

$$P(\beta_1, \ldots, \beta_m) = 0$$

where P is a polynomial with rational integer coefficients constructed in part (a), and where $\beta_1, \ldots, \beta_m$ are certain algebraic numbers of degree d satisfying (4.5). In general the degree of the field

$$\mathbf{Q}(\beta_1, \ldots, \beta_m)$$

generated by $\beta_1, \ldots, \beta_m$ may be as large as d^m.

Suppose now that this is the case. The number

$$(b_1 \ldots b_m)\, d^{m-1}\, \mathcal{N}\left(P(\beta_1, \ldots, \beta_m)\right), \tag{4.6}$$

where $b_1, \ldots, b_m$ are the leading coefficients of the defining polynomials of $\beta_1, \ldots, \beta_m$ and where $\mathcal{N}$ denotes the norm of $\mathbf{Q}(\beta_1, \ldots, \beta_m)$ over $\mathbf{Q}$, is rational. The conjugates of $P(\beta_1, \ldots, \beta_m)$ are $P(\beta_1^{(i_1)}, \ldots, \beta_m^{(i_m)})$ where $1 \leqq i_h \leqq d$ $(h = 1, \ldots, m)$ and where $\beta_h = \beta_h^{(1)}, \beta_h^{(2)}, \ldots, \beta_h^{(m)}$ are the distinct conjugates of β_h. Since each number β_h and each of its conjugates occurs at most to the power d^{m-1}, and since $b_h \beta_h^{(i_1)} \ldots \beta_h^{(i_t)}$ is an algebraic integer if $i_1, \ldots, i_t$ are

distinct (see Schneider (1957), Hilfssatz 17 or LeVeque (1955), vol. 2, p. 64), the number (4.6) is a rational integer. One can estimate a factor $P(\beta_1^{(i_1)}, \ldots, \beta_m^{(i_m)})$ of (4.6) well only if many of the superscripts $i_1, \ldots, i_m$ equal 1. Wirsing in his proof used the fact that for most of the d^m factors, about $\frac{m}{d}$ of these superscripts equal 1. This enabled him to show that under suitable conditions the number (4.6) has absolute value less than 1, and hence is zero.

4.4. Let K be a number field of degree t and let $\alpha_1, \ldots, \alpha_t$ be arbitrary real or complex algebraic numbers. For $\beta \in K$, Mahler puts

$$f(\beta) = \prod_{j=1}^{t} \min(1, |\alpha_j - \beta^{(j)}|),$$

where $\beta^{(1)}, \ldots, \beta^{(t)}$ are the conjugates of β corresponding to the conjugates $\omega^{(1)} = \omega, \ldots, \omega^{(t)}$ of a fixed generator (i.e. primitive element) ω of K.

THEOREM 4C (*Mahler* 1963).[1]) *Suppose $K, \alpha_1, \ldots, \alpha_t, f(\beta)$ are as above, and suppose that $\delta > 0$. There are only finitely many β in K with*

$$f(\beta) < H_K(\beta)^{-2-\delta}.$$

Since $f(\beta) \leqq |\alpha_1 - \beta|$, it is easy to see that Theorem 4C sharpens Theorem 4A. Suppose now that K is complex and that for every β of K, $\beta^{(2)}$ is the complex conjugate $\bar{\beta}$ of β. Also suppose that $\alpha_2 = \bar{\alpha}_1$. By Theorem 4C there are only finitely many β in K with

$$|\alpha_1 - \beta^{(1)}|\,|\alpha_2 - \beta^{(2)}| < H_K(\beta)^{-2-\delta},$$

i.e. with

$$|\alpha_1 - \beta| < H_K(\beta)^{-1-(\delta/2)}.$$

Hence if K is complex, then the exponent $-2-\delta$ in (4.2) may be replaced by $-1-\delta$. It follows that in general if α is a complex (non-real) algebraic number, then the exponent $-2-\delta$ in (4.1) may be replaced by $-1-\delta$. By (4.4) the exponent $-1-\delta$ is best possible in this case.

Suppose β is an element of K of degree d. If b_0 is the leading coefficient of the defining polynomial P of β, then $c_0 = b_0^{t/d}$ is the leading coefficient

[1]) See also Mahler (1961), Appendix C, Assertion (2.II).

of the polynomial $P^{t/d}$ used in the definition of $H_K(\beta)$. For every σ in $0 \leqq \sigma \leqq 1$, let $\mathscr{C}(\sigma)$ be the class of β in K with

$$|c_0| \leqq H_K(\beta)^\sigma .$$

Mahler (1963) proved a result which contains Theorem 4C, namely

THEOREM 4D. *Suppose K, $\alpha_1, \ldots, \alpha_t, f(\beta)$, δ are as above, and suppose $0 \leqq \sigma \leqq 1$. There are only finitely many β in $\mathscr{C}(\sigma)$ with*

$$f(\beta) < H_K(\beta)^{-1-\sigma-\delta} .$$

Of particular interest is the case when $\sigma = 0$, because $\mathscr{C}(0)$ consists precisely of the integers of K. Therefore given an algebraic number α, there are only finitely many integers β of K with $|\alpha - \beta| < H_K(\beta)^{-1-\delta}$, and by applying this to K and its subfields it follows that there are only finitely many integers β of K with

$$|\alpha - \beta| < H(\beta)^{-1-\delta} . \tag{4.7}$$

The exponent $-1-\delta$ in (4.7) may be replaced by $-\frac{1}{2}-\delta$ if α is complex.

Mahler also proved some "inhomogeneous" theorems, which are contained in more recent and more general results to be stated in §7.4.

4.5. The first one to recognize the importance of p-adic diophantine approximations was K. Mahler. He developed an extensive theory and in particular he proved (1933a, b) the following

THEOREM 4E (Mahler (1933a). *Suppose $F(x, y)$ is a binary form as in Theorem 2C and suppose $p_1, \ldots, p_r$ are distinct rational primes. There are only finitely many rational integers $x, y, z_1, \ldots, z_r$ with*

$$F(x, y) = p_1^{z_1} \ldots p_r^{z_r} .$$

Mahler (1933c) gave an asymptotic formula for the number $N(m)$ of solutions of

$$|F(x, y)|\,|F(x, y)|_{p_1} \ldots |F(x, y)|_{p_r} \leqq m$$

where $|\ |_p$ denotes the p-adic valuation. His results were generalized to algebraic number fields by Parry (1940, 1950). For a sharper version of Theorem 4E see Theorem 5E and the remarks below it. Using Roth's method, Ridout (1957) proved (but see also Schneider (1957), Satz 6) a result which can be formulated as follows.

THEOREM 4F. *Let α be a real algebraic number distinct from zero and let $p_1, \ldots, p_r, q_1, \ldots, q_r$ be distinct rational primes. Suppose $\delta > 0$. There are only finitely many rationals p/q* with

$$p = p_1^{a_1} \cdots p_r^{a_r} p' , \quad q = q_1^{b_1} \cdots q_s^{b_s} q'$$

where $a_1, \ldots, a_r, b_1, \ldots, b_s$ are non-negative integers and where p', q' are non-zero integers such that

$$\left| \alpha - \frac{p}{q} \right| < \frac{1}{|p'q'| \cdot |pq|^{\delta}} . \tag{4.8}$$

The case $p' = q' = 1$ of this theorem is due to Mahler (1936). In this paper Mahler also proved a weaker version of Theorem 4F, of the type of Schneider's Theorem mentioned in §2.4. For a rather better recent estimate in this case see Theorem 5B. Bounds for the number of solutions of inequalities of the type (4.8) were given by Fraenkel (1962). More general versions of Theorem 4F were given by Stepanov (1967) and Walliser (1969).

Now let α be a real algebraic irrational. Recall that the convergents p_n/q_n to α satisfy $|\alpha - p_n/q_n| < q_n^{-2}$. It follows from Mahler's (1936) result that the greatest prime factor of $p_n q_n$ tends to infinity, and it follows from Theorem 4F that in fact the greatest prime factor of p_n as well as that of q_n tends to infinity.

Let $K = \mathbf{Q}(\omega)$ be a number field of degree t. We shall recall some well known facts about valuations of K. Suppose that $t = r + 2s$ and that $\omega^{(1)}, \ldots, \omega^{(r)}$ are real and $\omega^{(r+1)}, \ldots, \omega^{(r+s)}, \omega^{(r+s+1)}, \ldots, \omega^{(t)}$ are complex with $\omega^{(r+s+j)}$ the complex conjugate of $\omega^{(r+j)}$ $(j=1, \ldots, s)$. Let Ω be the set consisting of the integers $1, 2, \ldots, r + s$ and of the prime ideals of the ring of integers of K. If $v \in \Omega$, $1 \leqq v \leqq r + s$ and if $\alpha \in K$, then we put $|\alpha|_v = |\alpha^{(v)}|$ where $|\ |$ denotes the ordinary absolute value. Now suppose v is a prime ideal $\mathfrak{p}$. The norm $\mathcal{N}(\mathfrak{p})$ equals $p^{N_{\mathfrak{p}}}$ where p is a rational prime and where $N_{\mathfrak{p}}$ is a positive rational integer. If $\alpha \in K$, $\alpha \neq 0$, then the fractional ideal (α) may uniquely be written $(\alpha) = \mathfrak{p}^a \mathfrak{p}_2^{a_2} \ldots \mathfrak{p}_k^{a_k}$ where $a, a_2, \ldots, a_k$ are rational integers and where $\mathfrak{p}, \mathfrak{p}_2, \ldots, \mathfrak{p}_k$ are distinct prime ideals. We now put $|\alpha|_v = p^{-a}$ and we put $|0|_v = 0$. It is clear that $|\alpha\beta|_v = |\alpha|_v |\beta|_v$ and that if $\alpha \neq 0$, then $|\alpha|_v = 1$ for all but finitely many $v \in \Omega$. The mappings $\alpha \to |\alpha|_v$ where $v \in \Omega$ are all the inequivalent valuations of K, and the Archimedean valuations are those where $v = 1, 2, \ldots, r + s$. For every $v \in \Omega$, there is a completion K_v of K with respect to $|\ |_v$.

Put $N_v = 1$ or $N_v = 2$ if $1 \leqq v \leqq r$ or if $r + 1 \leqq v \leqq r + s$, respectively; N_v was defined above when v is a prime ideal. Put $\| \alpha \|_v = | \alpha |_v^{N_v}$. It is clear that the definition of $| \alpha |_v$ and of $\| \alpha \|_v$ can be extended to $\alpha \in K_v$. The product formula

$$\prod_{v \in \Omega} \| \alpha \|_v = 1$$

holds for every non-zero α in K. One can show that

$$c_1 (K) H_K (\beta) \leqq \prod_{v \in \Omega} \max (1, \| \beta \|_v) \leqq c_2 (K) H_K (\beta) . \tag{4.9}$$

THEOREM 4G. *Let S be a finite subset of Ω. For each $v \in S$, let α_v be an element of K_v which is algebraic over K. For $\beta \in K$ put*

$$g (\beta) = \prod_{v \in S} \min (1, \| \alpha_v - \beta \|_v) .$$

Then for every $\delta > 0$ there are only finitely many $\beta \in K$ with

$$g (\beta) < H_K (\beta)^{-2-\delta} .$$

A more general version of this theorem may be found in Mahler ((1961), Appendix C, Assertion (2,I)). See also Lang ((1962), ch. 6). In its present form Theorem 4G does not contain Theorem 4F, but Mahler's generalization of it does. The case of Theorem 4G when K is the field of rationals is due to Ridout (1958). It may be seen that Mahler's Theorem 4C is equivalent with the case of Theorem 4G when $S = \{1, 2, ..., r + s\}$, i.e. when we are considering only Archimedean valuations. Lang and Ridout also gave p-adic versions of Theorem 2C and thus sharpened Theorem 4E. Like Roth's Theorem, the results discussed here do not permit to give an estimate for the " size " (say $H (\beta)$) of the solutions, and hence they are non-effective. For weaker but effective p-adic results see Theorem 5B, 5D and 5E. For an effective weaker version of Theorem 4G see Sprindžuk (1970b, 1971a).

5. EFFECTIVE METHODS. BAKER'S THEOREM

5.1. All the results obtained by the method of Thue, Siegel and Roth share the disadvantage that they are non-effective. Although they show that certain inequalities and equations have only finitely many integer solutions, they do not give bounds for the size of the solutions and hence give no method to compute all the solutions.

Effective bounds, which however do not imply Roth's Theorem, and which do not imply Thue's or Siegel's Theorem unless α is of a special type, were given by Baker. He first used hypergeometric series (see also Siegel (1937)) to deal with algebraic numbers of the type $\alpha = \sqrt[d]{a/b}$. He showed (1964a) that if $\kappa > 2$, $d \geqq 3$ and if a, b are integers with $b > 0$, $a > (a-b)^{c_1} c_2$ where $c_i = c_i(\kappa, d)$ $(i=1, 2)$, then all rational numbers p/q with $q > 0$ satisfy

$$\left| \sqrt[d]{\frac{a}{b}} - \frac{p}{q} \right| > c_3 q^{-\kappa}$$

where $c_3 = c_3(d, \kappa, a, b)$. The constants c_1, c_2, c_3 are computable here. In another paper (1964b), Baker proved among other results that

$$\left| \sqrt[3]{2} - \frac{p}{q} \right| > \frac{10^{-6}}{q^{2.955}}.$$

These results of Baker improve the exponent in Liouville's Theorem for certain algebraic numbers. Using his estimates of linear forms whose coefficients are logarithms of algebraic numbers, Baker also proved a result which holds for *all* real algebraic numbers and which improves Liouville's Theorem by a factor which is smaller than any positive power of q:

THEOREM 5A (Baker 1968b). *Suppose α is a real algebraic number of degree $d \geqq 3$ and suppose $\kappa > d$. Then there is a computable $c_4 = c_4(\alpha, \kappa) > 0$ such that*

$$\left| \alpha - \frac{p}{q} \right| > c_4 e^{(\log q)^{1/\kappa}} q^{-d}$$

for every rational $\frac{p}{q}$ with $q > 0$.

Hence if $f(q)$ is of smaller order of magnitude than $e^{(\log q)^{1/\kappa}}$ for some $\kappa > d$, say if $f(q) \leqq e^{(\log q)^{1/(d+\delta)}}$ where $\delta > 0$, then the solutions $\frac{p}{q}$ of $\left| \alpha - \frac{p}{q} \right| < f(q) q^{-d}$ must have $q \leqq q_1 = q_1(\alpha, \delta)$ where q_1 is computable. Recently Baker and Stark (to appear) could replace $\kappa > d$ by the milder condition $\kappa > 1$.

Feldman (1968a, 1968b) proved a result which contains the following.

THEOREM 5B. *Suppose α is an irrational algebraic number and let $p_1, \ldots, p_r$ be distinct rational primes. Then for all rationals $\frac{p}{q}$ with $p \geqq 3$, $q \geqq 3$, $(p, q) = 1$, of the type*

$$\frac{p}{q} = p_1^{a_1} \ldots p_r^{a_r}$$

with rational integers $a_1, \ldots, a_r$, one has

$$\left| \alpha - \frac{p}{q} \right| > \frac{1}{(\log q)^{c_5}},$$

where c_5 is an effectively computable constant depending on $\alpha, p_1, \ldots, p_r$.

This sharpens the case $p' = q' = 1$ of Theorem 4F. To prove Theorem 5B, Feldman used the method of Baker and refined it in the special case needed here. Baker's (1966, 1967b) papers would have yielded $|\alpha - (p/q)| > c_6 e^{-(\log \log q)^{\kappa}}$.

5.2. As for Thue's equation, the following effective theorem holds.

THEOREM 5C (Baker 1968b, 1968c). *Suppose the form $F(x, y)$ in Thue's equation*

$$F(x, y) = m \tag{5.1}$$

is of degree $d \geqq 3$, it has rational integer coefficients and is irreducible over the rationals. Then every integer solution (x, y) of this equation satisfies

$$\max(|x|, |y|) < \exp\left((dH)^{(10d)^5} + (\log m)^{2d+2}\right),$$

where H is the height of F.

Baker also gave explicit bounds for the solutions of elliptic and hyperelliptic equations (1968c, 1968d, 1969) and Baker and Coates (1970) did the same for equations which define curves of genus 1. Vinogradov and Sprindžuk (1968), Coates (1969, 1970a, 1970b) and Sprindzuk (1970b) used Baker's method to prove effective p-adic theorems.

Sprindžuk (1969, 1970a) used a p-adic method for estimating the size of the integer solutions $x, y, z_1, \ldots, z_r$ of the equation

$$F(x, y) = m p_1^{z_1} \ldots p_r^{z_r} \qquad (x, y) = 1, z_1 \geqq 0, \ldots, z_r \geqq 0 \tag{5.2}$$

where $F(x, y)$ is an irreducible form of degree $d \geqq 3$ and where $p_1, \ldots, p_r$ are rational primes. In (1970b) he improved these results. He defined *exceptional* forms $F(x, y)$ and showed that if $F(x, y)$ is not exceptional then $\max(|x|, |y|) < c_1 \exp(\log |m|)^{\kappa}$ where $\kappa > 2$ and $c_1 = c_1(F, \kappa, p_1, \ldots, p_r)$ is an effective constant independent of m. He further improved his results in (1971a). He showed that there are no exceptional forms of degree $d \geqq 5$ and he proved the following:

THEOREM 5D. *Suppose $F(x, y)$ is not an exceptional form. Then all the integer solutions of* (5.2) *satisfy*

$$\max(|x|, |y|) < c_2 \mid m \mid^{(\log \log |m|)^{4(d+r+1)}}. \tag{5.3}$$

Here $c_2 = c_2(F, p_1, \ldots, p_r)$ is effective.

