

Editors:

A. Dold, Heidelberg

B. Eckmann, Zürich

Springer

Berlin

Heidelberg

New York

Barcelona

Budapest

Hong Kong

London

Milan

Paris

Santa Clara

Singapore

Tokyo

Wolfgang M. Schmidt

Diophantine Approximation



Springer

Author

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

1st edition 1980
2nd printing 1996 (with minor corrections)

Mathematics Subject Classification (1980): 10B16, 10E05, 10E15, 10E40,
10F05, 10F10, 10F20, 10F30, 10K15

Library of Congress Cataloging in Publication Data.

Schmidt, Wolfgang M., Diophantine approximation. (Lecture notes in mathematics ; 785)
Bibliography: p. Includes index. 1. Algebraic number theory. 2. Approximation, Diophantine.
I. Title. II. Series: Lecture notes in mathematics (Berlin) ; 785. QA3.L28 no. 785 [QA247]
510s [512'.74] 80-11695
ISBN 0-387-09762-7

ISBN 3-540-09762-7 Springer-Verlag Berlin Heidelberg New York

This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically those of translation, reprinting, re-use of illustrations, broadcasting, reproduction by photocopying machine or similar means, and storage in data banks. Under § 54 of the German Copyright Law, where copies are made for other than private use, a fee is payable to the publisher, the amount of the fee to be determined by agreement with the publisher.

© Springer-Verlag Berlin Heidelberg 1980
Printed in Germany

SPIN: 10530138 46/2142-543210 - Printed on acid-free paper

Preface

In spring 1970 I gave a course in Diophantine Approximation at the University of Colorado, which culminated in simultaneous approximation to algebraic numbers. A limited supply of mimeographed Lecture Notes was soon gone. The completion of these new Notes was greatly delayed by my decision to add further material.

The present chapter on simultaneous approximations to algebraic numbers is much more general than the one in the original Notes. This generality is necessary to supply a basis for the subsequent chapter on norm form equations. There is a new last chapter on approximation by algebraic numbers. I wish to thank all those, in particular Professor C.L. Siegel, who have pointed out a number of mistakes in the original Notes. I hope that not too many new mistakes have crept into these new Notes.

The present Notes contain only a small part of the theory of Diophantine Approximation. The main emphasis is on approximation to algebraic numbers. But even here not everything is included. I follow the approach which was initiated by Thue in 1908, and further developed by Siegel and by Roth, but I do not include the effective results due to Baker. Not included is approximation in p -adic fields, for which see e.g. Schlickewei [1976, 1977], or approximation in power series fields, for which see e.g., Osgood [1977] and Ratliff [1978]. Totally missing are Pisot-Vijayaraghavan Numbers, inhomogeneous approximation and uniform distribution. For these see e.g. Cassels [1957] and Kuipers and Niederreiter [1974]. Also excluded are Weyl Sums, nonlinear approxi-

mation and diophantine inequalities involving forms in many variables.

My pace is in general very leisurely and slow. This will be especially apparent when comparing Baker's [1975] chapter on approximation to algebraic numbers with my two separate chapters, one dealing with Roth's Theorem on approximation to a single algebraic number, the other with simultaneous approximation to algebraic numbers.

Possible sequences are chapters

I, II, III, for a reader who is interested in game and measure theoretic results, or

I, II, V, for a reader who wants to study Roth's Theorem, or

I, II, IV, V, VI, VII (§ 11, 12), VIII (§ 7-10), for a general theory of simultaneous approximation to algebraic numbers, or

I, II, IV, V, VI, VII, if the goal is norm form equations, or

I, II, VIII (§ 1-6, §11), if the emphasis is on approximation by algebraic numbers.

December 1979

W.M. Schmidt

Notation

A real number ξ may uniquely be written as

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

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

$\|\xi\| = \min(\{\xi\}, 1 - \{\xi\})$ is the distance from ξ to the nearest integer,

U denotes the unit interval $0 \leq \xi < 1$.

\mathbb{R}^n denotes the n -dimensional real space,

E^n denotes Euclidean n -space.

$\underline{x}, \underline{y}, \dots$ will denote vectors; so $\underline{x} = (x_1, \dots, x_n) \in \mathbb{R}^n$, or

$$\underline{x} = (x_1, \dots, x_m) \in \mathbb{R}^m , \text{ etc.}$$

Addition and multiplication of vectors by scalars is obvious.

$\underline{e}_1, \dots, \underline{e}_n$ will denote basis vectors.

λK , where $\lambda > 0$ and where K is in \mathbb{R}^n , is the set of elements

$$\lambda \underline{x} \text{ with } \underline{x} \in K .$$

δ_{ij} is the Kronecker Symbol.

X, Y, \dots , in general will be variables, while x, y, \dots will be

real, usually rational integers. But this rule is sometimes hard to

follow: In chapter IV, the symbols X, Y, \dots will also be used to denote coordinates in compound spaces.

$$|\underline{x}| = \max(|x_1|, \dots, |x_n|) \text{ if } \underline{x} = (x_1, \dots, x_n) . \text{ However}$$

$|\underline{\beta}|$, where $\underline{\beta} = (\beta_1, \dots, \beta_n)$ has coordinates in an algebraic number field K , is given by $|\underline{\beta}| = \max(|\beta_1^{(1)}|, \dots, |\beta_n^{(1)}|, \dots, |\beta_1^{(k)}|, \dots, |\beta_n^{(k)}|)$, if $\beta^{(1)} = \beta$, $\beta^{(2)}, \dots, \beta^{(k)}$ are the conjugates of an elements β

(But, on p. 173 , $|\gamma|$ for a single element γ has a different meaning.)

\overline{P} is the maximum absolute value of the coefficients of a polynomial P ,

\mathbb{Q} is the field of rationals,

\mathbb{R} is the field of reals,

\mathbb{C} is the field of complex numbers.

$[L:K]$ is the degree of a field extension L over K .

$\{a,b,\dots,w\}$ denotes the set consisting of a,b,\dots,w , and

\sim denotes a set theoretic difference.

\ll is the Vinogradov symbol. Thus e.g. $f(\underline{x}) \ll g(\underline{x})$ means that $|f(\underline{x})| \leq c|g(\underline{x})|$ with a constant c . Often this "implied" constant c may depend on extra parameters, such as the dimension, etc.

$\gg \ll$, in the context $f \ll g$, means that both $f \ll g$ and $g \ll f$.

o , the "little o" , in the context $f(n) = o(g(n))$, means that $f(n)/g(n)$ tends to 0 as $n \rightarrow \infty$.

g.c.d. denotes the greatest common divisor of integers.

Starred Theorems, such as Theorem 6A* , are not proved in these Notes.

Table of Contents

I.	Approximation to Irrational Numbers by Rationals.	
1.	Dirichlet's Theorem.	1
2.	Farey Series	2
3.	Continued Fractions: Algebraic Theory.	7
4.	Simple Continued Fractions	11
5.	Continued Fractions and Approximation to Irrationals by Rationals	16
6.	Further results.	23
II.	Simultaneous Approximation.	
1.	Dirichlet's Theorem on Simultaneous Approximation.	27
2.	Theorems of Blichfeldt and Minkowski	29
3.	Improvement of the Simultaneous Approximation Constants.	36
4.	Badly Approximable Systems of Linear Forms	41
III.	Games and Measures.	
	First Part: Games	
1.	The (α, β) - Game.	48
2.	Badly Approximable n -tuples and (α, β) - Games.	52
	Second Part: Measures	
3.	Statement of Results	60
4.	The convergence part of Theorem 3A	63
5.	The idea of the proof of Theorem 3B.	63
6.	On certain intervals	65
7.	Sums involving a function $\Psi(k, q)$	66
8.	Bounds for certain integrals	69

9. Proof of Theorem 3B.	74
10. The case $n \geq 2$	77

IV. Integer Points in Parallelepipeds.

1. Minkowski's Theorem on Successive Minima	80
2. Jordan's Theorem	87
3. Davenport's Lemma.	89
4. Reciprocal Parallelepipeds	92
5. Khintchine's Transference Principle.	95
6. The Grassman Algebra	102
7. Mähler's Theory of Compound Sets	108
8. Point Lattices	111

V. Roth's Theorem.

1. Liouville's Theorem.	114
2. Roth's Theorem and its History	115
3. Thue's Equation.	118
4. Combinatorial Lemmas	121
5. Further auxiliary Lemmas	125
6. The Index of a Polynomial.	129
7. The Index Theorem.	132
8. The Index of $P(X_1, \dots, X_m)$ at Rational Points near $(\alpha, \alpha, \dots, \alpha)$	134
9. Generalized Wronskians	137
10. Roth's Lemma	141
11. Conclusion of the proof of Roth's Theorem.	148

VI. Simultaneous Approximation to Algebraic Numbers.

1. Basic Results.	151
---------------------------	-----

2. Roth Systems.	155
3. The Strong Subspace Theorem	162
4. The Index of a Polynomial	166
5. Some Auxiliary Lemmas	172
6. The Index Theorem	176
7. The Polynomial Theorem.	180
8. Grids	183
9. The Index of P with respect to certain Rational Linear Forms	187
10. An Analogue of Roth's Lemma	190
11. The size of \underline{g}_n^*	195
12. The Next to Last Minimum.	197
13. The Constancy of \underline{g}_n^*	200
14. The Last Two Minima	202
15. Proof of the Strong Subspace Theorem.	205

VII. Norm Form Equations.

1. Norm Form Equations	208
2. Full Modules.	212
3. An Example.	213
4. The General Case.	215
5. Induction on the rank of \mathfrak{M}	219
6. Linear Inequalities in a Simplex.	221
7. Constuction of a field L	223
8. The Main Lemma.	228
9. Proof of the Main Theorem	234
10. Equations $\mathfrak{N}(M(\underline{x})) = P(\underline{x})$	236

11. Another Theorem on Linear Forms	240
12. Proof of the Theorem on Linear Forms	242
13. Proof of Theorem 10A	247
14. Proof of Theorem 10C	248

VIII. Approximation By Algebraic Numbers

1. The Setting	251
2. Field Height and Approximation by Elements of a Given Number Field	252
3. Absolute Height and Approximation by Algebraic Numbers of Bounded Degree.	255
4. Approximation by Quadratic Irrationals	260
5. Approximation by Quadratic Irrationals, Continued.	264
6. Proof of Wirsing's Theorem	268
7. A Subspace Theorem for Number Fields	272
8. Approximation to Algebraic Numbers by Elements of a Number Field	275
9. Approximation to Algebraic Numbers by Algebraic Numbers of Bounded Degree.	278
10. Mahler's Classification of Transcendental Numbers.	280
11. A Theorem of Mignotte.	281
References	289

I. Approximation to Irrational Numbers by Rationals.

References: Dirichlet (1842), Hurwitz (1891), Perron (1954), Cassels (1957).

§1. Dirichlet's Theorem.

Given a real number α , let $[\alpha]$, the integer part of α , denote the greatest integer $\leq \alpha$, and let $\{\alpha\} = \alpha - [\alpha]$. Then $\{\alpha\}$ is the fractional part of α , and satisfies $0 \leq \{\alpha\} < 1$. Also, let $\|\alpha\|$ denote the distance from α to the nearest integer. Then always $0 \leq \|\alpha\| \leq \frac{1}{2}$.

THEOREM 1A. (Dirichlet (1842)). Let α and Q be real numbers with $Q > 1$. Then there exist integers p, q such that $1 \leq q < Q$ and $|\alpha q - p| \leq \frac{1}{Q}$.

Proof. First assume that Q is an integer. Consider the following $Q + 1$ numbers:

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

They all lie in the unit interval $0 \leq x \leq 1$. We divide the unit interval into Q subintervals

$$\frac{u}{Q} \leq x < \frac{u+1}{Q} \quad (u = 0, 1, \dots, Q-1),$$

but with $<$ replaced by \leq if $u = Q-1$. At least one such subinterval contains two (or more) of the $Q+1$ numbers above. Hence there are integers r_1, r_2, s_1, s_2 with $0 \leq r_i < Q$ ($i=1, 2$) and $r_1 \neq r_2$ such that

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

If, say, $r_1 > r_2$, put $q = r_1 - r_2$, $p = s_1 - s_2$. Then $1 \leq q < Q$ and $|q\alpha - p| \leq \frac{1}{Q}$, proving the theorem when Q is an integer.

Next, suppose Q is not an integer. Apply what has already been proved to $Q' = [Q] + 1$. Then $1 \leq q < Q'$ implies $1 \leq q \leq [Q]$, whence $1 \leq q < Q$, and the theorem is true for Q .

Remark. The two inequalities in Dirichlet's Theorem yield

$$|\alpha - \frac{p}{q}| \leq \frac{1}{Qq} < \frac{1}{q^2}.$$

COROLLARY 1B. Suppose that α is irrational. Then there exist infinitely many pairs p, q of relatively prime integers with

$$(1.1) \quad |\alpha - \frac{p}{q}| < \frac{1}{q^2}.$$

Proof. Dirichlet's Theorem obviously remains true if we ask for relatively prime integers p, q satisfying $1 \leq q < Q$ and $|q\alpha - p| \leq \frac{1}{Q}$. Since α is irrational, $q\alpha - p$ is never zero, and hence for any given p, q , the inequality $|q\alpha - p| \leq \frac{1}{Q}$ can only be satisfied for $Q \leq Q_0(p, q)$. Hence as $Q \rightarrow \infty$, there will be infinitely many distinct pairs p, q of relatively prime integers occurring in Dirichlet's Theorem.

Remark. This corollary is not true if α is rational. For suppose that $\alpha = \frac{u}{v}$. If $\alpha \neq \frac{p}{q}$, then $|\alpha - \frac{p}{q}| = |\frac{u}{v} - \frac{p}{q}| = |\frac{qu - pv}{vq}| \geq \frac{1}{vq}$, and therefore (1.1) can be satisfied by only finitely many pairs p, q of relatively prime integers.

§2. Farey Series.

Definition. The Farey series \mathcal{F}_n of order $n (\geq 1)$ is the sequence of rationals in their lowest terms between 0 and 1 with

denominators $\leq n$, written in ascending order. For example,

$$\mathcal{T}_5: 0, \frac{1}{5}, \frac{1}{4}, \frac{1}{3}, \frac{2}{5}, \frac{1}{2}, \frac{3}{5}, \frac{2}{3}, \frac{3}{4}, \frac{4}{5}, 1.$$

THEOREM 2A. If $\frac{h}{k}, \frac{h'}{k'}$ are successive terms in \mathcal{T}_n , then
 $h'k - hk' = 1$.

We need

LEMMA 2B. Suppose that $\underline{x} = (x_1, x_2)$ and $\underline{y} = (y_1, y_2)$ are integer points in the plane, with $\underline{0} = (0, 0), \underline{x}, \underline{y}$ not on a line. Suppose further that the closed triangle with vertices $\underline{0}, \underline{x}, \underline{y}$ contains no integer points but its vertices. Then

$$x_1 y_2 - x_2 y_1 = \pm 1.$$

Proof of the Lemma. Let \mathcal{T} be the triangle mentioned above, and let \mathcal{O} be the closed parallelogram with vertices $\underline{0}, \underline{x}, \underline{y}$ and $\underline{x} + \underline{y}$. Then \mathcal{O} contains no integer points but its vertices: for suppose that \underline{z} is an integer point in \mathcal{O} , $\underline{z} \notin \mathcal{T}$. Then $\underline{x} + \underline{y} - \underline{z} \in \mathcal{T}$, hence $\underline{x} + \underline{y} - \underline{z} = \underline{0}, \underline{x}$ or \underline{y} , and therefore $\underline{z} = \underline{x} + \underline{y}, \underline{y}$ or \underline{x} .

If \underline{p} is any integer point, we may write $\underline{p} = \lambda \underline{x} + \mu \underline{y}$ with real coefficients λ, μ since $\underline{0}, \underline{x}, \underline{y}$ are not collinear. Then $\underline{p} = \underline{p}' + \underline{p}''$, where

$$\underline{p}' = [\lambda] \underline{x} + [\mu] \underline{y} \text{ and } \underline{p}'' = \{\lambda\} \underline{x} + \{\mu\} \underline{y}.$$

Both \underline{p} and \underline{p}' are integer points, hence so is \underline{p}'' . Also $\underline{p}'' \in \mathcal{O}$. Since $\underline{p}'' \neq \underline{x}, \underline{y}$ and $\underline{x} + \underline{y}$, we have $\underline{p}'' = \underline{0}$. Therefore $\underline{p} = \lambda \underline{x} + \mu \underline{y}$ with integer coefficients λ, μ .

In particular,

$$(1,0) = \lambda \underline{x} + \mu \underline{y} = (\lambda x_1 + \mu y_1, \lambda x_2 + \mu y_2) ,$$

$$(0,1) = \lambda' \underline{x} + \mu' \underline{y} = (\lambda' x_1 + \mu' y_1, \lambda' x_2 + \mu' y_2)$$

for certain integers $\lambda, \mu, \lambda', \mu'$. It follows that

$$1 = \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = \begin{vmatrix} \lambda & \mu \\ \lambda' & \mu' \end{vmatrix} \cdot \begin{vmatrix} x_1 & x_2 \\ y_1 & y_2 \end{vmatrix} ,$$

whence

$$\begin{vmatrix} x_1 & x_2 \\ y_1 & y_2 \end{vmatrix} = \pm 1$$

as claimed.

Proof of the Theorem. Put $\underline{x} = (h, k)$, $\underline{y} = (h', k')$. Then $\underline{0}, \underline{x}, \underline{y}$ are not collinear since $\gcd(h, k) = \gcd(h', k') = 1$ and $\underline{x} \neq \underline{y}$. Let \mathcal{T} denote the closed triangle with vertices $\underline{0}, \underline{x}, \underline{y}$. Then there is no integer point in \mathcal{T} besides $\underline{0}, \underline{x}, \underline{y}$. For if there were such a point (h'', k'') , then there also would be a point with $\gcd(h'', k'') = 1$. Then $(h'', k'') = \lambda(h, k) + \mu(h', k')$ with $\lambda \geq 0, \mu \geq 0, 0 < \lambda + \mu \leq 1$ and (λ, μ) not equal to $(1, 0)$ or to $(0, 1)$. This implies that $k'' \leq \lambda n + \mu n \leq n$. We have $\lambda > 0, \mu > 0$ (since $\gcd(h, k) = \gcd(h', k') = 1$), whence $\frac{h}{k} < \frac{h''}{k''} < \frac{h'}{k'}$. Thus $\frac{h''}{k''}$ would belong to \mathfrak{F}_n , contradicting the supposition that $\frac{h}{k}$ and $\frac{h'}{k'}$ are consecutive elements of \mathfrak{F}_n . The hypotheses of Lemma 2B are now satisfied, and we conclude that $h'k - hk' = \pm 1$. Since $\frac{h}{k} < \frac{h'}{k'}$, we have $h'k - hk' = 1$.

COROLLARY 2C. If $\frac{h}{k}, \frac{h''}{k''}, \frac{h'}{k'}$ are consecutive elements of \mathfrak{F}_n , then

$$\frac{h''}{k''} = \frac{h+h'}{k+k'} .$$

Proof. By the theorem, $h''k - hk'' = 1$ and $h'k'' - h''k' = 1$,
so that $h''(k+k') - k''(h+h') = 0$.

LEMMA 2D. Suppose that $\frac{h}{k}, \frac{h'}{k'}$ are successive terms in the Farey series \mathfrak{F}_n , and put $h'' = h + h', k'' = k + k'$. (Note that $\frac{h''}{k''}$ does NOT belong to \mathfrak{F}_n). Then for every α in $\frac{h}{k} \leq \alpha \leq \frac{h'}{k'}$, at least one of the following three inequalities holds:

$$(2.1) \quad \left| \alpha - \frac{h}{k} \right| < \frac{1}{\sqrt{5}k^2}, \quad \left| \alpha - \frac{h''}{k''} \right| < \frac{1}{\sqrt{5}k''^2}, \quad \left| \alpha - \frac{h'}{k'} \right| < \frac{1}{\sqrt{5}k'^2} .$$

Proof. We may assume that $\alpha \geq \frac{h''}{k''}$. Namely, otherwise replace α by $1 - \alpha$, $\frac{h}{k}$ by $1 - \frac{h'}{k'}$, etc. If none of the inequalities above hold, then

$$\alpha - \frac{h}{k} \geq \frac{1}{\sqrt{5}k^2}, \quad \alpha - \frac{h''}{k''} \geq \frac{1}{\sqrt{5}k''^2}, \quad \frac{h'}{k'} - \alpha \geq \frac{1}{\sqrt{5}k'^2} .$$

Adding the first and third inequalities, we obtain

$$\frac{h'}{k'} - \frac{h}{k} = \frac{1}{kk'} \geq \frac{1}{\sqrt{5}} \left(\frac{1}{k^2} + \frac{1}{k'^2} \right);$$

adding the second and third inequalities, we obtain

$$\frac{h'}{k'} - \frac{h''}{k''} = \frac{1}{k'k''} \geq \frac{1}{\sqrt{5}} \left(\frac{1}{k'^2} + \frac{1}{k''^2} \right) .$$

Then $\sqrt{5}kk' \geq k^2 + k'^2$ and $\sqrt{5}k'k'' \geq k'^2 + k''^2$, so that

$$\sqrt{5}k'(k+k'') \geq k^2 + 2k'^2 + k''^2, \text{ and therefore } \sqrt{5}k'(2k+k') \geq 2k^2 + 3k'^2 + 2kk'.$$

It follows that

$$0 \geq \frac{1}{2}((\sqrt{5}-1)k' - 2k)^2 .$$

But this is impossible, since k and k' are nonzero integers.