A full account of this work is given by Sprindžuk (to appear).

Baker (to appear) further improves this estimate for the more special equation $F(x, y) = m$ but for all irreducible forms F of degree $d \geqq 3$ without exception, and derives the estimate

$$\max(|x|, |y|) < c_3 \mid m \mid^{\log \log |m|}.$$

It is almost certain that this estimate can be extended to the more general equation (5.2). On the other hand Sprindžuk at the end of his (1971a) paper indicates that his method can be used to replace (5.3) by the still sharper inequality

$$\max(|x|, |y|) < c_4 \mid m \mid^{c_5}.$$

THEOREM 5E (Sprindžuk 1971b). *Suppose the binary form $F(x, y)$ of degree $d \geqq 3$ is not exceptional. Let x, y be coprime integers with $X = \max(|x|, |y|) > 10$. Then the greatest prime factor of $F(x, y)$ is*

$$> c_6 \log \log X / \log \log \log X \tag{5.4}$$

where $c_6 = c_6(F)$ is effectively computable.

Earlier Coates (1970a) had given the lower bound $c_7 (\log \log X)^{1/4}$ which holds for all irreducible forms of degree $d \geqq 3$. Probably it is possible to generalize Baker's paper (to appear) to the p-adic case, and then to prove the estimate of Theorem 5E for all irreducible forms of degree $d \geqq 3$. Mahler's Theorem 4E had said that the greatest prime factor of $F(x, y)$ tends to infinity as $X \to \infty$.

Now suppose $F(x)$ is a polynomial in one variable x with rational integer coefficients, of degree $d \geqq 2$ and with distinct roots. Keates (1969) shows that the greatest prime factor $p(x)$ of $F(x)$ is $> c_8 \log\log|x|$ if $F(x)$ is of some special type, e.g. if $F(x)$ is of degree $d = 2$ or 3. Combining Keates' argument with recent papers of Baker and Sprindžuk it might be possible to show that $p(x) > c_8 \log\log|x|$ in general. It is interesting that this bound is only slightly better than the inequality (5.4) for forms in two variables. Further references on $p(x)$ are given in Keates (1969).

For an effective version of Theorem 4G see Sprindžuk (1970b, 1971a).

5.3. Baker derived Theorem 5A from his deep lower bounds for expressions of the type

$$|\beta_1 \log \alpha_1 + \ldots + \beta_n \log \alpha_n| \tag{5.5}$$

where $\alpha_1, \ldots, \alpha_n, \beta_1, \ldots, \beta_n$ are non-zero algebraic numbers such that $\log \alpha_1, \ldots, \log \alpha_n$ are linearly independent over the rationals, which he developed in (1966, 1967b, 1967c, 1968a).

Namely, suppose $\left|\alpha - \frac{p}{q}\right|$ is small and put $\beta = q\alpha - p$. Let $\mathbf{Q}(\alpha)$ be the field obtained by adjoining α to the field $\mathbf{Q}$ of rationals, let $\alpha^{(1)} = \alpha, \alpha^{(2)}, \ldots, \alpha^{(d)}$ be the conjugates of α and for $\omega \in \mathbf{Q}$ let $\omega^{(1)}, \ldots, \omega^{(d)}$ be the conjugates of ω corresponding to $\alpha^{(1)}, \ldots, \alpha^{(d)}$. We have

$$(\alpha^{(j)} - \alpha^{(k)})\beta^{(l)} + (\alpha^{(k)} - \alpha^{(l)})\beta^{(j)} + (\alpha^{(l)} - \alpha^{(j)})\beta^{(k)} = 0$$

for any integers j, k, l with $1 \leqq j, k, l \leqq d$. Let γ be an associate of β, of the type

$$\gamma = \beta\, \eta_1^{b_1} \ldots \eta_r^{b_r}$$

where $\eta_1, \ldots, \eta_r$ is a fixed set of fundamental units of $\mathbf{Q}(\alpha)$ and where $b_1, \ldots, b_r$ are rational integers. We have

$$\frac{(\alpha^{(k)} - \alpha^{(l)})\beta^{(j)}}{(\alpha^{(j)} - \alpha^{(l)})\beta^{(k)}} - 1 = \frac{(\alpha^{(k)} - \alpha^{(j)})\beta^{(l)}}{(\alpha^{(j)} - \alpha^{(l)})\beta^{(k)}},$$

whence

$$\alpha_1^{b_1} \ldots \alpha_r^{b_r} \alpha_{r+1}^{-1} - 1 = \sigma \tag{5.6}$$

where

$$\alpha_s = \eta_s^{(k)}/\eta_s^{(j)}\ (1 \leqq s \leqq r), \quad \alpha_{r+1} = \frac{(\alpha^{(j)} - \alpha^{(l)})\gamma^{(k)}}{(\alpha^{(k)} - \alpha^{(l)})\gamma^{(j)}}$$

and

$$\sigma = \frac{(\alpha^{(k)} - \alpha^{(j)})\,\beta^{(l)}}{(\alpha^{(j)} - \alpha^{(l)})\,\beta^{(k)}}.$$

Now $|\beta| = |\beta^{(1)}|$ is small by hypothesis, and it is clear that there is a conjugate $\beta^{(k)}$ with $|\beta^{(k)}| \geqq c_0(\alpha)|q|$. We now put $l = 1$ and pick j distinct from k, l. The quotient $|\beta^{(l)}/\beta^{(k)}|$ and hence $|\sigma|$ is then small. Therefore the left hand side of (5.6) will be small and

$$|b_1 \log \alpha_1 + \ldots + b_r \log \alpha_r - \log \alpha_{r+1} - k\pi i|$$

will be small for some integer k. Since $\pi i = \log(-1)$, this expression is of the type (5.5). One can choose the associate γ of β such that all the quotients $|\gamma^{(k)}/\gamma^{(j)}|$ $(1 \leqq k, j \leqq d)$ are bounded independently of p, q, and hence α_{r+1} as well as $\alpha_1, \ldots, \alpha_r$ and their conjugates are bounded. Substituting explicit values for the estimates and using his lower bounds for (5.5), Baker obtains a contradiction if $\beta = q\alpha - p$ is too small, and thereby he proves Theorem 5A.

A more quantitative discussion of this argument as it applies in the proof of Theorem 5C is given by Baker (1971). There is an anticipation of the argument at the end of Gelfond's (1952) book. Gelfond dealt with certain cubic Thue equations $F(x, y) = 1$ and pointed out that a lower bound for (5.5) (which then was not known) would provide upper bounds for the size of solutions of these equations.

6. Simultaneous approximation to real numbers by rationals

6.1. In this section we shall provide the background for the more special problem of simultaneous approximation to real algebraic numbers, which will be discussed in §7. Using the same general principles that were used in the proof of Theorem 1A and its corollary, Dirichlet (1842) proved the following two theorems and their corollaries.

THEOREM 6A. *Let $\alpha_1, \ldots, \alpha_l$ be real numbers and suppose Q is an integer greater than* 1. *Then there exist integers $q, p_1, \ldots, p_l$ with*

$$(6.1) \qquad 1 \leqq q < Q^l \quad \textit{and} \quad |\alpha_i q - p_i| \leqq Q^{-1} \qquad (i = 1, \ldots, l).$$

COROLLARY 6B. *Suppose at least one of* $\alpha_1, \ldots, \alpha_l$ *is irrational. Then there are infinitely many rational l-tuples* $(p_1/q, \ldots, p_l/q)$ *with* $q > 0$ *and* $(q, p_1, \ldots, p_l)$ [1] $= 1$ *and such that*

$$\left| \alpha_i - \frac{p_i}{q} \right| < q^{-1-(1/l)} \qquad (i = 1, \ldots, l). \tag{6.2}$$

The restriction in Theorem 6A that Q is an integer can be removed by using a slightly different proof. Essentially the theorem says that if $\alpha_0, \alpha_1, \ldots, \alpha_l$ are real numbers, not all zero, then there exist non-zero integer $(l+1)$-tuples $(q, p_1, \ldots, p_l)$ which are fairly proportional to $(\alpha_0, \alpha_1, \ldots, \alpha_l)$. Put differently, it says that if $(\alpha_0, \alpha_1, \ldots, \alpha_l)$ is a non-zero vector in $(l+1)$-dimensional space, then there are non-zero integer vectors in that space whose direction is fairly close to that of $(\alpha_0, \alpha_1, \ldots, \alpha_l)$.

THEOREM 6C. *Suppose* $\alpha_1, \ldots, \alpha_l$ *and* Q *are as in Theorem* 6A. *Then there exist integers* $q_1, \ldots, q_l, p$ *with*

$$\begin{aligned} &1 \leqq \max(|q_1|, \ldots, |q_l|) < Q^{1/l} \quad \textit{and} \\ &|\alpha_1 q_1 + \ldots + \alpha_l q_l + p| \leqq Q^{-1}. \end{aligned} \tag{6.3}$$

COROLLARY 6D. *Suppose* $1, \alpha_1, \ldots, \alpha_l$ *are linearly independent over the rationals. Then there are infinitely many* $(l+1)$*-tuples of coprime integers* $q_1, \ldots, q_l, p$ *with* $q = \max(|q_1|, \ldots, |q_l|) > 0$ *and with*

$$|\alpha_1 q_1 + \ldots + \alpha_l q_l + p| < q^{-l}. \tag{6.4}$$

Again the restriction in Theorem 6C that Q is an integer can be removed. A geometric interpretation is that if we have a hyperplane in $(l+1)$-dimensional space defined by an equation $\alpha_0 x_0 + \alpha_1 x_1 + \ldots + \alpha_l x_l = 0$, then there are integer points $(p, q_1, \ldots, q_l)$ which almost satisfy this equation and which therefore in some sense are fairly close to the hyperplane.

6.2. Let us say that $(\alpha_1, \ldots, \alpha_l)$ is *badly approximable of the first type* if Corollary 6B cannot be improved by an arbitrary factor, i.e. if there is a constant $c = c(\alpha_1, \ldots, \alpha_l) > 0$ such that

$$\max\left(\left| \alpha_1 - \frac{p_1}{q} \right|, \ldots, \left| \alpha_l - \frac{p_l}{q} \right| \right) > c q^{-1-(1/l)}$$

[1]) I.e. the greatest common divisor of $q, p_1, \ldots, p_l$.

for every rational l-tuple $(p_1/q, \ldots, p_l/q)$. Let us say that $(\alpha_1, \ldots, \alpha_l)$ is *very well approximable of the first type* if the exponent in Corollary 6B can be improved, i.e. if there is a $\delta = \delta(\alpha_1, \ldots, \alpha_l) > 0$ such that the system of inequalities

$$\left| \alpha_i - \frac{p_i}{q} \right| < q^{-1-(1/l)-\delta} \qquad (i = 1, \ldots, l)$$

has infinitely many solutions $(p_1/q, \ldots, p_l/q)$. Similarly we shall say that $(\alpha_1, \ldots, \alpha_l)$ is *badly approximable of the second type* if there is a $c' = c'(\alpha_1, \ldots, \alpha_l) > 0$ such that

$$|\alpha_1 q_1 + \ldots + \alpha_l q_l + p| > c' q^{-l}$$

for any integers $q_1, \ldots, q_l, p$ with $q = \max(|q_1|, \ldots, |q_l|) > 0$, and that it is *very well approximable of the second type* if there is a $\delta' = \delta'(\alpha_1, \ldots, \alpha_l) > 0$ such that the inequality

$$|\alpha_1 q_1 + \ldots + \alpha_l q_l + p| < q^{-l-\delta'}$$

has infinitely many solutions.

Theorems 6A and 6C and their corollaries are usually considered *dual* to each other, and usually if one has a refinement of one of them one can prove a refinement of the other. In fact Khintchine (1925, 1926a) proved a *transference principle* which contains the following theorem as a special case.

THEOREM 6E. *An l-tuple $(\alpha_1, \ldots, \alpha_l)$ is badly approximable of the first type if and only if it is badly approximable of the second type. It is very well approximable of the first type precisely if it is very well approximable of the second type.*

The first of the four assertions of this theorem had earlier been proved by Perron (1921). In view of the theorem we may speak of *badly approximable* and of *very well approximable* l-tuples.

6.3. Now suppose that $\alpha_1, \ldots, \alpha_l$ are real algebraic numbers and that $1, \alpha_1, \ldots, \alpha_l$ is a basis of a number field K of degree $n = l + 1$. There is a rational integer $a > 0$ such that $a\alpha_1, \ldots, a\alpha_l$ are algebraic integers, and hence for any rational integers $q_1, \ldots, q_l, p$ which are not all zero, the norm

$$\mathcal{N}\left(a(\alpha_1 q_1 + \ldots + \alpha_l q_l + p)\right)$$

is a non-zero rational integer and hence has absolute value at least 1. The conjugate factors $\alpha_1^{(i)} q_1 + \ldots + \alpha_l^{(i)} q_l + p$ have absolute values $\leqq c_1 \max(|q_1|. \ldots, |q_l|, |p|)$, and this implies that

$$|\alpha_1 q_1 + \ldots + \alpha_l q_l + p| \geqq c_2 (\max(|q_1|, \ldots, |q_l|, |p|))^{-l}.$$

Now if $|\alpha_1 q_1 + \ldots + \alpha_l q_l + p|$ is small, then $q = \max(|q_1|, \ldots, |q_l|) \geqq c_3 \max(|q_1|, \ldots, |q_l|, |p|)$, and we get $|\alpha_1 q_1 + \ldots + \alpha_l q_l + p| \geqq c_4 q^{-l}$. Thus we have

THEOREM 6F. *l-tuples $(\alpha_1, \ldots, \alpha_l)$ such that $1, \alpha_1, \ldots, \alpha_l$ is a basis of a real number field, are badly approximable.*

In particular badly approximable l-tuples exist, and Corollaries 6B and 6D can be improved at most by constant factors. In fact they can be improved by respective factors $c_5(l) < 1$, $c_6(l) < 1$, but the best value for these factors is known only when $l = 1$, when $c_5(l) = c_6(l) = \frac{1}{\sqrt{5}}$ by Theorem 1C. It is possible but there is no strong evidence that the extreme cases are attained by the l-tuples of Theorem 6F, and that therefore the optimal values of $c_5(l)$, $c_6(l)$ are algebraic of degree $l + 1$. The latest information on $c_5(l)$, $c_6(l)$ may be found in Cassels (1955) and the references given there.

The following " metrical " theorem is a consequence of a more general theorem of Khintchine (1926b).

THEOREM 6G. *Almost no l-tuple $(\alpha_1, \ldots, \alpha_l)$ (in the sense of Lebesgue measure) is either badly approximable or very well approximable.*

We saw that Corollary 6B cannot be improved by more than a constant factor. Combining the inequalities (6.2) we obtain

$$|\alpha_1 - (p_1/q)| \ldots |\alpha_l - (p_l/q)| < q^{-l-1}$$

or

$$|q| \cdot |\alpha_1 q - p_1| \ldots |\alpha_l q - p_l| < 1,$$

and therefore

$$|q| \cdot \|\alpha_1 q\| \ldots \|\alpha_l q\| < 1, \tag{6.5}$$

where $\| \xi \|$ denotes the distance from a real number ξ to the nearest integer. It is possible that (6.5) can be improved by more than a constant factor if $l > 1$. This is in fact a famous conjecture of Littlewood, which is usually stated in the form that for $l > 1$ and arbitrary real numbers $\alpha_1, \ldots, \alpha_l$,

$$\liminf_{q \to \infty} q \| \alpha_1 q \| \ldots \| \alpha_l q \| = 0 .$$

6.4. Dirichlet's principle in the proof of Theorems 6A, 6C may be replaced by Minkowski's Convex Body Theorem:

THEOREM 6H (Minkowski 1896). *Suppose K is a convex set in Euclidean E^n, symmetric at* $\mathbf{0}$ *(i.e. if a point $\mathbf{x} \in K$ then also $-\mathbf{x} \in K$) and with volume $V(K) > 2^n$. Then K contains an integer point different from* $\mathbf{0}$.