LEMMA 2E. Suppose α is a real quadratic irrational which is a root of a non-zero polynomial

$$P(X) = aX^2 + bX + c$$

with rational integer coefficients and discriminant $D = b^2 - 4ac$.

Then for $A > \sqrt{D}$, the inequality

$$(2.2) \quad \left| \alpha - \frac{p}{q} \right| < \frac{1}{Aq^2}$$

has only finitely many solutions.

Proof. Write $P(X) = a(X - \alpha)(X - \alpha')$, so that $D = a^2(\alpha - \alpha')^2$.

Given p/q with (2.2) we have

$$\frac{1}{q} \leq \left| P\left(\frac{p}{q}\right) \right| = \left| \alpha - \frac{p}{q} \right| \left| a(\alpha' - \frac{p}{q}) \right| < \frac{1}{Aq^2} \left| a(\alpha' - \alpha + \alpha - \frac{p}{q}) \right| < \frac{\sqrt{D}}{Aq^2} + \frac{|a|}{A^2q^4},$$

which clearly is impossible if $A > \sqrt{D}$ and if q is large.

THEOREM 2F. (Hurwitz (1891)).

(i) For every irrational number α there are infinitely many distinct rationals $\frac{p}{q}$ with

$$(2.3) \quad \left| \alpha - \frac{p}{q} \right| < \frac{1}{\sqrt{5}q^2}.$$

(ii) This would be wrong if $\sqrt{5}$ were replaced by a constant $A > \sqrt{5}$.

Proof. We may suppose that $0 < \alpha < 1$. If $\frac{h}{k}$ and $\frac{h'}{k'}$ are the successive terms in the Farey series \mathfrak{F}_n with $\frac{h}{k} < \alpha < \frac{h'}{k'}$, then

according to Lemma 2D, at least one of $\frac{h}{k}$, $\frac{h'}{k'}$, and $\frac{h''}{k''}$ with $h'' = h + h'$, $k'' = k + k'$ satisfies (2.3). Then by the construction of the Farey Series, $\left| \frac{h}{k} - \frac{h'}{k'} \right| \leq 1/n$ and therefore $\left| \alpha - \frac{p}{q} \right| \leq 1/n$. Since α is irrational, this can for fixed $\frac{p}{q}$ only be true for small values of n . Hence as n goes to infinity, we obtain infinitely many $\frac{p}{q}$ with (2.3).

Assertion (ii) follows from the case $\alpha = \frac{1}{2}(\sqrt{5}-1)$ and $P(X) = X^2 + X - 1$ of Lemma 2E, since $D = 5$ in this case.

§3. Continued Fractions: Algebraic Theory.

Let a_0, a_1, a_2, \dots be variables. We define polynomials $p_0, q_0, p_1, q_1, p_2, q_2, \dots$, with p_n and q_n being polynomials in a_0, a_1, \dots, a_n , as follows:

First, we define $p_0 = a_0$ and $q_0 = 1$. Next, suppose that $p_0, q_0, \dots, p_{n-1}, q_{n-1}$ have already been defined. Using the abbreviations

$$p'_k = p_k(a_1, a_2, \dots, a_{k+1}) \text{ and } q'_k = q_k(a_1, a_2, \dots, a_{k+1}),$$

we define

$$p_n = a_0 p'_{n-1} + q'_{n-1} \text{ and } q_n = p'_{n-1}.$$

$\frac{p_n}{q_n}$ is a rational function of a_0, a_1, \dots, a_n , and we write

$$\frac{p_n}{q_n} = [a_0, a_1, \dots, a_n].$$

In particular, $[a_0] = \frac{p_0}{q_0} = a_0$. Unfortunately, if a_0 is not an integer, this notation conflicts with the greatest integer notation introduced in §1. However, the meaning of $[a_0]$ will be clear from the context.

For $n > 0$, we have

$$[a_0, a_1, \dots, a_n] = \frac{p_n}{q_n} = \frac{a_0 p'_{n-1} + q'_{n-1}}{p'_{n-1}} = a_0 + \frac{1}{p'_{n-1}/q'_{n-1}} = a_0 + \frac{1}{[a_1, \dots, a_n]} .$$

Applying this repeatedly, we obtain

$$[a_0, a_1, \dots, a_n] = a_0 + \frac{1}{a_1 + \frac{1}{a_2 + \frac{1}{\ddots a_{n-1} + \frac{1}{a_n}}}} .$$

A rational function of this type is called a continued fraction.

LEMMA 3A. For $n \geq 2$, we have

$$p_n = a_n p_{n-1} + p_{n-2} ,$$

$$q_n = a_n q_{n-1} + q_{n-2} .$$

Proof. For $n = 2$, the assertion is verified directly. Assume, then that $n > 2$, and that the assertion is true for $n - 1$. Then

$$p'_{n-1} = a_n p'_{n-2} + p'_{n-3} ,$$

$$q'_{n-1} = a_n q'_{n-2} + q'_{n-3} .$$

We obtain

$$\begin{aligned} p_n &= a_0 p'_{n-1} + q'_{n-1} = a_0 (a_n p'_{n-2} + p'_{n-3}) + a_n q'_{n-2} + q'_{n-3} \\ &= a_n (a_0 p'_{n-2} + q'_{n-2}) + a_0 p'_{n-3} + q'_{n-3} \\ &= a_n p_{n-1} + p_{n-2} , \end{aligned}$$

and

$$\begin{aligned}
 q_n &= p'_{n-1} = a_n p'_{n-2} + p'_{n-3} \\
 &= a_n q_{n-1} + q_{n-2} ,
 \end{aligned}$$

which proves the lemma.

For convenience, we also define $p_{-2}=0, p_{-1}=1, q_{-2}=1, q_{-1}=0$.

It is easily checked that Lemma 3A then holds for all $n \geq 0$.

LEMMA 3B. For each k with $1 \leq k \leq n$, let

$$r_k = [a_k, a_{k+1}, \dots, a_n] .$$

Then

$$[a_0, a_1, \dots, a_k, \dots, a_n] = [a_0, a_1, \dots, a_{k-1}, [a_k, \dots, a_n]]$$

$$= \frac{p_{k-1} r_k + p_{k-2}}{q_{k-1} r_k + q_{k-2}} .$$

Proof. The second equality follows from Lemma 3A. The first equality may be proved by induction on k . The lemma is true for $k=1$, since

$$[a_0, a_1, \dots, a_n] = a_0 + \frac{1}{[a_1, \dots, a_n]} = [a_0, [a_1, \dots, a_n]] .$$

Now suppose that the lemma holds for $k-1$, where $1 < k \leq n$. Then

$$\begin{aligned}
 [a_0, a_1, \dots, a_n] &= a_0 + \frac{1}{[a_1, \dots, a_k, \dots, a_n]} = a_0 + \frac{1}{[a_1, \dots, a_{k-1}, [a_k, \dots, a_n]]} \\
 &= [a_0, a_1, \dots, a_{k-1}, [a_k, \dots, a_n]] ,
 \end{aligned}$$

which proves the lemma.

LEMMA 3C. For $n \geq -1$, we have

$$q_n p_{n-1} - p_n q_{n-1} = (-1)^n .$$

Proof. We proceed by induction on n . Now $q_{-1} p_{-2} - p_{-1} q_{-2} = 0^2 - (-1)^2 = (-1)^{-1}$, so the lemma is true for $n = -1$. Assuming that the lemma is true for $n-1$, we have

$$\begin{aligned} q_n p_{n-1} - p_n q_{n-1} &= (a_n q_{n-1} + q_{n-2}) p_{n-1} - (a_n p_{n-1} + p_{n-2}) q_{n-1} \\ &= -(q_{n-1} p_{n-2} - p_{n-1} q_{n-2}) = -(-1)^{n-1} = (-1)^n , \end{aligned}$$

and the lemma is true for n .

LEMMA 3D. For $n \geq 0$, $q_n p_{n-2} - p_n q_{n-2} = (-1)^{n-1} a_n$.

Proof. Using Lemma 3A and 3C, we have

$$\begin{aligned} q_n p_{n-2} - p_n q_{n-2} &= (a_n q_{n-1} + q_{n-2}) p_{n-2} - (a_n p_{n-1} + p_{n-2}) q_{n-2} \\ &= (q_{n-1} p_{n-2} - p_{n-1} q_{n-2}) a_n = (-1)^{n-1} a_n . \end{aligned}$$

LEMMA 3E. Put $\alpha = [a_0, a_1, \dots, a_{n+1}]$. Then

$$q_n \alpha - p_n = \frac{(-1)^n}{a_{n+1} q_n + q_{n-1}} .$$

$$\text{Proof. } q_n \alpha - p_n = q_n \frac{p_{n+1}}{q_{n+1}} - p_n = \frac{-(q_{n+1} p_n - p_{n+1} q_n)}{q_{n+1}} = \frac{(-1)^n}{a_{n+1} q_n + q_{n-1}}$$

by Lemmas 3A and 3C.

LEMMA 3F. For $n \geq 1$, we have

$$\frac{q_n}{q_{n-1}} = [a_n, a_{n-1}, \dots, a_1] .$$

Proof. We use induction on n . $\frac{q_1}{q_0} = \frac{a_1}{1} = [a_1]$, so the lemma is true for $n = 1$. Now suppose that the lemma is true for $n-1$ in place of n . Then

$$\begin{aligned}\frac{q_n}{q_{n-1}} &= \frac{a_n q_{n-1} + q_{n-2}}{q_{n-1}} = a_n + \frac{1}{q_{n-1}/q_{n-2}} = a_n + \frac{1}{[a_{n-1}, \dots, a_1]} \\ &= [a_n, a_{n-1}, \dots, a_1],\end{aligned}$$

and the lemma is proved.

§4. Simple continued fractions.