Sometimes one needs a more general version of this result in which the set of integer points is replaced by a *point lattice* Λ. Namely, such a lattice Λ is any discrete subgroup of the vector space E^n which contains n linearly independent vectors. It is easy to see that Λ is obtained from the set of integer points by a non-singular linear transformation A, and although A is not determined by Λ, the absolute value of the determinant of A is. This absolute value is called the *determinant* of the lattice Λ and will be denoted by $d(\Lambda)$. Theorem 6H remains true if the integer points are replaced by a lattice Λ and if the inequality $V(K) > 2^n$ is replaced by $V(K) > 2^n d(\Lambda.)$

A special case of Theorem 6H is when K is a parallelepiped given by inequalities

$$(6.6) \qquad | L_i(\mathbf{x}) | < R_i \qquad (i = 1, \ldots, n)$$

where $L_i(\mathbf{x}) = c_{i1}x_1 + \ldots + c_{in}x_n$ $(i = 1, \ldots, n)$ are linear forms of determinant 1 and where the R_i's are positive constants with $R_1 R_2 \ldots R_n > 1$. Continuity arguments show that the conclusion is still true if $R_1 R_2 \ldots R_n = 1$ and if one of the inequalities in (6.6) is replaced by $\leqq$. Thus we have

THEOREM 6I (Minkowski's Linear Forms Theorem). *Suppose $L_1, \ldots, L_n$ are linear forms with determinant* 1 *and suppose that $R_1, \ldots, R_n$ are positive with $R_1 \ldots R_n \geqq 1$. There is an integer point $\mathbf{x} \neq \mathbf{0}$ with*

$$| L_1(\mathbf{x}) | \leqq R_1 , \ | L_2(\mathbf{x}) | < R_2, \ldots, | L_n(\mathbf{x}) | < R_n .$$

Now suppose $l \geqq 1$ and put $n = l + 1$, and for vectors $\mathbf{x} = (q, p_1, \ldots, p_l)$ put $L_1(\mathbf{x}) = \alpha_1 q - p_1, \ldots, L_l(\mathbf{x}) = \alpha_l q - p_l$, but $L_n(\mathbf{x}) = q$. We obtain

Theorem 6A by applying Minkowski's Linear Forms Theorem to these linear forms and to $R_1 = \ldots = R_l = Q^{-1}$ and $R_n = Q^l$.

6.5. For later applications it will be convenient to state explicitly two other simple applications of Minkowski's Linear Forms Theorem. Suppose $1 \leqq m < n$ and let $L_1(\mathbf{x}), \ldots, L_m(\mathbf{x})$ be linearly independent linear forms. Assume without loss of generality that $L_1, \ldots, L_m, x_1, \ldots, x_{n-m}$ are linearly independent, and that these n linear forms have determinant d. By Theorem 6I there is for every $Q > 1$ an integer point $\mathbf{x} \neq \mathbf{0}$ with

$$|d|^{-1/n} |L_i(\mathbf{x})| \leqq Q^{-(n-m)} \qquad (i = 1, \ldots, m)$$

and

$$|d|^{-1/n} |x_j| \leqq Q^m \qquad (j = 1, \ldots, n-m).$$

Then if the norm $|\mathbf{x}|$ of $\mathbf{x} = (x_1, \ldots, x_n)$ is defined by

$$|\mathbf{x}| = \max(|x_1|, \ldots, |x_n|), \tag{6.7}$$

we have $|\mathbf{x}| \leqq c_1 Q^m$ and

$$|L_i(\mathbf{x})| \leqq c_2 |\mathbf{x}|^{-(n-m)/m} \qquad (i = 1, \ldots, m) \tag{6.8}$$

where c_1, c_2 depend on $L_1, \ldots, L_m$ only. Since Q may be chosen arbitrarily large, it follows that there are infinitely many integer points $\mathbf{x} \neq \mathbf{0}$ with (6.8). More generally, it can be shown that if $L_1(\mathbf{x}), \ldots, L_m(\mathbf{x})$ are linear forms of rank r (i.e. there are r but not $r+1$ linearly independent ones among them) with $1 \leqq r < n$, then the exponent in (6.8) may be replaced by $-(n-r)/r$. Therefore the following holds.

COROLLARY 6J. *Suppose $L_1, \ldots, L_m$ are linear forms of rank r with $1 \leqq r < n$. There is a $c_3 = c_3(L_1, \ldots, L_m)$ such that there are infinitely many integer points $\mathbf{x} \neq \mathbf{0}$ with*

$$|L_i(\mathbf{x})| \leqq c_3 |\mathbf{x}|^{-(n-r)/r} \qquad (i = 1, \ldots, m).$$

Corollary 6J essentially implies Corollaries 6B, 6D, i.e. it implies versions of these corollaries involving constants such as c_3. Finally Theorem 6I yields

COROLLARY 6K. *Suppose $L_1(\mathbf{x}), \ldots, L_n(\mathbf{x})$ are linear forms of determinant $d \neq 0$. Suppose $\gamma_1, \ldots, \gamma_n$ are reals with $\gamma_1 + \ldots + \gamma_n = 0$. For any $Q > 0$ there is an integer point $\mathbf{x} \neq \mathbf{0}$ with*

$$|L_i(\mathbf{x})| \leqq |d|^{1/n} Q^{\gamma_i} \qquad (i=1, \ldots, n).$$

6.6. An important special case of Theorem 6C is when $\alpha_1 = \alpha$, $\alpha_2 = \alpha^2, \ldots, \alpha_l = \alpha^l$. Changing our notation from l to d, we obtain a solution of the inequalities

$$1 \leqq \max(|q_1|, \ldots, |q_d|) < Q^{1/d}, \ |q_d\alpha^d + \ldots + q_1\alpha + q_0| \leqq Q^{-1}.$$

The polynomial $P(x) = q_d x^d + \ldots + q_1 x + q_0$ has height $H(P) \leqq c_1 Q^{1/d}$ and $|P(\alpha)| \leqq Q^{-1}$, whence

$$|P(\alpha)| \leqq c_2 H(P)^{-d},$$

where c_1, c_2 depend on α and on d only. Now one can show that unless $|P'(\alpha)|$ is extremely small, there is a real root β of P with

$$|\alpha - \beta| \leqq c_3 |P(\alpha)| / |P'(\alpha)| \leqq c_2 c_3 H(P)^{-d} |P'(\alpha)|^{-1}.$$

In general it is likely that $|P'(\alpha)|$ is of about the same order of magnitude as $H(P)$, and then we obtain

$$|\alpha - \beta| \leqq c_4 H(P)^{-d-1} \leqq c_5 H(\beta)^{-d-1}.$$

(The defining polynomial of β is a divisor of P, and this implies that $H(\beta) \leqq c_6 H(P)$ by, e.g., Theorem 4-3 in vol. 2 of LeVeque (1955)). Unfortunately we don't know whether $|P'(\alpha)|$ is large. At any rate one is tempted to conjecture that for every real α which is not itself algebraic of degree $\leqq d$, there are infinitely many real algebraic β of degree $\leqq d$ such that

$$|\alpha - \beta| \leqq c_7 H(\beta)^{-d-1}. \tag{6.9}$$

A weaker conjecture is that for every α as above and every $\varepsilon > 0$ there are infinitely many real algebraic numbers β of degree $\leqq d$ with

$$|\alpha - \beta| < H(\beta)^{-(d+1-\varepsilon)}. \tag{6.10}$$

The conjecture related to (6.9) is true for $d = 1$ by Dirichlet's Theorem, and it was shown to be true for $d = 2$ by Davenport and Schmidt (1967). For general d, Wirsing (1961) showed that there are infinitely many β of degree $\leqq d$ with

$$|\alpha - \beta| \leqq c_8 H(\beta)^{-(d+3)/2}.$$

He also showed that if $(\alpha, \alpha^2, \ldots, \alpha^d)$ is not very well aproximable, then (6.10) does have infinitely many solutions for every $\varepsilon > 0$.

Inequalities as above in which β is an algebraic *integer* are more difficult. Here one has to deal with polynomials $x^d + q_{d-1} x^{d-1} + \dots + q_1 x + q_0$, and hence one has to deal with an inhomogeneous approximation problem. One might conjecture that if $d \geqq 2$ and if α is not an algebraic integer of degree d and is not algebraic of degree $\leqq d - 1$, then for every $\varepsilon > 0$ there are infinitely many real algebraic integers β of degree $\leqq d$ with

$$|\alpha - \beta| < H(\beta)^{-(d-\varepsilon)}. \tag{6.11}$$

This conjecture is true if $(\alpha, \alpha^2, \dots, \alpha^{d-1})$ is not very well approximable. Davenport and Schmidt (1969) showed a result with (6.11) replaced by

$$|\alpha - \beta| \leqq c_9 H(\beta)^{-[(d+1)/2]}.$$

6.7. We have discussed approximation properties of general l-tuples $\alpha_1, \dots, \alpha_l$ and of l-tuples $\alpha, \alpha^2, \dots, \alpha^l$. Interesting questions arise if one asks about approximation properties of special l-tuples. For example, $(e, e^2, \dots, e^l)$ is not very well approximable (Popken (1929); see Schneider (1957), Kap. 4). A more general result (which is analogous to Theorem 7A below) concerning the l-tuple $\alpha_1 = e^{r_1}, \dots, \alpha_l = e^{r_l}$ with distinct non-zero rationals $r_1, \dots, r_l$ was proved by Baker (1965). For the behavior of l-tuples $\log \alpha_1, \dots, \log \alpha_l$ where $\alpha_1, \dots, \alpha_l$ are algebraic, see Baker (1966, 1967b, 1967c, 1968a) and Feldman (1968a, 1968b). In the next section we shall turn to l-tuples of real algebraic numbers.

7. Simultaneous approximation to algebraic numbers by rationals

7.1. We have already seen (Theorem 6F) that $(\alpha_1, \dots, \alpha_l)$ is badly approximable if $1, \alpha_1, \dots, \alpha_l$ is a basis of a real algebraic number field. In the same way one can show that if $1, \alpha_1, \dots, \alpha_l$ are linearly independent over the field of rationals and if they generate a field of degree d, then

$$|\alpha_1 q_1 + \dots + \alpha_l q_l + p| \geqq c_1 |\mathbf{q}|^{-d+1}$$

for every non-zero integer point $\mathbf{q} = (q_1, \dots, q_l, p)$. Here $c_1 = c_1(\alpha_1, \dots, \alpha_l) > 0$ is easily computable. The case $l = 1$ of this inequality yields Liouville's Theorem 2A.

Cassels and Swinnerton-Dyer (1955) have shown that Littlewood's conjecture is true for l-tuples $(\alpha_1, \dots, \alpha_l)$ such that $1, \alpha_1, \dots, \alpha_l$ is a basis of a real number field. (This conjecture applies only if $l > 1$.) Peck (1961) showed

for such l-tuples with $l > 1$ that there are infinitely many rational l-tuples $\left(\frac{p_1}{q}, \ldots, \frac{p_l}{q}\right)$ with

$$\left|\alpha_1 - \frac{p_1}{q}\right| < c_2 q^{-1-(1/l)},$$

$$\left|\alpha_i - \frac{p_i}{q}\right| < c_2 q^{-1-(1/l)} (\log q)^{-1/(l-1)} \qquad (i = 2, \ldots, l).$$

Schmidt (1966) derived an asymptotic formula for the number $v(N)$ of solutions with $q \leqq N$ of $\left|\alpha_i - \frac{p_i}{q}\right| < \psi(q)$ $(i = 1, \ldots, l)$ for such l-tuples and for certain functions $\psi(q)$. Earlier Lang (1965b, 1965c, 1966a) had done this for $l = 1$ and for a wider class of numbers α_1. Adams (1967) replaced our special l-tuples by badly approximable l-tuples and proved (1969a, 1969b, to appear) other results of this type.

7.2. As in §6.3, $\| \xi \|$ will denote the distance from a real number ξ to the nearest integer.

THEOREM 7A. *Suppose $\alpha_1, \ldots, \alpha_l$ are real algebraic numbers such that $1, \alpha_1, \ldots, \alpha_l$ are linearly independent over the rationals, and suppose $\delta > 0$. There are only finitely many positive integers q with*

$$q^{1+\delta} \| \alpha_1 q \| \ldots \| \alpha_l q \| < 1. \tag{7.1}$$

The inequalities

$$\left|\alpha_i - \frac{p_i}{q}\right| < q^{-1-(1/l)-\delta} \qquad (i = 1, \ldots, l) \tag{7.2}$$

imply that $\| \alpha_i q \| < q^{-(1/l)-\delta}$ $(i = 1, \ldots, l)$ and hence they imply (7.1). Therefore (7.2) has only finitely many solutions, and we obtain

COROLLARY 7B. *Suppose $\alpha_1, \ldots, \alpha_l$ and δ are as in Theorem 7A. Then there are only finitely many rational l-tuples $\left(\frac{p_1}{q}, \ldots, \frac{p_l}{q}\right)$ with* (7.2).

THEOREM 7C. *Again assume that $\alpha_1, \ldots, \alpha_l$ and δ are as in Theorem 7A. Then there are only finitely many l-tuples of non-zero integers $q_1, \ldots, q_l$ with*

$$(7.3) \qquad |q_1 q_2 \dots q_l|^{1+\delta} \, \| \alpha_1 q_1 + \dots + \alpha_l q_l \| < 1 .$$

By applying Theorem 7C to all the non-empty subsets of $\alpha_1, \dots, \alpha_l$ one deduces

COROLLARY 7D. *If again* $\alpha_1, \dots, \alpha_l$; δ *are as in Theorem* 7A, *then there are only finitely many* $(l+1)$*-tuples of integers* $q_1, \dots, q_l, p$ *with* $q = \max(|q_1|, \dots, |q_l|) > 0$ *and with*

$$(7.4) \qquad |\alpha_1 q_1 + \dots + \alpha_l q_l + p| < q^{-l-\delta} .$$

By Corollaries 6B and 6D the exponents in Corollaries 7B and 7D are best possible. In view of Khintchine's Transference Principle (Theorem 6E), the Corollaries 7B and 7D say the same, namely that $\alpha_1, \dots, \alpha_l$ is not very well approximable. The case $l = 1$ of these corollaries is Roth's Theorem. Theorems 7A and 7C and their corollaries were proved by Schmidt (1970). They had been anticipated by a weaker version of the case $l = 2$ and by the case $l = 2$ itself (Schmidt 1965 and 1967a, respectively).

Before Roth's Theorem was known Hasse (1939) used Siegel's method to derive estimates for simultaneous approximation. Baker (1967a), Feldman (1970a) and Osgood (1970) proved weaker but effective versions of Corollary 7D for special algebraic numbers $\alpha_1, \dots, \alpha_l$.

7.3. Corollary 7B shows that the exponent in Corollary 6B is best possible for algebraic numbers $\alpha_1, \dots, \alpha_l$, and Corollary 7D does the same for Corollary 6D. We shall now examine Corollaries 6J and 6K in the special case when the coefficients of the linear forms involved are algebraic. Suppose $1 \leqq m < n$ and $L_1(\mathbf{x}), \dots, L_m(\mathbf{x})$ are linear forms with real algebraic coefficients. We shall call $L_1, \dots, L_m$ a *Roth System* if for every $\delta > 0$ the inequalities

$$(7.5) \qquad |L_i(\mathbf{x})| < |\mathbf{x}|^{-((n-m)/m)-\delta} \qquad (i = 1, \dots, m)$$

have only finitely many solutions in integer points $\mathbf{x} \neq \mathbf{0}$. Roth's Theorem says that for $n = 2$, $m = 1$, the linear form $L(\mathbf{x}) = \alpha x_1 - x_2$ with a real algebraic irrational α is a Roth System.

THEOREM 7E (Schmidt (1971a)). *Linear forms* $L_1(\mathbf{x}), \dots, L_m(\mathbf{x})$ *with real algebraic coefficients and with* $m < n$ *are a Roth System if and only if their restrictions to every rational subspace* S^d *of dimension* d *with* $1 \leqq d \leqq n$ *have rank* r *satisfying*

$$r \geqq dm/n\,. \tag{7.6}$$

This theorem contains Corollaries 7B and 7D. For suppose $m = 1$, $n = l + 1$ and $L(\mathbf{x}) = L_1(\mathbf{x}) = \alpha_1 x_1 + \ldots + \alpha_l x_l + x_n$ where the numbers $\alpha_1, \ldots, \alpha_l, 1$ are algebraic and linearly independent over the rationals. Then $L(\mathbf{x}) \neq 0$ for every integer point $\mathbf{x} \neq \mathbf{0}$, and hence L has rank $r = 1$ on every rational subspace $S^d \neq \mathbf{0}$. Since $dm/n = d/n \leqq 1$, the inequality (7.6) is always satisfied, and $L(\mathbf{x})$ is a Roth System. Hence there are only finitely many integer points $\mathbf{x} \neq \mathbf{0}$ with $|L(\mathbf{x})| < |\mathbf{x}|^{-(n-1)-\delta} = |\mathbf{x}|^{-l-\delta}$, and Corollary 7D follows. Corollary 7B can be similarly derived.

The necessity of the condition (7.6) in Theorem 7E is easy to see: A rational subspace S^d is a d-dimensional Euclidean space, and the integer points in such a space form a lattice Λ. By applying a result analogous to Corollary 6J to the restrictions of $L_1, \ldots, L_m$ to S^d and to the lattice Λ, we obtain infinitely many integer points $\mathbf{x} \neq \mathbf{0}$ in S^d with

$$|L_i(\mathbf{x})| \leqq c_1 |\mathbf{x}|^{-(d-r)/r} = c_1 |\mathbf{x}|^{1-(d/r)} \qquad (i = 1, \ldots, m)\,.$$

Now if $r < dm/n$, say if $r = dmn^{-1}(1+\delta)^{-1}$, then

$$|L_i(\mathbf{x})| \leqq c_1 |\mathbf{x}|^{1-(n/m)(1+\delta)} \leqq c_1 |\mathbf{x}|^{-((n-m)/m)-\delta} \qquad (i = 1, \ldots, m)\,,$$

and we don't have a Roth System.