LEMMA 4A. Suppose that a_0, a_1, a_2, \dots , are real numbers with
 a_1, a_2, \dots positive. Then

$$\frac{p_0}{q_0} < \frac{p_2}{q_2} < \frac{p_4}{q_4} < \dots; \dots < \frac{p_5}{q_5} < \frac{p_3}{q_3} < \frac{p_1}{q_1}.$$

Proof. By Lemma 3D,

$$\frac{p_{n-2}}{q_{n-2}} - \frac{p_n}{q_n} = \frac{(-1)^{n-1} a_n}{q_{n-2} q_n}.$$

Applying this with $n \geq 2$, n even, we get $\frac{p_{n-2}}{q_{n-2}} - \frac{p_n}{q_n} < 0$; applying

this with $n \geq 3$, n odd, we get $\frac{p_{n-2}}{q_{n-2}} - \frac{p_n}{q_n} > 0$.

It remains to show that $\frac{p_n}{q_n} < \frac{p_m}{q_m}$ if n is even and m is odd.

Suppose, say, that $n < m$. Then since $\frac{p_n}{q_n} \leq \frac{p_{m-1}}{q_{m-1}}$, it will suffice

to show that $\frac{p_{m-1}}{q_{m-1}} < \frac{p_m}{q_m}$. But the last inequality is true since

$q_m p_{m-1} - p_m q_{m-1} = (-1)^m < 0$. (Recall that m is odd.) The case

$n > m$ is proved similarly.

LEMMA 4B. Suppose that $a_0, a_1 \geq 1, \dots, a_n \geq 1$ are integers. Then
 $[a_0, a_1, \dots, a_n]$ is rational. Conversely, given a rational number $\frac{u}{v}$,
there will be an n and integers a_0, a_1, \dots, a_n as above with

$$(4.1) \quad \frac{u}{v} = [a_0, a_1, \dots, a_n] .$$

Moreover, $a_0 \geq 1$ if $\frac{u}{v} \geq 1$.

Proof. Only the second part needs a proof. Without loss of generality, we assume that $v > 0$ and $(u, v) = 1$. The proof is by induction on v . If $v = 1$, $\frac{u}{v}$ is an integer and we have $\frac{u}{v} = [a_0]$ with $a_0 = \frac{u}{v}$. Suppose, then, that $v > 1$ (so that $\frac{u}{v}$ is not an integer). By the Division Algorithm, there exist integers q, r with $u = vq + r$ and $1 \leq r < v$. By induction, we may write $\frac{v}{r}$ as a continued fraction, say

$$\frac{v}{r} = [a_1, \dots, a_n] .$$

Since $\frac{v}{r} > 1$, we have $a_1 \geq 1, \dots, a_n \geq 1$. Therefore

$$\frac{u}{v} = q + \frac{1}{v/r} = q + \frac{1}{[a_1, \dots, a_n]} = [q, a_1, \dots, a_n] ,$$

and (4.1) is established with $a_0 = q$. Clearly $a_0 = q \geq 1$ if $\frac{u}{v} \geq 1$, and the proof of the lemma is complete.

Definitions. If $a_0, a_1 \geq 1, \dots, a_n \geq 1$ are integers and if $r = [a_0, a_1, \dots, a_n]$, then this is an expansion of r into a finite simple continued fraction. $\frac{p_i}{q_i}$ is the i^{th} convergent to α , a_i is the i^{th} partial quotient of α , and $\alpha_i = [a_i, a_{i+1}, \dots, a_n]$ is the

i^{th} complete quotient of α . Note that $\frac{p_i}{q_i}$ is in reduced terms by Lemma 3C.

LEMMA 4C. (i) Suppose that r is an integer. Then there are precisely two expansions of r into a simple continued fraction as above, namely $r = [r]$ and $r = [r-1, 1]$. (ii) For r rational but not integral, r has precisely two simple continued fraction expansions: one is of the form $[a_0, a_1, \dots, a_{n-1}, a_n]$ with $a_n \geq 2$, and the other is $[a_0, a_1, \dots, a_{n-1}, a_n - 1, 1]$.

Proof. In either case, r is rational, and so by Lemma 4B, r has an expansion as a simple continued fraction

$$r = [a_0, a_1, \dots, a_n],$$

where $a_0, a_1 \geq 1, \dots, a_n \geq 1$ are integers.

Suppose that r is an integer. If $n = 0$, then $r = [a_0] = a_0$, so $r = [r]$. If $n > 0$, then

$$r = a_0 + \frac{1}{[a_1, \dots, a_n]}$$

and $[a_1, \dots, a_n] \geq 1$. Since $r - a_0$ is an integer, $[a_1, \dots, a_n] = 1$. Because $a_1 \geq 1$, it follows that $n = 1$ and $a_1 = 1$. Then $a_0 = r - 1$, and $r = [r - 1, 1]$. This proves (i).

Now suppose that $r = \frac{u}{v}$ with $(u, v) = 1$ and $v > 0$. The case when $v = 1$ is treated in part (i) of the lemma. To prove the lemma in general we use induction on v . If $v > 1$, then a_0 in the continued fraction expansion of $\frac{u}{v}$ is uniquely determined: it is the integer part of $\frac{u}{v}$. We have

$$\frac{u}{v} = a_0 + \frac{\frac{u}{v} - a_0}{1} = a_0 + \frac{1}{\frac{1}{\frac{u}{v} - a_0}}$$

where $\alpha_1 = \frac{v}{u_1} > 1$ has a denominator less than v . So by induction, α_1 has two expansions into a simple continued fraction, of the form $[a_1, \dots, a_{n-1}, a_n]$ with $a_n \geq 2$ and $[a_1, \dots, a_{n-1}, a_n - 1, 1]$. The lemma follows.

LEMMA 4D. Suppose that $a_0, a_1 \geq 1, a_2 \geq 1, \dots$ are integers. Then $\lim_{n \rightarrow \infty} \frac{p_n}{q_n}$ exists and is irrational. Conversely, if α is irrational, then there exist unique integers a_0, a_1, a_2, \dots as above with $\alpha = \lim_{n \rightarrow \infty} \frac{p_n}{q_n}$.

Notation. $\lim_{n \rightarrow \infty} [a_0, a_1, \dots, a_n] = [a_0, a_1, a_2, \dots]$. If

$\alpha = [a_0, a_1, a_2, \dots]$, then this is an expansion of α into a simple continued fraction. We retain the terminology used above for finite simple continued fractions: namely, $\frac{p_i}{q_i}$ is the i^{th} convergent to α , a_i is the i^{th} partial quotient of α , $\alpha_i = [a_i, a_{i+1}, \dots]$ is the i^{th} complete quotient of α .

Proof of the lemma. (First part) By Lemma 3C,

$$\frac{p_{n-1}}{q_{n-1}} - \frac{p_n}{q_n} = \frac{(-1)^n}{q_{n-1}q_n}.$$

Since $\frac{p_0}{q_0} < \frac{p_2}{q_2} < \dots$, and since these are bounded above by $\frac{p_1}{q_1}$, $\lim_{\substack{n \rightarrow \infty \\ n \text{ even}}} \frac{p_n}{q_n}$ exists. Similarly, $\lim_{\substack{n \rightarrow \infty \\ n \text{ odd}}} \frac{p_n}{q_n}$ exists. But these two limits are

equal, since $\lim_{n \rightarrow \infty} \frac{(-1)^n}{q_{n-1}q_n} = 0$. (Here we use that $q_n = a_n q_{n-1} + q_{n-2}$,

hence $q_n \geq q_{n-1} + 1$ if $n \geq 1$, and therefore $q_n \rightarrow \infty$ as $n \rightarrow \infty$.)

Let $\alpha = \lim_{n \rightarrow \infty} \frac{p_n}{q_n}$. Since α lies strictly between $\frac{p_n}{q_n}$ and

$\frac{p_{n+1}}{q_{n+1}}$, we have

$$\left| \alpha - \frac{p_n}{q_n} \right| < \frac{1}{q_n q_{n+1}} < \frac{1}{q_n^2}$$

by Lemma 3C. Now p_n and q_n are relatively prime for each n by Lemma 3C, so that there are infinitely many "reduced fractions" $\frac{p}{q}$ with $\left| \alpha - \frac{p}{q} \right| < \frac{1}{2q}$. By the remark at the end of §1, α must be irrational.

(Second part). We suppose that α is irrational. Let a_0 be the integer part of α , and define α_1 by $\alpha = a_0 + \frac{1}{\alpha_1}$. Since α is irrational, α_1 is irrational; also $\alpha_1 > 1$. In general, having defined integers $a_0, a_1 \geq 1, \dots, a_{k-1} \geq 1$ and irrationals $\alpha_1 > 1, \alpha_2 > 1, \dots, \alpha_k > 1$, let a_k be the integer part of α_k and define α_{k+1} by $\alpha_k = a_k + \frac{1}{\alpha_{k+1}}$. Then $a_k \geq 1$ and $\alpha_{k+1} > 1$; also, α_{k+1} is irrational since α_k is irrational. We now show that

$$\alpha = [a_0, a_1, a_2, \dots]$$

For each $n \geq 0$, we have $\alpha = [a_0, a_1, \dots, a_n, \alpha_{n+1}]$. By Lemma 3E,

$$q_n \alpha - p_n = \frac{(-1)^n}{\alpha_{n+1} q_n + q_{n-1}},$$

whence

$$\left| \alpha - \frac{p_n}{q_n} \right| < \frac{1}{2q_n}.$$

This implies that $\lim_{n \rightarrow \infty} \frac{p_n}{q_n} = \alpha$.

It remains to prove that the integers $a_0, a_1 \geq 1, a_2 \geq 1, \dots$ are unique. We have

$$\alpha = [a_0, a_1, a_2, \dots] = a_0 + \frac{1}{[a_1, a_2, \dots]} :$$

$0 \leq \alpha - a_0 < 1$, so a_0 is the integer part of α , and is therefore unique. It follows, then, that $\alpha_1 = [a_1, a_2, \dots]$ is unique. Then a_1 is the integer part of α_1 and is unique, etc.

§5. Continued Fractions and Approximation to Irrationals by Rationals.