Suppose $L_1(\mathbf{x}), \ldots, L_n(\mathbf{x})$ are linear forms with real algebraic coefficients and suppose $\gamma_1, \ldots, \gamma_n$ are reals with $\gamma_1 + \ldots + \gamma_n = 0$. In view of Corollary 6K the following definition is natural. We shall call $(L_1, \ldots, L_n; \gamma_1, \ldots, \gamma_n)$ a *General Roth System* if for every $\delta > 0$ there is a $Q_0 = Q_0(L_1, \ldots, L_n; \gamma_1, \ldots, \gamma_n; \delta)$ such that for $Q > Q_0$ there is no integer point $\mathbf{x} \neq \mathbf{0}$ with

$$|L_i(\mathbf{x})| < Q^{\gamma_i - \delta} \qquad (i = 1, \ldots, n)\,.$$

Roth's Theorem says that for $n = 2$ and an algebraic irrational α, the system $(L_1(\mathbf{x}) = \alpha x_1 - x_2, L_2(\mathbf{x}) = x_1; \gamma_1 = -1, \gamma_2 = 1)$ is a General Roth System. Schmidt (1971a) derives necessary and sufficient conditions for General Roth Systems which contain Theorem 7E as a special case.

7.4. We shall briefly discuss an inhomogeneous approximation problem. Suppose $l > 1$ and suppose $1, \alpha_1, \ldots, \alpha_l$ are algebraic and linearly independent over the rational field $\mathbf{Q}$. The special case $q_l = 1$ of Theorem 7C shows that there are only finitely many integer l-tuples $q_1, \ldots, q_{l-1}, p$ with $q = \max(|q_1|, \ldots, |q_{l-1}|) > 0$ and with

$$|\alpha_1 q_1 + \ldots + \alpha_{l-1} q_{l-1} + p + \alpha_l| < q^{-(l-1)-\delta}\,.$$

One can easily show that more generally this is still true if α_l is not of the type $\alpha_l = \alpha_1 x_1 + \ldots + \alpha_{l-1} x_{l-1} + x_0$ with rational integers $x_0, x_1, \ldots, x_{l-1}$. Changing the notation we obtain

COROLLARY 7F. *Suppose $\alpha_1, \ldots, \alpha_l, \beta$ are real algebraic numbers such that β is not a linear combination of $\alpha_1, \ldots, \alpha_l$ with rational integer coefficients. Then for every $\delta > 0$ there are only finitely many integer l-tuples $q_1, \ldots, q_l$ with $q = \max(|q_1|, \ldots, |q_l|) > 0$ and with*

$$|\alpha_1 q_1 + \ldots + \alpha_l q_l + \beta| < q^{-(l-1)-\delta}.$$

This holds also when $l = 1$, but is trivial in this case. The case when $l = 2$ and α_1/α_2 is quadratic was proved by Mahler (1963). Combining Corollary 7D with certain transference theorems (see, e.g., Cassels (1957), ch. V) one obtains

COROLLARY 7G. *Suppose $\alpha_1, \ldots, \alpha_l$ are real, algebraic and linearly independent over* **Q**. *Then for every real β and every $\varepsilon > 0$ there are infinitely many integer l-tuples $(q_1, \ldots, q_l)$ with $q = \max(|q_1|, \ldots, |q_l|) > 0$ and*

$$|\alpha_1 q_1 + \ldots + \alpha_l q_l + \beta| < q^{-(l-1-\varepsilon)}.$$

7.5. Suppose α is a real algebraic number. Assume at first that it is not algebraic of degree $\leqq d$ where d is a given positive integer. Then $1, \alpha, \ldots, \alpha^d$ are linearly independent over the rationals, and by Corollary 7D there are only finitely many integer solutions of

$$|q_d \alpha^d + \ldots + q_1 \alpha + q_0| < q^{-d-\delta} \; (q = \max(|q_1|, \ldots, |q_d|) > 0)$$

for any given $\delta > 0$. Thus there are only finitely many polynomials $P(x)$ of degree at most d with rational integer coefficients and with

$$|P(\alpha)| < H(P)^{-d-\delta}.$$

Now if β is a root of $P(x)$ and if, say, $P(x) = a(x-\beta)(x-\beta_2)\ldots(x-\beta_e)$, then $|P(\alpha)| = |\alpha - \beta|\,|a(\alpha-\beta_2)\ldots(\alpha-\beta_e)| \leqq |\alpha - \beta|\,(|\alpha|+1)^{e-1} c_1 H(P)$ by the well known inequality $|a|(1+|\beta|)(1+|\beta_2|)\ldots(1+|\beta_e|) \leqq c_1 H(P)$ where $c_1 = c_1(e)$. (See e.g. LeVeque (1955), vol. 2, Theorem 4.2.) Thus $|\alpha - \beta| < H(P)^{-d-1-\delta}$ would imply that $|P(\alpha)| < c_2 H(P)^{-d-\delta}$, which has only finitely many solutions. Thus we see that the inequality

$$|\alpha - \beta| < H(\beta)^{-d-1-\delta} \tag{7.7}$$

has for every $\delta > 0$ only finitely many solutions in algebraic numbers β of degree $\leqq d$. It can be shown that the assumption on the degree of α can be removed, and we obtain

THEOREM 7H. *Suppose α is a real algebraic number, d a positive integer, $\delta > 0$. There are only finitely many (real or complex) algebraic numbers β of degree at most d with* (7.7).

This supersedes Wirsing's Theorem 4B. Suppose α is real and algebraic but not algebraic of degree $\leqq d$. Then by Corollary 7D the d-tuple $(\alpha, \alpha^2, \ldots, \alpha^d)$ is not very well approximable. Using a result of Wirsing (1961) mentioned in §6.6, we obtain a theorem which complements Theorem 7H.

THEOREM 7I. *Suppose α is algebraic of some degree greater than d. Then for every $\varepsilon > 0$ there are infinitely many real algebraic numbers β of degree $\leqq d$ with* (6.10), *i.e. with*

$$|\alpha - \beta| < H(\beta)^{-d-1+\varepsilon}.$$

In order to obtain results about approximation by algebraic *integers* β, one has to apply Corollary 7F with $l = d$ and $\alpha_1 = 1, \alpha_2 = \alpha, \ldots, \alpha_{d-1} = \alpha^{d-2}, \alpha_d = \alpha^{d-1}, \beta = \alpha^d$.

THEOREM 7J. *Suppose α, d, δ are as in Theorem* 7H. *There are only finitely many (real or complex) algebraic integers β of degree at most d with*

$$|\alpha - \beta| < H(\beta)^{-d-\delta}.$$

Using certain transference principles (see Davenport and Schmidt (1969)) together with the results of this section one can prove

THEOREM 7K. *Suppose $d \geqq 2$ and α is a real algebraic number of some degree $\geqq d$ but is not an algebraic integer of degree d. Then for every $\varepsilon > 0$ there are infinitely many real algebraic integers β of degree $\leqq d$ with*

$$|\alpha - \beta| < H(\beta)^{-d+\varepsilon}.$$

7.6. In the course of his classification of algebraic and transcendental real numbers, Mahler (1932) defines $\omega_d = \omega_d(\alpha)$ as the supremum of the

numbers ω such that there are infinitely many polynomials P with rational integer coefficients of degree $\leqq d$ and with

$$0 < |P(\alpha)| < H(P)^{-\omega}.$$

By Corollary 6D it is clear that $\omega_d \geqq d$ unless α is algebraic of degree $\leqq d$. Furthermore if α is algebraic of degree n, then one can show using the norm of $P(\alpha)$ that $\omega_d \leqq n - 1$ $(d=1, 2, \ldots)$. Thus Mahler could characterize the algebraic numbers α by the property that $\omega_d(\alpha)$ $(d=1, 2, \ldots)$ remains bounded.

Koksma (1939) defines $\omega_d^* = \omega_d^*(\alpha)$ as the supremum of the numbers ω^* such that there are infinitely many algebraic numbers β of degree $\leqq d$ with

$$|\alpha - \beta| < H(\beta)^{-1-\omega^*}.$$

It is easy to see that $\omega_d^* \leqq \omega_d$ and Wirsing (1961) showed that $\omega_d^* \geqq \frac{1}{2}(\omega_d + 1)$ if α is transcendental. Hence the algebraic numbers can also be characterized by the property that $\omega_d^*(\alpha)$ $(d=1, 2, \ldots)$ is bounded. We have $\omega_d^* \leqq \omega_d \leqq n - 1$ if α is algebraic of degree n, and the results of the last section show that $\omega_d^* = d$ if $d \leqq n - 1$. Since ω_d^* and ω_d increase with d, we have for algebraic α of degree n,

$$\omega_d = \omega_d^* = \begin{cases} d & \text{if } d \leqq n - 1 \\ n - 1 & \text{if } d \geqq n. \end{cases}$$

Thus the exponent in Theorem 7H is best possible precisely if $d < n$.

Another characterization of algebraic numbers by approximation properties was given by Gelfond (1952, §III.4, Lemma VII) and refined by Lang (1965a) and Tijdeman (1971, Lemma 6). This lemma was slightly improved by D. Brownawell (unpublished).

8. Tools from the Geometry of Numbers

8.1. To prove the theorems enunciated in the last section one needs certain results from the Geometry of Numbers. This field was first investigated under this name by Minkowski (1896). Other books on the Geometry of Numbers are Cassels (1959) and Lekkerkerker (1969).

Let K be a symmetric [1]) convex set in Euclidean E^n. For convenience let us assume that K is compact and has a non-empty interior. For $\lambda > 0$ let λK be the set consisting of the points $\lambda \mathbf{x}$ with $\mathbf{x} \in K$. Minkowski defines

[1]) I.e. if $\mathbf{x} \in K$, then also $-\mathbf{x} \in K$.

the *first minimum* λ_1 as the least positive value of λ such that λK contains an integer point $\mathbf{x} \neq \mathbf{0}$. More generally, for $1 \leqq j \leqq n$, the j-th minimum λ_j is the least positive value of λ such that λK contains j linearly independent integer points. It is clear that $0 < \lambda_1 \leqq \lambda_2 \leqq \ldots \leqq \lambda_n < \infty$, and that there are linearly independent integer points $\mathbf{x}_1, \ldots, \mathbf{x}_n$ with

$$\mathbf{x}_j \in \lambda_j K \qquad (j = 1, \ldots, n) . \tag{8.1}$$

Minkowski's Theorem 6H is easily seen to be equivalent with the inequality

$$\lambda_1^n V(K) \leqq 2^n .$$

Later Minkowski could refine this to the much stronger

THEOREM 8A (Minkowski's Theorem on Successive Minima).

$$2^n/n\,! \leqq \lambda_1 \ldots \lambda_n V(K) \leqq 2^n . \tag{8.2}$$

Like Theorem 6H this result can be generalized to arbitrary lattices Λ, and then (8.2) is to be replaced by

$$d(\Lambda)\, 2^n/n\,! \leqq \lambda_1 \ldots \lambda_n V(K) \leqq d(\Lambda)\, 2^n . \tag{8.3}$$

Of particular interest to us will be the situation when $L_1(\mathbf{x}), \ldots, L_n(\mathbf{x})$ are linearly independent linear forms and $R_1, \ldots, R_n$ are positive numbers, and when K is the parallelepiped defined by [1])

$$|L_i(\mathbf{x})| \leqq R_i \qquad (i = 1, \ldots, n) . \tag{8.4}$$

In the special case when $R_1 \ldots R_n = 1$ and when $|\det(L_1, \ldots, L_n)| = \Delta$, say, we have $V(K) = 2^n/\Delta$, whence $\Delta/n! \leqq \lambda_1 \ldots \lambda_n \leqq \Delta$. In particular we have

$$1 \ll \lambda_1 \ldots \lambda_n \ll 1 , \tag{8.5}$$

where the notation $A \ll B$ means that $A \leqq cB$ with $c = c(n, \Delta)$. Later on the notation $A \gg \ll B$ will mean that both $A \ll B$ and $B \ll A$.

8.2. We shall need three so-called " transference theorems " which relate the successive minima of certain parallelepipeds to the successive minima of other parallelepipeds.

---

[1]) The case when $R_1 = \ldots = R_n = 1$ is just as general, but the factors $R_1, \ldots, R_n$ will be convenient for later applications.

THEOREM 8B (" Davenport's Lemma " (Davenport, 1937)). *Let* $\lambda_1, \ldots, \lambda_n$ *be the successive minima of the parallelepiped* Π *given by* (8.4). *Let* $\rho_1, \ldots, \rho_n$ *be numbers with*

$$\rho_1 \geqq \rho_2 \geqq \ldots \geqq \rho_n > 0 \quad \textit{and} \quad \rho_1\lambda_1 \leqq \ldots \leqq \rho_n\lambda_n .$$

Then there is a permutation $(t_1, t_2, \ldots, t_n)$ *of* $(1, 2, \ldots, n)$ *such that the successive minima* $\lambda_1', \ldots, \lambda_n'$ *of the new parallelepiped* Π' *given by*

$$| L_i(\mathbf{x}) | \leqq R_i \rho_{t_i}^{-1} \qquad (i = 1, \ldots, n)$$

satisfy

$$\lambda_j' \gg \ll \rho_j \lambda_j \qquad (j = 1, \ldots, n) . \tag{8.6}$$

Moreover, let $\mathbf{x}_1, \ldots, \mathbf{x}_n$ *be linearly independent integer points with* (8.1), *i.e. with* $R_i^{-1} \, | L_i(\mathbf{x}_j) | \leqq \lambda_j$ $(i, j = 1, \ldots, n)$. *Let* T_0 *be the subspace consisting of* $\mathbf{0}$, *and for* $1 \leqq j \leqq n$ *let* T_j *be the subspace spanned by* $\mathbf{x}_1, \ldots, \mathbf{x}_j$. *Then every integer point* $\mathbf{x}$ *outside the subspace* T_{j-1} *where* $1 \leqq j \leqq n$ *satisfies*

$$\max \left(R_1^{-1} \rho_{j_1} | L_1(\mathbf{x}) |, \ldots, R_n^{-1} \rho_{j_n} | L_n(\mathbf{x}) | \right) \gg \lambda_j' .$$

Note that the ratios of $\rho_1\lambda_1, \ldots, \rho_n\lambda_n$ are equal to or smaller than the ratios of $\lambda_1, \ldots, \lambda_n$, so that the successive minima have been " pushed closer together ". Usually in transference theorems only inequalities such as (8.6) are given. But the last statement of the theorem will also be needed.

8.3. Every linear form $L(\mathbf{x})$ is of the type $L(\mathbf{x}) = \mathbf{ax}$ where $\mathbf{a}$ is a fixed vector and where $\mathbf{ax}$ denotes the inner product. Now suppose that $L_1(\mathbf{x}), \ldots, L_n(\mathbf{x})$ are linearly independent linear forms. Then if $L_i(\mathbf{x}) = \mathbf{a}_i\mathbf{x}$ $(i = 1, \ldots, n)$, the vectors $\mathbf{a}_1, \ldots, \mathbf{a}_n$ are linearly independent. There are unique vectors $\mathbf{a}_1^*, \ldots, \mathbf{a}_n^*$ with

$$\mathbf{a}_i\mathbf{a}_j^* = \delta_{ij} = \begin{cases} 1 \text{ if } i = j , \\ 0 \text{ otherwise} . \end{cases}$$

The linear forms $L_1^*, \ldots, L_n^*$ given by $L_i^*(\mathbf{x}) = \mathbf{a}_i^*\mathbf{x}$ $(i = 1, \ldots, n)$ are called *dual* to $L_1, \ldots, L_n$; they satisfy the identity $L_1(\mathbf{x}) L_1^*(\mathbf{y}) + \ldots + L_n(\mathbf{x}) L_n^*(\mathbf{y}) = \mathbf{xy}$. The dual linear forms are again linearly independent, and they have determinant 1 if $L_1, \ldots, L_n$ have determinant 1. The parallelepiped

$$\Pi^* : \; | L_i^*(\mathbf{x}) | \leqq R_i^{-1} \qquad (i = 1, \ldots, n)$$

is called the *dual* of the parallelepiped Π defined by (8.4).

Remark. One can define the *polar* set of any convex symmetric set, and the dual of a parallelepiped is closely related to its polar set. But the polar set of a parallelepiped has the disadvantage that it need not be a parallelepiped.

THEOREM 8C (Mahler 1939). *Let $\lambda_1, \ldots, \lambda_n$ and $\lambda_1^*, \ldots, \lambda_n^*$ be the successive minima of a parallelepiped Π and of its dual Π^*, respectively. Then*

$$\lambda_j^* \gg \ll \lambda_{n+1-j}^{-1} \qquad (j=1, \ldots, n).$$

Moreover, if $\mathbf{x}_1, \ldots, \mathbf{x}_n$ are linearly independent points with (8.1), *i.e. with $|L_i(\mathbf{x}_j)| \leqq \lambda_j R_i$ $(i, j=1, \ldots, n)$, and if $\mathbf{x}_1^*, \ldots, \mathbf{x}_n^*$ are defined by $\mathbf{x}_i \mathbf{x}_j^* = \delta_{ij}$ $(i, j=1, \ldots, n)$, then*

$$|L_i^*(\mathbf{x}_{n+1-j}^*)| \ll \lambda_j^* R_i^{-1} \qquad (i, j=1, \ldots, n). \tag{8.7}$$

8.4. Suppose $1 \leqq p \leqq n$ and put $l = \binom{n}{p}$. Vectors in E^n will be denoted as usual by $\mathbf{a}, \mathbf{b}, \ldots$, and vectors in E^l will be denoted by $\mathbf{A}, \mathbf{B}, \ldots$. By

$$\mathbf{a}_1 \wedge \ldots \wedge \mathbf{a}_p$$

we shall denote the exterior product of the vectors $\mathbf{a}_1, \ldots, \mathbf{a}_p$, i.e. the vector in E^l whose coordinates are the $(p \times p)$-determinants formed from the matrix with rows $\mathbf{a}_1, \ldots, \mathbf{a}_p$, and arranged in lexicographic order. For example if $n = 4$ and $p = 2$, then $l = 6$, and if $\mathbf{a} = (\alpha_1, \alpha_2, \alpha_3, \alpha_4)$, $\mathbf{b} = (\beta_1, \beta_2, \beta_3, \beta_4)$, then

$$\mathbf{a} \wedge \mathbf{b} = \left(\begin{vmatrix} \alpha_1 & \alpha_2 \\ \beta_1 & \beta_2 \end{vmatrix}, \begin{vmatrix} \alpha_1 & \alpha_3 \\ \beta_1 & \beta_3 \end{vmatrix}, \begin{vmatrix} \alpha_1 & \alpha_4 \\ \beta_1 & \beta_4 \end{vmatrix}, \begin{vmatrix} \alpha_2 & \alpha_3 \\ \beta_2 & \beta_3 \end{vmatrix}, \begin{vmatrix} \alpha_2 & \alpha_4 \\ \beta_2 & \beta_4 \end{vmatrix}, \begin{vmatrix} \alpha_3 & \alpha_4 \\ \beta_3 & \beta_4 \end{vmatrix} \right).$$

Let $C(n, p)$ be the set of all p-tuples of integers $i_1, \ldots, i_p$ with $1 \leqq i_1 < \ldots < i_p \leqq n$. There are l such p-tuples.

Now suppose that $L_1(\mathbf{x}) = \mathbf{a}_1 \mathbf{x}, \ldots, L_n(\mathbf{x}) = \mathbf{a}_n \mathbf{x}$ are independent linear forms. For $\sigma = \{i_1, \ldots, i_p\}$ in $C(n, p)$, let $\mathbf{A}_\sigma$ be the vector

$$\mathbf{A}_\sigma = \mathbf{a}_{i_1} \wedge \ldots \wedge \mathbf{a}_{i_p}.$$

Let $L_\sigma^{(p)}$ be the linear form in E^l defined by $L_\sigma^{(p)}(\mathbf{X}) = \mathbf{A}_\sigma \mathbf{X}$. The l linear forms $L_\sigma^{(p)}$ with $\sigma \in C(n, p)$ are again linearly independent, and they have determinant 1 if $L_1, \ldots, L_n$ have determinant 1. Let $R_1, \ldots, R_n$ be positive constants with $R_1 R_2 \ldots R_n = 1$ and define R_σ by $R_\sigma = \prod_{i \in \sigma} R_i$. The inequalities

$$|L_\sigma^{(p)}(\mathbf{X})| \leqq R_\sigma \qquad (\sigma \in C(n, p))$$

define a parallelepiped $\Pi^{(p)}$ in E^l which we shall call the p-th *pseudocompound* of the parallelepiped Π defined by (8.4).

Remarks. Mahler (1955) defined the p-th *compound* of any symmetric convex set, and the pseudocompound of a parallelepiped is closely related to its compound. But the compound of a parallelepiped is not necessarily a parallelepiped. Except for the notation, the $(n-1)$-st pseudocompound is the same as the dual of a parallelepiped, and hence the results of the last subsection may be interpreted as special cases of the results of the present subsection.

THEOREM 8D (Mahler 1955). *Let $\lambda_1, \ldots, \lambda_n$ and $v_1, \ldots, v_l$ be the successive minima of a parallelepiped Π and of its p-th pseudocompound $\Pi^{(p)}$, respectively. For $\sigma \in C(n, p)$ put $\lambda_\sigma = \prod_{i \in \sigma} \lambda_i$ and order the elements of $C(n, p)$ as $\sigma_1, \ldots, \sigma_l$ such that $\lambda_{\sigma_1} \leqq \ldots \leqq \lambda_{\sigma_l}$. Then*

$$v_j \gg \ll \lambda_{\sigma_j} \qquad (j = 1, \ldots, l).$$

Moreover, if $\mathbf{x}_1, \ldots, \mathbf{x}_n$ are linearly independent integer points with (8.1), *i.e. with $|L_i(\mathbf{x}_j)| \leqq \lambda_j R_i$ $(i, j = 1, \ldots, n)$, and if for $\tau = \{j_1, \ldots, j_p\}$ in $C(n, p)$ we put $\mathbf{X}_\tau = \mathbf{x}_{j_1} \wedge \ldots \wedge \mathbf{x}_{j_p}$, then*

$$|L_\sigma^{(p)}(\mathbf{X}_\tau)| \ll \lambda_\tau R_\sigma \qquad (\sigma, \tau \in C(n, p)).$$

9. OUTLINE OF THE PROOF OF THE THEOREMS ON SIMULTANEOUS APPROXIMATION TO ALGEBRAIC NUMBERS

9.1. Let us see what happens if we try to generalize Roth's proof to prove, say, Corollary 7B. In Roth's proof we constructed a polynomial $P(x_1, \ldots, x_m)$ in m variables $x_1, \ldots, x_m$ which had a zero of high order at $(\alpha, \ldots, \alpha)$. Hence the natural thing to try would be

(a) to construct a polynomial $P(x_{11}, \ldots, x_{1l}; \ldots; x_{m1}, \ldots, x_{ml})$ in ml variables of total degree $\leqq r_h$ in each block of variables $x_{h1}, \ldots, x_{hl}$ $(h = 1, \ldots, m)$ with a zero of high order at $(\alpha_1, \ldots, \alpha_l; \ldots; \alpha_1, \ldots, \alpha_l)$. Then

(b) one would have to show that if each of m given rational l-tuples $\left(\frac{p_{h1}}{q_h}, \ldots, \frac{p_{hl}}{q_h}\right)$ $(h = 1, \ldots, m)$ satisfies (7.2), then P also has a zero of high order at

$$\left(\frac{p_{11}}{q_1}, \ldots, \frac{p_{1l}}{q_1}; \ldots; \frac{p_{m1}}{q_m}, \ldots, \frac{p_{ml}}{q_m}\right).$$

Finally

(c) one would have to show that under suitable conditions P cannot have a high zero at such a rational point.

If we proceed in this fashion, we encounter difficulties in (c). In Roth's Lemma 3C it was essential that P had rather different degrees in its variables and that the denominators in $\frac{p_1}{q_1}, \ldots, \frac{p_m}{q_m}$ increased very fast. In our present situation the first l denominators are equal, so that Roth's Lemma does not apply. The example $m = 1$, $l = 2$, $P(x_1, x_2) = (x_1 - x_2)^r$ shows that we cannot expect to have a lemma similar to Roth's in our present context, since P has a zero of order as high as r at every point (ξ, ξ).

The polynomial P is defined on $E^l \times \ldots \times E^l$ (m copies). While it is difficult to say much about the order of vanishing of P at rational points $\mathbf{r}_1 \times \ldots \times \mathbf{r}_m$, it is easier to show that P cannot have a zero of high order on certain linear manifolds $\mathscr{M}_1 \times \ldots \times \mathscr{M}_m$ where each $\mathscr{M}_h$ is a rational (i.e. defined by a linear equation with rational coefficients) hyperplane in E^l. We can illustrate this when $m = 1$. Namely, $\mathscr{M}_1$ is defined by an equation $a_0 + a_1 x_1 + \ldots + a_l x_l = 0$ which can be normalized such that $a_0, a_1, \ldots, a_l$ are coprime rational integers. If $P(x_1, \ldots, x_l)$ has a zero of order $\geqq i$ on $\mathscr{M}_1$ (i.e. P has a zero of order $\geqq i$ at every point of $\mathscr{M}_1$), then $P(x_1, \ldots, x_l) = (a_0 + a_1 x_1 + \ldots + a_l x_l)^i R(x_1, \ldots, x_l)$, where R has integer coefficients by Gauss' Lemma. It follows that

$$(H(M))^i \leqq H(P) \tag{9.1}$$

where $H(M)$ is the height of $M(\mathbf{x}) = a_0 + a_1 x_1 + \ldots + a_l x_l$. This inequality provides a good upper bound for i if $H(M)$ is large.

9.2. It will be more convenient to deal with hyperplanes through the origin in E^{l+1} than with hyperplanes in E^l. Hence we shall put

$$n = l + 1 \tag{9.2}$$

and we shall consider polynomials $P(x_{11}, \ldots, x_{1n}; \ldots; x_{m1}, \ldots, x_{mn})$ which are homogeneous of degree r_h in each block of variables $x_{h1}, \ldots, x_{hn}$ $(h = 1, \ldots, m)$. The manifold $\mathscr{M}_1 \times \ldots \times \mathscr{M}_m$ now becomes a subspace defined by $L_1(x_{11}, \ldots, x_{1n}) = \ldots = L_m(x_{m1}, \ldots, x_{mn}) = 0$, where each L_h is a not

identically vanishing linear form in $x_{h1}, \ldots, x_{hn}$ $(h=1, \ldots, m)$. The polynomial P vanishes on $\mathscr{M}_1 \times \ldots \times \mathscr{M}_m$ precisely if it lies in the ideal generated by $L_1, \ldots, L_m$. A suitable definition of the index is now as follows.

Let $L_h = L_h(x_{h1}, \ldots, x_{hn})$ $(h=1, \ldots, m)$ be not identically vanishing linear forms. For positive integers $r_1, \ldots, r_m$ and for $c \geqq 0$ let $\mathscr{T}(c)$ be the ideal generated by the products $L_1^{i_1} \ldots L_m^{i_m}$ with

$$\frac{i_1}{r_1} + \ldots + \frac{i_m}{r_m} \geqq c .$$

The index of P with respect to $(L_1, \ldots, L_m; r_1, \ldots, r_m)$ is the largest value of c such that $P \in \mathscr{T}(c)$ if P is not identically zero, and it is $+\infty$ if P is identically zero.

9.3. Now suppose that $L(\mathbf{x}) = \alpha_1 x_1 + \ldots + \alpha_n x_n$ has real algebraic coefficients. In analogy with Lemma 3A in step (a) in the proof of Roth's Theorem, one can construct a polynomial P as above which is not identically zero and which has not too large rational integer coefficients, such that P has index at least

$$\left(\frac{1}{n} - \varepsilon\right) m ,$$

with respect to $(L, \ldots, L; r_1, \ldots, r_m)$. Here L really occurs with m different meanings; namely, the h-th copy of L means $\alpha_1 x_{h1} + \ldots + \alpha_n x_{hn}$ $(h=1, \ldots, m)$. Perhaps it should be explained why the factor $\frac{1}{2} - \varepsilon$ in Lemma 3A is now replaced by $\frac{1}{n} - \varepsilon$. A form P in mn variables $x_{11}, \ldots, x_{1n}; \ldots; x_{m1}, \ldots, x_{mn}$ is also a form in $L, x_{12}, \ldots, x_{1n}; \ldots; L, x_{m2}, \ldots, x_{mn}$ provided $\alpha_1 \neq 0$ (and where L occurs with different meanings again). Now for " most " monomials in $L, x_{12}, \ldots, x_{1n}; \ldots; L, x_{m2}, \ldots, x_{mn}$ the degree in L will be about $\frac{1}{n}$ times the total degree of the monomial, and hence will be greater than $\left(\frac{1}{n} - \varepsilon\right)$ times the total degree of the monomial.

But a result with only one linear form L is not enough. In general, say when dealing with General Roth Systems, one has n linear forms $L_1, \ldots, L_n$ to start with, and one can deal with them simultaneously. The following result now replaces Lemma 3A.

LEMMA 9A. *Let $L_1, \ldots, L_n$ be not identically vanishing linear forms with real algebraic coefficients. Suppose $\varepsilon > 0$. Then if $m > m_0(L_1, \ldots, L_n; \varepsilon)$ and if $r_1, \ldots, r_m$ are positive integers, there is a polynomial $P(x_{11}, \ldots, x_{1n}; \ldots; x_{m1}, \ldots, x_{mn}) \not\equiv 0$ with rational integer coefficients such that*

(i) *P is homogeneous in $x_{h1}, \ldots, x_{hn}$ of degree r_h $(h=1, \ldots, m)$.*

(ii) *P has index $\geqq \left(\frac{1}{n} - \varepsilon\right) m$ with respect to $(L_i, \ldots, L_i; r_1, \ldots, r_m)$ $(i=1, \ldots, n)$.*

(iii) *$H(P) \leqq B^{r_1 + \ldots + r_m}$ where $B = B(L_1, \ldots, L_m)$.*

This takes care of generalizing part (a) of Roth's proof. We have chosen our definition of the index such that (c) has a chance of going through, and in fact one can derive from Roth's Lemma 3C a more general lemma that applies in our situation. Namely, if $M_1(\mathbf{x}), \ldots, M_m(\mathbf{x})$ are linear forms with rational integer coefficients, then under suitable conditions the index of P with respect to $(M_1, \ldots, M_m; r_1, \ldots, r_m)$ is $\leqq \varepsilon$.

9.4. If thus remains to deal with part (b). Suppose, say, that we want to derive a criterion for General Roth Systems as defined in §7.3. Suppose $L_1, \ldots, L_n$ are linear forms with real algebraic coefficients and suppose $\gamma_1 + \ldots + \gamma_n = 0$. Suppose there is a $\delta > 0$ and there are arbitrarily large values of Q for which there is an integer point $\mathbf{x} \neq \mathbf{0}$ with $|L_i(\mathbf{x})| < Q^{\gamma_i - \delta}$ $(i=1, \ldots, n)$. Assume in particular that this is true for $Q = Q_1, \ldots, Q_m$ and with integer points $\mathbf{x}_1, \ldots, \mathbf{x}_m$, respectively. An argument like the one used in the proof of Lemma 3B shows that if suitable auxiliary conditions are satisfied, then the polynomial P of Lemma 9A does in fact have

$$P(\mathbf{x}_1, \ldots, \mathbf{x}_m) = 0 .$$

But this is not what we really need. Namely, we need a rational subspace of the type $\mathscr{M}_1 \times \ldots \times \mathscr{M}_m$ where each $\mathscr{M}_h$ is a hyperplane of E^n, such that P vanishes on this subspace.

There is a way out of this difficulty, although it is a rather costly one. Namely, we have to assume that for each Q_h $(h=1, \ldots, m)$ there is not just one but there are

$$l = n - 1$$

linearly independent integer points $\mathbf{x}_h^{(1)}, \ldots, \mathbf{x}_h^{(l)}$ with

$$(9.3) \qquad |L_i(\mathbf{x}_h^{(j)})| \leqq Q_h^{\gamma_i - \delta} \quad (i=1, \ldots, n; j=1, \ldots, l; h=1, \ldots, m).$$

Now if $\mathscr{M}_h$ is the hyperplane through $\mathbf{0}$ spanned by $\mathbf{x}_h^{(1)}, \ldots, \mathbf{x}_h^{(l)}$ ($h = 1, \ldots, m$), then one can show that P vanishes on $\mathscr{M}_1 \times \ldots \times \mathscr{M}_m$. In fact one can show that if M_h is the linear form defining $\mathscr{M}_h$ ($h=1, \ldots, m$), then the index of P with respect to $(M_1, \ldots, M_m; r_1, \ldots, r_m)$ is $\geqq m\varepsilon$, which in conjunction with (c) gives the desired contradiction.