Recall that by a formula of the last section, every convergent $\frac{p}{q}$ of α satisfies

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

THEOREM 5A. (Vahlen (1895)). Let $\frac{p_{n-1}}{q_{n-1}}, \frac{p_n}{q_n}$ be consecutive convergents to α . Then at least one of them satisfies

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

Proof. The numbers $\alpha - \frac{p_n}{q_n}$, $\alpha - \frac{p_{n-1}}{q_{n-1}}$ are of opposite sign, hence

we have

$$\left| \alpha - \frac{p_n}{q_n} \right| + \left| \alpha - \frac{p_{n-1}}{q_{n-1}} \right| = \left| \frac{p_n}{q_n} - \frac{p_{n-1}}{q_{n-1}} \right| = \frac{1}{q_n q_{n-1}} < \frac{1}{2q_n^2} + \frac{1}{2q_{n-1}^2},$$

since $ab < \frac{1}{2}(a^2 + b^2)$ if $a \neq b$. Hence, either

$$\left| \alpha - \frac{p_n}{q_n} \right| < \frac{1}{2q_n^2} \quad \text{or} \quad \left| \alpha - \frac{p_{n-1}}{q_{n-1}} \right| < \frac{1}{2q_{n-1}^2}.$$

THEOREM 5B. (E. Borel (1903a,b)). Let $\frac{p_{n-2}}{q_{n-2}}, \frac{p_{n-1}}{q_{n-1}}, \frac{p_n}{q_n}$ be three consecutive convergents to α . Then at least one of them satisfies

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

Proof. Put $\alpha = [a_0, a_1, \dots]$, $\alpha_i = [a_i, a_{i+1}, \dots]$ and

$\beta_i = \frac{q_{i-2}}{q_{i-1}}$ for $i = n-1, n, n+1$. We have

$$\alpha = [a_0, a_1, \dots, a_n, \alpha_{n+1}] ,$$

so

$$\alpha q_n - p_n = \frac{(-1)^n}{\alpha_{n+1} q_n + q_{n-1}}$$

by Lemma 3E. Thus

$$\left| \alpha - \frac{p_n}{q_n} \right| = \frac{1}{q_n (\alpha_{n+1} q_n + q_{n-1})} = \frac{1}{q_n^2 (\alpha_{n+1} + \beta_{n+1})} .$$

To complete the proof, it suffices to show that there cannot be three integers $i = n-1, n, n+1$ with

$$(5.1) \quad \alpha_i + \beta_i \leq \sqrt{5} .$$

Suppose that (5.1) were true for $i = n-1, n$. Now

$$\alpha_{n-1} = a_{n-1} + \frac{1}{\alpha_n} \quad \text{and} \quad \frac{1}{\beta_n} = \frac{q_{n-1}}{q_{n-2}} = a_{n-1} + \frac{q_{n-3}}{q_{n-2}} = a_{n-1} + \beta_{n-1} , \text{ so}$$

$$\frac{1}{\alpha_n} + \frac{1}{\beta_n} = \alpha_{n-1} + \beta_{n-1} \leq \sqrt{5} .$$

Therefore

$$1 = \alpha_n \left(\frac{1}{\alpha_n} \right) \leq (\sqrt{5} - \beta_n) \left(\sqrt{5} - \frac{1}{\beta_n} \right) ,$$

which is equivalent to

$$\beta_n^2 - \sqrt{5} \beta_n + 1 \leq 0 .$$

It follows that $\beta_n \geq \frac{1}{2}(\sqrt{5} - 1)$; since β_n is rational,

$$\beta_n > \frac{1}{2}(\sqrt{5} - 1) .$$

If (5.1) also were true for $i = n, n+1$, then

$$\beta_{n+1} > \frac{1}{2}(\sqrt{5} - 1) ,$$

and therefore

$$1 \leq a_n = \frac{q_n}{q_{n-1}} - \frac{q_{n-2}}{q_{n-1}} = \frac{1}{\beta_{n+1}} - \beta_n < \frac{2}{\sqrt{5}-1} - \frac{\sqrt{5}-1}{2} = 1 ,$$

a contradiction. This completes the proof of the theorem.

THEOREM 5C (Legendre) Suppose p, q are relatively prime integers
with $q > 0$ and

$$|\alpha - \frac{p}{q}| < \frac{1}{2q^2} .$$

Then $\frac{p}{q}$ is a convergent to α .

Proof. We assume that $\alpha \neq \frac{p}{q}$, else the theorem is trivially true. Then we may write

$$\alpha - \frac{p}{q} = \frac{\varepsilon \vartheta}{q^2} ,$$

where $0 < \vartheta < \frac{1}{2}$ and $\varepsilon = \pm 1$. From Lemma 4C there is an expansion of $\frac{p}{q}$ into a simple continued fraction

$$\frac{p}{q} = [b_0, b_1, \dots, b_{n-1}] ,$$

where n is chosen so that $(-1)^{n-1} = \varepsilon$.

Define ω by

$$(5.2) \quad \alpha = \frac{\omega p_{n-1} + p_{n-2}}{\omega q_{n-1} + q_{n-2}} ,$$

so that

$$\alpha = [b_0, b_1, \dots, b_{n-1}, \omega] .$$

(Note that (5.2) is equivalent to $(\alpha q_{n-1} - p_{n-1})\omega = p_{n-2} - \alpha q_{n-2}$. We may

assume that $\alpha q_{n-1} - p_{n-1} \neq 0$, else $\alpha = \frac{p_{n-1}}{q_{n-1}} = \frac{p}{q}$.) Then

$$\frac{\varepsilon \vartheta}{q} = \alpha - \frac{p}{q} = \frac{1}{q_{n-1}} (\alpha q_{n-1} - p_{n-1}) = \frac{1}{q_{n-1}} \cdot \frac{(-1)^{n-1}}{\omega q_{n-1} + q_{n-2}}$$

from Lemma 3E, and therefore

$$\vartheta = \frac{q_{n-1}}{\omega q_{n-1} + q_{n-2}} .$$

Solving this for ω , we obtain

$$\omega = \frac{q_{n-1} - \vartheta q_{n-2}}{\vartheta q_{n-1}} = \frac{1}{\vartheta} - \frac{q_{n-2}}{q_{n-1}} ;$$

it follows that $\omega > 2 - 1 = 1$. Expand ω into a (finite or infinite) simple continued fraction:

$$\omega = [b_n, b_{n+1}, \dots] .$$

Since $\omega > 1$, each of the integers b_j ($j = n, n+1, \dots$) is positive and therefore

$$\begin{aligned} \alpha &= [b_0, b_1, \dots, b_{n-1}, [b_n, b_{n+1}, \dots]] \\ &= [b_0, b_1, \dots, b_{n-1}, b_n, b_{n+1}, \dots] \end{aligned}$$

by Lemma 3B, passing to the limit if necessary. This is a simple continued fraction for α , and

$$\frac{p}{q} = \frac{p_{n-1}}{q_{n-1}} = [b_0, b_1, \dots, b_{n-1}]$$

is a convergent to α , so the proof is finished.

LEMMA 5D. Suppose α has a continued fraction expansion of the type

$$\alpha = [a_0, a_1, \dots, a_N, 1, 1, \dots] .$$

Then

$$\lim_{n \rightarrow \infty} q_n^2 \left| \alpha - \frac{p_n}{q_n} \right| = \frac{1}{\sqrt{5}} .$$

Proof. Using the notation introduced in the proof of Theorem 5A,

$$\left| \alpha - \frac{p_n}{q_n} \right| = \frac{1}{q_n^2 (\alpha_{n+1} + \beta_{n+1})} .$$

Here if n is sufficiently large,

$$\alpha_{n+1} = [1, 1, 1, \dots] = \frac{1}{2}(\sqrt{5} + 1),$$

and

$$\begin{aligned} \frac{1}{\beta_{n+1}} &= \frac{q_n}{q_{n-1}} = [a_n, a_{n-1}, \dots, a_N, \dots, a_1, a_0] \\ &= [1, 1, \dots, 1, a_N, \dots, a_1, a_0] . \end{aligned}$$

$\xleftrightarrow[n-N]{\quad}$

Since $\xleftrightarrow[n-N-1]{[1, 1, \dots, 1]}$ and $\xleftrightarrow[n-N]{[1, 1, \dots, 1]}$ are consecutive convergents to

$1/\beta_{n+1}$, the number $1/\beta_{n+1}$ lies between these convergents, and therefore $1/\beta_{n+1}$ approaches $[1, 1, \dots] = \frac{1}{2}(\sqrt{5} + 1)$ as $n \rightarrow \infty$.

Hence $\beta_{n+1} \rightarrow (\frac{1}{2}(\sqrt{5} + 1))^{-1} = \frac{1}{2}(\sqrt{5} - 1)$, and $\alpha_{n+1} + \beta_{n+1} \rightarrow \sqrt{5}$.

Continued fractions may be used to give

Another proof of Theorem 2F. Namely, assertion (i) follows easily from Theorem 5B, while assertion (ii) is a consequence of Theorem 5C and Lemma 5D.

THEOREM 5E. (Lagrange (1770a). Suppose that α is irrational,
and let $\frac{p_0}{q_0}, \frac{p_1}{q_1}, \dots$ be the convergents to α . Then

$$(i) \quad |\alpha q_0 - p_0| > |\alpha q_1 - p_1| > |\alpha q_2 - p_2| > \dots$$

$$(ii) \quad \text{If } n \geq 1 \text{ and } 1 \leq q \leq q_n, \text{ and if } (p, q) \neq (p_n, q_n), \\ (p, q) \neq (p_{n-1}, q_{n-1}), \text{ then } |\alpha q - p| > |\alpha q_{n-1} - p_{n-1}|.$$

Remark. It follows from (i) and (ii) that if $1 \leq q \leq q_n$, $(p, q) \neq (p_n, q_n)$ and $n \geq 1$, then $|\alpha q - p| > |\alpha q_n - p_n|$. This result is sometimes called the "Law of Best Approximation."

Proof of the theorem. From Lemma 3E,

$$|\alpha q_n - p_n| = \frac{1}{\alpha_{n+1} q_n + q_{n-1}} < \frac{1}{q_n + q_{n-1}},$$

and

$$|\alpha q_{n-1} - p_{n-1}| = \frac{1}{\alpha_n q_{n-1} + q_{n-2}} > \frac{1}{(a_n + 1) q_{n-1} + q_{n-2}} = \frac{1}{q_n + q_{n-1}},$$

which proves (i).

To prove (ii), we first define μ, ν by the equations

$$\mu p_n + \nu p_{n-1} = p,$$

$$\mu q_n + \nu q_{n-1} = q.$$

The matrix determined by these two equations has determinant ± 1 , hence μ, ν are integers.

If $v = 0$, then $(p, q) = \mu(p_n, q_n)$. But this is impossible since $0 < q \leq q_n$ and $(p, q) \neq (p_n, q_n)$. If $\mu = 0$, then $(p, q) = v(p_{n-1}, q_{n-1})$. Since $(p, q) \neq (p_{n-1}, q_{n-1})$, we have $v \geq 2$, and therefore

$$|\alpha q - p| \geq 2|\alpha q_{n-1} - p_{n-1}| > |\alpha q_{n-1} - p_{n-1}|.$$

If both $\mu \neq 0$ and $v \neq 0$, then since $1 \leq q \leq q_n$, μ and v are of opposite sign. Hence $\mu(\alpha q_n - p_n)$ and $v(\alpha q_{n-1} - p_{n-1})$ are of the same sign, and therefore

$$|\alpha q - p| = |\mu(\alpha q_n - p_n)| + |v(\alpha q_{n-1} - p_{n-1})|.$$

Thus $|\alpha q - p| > |\alpha q_{n-1} - p_{n-1}|$ since $\mu v \neq 0$, and (ii) is established.

Definition. An irrational number α is badly approximable if there is a constant $c = c(\alpha) > 0$ such that

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

for every rational $\frac{p}{q}$. (In view of Theorem 2F, such a constant c must satisfy $0 < c < \frac{1}{\sqrt{5}}$.)

THEOREM 5F. α is badly approximable if and only if the partial quotients in its continued fraction expansion are bounded:

$$|a_n| \leq K(\alpha).$$

Proof. In order to study the inequality

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

where $0 < c(\alpha) < \frac{1}{2}$, we may restrict ourselves to convergents $\frac{p_n}{q_n}$ by

Theorem 5C. From the proof of Theorem 5B and from Lemma 3F, we have

$$\left| \alpha - \frac{p_n}{q_n} \right| = \frac{1}{q_n^2 (\alpha_{n+1} + \beta_{n+1})} = \frac{1}{q_n^2 [a_{n+1}, a_{n+2}, \dots] + \frac{1}{[a_n, a_{n-1}, \dots, a_1]}} ,$$

and hence

$$\frac{1}{q_n^2 (a_{n+1} + 2)} \leq \left| \alpha - \frac{p_n}{q_n} \right| \leq \frac{1}{q_n^2 a_{n+1}} .$$

The theorem follows.

COROLLARY 5G. There exist continuum many badly approximable numbers,
and there exist continuum many numbers which are not badly approximable.

§6. Further results.

We state results without proofs. For proofs see Perron (1954), Cassels (1957), Hardy and Wright (1954), and the references given below.

Definition. An infinite continued fraction $[a_0, a_1, a_2, \dots]$ is periodic if there are integers $k \geq 0$, $m \geq 1$ such that

$$a_{n+m} = a_n \text{ provided } n \geq k .$$

In this case, we write the fraction as

$$[a_0, a_1, \dots, a_{k-1}, \overline{a_k, a_{k+1}, \dots, a_{k+m-1}}] .$$

THEOREM 6A^{*}. α has a periodic simple continued fraction expansion if and only if α is a quadratic irrational.

The necessity was proved by Euler (1737); see also (1748). The sufficiency was proved by Lagrange (1770b). See also Lagrange's Oeuvres, II.

Definition. Two irrational numbers α and β are equivalent if there exist integers a, b, c, d with $ad - bc = \pm 1$ such that

$$\beta = \frac{a\alpha + b}{c\alpha + d}.$$

(It is not difficult to verify that this is an equivalence relation.)

THEOREM 6B^{*}. (Serret, 1878)[†] Let $\alpha = [a_0, a_1, a_2, \dots]$ and $\beta = [b_0, b_1, b_2, \dots]$ be irrational. Then α and β are equivalent precisely if there are integers u and v such that $a_{u+n} = b_{v+n}$ ($n = 1, 2, \dots$).

THEOREM 6C^{*} (Hurwitz, 1891). Suppose that $\alpha = [a_0, a_1, a_2, \dots]$ with $a_n \geq 2$ for infinitely many n . Then there are infinitely many distinct rationals $\frac{p}{q}$ with

$$|\alpha - \frac{p}{q}| < \frac{1}{\sqrt{8} q^2}.$$

COROLLARY 6D^{*}. The inequality

$$|\alpha - \frac{p}{q}| < \frac{1}{\sqrt{8} q^2}$$

has infinitely many rational solutions $\frac{p}{q}$ whenever α is not equivalent to $\frac{1}{2}(\sqrt{5} + 1)$.

Definition. $v(\alpha) = \liminf_{q \rightarrow \infty} q \|q\alpha\|$.

Remarks. Trivially, $v(\alpha) = 0$ whenever α is rational. It is clear that $v(\alpha) > 0$ if and only if α is badly approximable. Theorem 2F implies that $v(\alpha) \leq \frac{1}{\sqrt{5}}$ for every real number α . There are

[†]) This is the date of the German translation. I was unable to find the original date.

numbers α with $v(\alpha) = \frac{1}{\sqrt{5}} : \alpha = \frac{1}{2}(\sqrt{5}+1)$, for example. Corollary

6D* implies that $v(\alpha) \leq \frac{1}{\sqrt{8}}$ whenever α is not equivalent to $\frac{1}{2}(\sqrt{5}+1)$.

THEOREM 6E* (Markoff (1879). See also Hurwitz (1906)). There exist numbers

$$\mu_1 = \frac{1}{\sqrt{5}} > \mu_2 = \frac{1}{\sqrt{8}} > \mu_3 > \mu_4 > \dots$$

with limit $\frac{1}{3}$ such that for every μ_i there are finitely many equivalence classes (as defined above Theorem 6B*) of numbers such that

$$v(\alpha) = \mu_i$$

precisely if α lies in such a class. Furthermore, if $\mu > \frac{1}{3}$ and $\mu \neq \mu_i$ ($i = 1, 2, 3, \dots$), then there is no α with $v(\alpha) = \mu$. Finally, there are continuum many α with $v(\alpha) = \frac{1}{3}$.

THEOREM 6F* (Perron, 1921a). There are no α with
 $\frac{1}{\sqrt{13}} < v(\alpha) < \frac{1}{\sqrt{12}}$, but there are α with $v(\alpha) = \frac{1}{\sqrt{13}}$ or $v(\alpha) = \frac{1}{\sqrt{12}}$.

THEOREM 6G* (M. Hall. The theorem follows from his (1947) paper). There is a constant $c > 0$ such that, for every number v satisfying $0 \leq v < c$, there is an α with $v(\alpha) = v$.

THEOREM 6H* (Hightower, 1970). There are countably many disjoint intervals I_1, I_2, \dots contained in $0 \leq x \leq \frac{1}{3}$ such that

(i) There is no α with $v(\alpha) \in \bigcup_{i=1}^{\infty} I_i$.

(ii) If I_i, I_j are any two distinct such intervals, there is an

α with $\nu(\alpha)$ "between" I_i and I_j .

The above results can be formulated in terms of the Lagrange spectrum, which is defined as the set of numbers ν of the type $\nu = \nu(\alpha)$ for some α . For recent references on this area, see Cusick (1975).

Definition. We say that Dirichlet's Theorem can be improved for some particular α if there is a constant $c(\alpha) < 1$ such that for every $Q > Q_0(\alpha)$, there exist integers p, q with $1 \leq q < Q$ and $|\alpha q - p| < \frac{c(\alpha)}{Q}$.

THEOREM 6I^{*} (Davenport and Schmidt (1968)). Dirichlet's Theorem can be improved for α precisely if α is badly approximable. Namely, put

$$\gamma(\alpha) = \liminf \frac{1}{[a_{n+1}, a_{n+2}, \dots][a_n, a_{n-1}, \dots, a_1]} .$$

Then one may pick $c(\alpha) > (1 + \gamma(\alpha))^{-1}$, but one may not pick $c(\alpha) < (1 + \gamma(\alpha))^{-1}$.

THEOREM 6J^{*} (Segre (1945)). Suppose that α is irrational and $\tau \geq 0$. Then there exist infinitely many rationals $\frac{h}{k}$ with

$$-\frac{1}{(1+4\tau)^{1/2} k^2} < \alpha - \frac{h}{k} < \frac{\tau}{(1+4\tau)^{1/2} k^2} .$$

Remark. If $\tau = 1$, the theorem reduces to Hurwitz' Theorem. A proof of the theorem, using Farey series, may be found, e.g., in Niven (1962).

II. Simultaneous Approximation

References: Dirichlet (1842), Minkowski (1896), Perron (1921b).

§1. Dirichlet's Theorem on Simultaneous Approximation

THEOREM 1A. Suppose that $\alpha_1, \dots, \alpha_n$ are n real numbers and that $Q > 1$ is an integer. Then there exist integers q, p_1, \dots, p_n with

$$(1.1) \quad 1 \leq q < Q^n \quad \text{and} \quad |\alpha_i q - p_i| \leq \frac{1}{Q} \quad (1 \leq i \leq n).$$

COROLLARY 1B. Suppose that at least one of $\alpha_1, \dots, \alpha_n$ is irrational.

Then there are infinitely many n -tuples $\frac{p_1}{q}, \dots, \frac{p_n}{q}$ with

$$(1.2) \quad \left| \alpha_i - \frac{p_i}{q} \right| < \frac{1}{q^{1+1/n}} \quad (1 \leq i \leq n).$$

Deduction of the Corollary. In Theorem 1A, we may require in addition that $\text{g.c.d.}(q, p_1, \dots, p_n) = 1$. The inequalities of the theorem clearly imply (1.2).

If, say, α_1 is irrational, then

$$|\alpha_1 q - p_1| \neq 0.$$

Hence for fixed q, p_1, \dots, p_n , (1.1) can hold only for $Q \leq Q_0$. Hence as $Q \rightarrow \infty$, we get infinitely many different solutions.

THEOREM 1C. Suppose that $\alpha_1, \dots, \alpha_n$ and Q are given as in Theorem 1A. Then there exist integers q_1, \dots, q_n, p with

$$(1.3) \quad 1 \leq \max(|q_1|, \dots, |q_n|) < Q^{1/n} \quad \text{and} \quad |a_1 q_1 + \dots + \alpha_n q_n - p| \leq \frac{1}{Q}.$$

COROLLARY 1D. Suppose that $1, \alpha_1, \dots, \alpha_n$ are linearly independent over the rationals. Then there are infinitely many coprime $(n+1)$ -tuples

(q_1, \dots, q_n, p) with

$$(1.4) \quad q = \max(|q_1|, \dots, |q_n|) > 0 \quad \text{and} \quad |\alpha_1 q_1 + \dots + \alpha_n q_n - p| < \frac{1}{q}.$$

Deduction of the Corollary. It is clear that (1.3) implies (1.4).

By linear independence, we always have

$$|\alpha_1 q_1 + \dots + \alpha_n q_n - p| \neq 0.$$

Hence for fixed q_1, \dots, q_n, p , (1.3) can hold only for $Q \leq Q_0$. Hence as $Q \rightarrow \infty$, we obtain infinitely many solutions.