9.5. But what have we really shown now? The inequalities

$$(9.4) \qquad |L_i(\mathbf{x})| \leqq Q^{\gamma_i} \qquad (i=1, \ldots, n)$$

define a parallelepiped. The presence of $l = n - 1$ linearly independent integer points $\mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(l)}$ with $|L_i(\mathbf{x}^{(j)})| \leqq Q^{\gamma_i - \delta}$ ($i=1, \ldots, n; j=1, \ldots, l$) means that the $(n-1)$ st minimum $\lambda_{n-1} = \lambda_{n-1}(Q)$ satisfies $\lambda_{n-1} \leqq Q^{-\delta}$. The inequalities (9.3) mean precisely that $\lambda_{n-1}(Q) \leqq Q^{-\delta}$ for $Q = Q_1, Q_2, \ldots, Q_m$. Thus we obtain a theorem about λ_{n-1}:

THEOREM 9B. *(Theorem on the next to last minimum). Suppose $n \geqq 2$ and $L_1, \ldots, L_n$ are linearly independent linear forms with real algebraic coefficients, and suppose $L_1^*, \ldots, L_n^*$ are their duals. Suppose $\delta > 0$, suppose $\gamma_1 + \ldots + \gamma_n = 0$, and let Σ be the set of integers i in $1 \leqq i \leqq n$ for which*

$$\gamma_i + \delta \geqq 0.$$

There is a $Q_0 = Q_0(L_1, \ldots, L_n; \gamma_1, \ldots, \gamma_n; \delta)$ with the following property: Let $\lambda_1 = \lambda_1(Q), \ldots, \lambda_n = \lambda_n(Q)$ be the successive minima of the parallelepiped $\Pi(Q)$ given by (9.4). Then for $Q > Q_0$ either

$$(9.5) \qquad \lambda_{n-1} > Q^{-\delta}$$

or

$$(9.6) \qquad L_i^*(\mathbf{x}_n^*) = 0 \text{ for every } i \in \Sigma,$$

where $\mathbf{x}_1^*, \ldots, \mathbf{x}_n^*$ are the duals [1]) *to linearly independent integer points $\mathbf{x}_1, \ldots, \mathbf{x}_n$ with $\mathbf{x}_j \in \lambda_j \Pi$ ($j=1, \ldots, n$).*

It was clear from the discussion above that some inequality such as (9.5) would result. The hyperplanes $\mathscr{M}$ of the discussion above were spanned by $\mathbf{x}_1, \ldots, \mathbf{x}_{n-1}$ (but with the notation $\mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(l)}$), and hence the coefficients

---

[1]) I.e. they satisfy $\mathbf{x}_i \mathbf{x}_j^* = \delta_{ij}$ ($i, j=1, \ldots, n$).

in the defining equation for $\mathscr{M}$ are proportional to $\mathbf{x}_n^*$. The alternative (9.6) had to be put in to allow for the possibility that $\mathscr{M}$ behaves in a somewhat degenerate fashion. In most cases, e.g., if the coefficients of some L_i^* with $i \in \Sigma$ are linearly independent over the rationals, then no integer point $\mathbf{x} \neq \mathbf{0}$ can satisfy (9.6), and then (9.5) must hold.

Theorem 9B gives information on λ_{n-1} rather than on λ_1. In what follows, transference theorems will be used to gain information on λ_1.

9.6. Theorem 9B says that if Q is large and $\lambda_{n-1} < Q^{-\delta}$, then $\mathbf{x}_n^*$ must lie in a certain subspace. The inequality (8.7) of Mahler's Theorem 8C further restricts the possibilities for $\mathbf{x}_n^*$. A combination of these results yields

COROLLARY 9C. *Suppose* $L_1, \ldots, L_n, \gamma_1, \ldots, \gamma_n, \delta, \mathbf{x}_1 = \mathbf{x}_1(Q), \ldots, \mathbf{x}_n = \mathbf{x}_n(Q), \mathbf{x}_1^* = \mathbf{x}_1^*(Q), \ldots, \mathbf{x}_n^* = \mathbf{x}_n^*(Q)$ *are as above. Suppose there are arbitrarily large values of* Q *with*

$$\lambda_{n-1} < Q^{-\delta}. \tag{9.7}$$

Then there is a fixed vector $\mathbf{c}$ *and there are arbitrarily large values of* Q *with* (9.7) *and with* $\mathbf{x}_n^*(Q) = \mathbf{c}$.

Next, the condition (9.7) will be replaced by

$$\lambda_{n-1} < Q^{-\delta} \lambda_n. \tag{9.8}$$

The latter condition usually is milder, since $\lambda_n \gg 1$ by (8.5).

THEOREM 9D. *(Theorem on the last two minima). Suppose* $L_1, \ldots, L_n, \gamma_1, \ldots, \gamma_n, \delta, \mathbf{x}_1, \ldots, \mathbf{x}_n, \mathbf{x}_1^*, \ldots, \mathbf{x}_n^*$ *are as above. Suppose there are arbitrarily large values of* Q *with* (9.8). *Then there are arbitrarily large values of* Q *with* (9.8) *and with* $\mathbf{x}_n^*(Q) = \mathbf{c}$, *where* $\mathbf{c}$ *is a fixed vector.*

To prove this theorem one needs Davenport's Lemma (Theorem 8B). Namely, put $\rho_0 = (\lambda_1 \ldots \lambda_{n-2}\lambda_{n-1}^2)^{1/n}$ and

$$\rho_1 = \rho_0/\lambda_1, \ldots, \rho_{n-1} = \rho_0/\lambda_{n-1}, \text{ but } \rho_n = \rho_0/\lambda_{n-1}.$$

By Davenport's Lemma we can compare the successive minima $\lambda_1, \ldots, \lambda_n$ of Π with the successive minima $\lambda_1', \ldots, \lambda_n'$ of another parallelepiped Π'. We have $\lambda_j' \gg \ll \rho_j \lambda_j$ $(j = 1, \ldots, n)$ and $\rho_0 \ll \lambda_1' \ll \ldots \ll \lambda_{n-1}' \ll \rho_0 \ll (\lambda_{n-1}/\lambda_n)^{1/n} \ll Q^{-\delta/n}$ by (8.5) and (9.8). Hence $\lambda_{n-1}' < Q^{-\delta/(2n)}$ if Q is large, and applying Corollary 9C to Π' we see that $\mathbf{x}_n^{*\prime}(Q)$ is the same

for arbitrarily large values of Q, which in turn (by the last assertion of Davenport's Lemma) implies that $\mathbf{x}_n^*(Q)$ is the same for certain arbitrarily large values of Q.

9.7. THEOREM 9E. *(Subspace Theorem). Suppose $L_1, \ldots, L_n, \gamma_1, \ldots, \gamma_n, \delta$, $\mathbf{x}_1(Q), \ldots, \mathbf{x}_n(Q)$ are as above. Suppose there is a d in $1 \leqq d \leqq n-1$ such that*

$$\lambda_d < \lambda_{d+1} Q^{-\delta} \tag{9.9}$$

for certain arbitrarily large values of Q. Then there is a fixed rational subspace S^d of dimension d such that for some arbitrarily large values of Q with (9.9), the points

$$\mathbf{x}_1(Q), \ldots, \mathbf{x}_d(Q) \text{ lie in } S^d .$$

For the proof put $p = n - d$ and construct the linear forms $L_\sigma^{(p)}$ as in §8.4. Also put $\Gamma_\sigma = \sum_{i \in \sigma} \gamma_i$. The inequalities

$$| L_\sigma^{(p)}(\mathbf{X}) | \leqq Q^{\Gamma_\sigma} \qquad (\sigma \in C(n, p))$$

define the p-th pseudocompound $\Pi^{(p)}$ of Π. By Mahler's Theorem 8D the last two minima v_{l-1}, v_l of this pseudocompound have

$$v_{l-1} \gg \ll \lambda_d \lambda_{d+2} \lambda_{d+3} \ldots \lambda_n , \quad v_l \gg \ll \lambda_{d+1} \lambda_{d+2} \lambda_{d+3} \ldots \lambda_n ,$$

whence $v_{l-1} < v_l Q^{-\delta/2}$ for large Q by (9.9). An application of Theorem 9D shows that $\mathbf{X}_l^*$ [1]) is the same for some arbitrarily large values of Q. Some algebra combined with the last assertion of Theorem 8D shows that (because of (9.9)) $\mathbf{X}_l^*$ is proportional to $\mathbf{x}_{d+1}^* \wedge \ldots \wedge \mathbf{x}_n^*$. It follows that the subspace S^* spanned by $\mathbf{x}_{d+1}^*, \ldots, \mathbf{x}_n^*$ is the same for some arbitrarily large values of Q. But for these values of Q the vectors $\mathbf{x}_1, \ldots, \mathbf{x}_d$ lie in the orthogonal complement S^d of S^*.

9.8. We shall illustrate the power of the Subspace Theorem by deducing Theorem 7E. Suppose we have $\delta > 0$, $1 \leqq m < n$, m linearly independent linear forms $L_1, \ldots, L_m$ with real algebraic coefficients, and infinitely many integer solutions $\mathbf{x} \neq \mathbf{0}$ of

---

[1]) $\mathbf{X}_l^*$ in E^l is defined in terms of $\Pi^{(p)}(Q)$ just as $\mathbf{x}_n^*$ in E^n was defined in terms of $\Pi(Q)$.

$$|L_i(\mathbf{x})| \leqq |\mathbf{x}|^{-((n-m)/m)-\delta} \qquad (i=1, \ldots, m).$$

We may assume without loss of generality that $L_1, \ldots, L_m, x_1, \ldots, x_{n-m}$ are linearly independent. Put $L_{m+1}(\mathbf{x}) = x_1, \ldots, L_n(\mathbf{x}) = x_{n-m}$. It is easy to see that there is a $\delta' > 0$ and there are arbitrarily large values of Q for which there are solutions $\mathbf{x} \neq \mathbf{0}$ of

$$|L_i(\mathbf{x})| \leqq Q^{\gamma_i - \delta'} \qquad (i=1, \ldots, n)$$

where $\gamma_1 = \ldots = \gamma_m = -(n-m)/m$ and $\gamma_{m+1} = \ldots = \gamma_n = 1$. For these values of Q one has $\lambda_1 = \lambda_1(Q) < Q^{-\delta'}$. Since $\lambda_1 \leqq \ldots \leqq \lambda_n$ and $1 \ll \lambda_1 \ldots \lambda_n \ll 1$, there is a d with $1 \leqq d \leqq n-1$ and a $\delta'' > 0$ such that

$$\lambda_d < \lambda_{d+1} Q^{-\delta''} \tag{9.10}$$

for arbitrarily large values of Q. Let S^d be the subspace in the conclusion of Theorem 9E.

Let $\Pi^*(Q)$ be the intersection of $\Pi(Q)$ and S^d; this is a symmetric convex set in S^d. Let $\lambda_1^*, \ldots, \lambda_d^*$ be the successive minima of $\Pi^*(Q)$ with respect to the lattice Λ of integer points in S^d, and let $V^* = V^*(Q)$ be the (d-dimensional) volume of $\Pi^*(Q)$. By applying (8.3) to the lattice Λ we obtain

$$1 \ll \lambda_1^* \ldots \lambda_d^* V^* \ll 1, \tag{9.11}$$

where the constants in $\ll$ may depend on S^d. There are arbitrarily large values of Q for which $\mathbf{x}_1(Q), \ldots, \mathbf{x}_d(Q)$ lie in S^d, and for these values we have $\lambda_1 = \lambda_1^*, \ldots, \lambda_d = \lambda_d^*$, whence by (8.5) and (9.10),

$$\begin{aligned} \lambda_1^* \ldots \lambda_d^* &= \lambda_1 \ldots \lambda_d = (\lambda_1 \ldots \lambda_d)^{d/n} (\lambda_1 \ldots \lambda_d)^{(n-d)/n} \\ &< (\lambda_1 \ldots \lambda_d)^{d/n} (\lambda_{d+1} \ldots \lambda_n)^{d/n} Q^{-\delta'' d(n-d)/n} \ll Q^{-\delta'' d(n-d)/n} = Q^{-\eta}, \end{aligned}$$

say. In conjunction with (9.11) this yields $V^* \gg Q^{\eta}$.

Now if $L_1, \ldots, L_m$ have rank r on S^d, then

$$V^* \ll Q^{-(r(n-m)/m)+d-r} = Q^{d-(rn/m)}.$$

It follows that $d - (rn/m) \geqq \eta > 0$ and that

$$r < dm/n.$$

This cannot happen if (7.6) holds, and hence $L_1, \ldots, L_m$ is a Roth System in this case. Since the case of linearly dependent forms $L_1, \ldots, L_m$ is trivial and since the other half of the theorem was proved in §7.3, Theorem 7E is established.

10. Norm forms

10.1. Let K be an algebraic number field of degree t. There are t isomorphisms of K into the complex numbers; denote the images of an element α of K under these isomorphisms by $\alpha^{(1)}, \ldots, \alpha^{(t)}$. Let $L(\mathbf{x}) = \alpha_1 x_1 + \ldots + \alpha_n x_n$ be a linear form with coefficients in K. For $1 \leqq i \leqq t$ put $L^{(i)}(\mathbf{x}) = \alpha_1^{(i)} x_1 + \ldots + \alpha_n^{(i)} x_n$. The norm

$$\mathcal{N}\left(L(\mathbf{x})\right) = L^{(1)}(\mathbf{x}) \ldots L^{(t)}(\mathbf{x})$$

is a form of degree t with rational coefficients. A form obtained in this way will be called a *norm form*. It is easy to see that every form $F(\mathbf{x})$ which has rational coefficients and is irreducible over the rationals but which is a product of linear forms with algebraic coefficients, is a constant times a norm form. In particular when $n = 2$, every form with rational coefficients which is irreducible over the rationals is essentially a norm form.

10.2. We may as well discuss more general products of linear forms with real or complex algebraic coefficients. For any real or complex number α we denote its complex conjugate by $\bar{\alpha}$. The complex conjugate of a linear form $L(\mathbf{x}) = \alpha_1 x_1 + \ldots + \alpha_n x_n$ is defined by $\bar{L}(\mathbf{x}) = \bar{\alpha}_1 x_1 + \ldots + \bar{\alpha}_n x_n$. We shall call linear forms $L_1, \ldots, L_t$ a *Symmetric System* if, except for the ordering, $\bar{L}_1, \ldots, \bar{L}_t$ are the same as the given forms.

Theorem 10A (Schmidt, 1971b). *Suppose $L_1, \ldots, L_t$ is a Symmetric System of linear forms with algebraic coefficients. Suppose $\eta > 0$. The following two conditions are equivalent:*

(a) *There is a constant $c_1 = c_1(L_1, \ldots, L_t; \eta)$ and there are infinitely many integer points $\mathbf{x}$ with*

$$|L_1(\mathbf{x}) \ldots L_t(\mathbf{x})| \leqq c_1 |\mathbf{x}|^{t-\eta}.$$

(b) *There is a rational subspace S^d of dimension d with $1 \leqq d \leqq n$ and there is a Symmetric System of linear forms $L_{i_1}, \ldots, L_{i_m}$ with $1 \leqq m \leqq t$ and $i_1 < \ldots < i_m$ whose restrictions to S^d have rank r with*

$$r \leqq dm/\eta \quad \text{and} \quad r < d. \tag{10.1}$$

This theorem again contains Roth's Theorem. It can be deduced from the Subspace Theorem.

10.3. We shall now discuss diophantine equations

$$\mathcal{N}(L(\mathbf{x})) = a \tag{10.2}$$

where a is a constant. As $\mathbf{x}$ runs through the integer points, $\mathcal{N}(L(\mathbf{x}))$ runs through certain rationals with bounded denominators. Hence there are constants a for which (10.2) has infinitely many integer solutions $\mathbf{x}$ precisely if there are constants b for which the inequality

$$|\mathcal{N}(L(\mathbf{x}))| \leqq b \tag{10.3}$$

has infinitely many solutions. This will in fact be the case if the coefficients of L are linearly dependent over the rationals, so that we shall assume in the sequel that the coefficients are linearly independent.

We shall say that a linear form with coefficients in K is *full* in K if its coefficients form a field basis of K. Suppose the linear form $L(\mathbf{x})$ is full in K where K is neither the rational field nor an imaginary quadratic field. Further assume for a moment that the coefficients of L form in fact an integer basis of K. By Dirichlet's unit theorem K contains infinitely many units, and hence there are infinitely many integer points $\mathbf{x}$ with $\mathcal{N}(L(\mathbf{x})) = 1$. By studying units of certain subrings of K one sees more generally that *if $L(\mathbf{x})$ is full in K where K is not rational or imaginary quadratic, then* (10.3) *has infinitely many solutions if b is large enough.* We shall say that a linear form $L(\mathbf{x})$ *represents* a linear form $L'(\mathbf{y})$ (where the number of components of $\mathbf{y}$ need not be n) if there is a constant c such that for every integer point $\mathbf{y}$ there is an integer point $\mathbf{x}$ with $L'(\mathbf{y}) = cL(\mathbf{x})$. Now suppose $L(\mathbf{x})$ is a linear form with coefficients in K. We shall call $L(\mathbf{x})$ *degenerate* if it represents a linear form $L'(\mathbf{y})$ which is full in a subfield K' of K which is neither rational nor imaginary quadratic. For example, $L(\mathbf{x}) = \sqrt{2}\,x_1 + \sqrt{3}\,x_2 + \sqrt{6}\,x_3$ is not full in $K = \mathbf{Q}(\sqrt{2}, \sqrt{3})$, but it represents the form $2y_1 + \sqrt{6}y_2 = \sqrt{2}(\sqrt{2}y_1 + \sqrt{3}y_2)$ which is full in $K' = \mathbf{Q}(\sqrt{6})$ and thus $L(\mathbf{x})$ is degenerate. From what we said above it follows that *for degenerate $L(\mathbf{x})$ and for large b the inequality* (10.3) *has infinitely many integer solutions.* A detailed proof may be found, e.g., in Borevich and Shafarevich (1966, ch. 2).

10.4. The converse also holds:

THEOREM 10B (Schmidt, 1971b). *Suppose $L(\mathbf{x})$ is a non-degenerate linear form with linearly independent coefficients in a number field K. Then for*

any fixed b, the inequality (10.3) *has only finitely many solutions in integer points* **x**.

When the number of variables $n = 2$, this becomes Thue's result on his equation $F(x, y) = m$ where $F(x, y)$ is a binary form. When $n = 3$ the theorem was shown by Skolem (1935) when K has degree $t = 5$ and by Chabauty (1938) for general degree, but these authors made the additional assumption that among the isomorphisms of K into the complex numbers there are at least two pairs of complex conjugates. Skolem and Chabauty used a p-adic method. The general case $n = 3$ was settled by Schmidt (1967b). Before the results of §7 were known, Győry (1968) assumed the hypothetical truth of Corollary 7B and derived results about norm forms in an arbitrary number of variables. Since he did not have Theorem 10A as a tool, his results are relatively weak. See also Győry (1969), where he proves some special cases of Theorem 10B. Ramachandra (1969) dealt with special norm forms and derived for them an asymptotic formula for the number of solutions of (10.3), thus generalizing Mahler's (1933c) result.

Theorem 10B and the theorems of Skolem and Chabauty are non-effective. Effective bounds for the size of the solutions of certain rather special equations with norm forms were given by Skolem (1937) (for further references see Skolem (1938)) and Feldman (1970b). See also the references given at the end of §7.2.

If $L(\mathbf{x})$ is full in K where K is neither rational nor imaginary quadratic, then the solutions of an equation (10.2) may be parametrized by using the group of units of K. More generally one can show that if $L(\mathbf{x})$ is degenerate, then all solutions of (10.2) with finitely many exceptions belong to finitely many parameter families. For example, in the equation

$$\mathcal{N}(\sqrt{2}x_1 + \sqrt{3}x_2 + \sqrt{6}x_3) = a\,, \tag{10.4}$$

all but finitely many solutions have $x_3 = 0$ or $x_2 = 0$ or $x_1 = 0$ and hence come from solutions of $\mathcal{N}(\sqrt{2}x_1 + \sqrt{3}x_2) = a$ or $\mathcal{N}(\sqrt{2}x_1 + \sqrt{6}x_3) = a$ or $\mathcal{N}(\sqrt{3}x_2 + \sqrt{6}x_3) = a$. Hence all but finitely many solutions come from one of the three equations

$$\mathcal{N}'(2x_1 + \sqrt{6}x_2) = \pm 2\sqrt{a}, \quad \mathcal{N}''(x_1 + \sqrt{3}x_3) = \pm\frac{1}{2}\sqrt{a},$$

$$\mathcal{N}'''(x_2 + \sqrt{2}x_3) = \pm\frac{1}{3}\sqrt{a}\,,$$

where $\mathcal{N}', \mathcal{N}'', \mathcal{N}'''$ are the norms from the fields $K' = \mathbf{Q}(\sqrt{6})$, $K'' = \mathbf{Q}(\sqrt{3})$ and $K''' = \mathbf{Q}(\sqrt{2})$. The solutions of these three equations can be easily described in terms of the units of the fields K', K'' and K'''. In particular (10.4) has only finitely many solutions unless a is a perfect square.

10.5. Roth's Theorem implied not only that for an irreducible binary form $F(x, y)$ of degree $t \geqq 3$ there are only finitely many solutions of $|F(x, y)| < a$, but according to Theorem 2C there are only finitely many solutions of $|F(x, y)| < (|x|+|y|)^v$ if $v < t - 2$. In the present context it is reasonable to expect that " in general " there are only finitely many integer points $\mathbf{x}$ with

$$|\mathcal{N}(L(\mathbf{x}))| < |\mathbf{x}|^v \tag{10.5}$$

if

$$v < t - n\,. \tag{10.6}$$

Using Minkowski's Linear Forms Theorem one can easily show that unless $n = 1$ or $n = 2$ and no conjugate of L has real coefficients, there are infinitely many $\mathbf{x}$ with $|\mathcal{N}(L(\mathbf{x}))| \leqq c\,|\mathbf{x}|^{t-n}$; hence $t - n$ in (10.6) is best possible.

Suppose $K = \mathbf{Q}(\alpha)$ is a number field of degree t and suppose $1 \leqq r \leqq t$. We shall say that K is *r times transitive* if for any r distinct conjugates $\alpha^{(i_1)}, \ldots, \alpha^{(i_r)}$ of α there is an element φ of the Galois group of $\mathbf{Q}(\alpha^{(1)}, \ldots, \alpha^{(t)})$ (i.e. the least normal extension of K) with $\varphi(\alpha^{(1)}) = \alpha^{(i_1)}, \ldots, \varphi(\alpha^{(r)}) = \alpha^{(i_r)}$. This definition is clearly independent of the primitive element α.

THEOREM 10C (Schmidt, in preparation). *Suppose the coefficients of $L(\mathbf{x}) = \alpha_1 x_1 + \ldots + \alpha_n x_n$ lie in a number field K and are linearly independent over the rationals. Suppose that K is generated by the quotients α_i/α_j $(1 \leqq i, j \leqq n)$ and that K is $(n-1)$-times transitive. Finally assume that any n of the conjugates of $L(\mathbf{x})$ are linearly independent. Then for every v with* (10.6) *there are only finitely many integer points $\mathbf{x}$ satisfying* (10.5).

COROLLARY 10D. *Suppose $L(\mathbf{x})$ is as above and suppose $G(\mathbf{x})$ is a polynomial of total degree $v < t - n$. Then the equation*

$$\mathcal{N}(L(\mathbf{x})) = G(\mathbf{x})$$

has only finitely many integer solutions.

This contains Corollary 2D.

10.6. Both Theorems 10B and 10C are derived from Theorem 10A. We shall briefly discuss the argument for Theorem 10B. We have to show that $L(\mathbf{x})$ is degenerate if the inequality

$$|\mathcal{N}(L(\mathbf{x}))| = |L^{(1)}(\mathbf{x}) \dots L^{(t)}(\mathbf{x})| \leqq c = c\,|\mathbf{x}|^{t-t}$$

has infinitely many solutions. By the case $\eta = t$ of the assertion (a) $\Rightarrow$ (b) of Theorem 10A there is a subspace S^d and there is a Symmetric System $L^{(i_1)}, \dots, L^{(i_m)}$ of forms whose restrictions to S^d have a rank r with

$$r \leqq dm/t \quad \textit{and} \quad r < d\,. \tag{10.7}$$

One can reduce the situation to the special case where $L(\mathbf{x}) = x_1 + \alpha_2 x_2 + \dots + \alpha_n x_n$ and $K = \mathbf{Q}(\alpha_2, \dots, \alpha_n)$, and where $d = n$. The conditions (10.7) now become $r \leqq nm/t$ and $r < n$, and with

$$q = t/n$$

they become

$$rq \leqq m \quad \textit{and} \quad r < n\,. \tag{10.8}$$

Now $rq < m$ is impossible (this would imply infinitely many solutions of $|\mathcal{N}(L(\mathbf{x}))| < |\mathbf{x}|^{-\delta}$ for some $\delta > 0$), and hence the rank r of every Symmetric System $L^{(i_1)}, \dots, L^{(i_m)}$ satisfies

$$m \leqq rq\,. \tag{10.9}$$

But by (10.8) there is a special Symmetric System $L^{(i_1)}, \dots, L^{(i_\mu)}$ of rank ρ with

$$\mu = \rho q \quad \textit{and} \quad \rho < n\,. \tag{10.10}$$

We choose μ and ρ as small as possible with this property. We may assume without loss of generality that the forms $L^{(i_1)}, \dots, L^{(i_\mu)}$ are $L^{(1)}, \dots, L^{(\mu)}$.

In what follows, α will be a primitive element of K, i.e. an element with $K = \mathbf{Q}(\alpha)$. We have to distinguish two cases.

(A) For every element φ of the Galois group of $\mathbf{Q}(\alpha^{(1)}, \dots, \alpha^{(t)})$, the two sets $\{\alpha^{(1)}, \dots, \alpha^{(\mu)}\}$ and $\{\varphi(\alpha^{(1)}), \dots, \varphi(\alpha^{(\mu)})\}$ are identical or disjoint.

(B) We have not (A).

In the case (A) it turns out that μ divides t and that $L(\mathbf{x})$ represents a full linear form $L'(\mathbf{y})$ in a field K' of degree t/μ, where K' is neither rational nor imaginary quadratic, and hence $L(\mathbf{x})$ is degenerate. Let us

discuss what happens in the case (B). For simplicity we shall assume that K is totally real.

There is an element φ of the Galois group such that the sets $\{\alpha^{(1)}, \ldots, \alpha^{(\mu)}\}$ and $\{\varphi(\alpha^{(1)}), \ldots, \varphi(\alpha^{(\mu)})\}$ are neither identical nor disjoint. We may assume without loss of generality that

$$\{\varphi(\alpha^{(1)}), \ldots, \varphi(\alpha^{(\mu)})\} = \{\alpha^{(1)}, \ldots, \alpha^{(l)}, \alpha^{(\mu+1)}, \ldots, \alpha^{(2\mu-l)}\}.$$

Here $1 \leqq l \leqq \mu - 1$. The forms $L^{(1)}, \ldots, L^{(\mu)}$ have rank ρ, and hence also $L^{(1)}, \ldots, L^{(l)}, L^{(\mu+1)}, \ldots, L^{(2\mu-l)}$ have rank ρ. Denote the rank of $L^{(1)}, \ldots, L^{(l)}$ by r_1 and the rank of $L^{(1)}, \ldots, L^{(\mu)}, \ldots, L^{(2\mu-l)}$ by r_2. It is easily seen that $r_2 \leqq 2\rho - r_1$, i.e. that

$$r_1 + r_2 \leqq 2\rho .$$

Since μ was chosen as small as possible with (10.10), and since $l \leqq \mu - 1$, we have $l < r_1 q$. The number $2\mu - l$ of elements of $L^{(1)}, \ldots, L^{(\mu)}, \ldots, L^{(2\mu-l)}$ satisfies $2\mu - l \leqq r_2 q$ by (10.9). Thus

$$2\mu = l + (2\mu - l) < r_1 q + r_2 q \leqq 2\rho q ,$$

which contradicts (10.10). Hence (B) is impossible if K is totally real.

We have in fact used the hypothesis that K is totally real, for in general $L^{(1)}, \ldots, L^{(l)}$ need not be a Symmetric System, and $l < r_1 q$ need not hold. The situation is therefore somewhat more complicated if K is not totally real.

11. Generalizations and open problems

11.1. The theorems of §7 and §10 can almost certainly be generalized to include p-adic valuations. I understand that work on this question is being done now. (p-adic versions of the results of §2 were discussed in §4.5). Next, suppose that K is an algebraic number field and that $\alpha_1, \ldots, \alpha_l$ are algebraic numbers such that $1, \alpha_1, \ldots, \alpha_l$ are linearly independent over K. It is likely that *for every $\delta > 0$ there are only finitely many l-tuples of elements $\beta_1, \ldots, \beta_l$ of K with*

$$|\alpha_i - \beta_i| < \mathscr{H}(\beta)^{-1-(1/l)-\delta} \quad (i = 1, \ldots, l), \tag{11.1}$$

where $\mathscr{H}(\beta)$ is a suitably defined height of $\beta = (\beta_1, \ldots, \beta_l)$. A possible definition for $\mathscr{H}(\beta)$ is

$$\mathscr{H}(\beta) = \prod_v \max(1, \|\beta_1\|_v, \ldots, \|\beta_l\|_v),$$

where v runs through the valuations of K and where $\| \|_v$ is defined as in §4.5. In view of (4.9), $\mathscr{H}(\beta)$ is almost the same as $H_K(\beta)$ if $l = 1$, and hence a theorem on (11.1) would generalize Le Veque's Theorem 4A. One could try to obtain a still more general theorem which would contain both the p-adic case and the case of a number field K. Such a result was put forward as a conjecture by Lang (1962, Ch. 6).

11.2. We have already said in §2.2 that it would be desirable to replace the factor q^δ in Roth's Theorem by something smaller, say by a power of $\log q$. The same is true of the generalizations of Roth's Theorem to simultaneous approximation.

The theorems of §2, 7 and 10 are non-effective. For approximation to a single algebraic number α there are the effective results of Baker (see §5), but for simultaneous approximation there are only the relatively special effective theorems of Baker (1967a), Feldman (1970a, 1970b) and Osgood (1970).

11.3. The following questions also appear to be very difficult. Suppose $(\alpha_1, \ldots, \alpha_l)$ is a point of transcendence degree $d < l$. The theorems of §7 deal with the case when the point is algebraic, i.e. when $d = 0$. What can one say for other values of d? Perron (1932) and Schmidt (1962) obtained results, of about the same level of sophistication as Liouville's Theorem, which can be used to show that certain given points have transcendence degree l.

A better question perhaps is how close rational points can come to a given algebraic variety. We may reformulate this question in a homogeneous setting. Let V be a homogeneous variety defined over the rationals (i.e. one defined by homogeneous polynomial equations with rational coefficients) in E^n with $n \geqq 2$. For every $\mathbf{x} \neq \mathbf{0}$ we put

$$\psi(V, \mathbf{x}) = \Delta(V, \mathbf{x}) \, |\mathbf{x}|^{-1}$$

where $\Delta(V, \mathbf{x})$ is the distance from $\mathbf{x}$ to V. It is clear that $\psi(V, \lambda\mathbf{x}) = \psi(V, \mathbf{x})$; the function $\psi(V, \mathbf{x})$ may be interpreted as the "angle" between V and the vector $\mathbf{x}$. We are interested in inequalities of the type

$$\psi(V, \mathbf{x}) < c \, |\mathbf{x}|^{-\omega} \tag{11.2}$$

where $\mathbf{x}$ runs through the integer points. We saw in §6 that Theorems 6A and 6C had such an interpretation. The best value of ω for which (11.2) has

infinitely many integer solutions can always be found if V is linear, i.e. is a subspace. For the non-linear case we have neither a good generalization of Dirichlet's Theorem nor anything like Roth's Theorem.

Suppose now that V is a hypersurface containing no integer point $\mathbf{x} \neq \mathbf{0}$ and defined by the equation $F(\mathbf{x}) = 0$ where F is a form of degree d with rational integer coefficients. For every integer point $\mathbf{x} \neq \mathbf{0}$ we have $|F(\mathbf{x})| \geqq 1$, and since $\left|\frac{\partial}{\partial x_i} F(\mathbf{x})\right| \leqq c_1 |\mathbf{x}|^{d-1}$ $(i=1, \ldots, n)$, the distance from $\mathbf{x}$ to V is $\geqq c_2 |\mathbf{x}|^{1-d}$, which in turn implies that

$$\psi(V, \mathbf{x}) \geqq c_3 |\mathbf{x}|^{-d},$$

where the constants depend only on V. This inequality may be interpreted as a generalization of Liouville's Theorem. Any improvement of this inequality, even though perhaps it may apply only to special classes of non-linear hypersurfaces, would be of great interest and would shed light on certain diophantine equations different from the equations with norm forms discussed in §10.

REFERENCES

ADAMS, W. W. (1967). Simultaneous asymptotic diophantine approximations. *Mathematika 14*, 173-180.

—— (1969a). Simultaneous asymptotic diophantine approximations to a basis of a real number field. *J. Number Theory 1*, 179-194.

—— (1969b). Simultaneous diophantine approximations and cubic irrationals. *Pacific J. Math. 30*, 1-14.

—— (To appear). Simultaneous Diophantine Approximations to a Basis of a Real Number Field. *Nagoya Math. J. 42.*

BAKER, A. (1962). Continued fractions of transcendental numbers. *Mathematika 9*, 1-8.

—— (1964a). Rational approximations to certain algebraic numbers. *Proc. London Math. Soc. (3) 14*, 385-398.

—— (1964b). Rational approximations to $\sqrt[3]{2}$ and other algebraic numbers. *Quart. J. Math. Oxford Ser.* (2) *15*, 375-383.

—— (1965). On some Diophantine inequalities involving the exponential function. *Can. J. Math. 17*, 616-626.

—— (1966). Linear forms in the logarithms of algebraic numbers. *Mathematika 13*, 204-216.

—— (1967a). Simultaneous approximations to certain algebraic numbers. *Proc. Camb. Phil. Soc. 63*, 693-702.

—— (1967b). Linear forms in the logarithms of algebraic numbers (II). *Mathematika 14*, 102-107.

BAKER, A. (1967c). Linear forms etc. (III). *Mathematika 14*, 220-224.

—— (1968a). Linear forms etc. (IV). *Mathematika 15*, 204-216.

—— (1968b). Contributions to the theory of diophantine equations (I). On the representation of integers by binary forms. *Phil. Trans. Royal Soc. London A 263*, 173-191.

—— (1968c). Contributions etc. (II). The diophantine equation $y^2 = x^3 + k$. *Ibid.*, 193-208.