THEOREM 1E. (Dirichlet (1842)) Suppose that α_{ij} ($1 \leq i \leq n$, $1 \leq j \leq m$) are nm real numbers and that $Q > 1$ is an integer. Then there exist integers $q_1, \dots, q_m, p_1, \dots, p_n$ with

$$(1.5) \quad 1 \leq \max(|q_1|, \dots, |q_m|) < Q^{n/m},$$

$$|\alpha_{i1} q_1 + \dots + \alpha_{im} q_m - p_i| \leq \frac{1}{Q} \quad (1 \leq i \leq n).$$

Remark. The hypothesis that " Q is an integer" is unnecessary. Our proof of Theorem 1E cannot immediately be adopted to this generalization, but see the remark following the proof of Theorem 2A. It is clear that Theorem 1E implies the two preceding theorems. We omit the obvious proof of the following

COROLLARY 1F. Suppose, for some i in $1 \leq i \leq n$, that $\alpha_{i1}, \dots, \alpha_{im}, 1$ are linearly independent over the rationals. Or suppose, more generally, that

$$(\alpha_{11} q_1 + \dots + \alpha_{1m} q_m, \dots, \alpha_{n1} q_1 + \dots + \alpha_{nm} q_m)$$

is never an integer point when (q_1, \dots, q_m) is a nonzero integer point. Then

there exist infinitely many coprime $(m+n)$ -tuples $(q_1, \dots, q_m, p_1, \dots, p_n)$ with

$$q = \max(|q_1|, \dots, |q_m|) > 0 \text{ and } |\alpha_{11}q_1 + \dots + \alpha_{1m}q_m - p_1| < \frac{1}{q^{n/m}} \quad (1 \leq i \leq n).$$

Proof of Theorem 1E. Consider points

$$(\{\alpha_{11}x_1 + \dots + \alpha_{1m}x_m\}, \dots, \{\alpha_{n1}x_1 + \dots + \alpha_{nm}x_m\}),$$

where each x_j is an integer satisfying

$$0 \leq x_j < Q^{n/m} \quad (1 \leq j \leq m).$$

There are at least Q^n such points, and each of these points lies in the closed unit cube \bar{U}^n consisting of points (t_1, \dots, t_n) with $0 \leq t_k \leq 1$ ($k = 1, \dots, n$). The point $(1, 1, \dots, 1)$ also lies in \bar{U}^n , so together there are at least $Q^n + 1$ points under consideration.

We divide \bar{U}^n into Q^n pairwise disjoint subcubes of side length $\frac{1}{Q}$. (Thus some of the cubes will contain some of their faces or edges and not others). Two of the points under consideration will be in the same subcube. These points are, say,

$$(\alpha_{11}x_1 + \dots + \alpha_{1m}x_m - y_1, \dots, \alpha_{n1}x_1 + \dots + \alpha_{nm}x_m - y_n),$$

$$(\alpha_{11}x'_1 + \dots + \alpha_{1m}x'_m - y'_1, \dots, \alpha_{n1}x'_1 + \dots + \alpha_{nm}x'_m - y'_n).$$

Here, $(x_1, \dots, x_m) \neq (x'_1, \dots, x'_m)$. Put $q_i = x_i - x'_i$ ($1 \leq i \leq m$) and $p_j = y_j - y'_j$ ($1 \leq j \leq n$). Then (1.5) is easily seen to be satisfied, and the proof is complete.

§2. Theorems of Blichfeldt and Minkowski.

Notation. In this section and throughout the remainder of these notes, we make the following conventions: E^n denotes real n -dimensional

Euclidean space, points of which are denoted by $\underline{x} = (x_1, \dots, x_n)$, \underline{y} , etc.; U is the closed-open unit interval $0 \leq x < 1$, and U^n is the unit cube of points $\underline{x} = (x_1, \dots, x_n)$ with each $x_k \in U$. \bar{U}^n is the closure of U^n .

Suppose that $\underline{x} = (x_1, \dots, x_n) \in E^n$. We put

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

If each x_k is an integer, we say that \underline{x} is an integer point.

Let \mathcal{S} be any subset of E^n . If $\underline{x} \in E^n$, we denote by $\mathcal{S} + \underline{x}$ the set[†] of all points $\underline{s} + \underline{x} = (s_1 + x_1, \dots, s_n + x_n)$ with $\underline{s} \in \mathcal{S}$. If λ is a real number, then $\lambda\mathcal{S}$ denotes the set of all points $\lambda\underline{s} = (\lambda s_1, \dots, \lambda s_n)$ with $\underline{s} \in \mathcal{S}$.

THEOREM 2A. (Blichfeldt (1914)) Let \mathcal{P} be a discrete (i.e., without limit point) set of points in E^n , invariant under translation by integer points, and with precisely N points in U^n . Let R be a measurable set of measure $\mu(R) > 0$. Then there is a point \underline{x} in U^n such that $R + \underline{x}$ contains at least $N\mu(R)$ points of \mathcal{P} . Moreover, if R is compact, then there is a translation of R containing more than $N\mu(R)$ points of \mathcal{P} .

Proof. Write $v(\mathcal{S})$ for the number of points of \mathcal{P} lying in a set \mathcal{S} . Now \mathcal{P} has N points in U^n , say $\underline{p}_1, \dots, \underline{p}_N$. Write $v_i(\mathcal{S})$ for the number of points $\underline{p}_i + \underline{g}$ (\underline{g} an integer point) which lie in \mathcal{S} ($1 \leq i \leq n$). Since \mathcal{P} is invariant under translation by integer points,

$$(2.1) \quad v(\mathcal{S}) = \sum_{i=1}^N v_i(\mathcal{S}).$$

[†]called the translation of \mathcal{S} by \underline{x} .

For each i ,

$$v_i(\mathcal{R} + \underline{x}) = \sum_{\underline{g}} \chi(\underline{p}_i + \underline{g} - \underline{x}) ,$$

where χ is the characteristic function of \mathcal{R} and the sum is over all integer points $\underline{g} \in E^n$. We obtain

$$\begin{aligned} \int_{U^n} v_i(\mathcal{R} + \underline{x}) d\underline{x} &= \int_{U^n} \sum_{\underline{g}} \chi(\underline{p}_i + \underline{g} - \underline{x}) d\underline{x} \\ &= \int_{E^n} \chi(\underline{z}) d\underline{z} \\ &= \mu(\mathcal{R}) . \end{aligned}$$

It follows from (2.1) that

$$\int_{U^n} v(\mathcal{R} + \underline{x}) d\underline{x} = N_\mu(\mathcal{R}) ,$$

and therefore $v(\mathcal{R} + \underline{x}) \geq N_\mu(\mathcal{R})$ for some $\underline{x} \in U^n$.

If \mathcal{R} is compact, there is nothing more to prove unless $N_\mu(\mathcal{R})$ is an integer, say v_0 . Suppose, then, that this is the case. It follows from the first part of the proof that for every positive integer k , there is a point $\underline{x}_k \in U^n$ such that $v(\mathcal{E}_k) \geq v_0 + 1$, where

$$\mathcal{E}_k = (1 + \frac{1}{k})\mathcal{R} + \underline{x}_k .$$

Since each \underline{x}_k belongs to the compact set \bar{U}^n , the sequence $(\underline{x}_k)_{k=1}^\infty$ contains a convergent subsequence $(\underline{x}_{k_j})_{j=1}^\infty$: say $\underline{x}_{k_j} \rightarrow \underline{x}_0$.

We have $v(\mathcal{E}_{k_j}) \geq v_0 + 1$ for each j ; since all of the sets \mathcal{E}_{k_j} are uniformly bounded, there are $v_0 + 1$ points $\underline{u}_1, \dots, \underline{u}_{v_0+1}$ which lie in infinitely many of the sets \mathcal{E}_{k_j} . Then $\underline{u}_1, \dots, \underline{u}_{v_0+1}$

also lie in $\mathcal{R} + \underline{x}_0$, in view of the compactness of \mathcal{R} . Hence

$$v(\mathcal{R} + \underline{x}_0) \geq v_0 + 1 > N_\mu(\mathcal{R}).$$

If \underline{x}_0 does not lie in U^n , we may replace it by a point in U^n which is obtained from it by a translation by an integer point.

The proof is now complete.

Remark. Theorem 2A may be used to extend Theorem 1E to hold for any $Q > 1$. Namely, consider points

$$(\alpha_{11}x_1 + \dots + \alpha_{1m}x_m, \dots, \alpha_{n1}x_1 + \dots + \alpha_{nm}x_m),$$

where each x_j is an integer satisfying

$$0 \leq x_j < Q^{n/m} \quad (1 \leq j \leq m).$$

Let \mathcal{O} be the set of such points, together with all of their translations by integer points. Note that the number N of points in \mathcal{O} lying in U^n satisfies $N \geq Q^n$. Let \mathcal{R} be the cube $0 \leq t_k \leq \frac{1}{Q}$ ($1 \leq k \leq n$). \mathcal{R} is compact and $\mu(\mathcal{R}) = 1/Q^n$.

By Theorem 2A, there is a translation $\mathcal{R} + \underline{x}$ with $v(\mathcal{R} + \underline{x}) > N_\mu(\mathcal{R}) \geq 1$: i.e., $v(\mathcal{R} + \underline{x}) \geq 2$. The proof is now completed as before.

THEOREM 2B. (Minkowski's Convex Body Theorem. Minkowski (1896)).

Let \mathcal{R} be a convex set in E^n , symmetric about $\underline{0}$, bounded and with volume $\mu(\mathcal{R})^\dagger$. Assume either that $\mu(\mathcal{R}) > 2^n$ or that \mathcal{R} is compact and $\mu(\mathcal{R}) \geq 2^n$. Then \mathcal{R} contains an integer point $\neq \underline{0}$.

By considering the cube of \underline{x} with $|x_i| < 1$ ($i = 1, \dots, n$) we see that Minkowski's Theorem is best possible.

[†]) It may be shown that such a set always does have a (Jordan)-volume. A reader who does not wish to use this fact may regard the existence of a volume as a further assumption on \mathcal{R} .

Proof. Either $\mu(\frac{1}{2}\mathcal{R}) > 1$, or \mathcal{R} is compact and $\mu(\frac{1}{2}\mathcal{R}) \geq 1$. In either case, we apply Theorem 2A to $\frac{1}{2}\mathcal{R}$ and \emptyset , where \emptyset is the set of all integer points in E^n . Then there is a translation $\frac{1}{2}\mathcal{R} + \underline{x}$ which contains two distinct integer points, say \underline{g}_1 and \underline{g}_2 .