—— (1968d). The diophantine equation $y^2 = ax^3 + bx^2 + cx + d$. *J. London Math. Soc. 43*, 1-9.

—— (1969). Bounds for the solutions of the hyperelliptic equation. *Proc. Camb. Phil. Soc. 65*, 439-444.

—— (1971). Effective methods in diophantine problems. *Proc. of Symp. in Pure Math. XX* (1969 Number Theory Institute), 195-205.

—— (to appear). A sharpening of the bounds for linear forms in logarithms. *Acta Arith.*

BAKER, A. and J. COATES (1970). Integer points on curves of genus 1. *Proc. Camb. Phil. Soc. 67*, 595-602.

BAKER, A. and H. M. STARK (1971). On a fundamental inequality in number theory. *Annals of Math. 94*, 190-199.

BOREVICH, Z. I. and I. R. SHAFAREVICH (1966). Number theory. (Translated from the Russian (1964) ed. *Academic Press*: New York and London.

BRYUNO, A. D. (1964). The expansion of algebraic numbers in continued fractions (Russian). *Zh. Vychisl. Mat. i Mat. Fiz. 4*, 211-221.

CASSELS, J. W. S. (1955). Simultaneous diophantine approximation. *J. London Math. Soc. 30*, 119-121.

—— (1957). An introduction to diophantine approximation. *Cambridge Tracts 45*, Cambridge University Press.

—— (1959). An introduction to the geometry of numbers. Grundlehren 99. *Springer Verlag*: Berlin-Göttingen-Heidelberg.

CASSELS, J. W. S. and H. P. F. SWINNERTON-DYER (1955). On the product of three homogeneous linear forms and indefinite ternary quadratic forms. *Philos. Trans. Roy. Soc. London Ser. A 248*, 73-96.

CHABAUTY, C. (1938). Sur les équations diophantinnes liées aux unités d'un corps de nombres algébriques fini. *Ann. Mat. Pura Appl. 17*, 127-168.

COATES, J. (1969). An effective p-adic analogue of a Theorem of Thue. *Acta Arith. 15*, 279-305.

—— (1970a). An effective etc. (II). The greatest prime factor of a binary form. *Acta Arith. 16*, 399-412.

—— (1970b). An effective etc. (III). The diophantine equation $y^2 = x^3 + k$. *Acta Arith. 16*, 425-435.

CUGIANI, M. (1959). Sulla approssimabilità dei numeri algebrici mediante numeri razionali. *Ann. Mat. Pura Appl. (4) 48*, 135-145.

DAVENPORT, H. (1937). Note on a result of Siegel. *Acta Arith. 2*, 262-265.

—— (1968). A note on Thue's Theorem. *Mathematika 15*, 76-87.

DAVENPORT, H. and K. F. ROTH (1955). Rational approximations to algebraic numbers. *Mathematika 2*, 160-167.

DAVENPORT, H. and W. M. SCHMIDT (1967). Approximation to real numbers by quadratic irrationals. *Acta Arith. 13*, 169-176.

—— (1969). Approximation to real numbers by algebraic integers. *Acta Arith. 15*, 393-416.

DIRICHLET, L. G. P. (1842). Verallgemeinerung eines Satzes aus der Lehre von den Kettenbrüchen nebst einigen Anwendungen auf die Theorie der Zahlen. *S. B. Preuss Akad. Wiss.* 93-95.

DYSON, F. J. (1947). The approximation to algebraic numbers by rationals. *Acta Math. 79*, 225-240.

FELDMAN, N. I. (1968a). Estimate for a linear form of logarithms of algebraic numbers (Russian). *Mat. Sbornik 76 (118)*, 304-319. *English Transl. Math. USSR Sbornik 5*, 291-307.

—— (1968b). An improvement of the estimate of a linear form in the logarithms of algebraic numbers (Russian). *Mat. Sbornik 77 (119)*, 423-436. *English Transl. Math. USSR Sbornik 6*, 393-406.

—— (1969). A certain inequality for a linear form in the logarithms of algebraic numbers (Russian). *Mat. Zametki 5*, 681-689.

—— (1970a). Bounds for linear forms of certain algebraic numbers (Russian). *Mat. Zametki 7*, 569-580. English transl. *Math. Notes 7* (1970) , 343-349.

—— (1970b). Effective bounds for the size of the solutions of certain diophantine equations (Russian). *Mat. Zametki 8*, 361-371.

FELDMAN, N. I. and A. B. SHIDLOVSKII (1967). The development and the present state of the theory of transcendental numbers. *Russian Math. Surveys 22*, 1-79.

FRAENKEL, A. S. (1962). On a theorem of Ridout in the theory of diophantine approximations. *Trans. Am. Math. Soc. 105*, 84-101.

GELFOND, A. O. (1952). Transcendental and algebraic numbers (Russian). (English transl. (1960), *Dover Publications*: New York.)

GYÖRY, K. (1968). Sur une classe des équations diophantinnes. *Publ. Math. Debrecen 15*, 165-179.

—— (1969). Représentation des nombres par des formes décomposables. I. *Publ. Math. Debrecen 16*, 253-263.

HASSE, H. (1939). Simultane Approximation algebraischer Zahlen durch algebraische Zahlen. *Monatsh. Math. 48*, 205-225.

HOOLEY, C. (1967). On binary cubic forms. *Journal f. Math. 226*, 30-87.

HURWITZ, A. (1891). Über die angenäherte Darstellung der Irrationalzahlen durch rationale Brüche. *Math. Ann. 39*, 279-284.

HYYRÖ, S. (1964). Über die Gleichung $ax^n - by^n = z$ und das Catalansche Problem. *Ann. Acad. Fennicae*, Ser. AI 355.

KEATES, M. (1969). On the greatest prime factor of a polynomial. *Proc. Edinb. Math. Soc. (2) 16*, 301-303.

KHINTCHINE, A. (1925). Zwei Bemerkungen zu einer Arbeit des Herrn Perron. *Math. Zeitschr. 22*, 274-284.

—— (1926a). Über eine Klasse linearer Diophantischer Approximationen. *Rend. Circ. Mat. Palermo 50*, 170-195.

—— (1926b). Zur metrischen Theorie der diophantischen Approximationen. *Math. Z. 24*, 706-714.

KOKSMA, J. F. (1936). Diophantische Approximationen. *Ergebnisse d. Math. u. Grenzgeb. 4*. Springer Verlag: Berlin.

—— (1939). Über die Mahlersche Klasseneinteilung der transzendenten Zahlen und die Approximation komplexer Zahlen durch algebraische Zahlen. *Mh. Math. Phys. 48*, 176-189.

LANG, S. (1962). Diophantine Geometry. Interscience tracts in pure and applied math. 11. J. Wiley & Sons: New York — London.

—— (1965a). Report on diophantine approximations. *Bull. de la Soc. Math. de France 93*, 117-192.

—— (1965b). Asymptotic approximation to quadratic irrationalities. *Am. J. Math. 87*, 481-487.

—— (1965c). Asymptotic approximation etc (II). *Ibid.*, 488-496.

LANG, S. (1966a). Asymptotic diophantine approximation. *Proc. of the Nat. Acad. of Sci. 55*, 31-34.

—— (1966b). Introduction to diophantine approximations. Addison-Wesley Publ. Co.: Reading, Mass.

—— (1971). Transcendental numbers and diophantine approximations. *Bull. Am. Math. Soc. 77*, 635-677.

LEKKERKERKER, C. G. (1969). Geometry of Numbers. Wolters-Noordhoff Publishing: Groningen.

LE VEQUE, W. J. (1955). Topics in number theory. Addison-Wesley Publ. Co.: Reading, Mass.

LIOUVILLE, J. (1844). Sur des classes très étendues de quantités dont la valeur n^e est ni algébrique, ni même réductible à des irrationnelles algébriques. *C. R. Acad. Sci. Paris 18*, 883-885 and 910-911.

LUTZ, E. (1955). Sur les approximations diophantinnes linéaires *P*-adiques. *Actualités scient. et ind. N° 1224*, Paris.

MAHLER, K. (1932). Zur Approximation der Exponentialfunktion und des Logarithmus I. *J. reine ang. Math. 166*, 118-136.

—— (1933a). Zur Approximation algebraischer Zahlen (I). Über den grössten Primteiler binärer Formen. *Math. Ann. 107*, 691-730.

—— (1933b). Zur Approximation etc. (II). Über die Anzahl der Darstellungen ganzer Zahlen durch Binärformen. *Math. Ann. 108*, 37-55.

—— (1933c). Zur Approximation etc. (III). Über die mittlere Anzahl der Darstellungen grosser Zahlen durch binäre Formen. *Acta Math. 62*, 91-166.

—— (1936). Ein Analogon zu einem Schneiderschen Satz. *Nederl. Akad. Wetensch. Proc. 39*, 633-640 and 729-737.

—— (1939). Ein Übertragungsprinzip für lineare Ungleichungen. *Cas. Pest Mat. 68*, 85-92.

—— (1953). On the approximation of π. *Proc. Akad. Wetensch. Ser. A 56*, 30-42.

—— (1955). On compound convex bodies (I). *Proc. London Math. Soc. (3) 5*, 358-379.

—— (1961). Lectures on diophantine approximation. Notre Dame University.

—— (1963). On the approximation of algebraic numbers by algebraic integers. *J. Austral. Math. Soc. 3*, 408-434.

MINKOWSKI, H. (1907). Diophantische Approximationen. Teubner: Leipzig u. Berlin.

—— (1896) und (1910). Geometrie der Zahlen. Teubner: Leipzig u. Berlin. (The 1910 ed. prepared posthumously by Hilbert and Speiser).

NEUMANN, J. VON and B. TUCKERMAN (1955). Continued fraction expansion of $2^{1/3}$. *Math. Tables Aids Comp. 9*, 23-24.

NIVEN, I. (1963). Diophantine Approximations. Interscience tracts in pure and applied math. 14. J. Wiley & Sons: New York — London.

OSGOOD, C. F. (1970). The simultaneous approximation of certain *k-th* roots. *Proc. Camb. Phil. Soc. 67*, 75-86.

PARRY, C. J. (1940). The *p*-adic generalization of the Thue-Siegel theorem. *J. London Math. Soc. 15*, 293-305.

—— (1950). The $\mathfrak{p}$-adic generalization of the Thue-Siegel theorem. *Acta Math. 83*, 1-100.

PECK, G. (1961). Simultaneous rational approximations to algebraic numbers. *Bull. Am. Math. Soc. 67*, 197-201.

PERRON, O. (1921). Über diophantische Approximationen. *Math. Ann. 83*, 77-84.

—— (1932). Über mehrfach transzendente Erweiterungen des natürlichen Rationalitätsbereiches. Sitzungsber. Bayer. Akad. Wiss. H 2, 79-86.

—— (1954). Die Lehre von den Kettenbrüchen. 3. Aufl. B. G. Teubner: Stuttgart.

POPKEN, J. (1929). Zur Transzendenz von e. *Math. Z. 29*, 525-541.

RAMACHANDRA, K. (1966). Approximation of algebraic numbers. Nachrichten d. Akad. d. Wiss. in Göttingen, *Math.-Phys. Kl*, 45-52.

—— (1969). A lattice point problem for norm forms in several variables. *J. Number Theory 1*, 534-555.

RICHTMYER, R. D., M. DEVANY and N. METROPOLIS (1962). Continued fraction expansions of algebraic numbers. *Numerische Math. 4*, 68-84.

RIDOUT, D. (1957). Rational approximations to algebraic numbers. *Mathematika 4*, 125-131.

—— (1958). The p-adic generalization of the Thue-Siegel-Roth Theorem. *Mathematika 5*, 40-48.

ROTH, K. F. (1955a). Rational approximations to algebraic numbers. *Mathematika 2*, 1-20.

—— (1955b). Corrigendum. *Ibid.*, 168.

SCHINZEL, A. (1967). Review of a paper by Hyyrö. *Zentralblatt Math. 137*, 257-258.

—— (1968). An improvement of Runge's Theorem on diophantine equations. *Commentarii Pontif. Acad. Soc. 2*, No. 20.

SCHMIDT, W. M. (1962). Simultaneous approximation and algebraic independence of numbers. *Bull. Am. Math. Soc. 68*, 475-478.

—— (1965). Über simultane Approximation algebraischer Zahlen durch rationale. *Acta Math. 114*, 159-206.

—— (1966). Simultaneous approximation to a basis of a real number field. *Amer. J. Math. 88*, 517-527.

—— (1967a). On simultaneous approximation of two algebraic numbers by rationals. *Acta Math. 119*, 27-50.

—— (1967b). Some diophantine equations in three variables with only finitely many solutions. *Mathematika 14*, 113-120.

—— (1970). Simultaneous approximation to algebraic numbers by rationals. *Acta Math. 125*, 189-201.

—— (1971a). Linear forms with algebraic coefficients. I. *J. of Number Theory 3*, 253-277.

—— (1971b). Linearformen mit algebraischen Koeffizienten. II. *Math. Ann. 191*, 1-20.

—— (in preparation). Norm form equations.

SCHNEIDER, Th. (1936). Über die Approximation algebraischer Zahlen. *J. reine angew. Math. 175*, 182-192.

—— (1957). Einführung in die transzendenten Zahlen. Grundlehren 81. Springer Verlag: Berlin-Göttingen-Heidelberg.

SIEGEL, C. L. (1921a). Approximation algebraischer Zahlen. *Math. Zeitschr. 10*, 173-213.

—— (1921b). Über Näherungswerte algebraischer Zahlen. *Math. Ann. 84*, 80-99.

—— (1929). Über einige Anwendungen diophantischer Approximationen. Abh. d. Preuss. Akad. d. Wiss., *Math. Phys. Kl.*, Nr. 1.

—— (1937). Die Gleichung $ax^n - by^n = c$. *Math. Ann. 114*, 57-88.

—— (1970). Einige Erläuterungen zu Thues Untersuchungen über Annäherungswerte algebraischer Zahlen und diophantische Gleichungen. Nachr. Akad. d. Wiss. Göttingen, *Math. Phys. Kl.*, Nr. 8.

SKOLEM, Th. (1935). Einige Sätze über p-adische Potenzreihen mit Anwendung auf gewisse exponentielle Gleichungen. *Math. Ann. 111*, 399-424.

—— (1937). Anwendung exponentieller Kongruenzen zum Beweis der Unlösbarkeit gewisser diophantischer Gleichungen. *Vid. Akad. Avh. Oslo I*, Nr. 12.

—— (1938). Diophantische Gleichungen. Ergebnisse d. Math. 5. Springer Verlag: Berlin.

SPRINDZUK, V. G. (1969). Effective estimates in " ternary " exponential diophantine equations (Russian). *Dokl. Akad. Nauk Belorusskoj SSR 13, No. 9*, 777-780.

—— (1970a). A new application of p-adic analysis on the representation of integers by binary forms (Russian). *Istvestia Akad. Nauk SSR, ser. math. 34, No. 5*, 1038-1063.

SPRINDZUK, V. G. (1970b). An effective estimate of rational approximations to algebraic numbers (Russian). *Dokl. Akad. Nauk Belorusskoj SSR 14, No. 8*, 681-684.

—— (1971a). An improvement of the estimate of rational approximations to algebraic numbers (Russian). *Dokl. Akad. Nauk Belorusskoj SSR 15, No. 2*, 101-104.

—— (1971b). On the greatest prime divisor of a binary form (Russian). *Ibid.*, No. 5, 389-391.

—— (1971c). Rational approximations to algebraic numbers (Russian). *Istvestia Akad. Nauk SSR 5.*

STEPANOW, S. A. (1967). The approximation of an algebraic number by algebraic numbers of a special form (Russian). Vestnik Moskov. Univ., Ser. I, *Math. Meh. 22, No. 6*, 78-86.

THUE, A. (1908). Bemerkungen über gewisse Näherungsbrüche algebraischer Zahlen. Über rationale Annäherungswerte der reellen Wurzel der ganzen Funktion dritten Grades $x^3 - ax - b$. On en general i store hele tal uløsbar ligning. Skrifter udgivne of Videnskabs-Selskabet i Christiania.

—— (1909). Über Annäherungswerte algebraischer Zahlen. *Journal f. Math. 135*, 284-305.

TIJDEMAN, R. (1971). On the algebraic independence of certain numbers. *Indag. Math. 33*, 146-162.

VINOGRADOV, A. I. and V. G. SPRINDZUK (1968). The representation of numbers by binary forms (Russian). *Mat. Zametki 3*, 369-376.

WALLISER, R. (1969). Zur Approximation algebraischer Zahlen durch arithmetisch charakterisierte algebraische Zahlen. *Arch. Math. (Basel) 20*, 384-391.

WIRSING, E. (1961). Approximation mit algebraischen Zahlen beschränkten Grades. *J. reine ang. Math. 206*, 67-77.

—— (1971). On approximations of algebraic numbers by algebraic numbers of bounded degree. *Proc. of Symp. in Pure Math. XX.* (1969 Number Theory Institute) 213-247.

(Reçu le 6 juillet 1971)

Wolfgang M. Schmidt
Department of Mathematics
University of Colorado
Boulder, Colorado 80302