Thus $\underline{g}_1 - \underline{x} \in \frac{1}{2}\mathcal{R}$ and $\underline{g}_2 - \underline{x} \in \frac{1}{2}\mathcal{R}$. By symmetry, $\underline{x} - \underline{g}_2 \in \frac{1}{2}\mathcal{R}$. Therefore, we may put $\underline{g}_1 - \underline{x} = \frac{1}{2}\underline{x}_1$ and $\underline{x} - \underline{g}_2 = \frac{1}{2}\underline{x}_2$, where \underline{x}_1 and \underline{x}_2 lie in \mathcal{R} . Now \mathcal{R} is convex, so $\underline{g}_1 - \underline{g}_2 = \frac{1}{2}\underline{x}_1 + \frac{1}{2}\underline{x}_2 \in \mathcal{R}$. Hence $\underline{g} = \underline{g}_1 - \underline{g}_2$ satisfies the conditions of the theorem.

Here is

Another proof of Minkowski's Theorem, due to Mordell (1934).

Suppose that $\mu(\mathcal{R}) > 2^n$. For positive integers m let \mathcal{R}_m be the set of points in \mathcal{R} which have rational coordinates with denominator m . As $m \rightarrow \infty$, the number of points in \mathcal{R}_m will be asymptotically equal to $\mu(\mathcal{R})m^n$, and hence for certain large m this number will be greater than $(2m)^n$. There will then be two points $(\frac{a_1}{m}, \dots, \frac{a_n}{m})$ and $(\frac{b_1}{m}, \dots, \frac{b_n}{m})$ in \mathcal{R}_m having

$$(2.2) \quad a_i \equiv b_i \pmod{2m} \quad (i = 1, \dots, n).$$

By the symmetry and convexity of \mathcal{R} , the point

$$\underline{g} = \frac{1}{2} \left(\frac{a_1}{m}, \dots, \frac{a_n}{m} \right) - \frac{1}{2} \left(\frac{b_1}{m}, \dots, \frac{b_n}{m} \right)$$

is a non-zero point in \mathcal{R} , which is an integer point by (2.2).

The case when $\mu(\mathcal{R}) = 2^n$ and \mathcal{R} is compact can easily be reduced to the case when $\mu(\mathcal{R}) > 2^n$.

THEOREM 2C. (Minkowski's Linear Forms Theorem) Suppose that β_{ij} ($1 \leq i, j \leq n$) are real numbers with determinant ± 1 . Suppose that

A_1, \dots, A_n are positive with $A_1 A_2 \dots A_n = 1$. Then there exists an integer point $\underline{x} = (x_1, \dots, x_n) \neq \underline{0}$ such that

$$(2.3) \quad |\beta_{11}x_1 + \dots + \beta_{1n}x_n| < A_1, \quad (1 \leq i \leq n-1),$$

and

$$(2.4) \quad |\beta_{n1}x_1 + \dots + \beta_{nn}x_n| \leq A_n.$$

Proof. Write

$$L_i(\underline{x}) = \beta_{i1}x_1 + \dots + \beta_{in}x_n \quad (1 \leq i \leq n),$$

and put

$$L'_i(\underline{x}) = \frac{1}{A_i} L_i(\underline{x}) \quad (1 \leq i \leq n).$$

Then (2.3) and (2.4) may be rewritten as

$$|L'_i(\underline{x})| < 1 \quad (1 \leq i \leq n-1)$$

and

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

The determinant of L'_1, \dots, L'_n is again equal to $\det(\beta_{ij}) = \pm 1$, so that we may restrict ourselves to the case in which $A_1 = \dots = A_n = 1$.

The basic idea of the proof is now as follows. The set \mathcal{R} of all $\underline{x} \in \mathbb{E}^n$ satisfying

$$|L'_i(\underline{x})| \leq 1 \quad (1 \leq i \leq n)$$

is obtained from \bar{U}^n by a linear transformation of determinant 1.

Hence \mathcal{R} is a symmetric, closed parallelepiped of volume 2^n . By

Minkowski's Convex Body Theorem, there is an integer point $\underline{x} \neq \underline{0}$ in this parallelepiped.

To get strict inequality in the first $n-1$ inequalities we modify this argument as follows. For each $\varepsilon > 0$, the system of inequalities $|L_i(\underline{x})| < 1$ ($1 \leq i \leq n-1$), $|L_n(\underline{x})| < 1+\varepsilon$ defines a symmetric parallelepiped Π_ε of volume $2^n(1+\varepsilon) > 2^n$. By Minkowski's Convex Body Theorem, there is an integer point $\underline{x}_\varepsilon \neq \underline{0}$ in Π_ε .

Now take a sequence $(\varepsilon_k)_{k=1}^\infty$, $\varepsilon_k \rightarrow 0^+$. This sequence gives rise to a sequence $(\underline{x}_{\varepsilon_k})_{k=1}^\infty$ of nonzero integer points, $\underline{x}_{\varepsilon_k} \in \Pi_{\varepsilon_k}$. Since all of the sets Π_{ε_k} are uniformly bounded, there is an integer point $\underline{x} \neq \underline{0}$ such that $\underline{x} = \underline{x}_{\varepsilon_k}$ for infinitely many k . Thus $\underline{x} \in \Pi_{\varepsilon_k}$ for infinitely many k , whence \underline{x} satisfies (2.3) and (2.4).

Another Proof of Theorem 1E. Put $l = m + n$, and consider the following l linear forms in $\underline{x} = (x_1, \dots, x_l)$:

$$L_i(\underline{x}) = x_i, \quad (1 \leq i \leq m)$$

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

Their determinant clearly is ± 1 .

Let $Q > 1$ be given, Q not necessarily an integer. By Lemma 2C, there is an integer point $\underline{x} \neq \underline{0}$ with

$$|L_i(\underline{x})| < Q^{n/m}, \quad (1 \leq i \leq m)$$

$$|L_{m+j}(\underline{x})| \leq Q^{-1} \quad (1 \leq j \leq n).$$

Put $q_i = x_i$ ($1 \leq i \leq m$) and $p_j = x_{m+j}$ ($1 \leq j \leq n$). Then

$$q = \max(|q_1|, \dots, |q_n|) < Q^{n/m},$$

and

$$(2.5) \quad |\alpha_{j1}q_1 + \dots + \alpha_{jm}q_m - p_j| \leq \frac{1}{Q} \quad (1 \leq j \leq n).$$

It remains to be shown that $q \geq 1$. If not, then all $q_i = 0$, and therefore

$$|p_j| \leq \frac{1}{Q} < 1 \quad (1 \leq i \leq n)$$

from (2.5). But then all $p_j = 0$, whence $\underline{x} = \underline{0}$, a contradiction.

§3. Improvement of the Simultaneous Approximation Constants.

THEOREM 3A. (Minkowski (1910) did the case $m = 1$). Consider
linear forms

$$L_i(\underline{x}) = \alpha_{i1}x_1 + \dots + \alpha_{im}x_m \quad (1 \leq i \leq n)$$

where $\underline{x} = (x_1, \dots, x_m)$. Put

$$f(\underline{x}) = (L_1(\underline{x}), \dots, L_n(\underline{x})).$$

Then there is an integer point $(\underline{x}, \underline{y}) = (x_1, \dots, x_m, y_1, \dots, y_n)$ in
 E^{m+n} with $\underline{x} \neq \underline{0}$ such that

$$(3.1) \quad |f(\underline{x}) - \underline{y}|^n < C_{m,n} \frac{1}{|\underline{x}|^m},$$

where

$$C_{m,n} = \frac{m^n n^n}{(m+n)^{m+n}} \cdot \frac{(m+n)!}{m!n!}.$$

Furthermore, suppose that whenever $\underline{x} \neq \underline{0}$ is an integer point in E^m ,
then $f(\underline{x})$ is not an integer point in E^n . Then there exist infinitely

many integer points $(\underline{x}, \underline{y})$ with $\underline{x} \neq \underline{0}$ and with coprime components satisfying (3.1) .

Remark. We make explicit the connection between Theorems 1E and 3A.

The conclusion of Theorem 1E may be restated as follows: There is an integer points $(\underline{x}, \underline{y}) = (x_1, \dots, x_m, y_1, \dots, y_n)$ in E^{m+n} with

$$1 \leq \left\lceil \underline{x} \right\rceil < Q^{n/m} \quad \text{and} \quad \left\lceil \mathcal{L}(\underline{x}) - \underline{y} \right\rceil \leq \frac{1}{Q} .$$

These inequalities yield

$$\left\lceil \mathcal{L}(\underline{x}) - \underline{y} \right\rceil^n \leq \frac{1}{Q^n} < \frac{1}{\left\lceil \underline{x} \right\rceil^m} .$$

Theorem 3A is an improvement over Theorem 1E since $C_{m,n} < 1$. Namely $C_{m,n} = \binom{m+n}{m} \left(\frac{m}{m+n}\right)^m \left(\frac{n}{m+n}\right)^n$ is just one of the $m+n+1$ terms in the binomial expansion of

$$1 = 1^{m+n} = \left(\frac{m}{m+n} + \frac{n}{m+n}\right)^{m+n} .$$

LEMMA 3B. Suppose that $m \geq 1$, $n \geq 1$ are integers and that $t > 0$.

Let $K_{m,n}$ be the set of points $(\underline{x}, \underline{y}) = (x_1, \dots, x_m, y_1, \dots, y_n)$ in E^{m+n} satisfying

$$(3.2) \quad t^{-n} \left\lceil \underline{x} \right\rceil + t^m \left\lceil \underline{y} \right\rceil \leq 1 .$$

Then $K_{m,n}$ is compact, symmetric (with respect to $\underline{0}$) and convex, with volume

$$V(m,n) = 2^{m+n} \frac{m!n!}{(m+n)!} .$$

Proof. Let us write $K_{m,n}(t)$ in place of $K_{m,n}$ for the moment. The transformation given by

$$\begin{aligned} x_i &\mapsto t^n x_i & (1 \leq i \leq m) \\ y_j &\mapsto t^{-m} y_j & (1 \leq j \leq n) \end{aligned}$$

is linear and has determinant $(t^n)^m (t^{-m})^n = 1$; furthermore, this transformation maps $K_{m,n}(1)$ onto $K_{m,n}(t)$. Since a linear transformation preserves compactness, symmetry and convexity, it is enough to verify that $K_{m,n}$ has these properties when $t = 1$.

The compactness and symmetry of $K_{m,n}$ are obvious. To prove convexity, suppose that $(\underline{x}, \underline{y})$ and $(\underline{x}', \underline{y}')$ belong to $K_{m,n}$ and that $\lambda \geq 0$, $\mu \geq 0$ satisfy $\lambda + \mu = 1$. Then $\lambda(\underline{x}, \underline{y}) + \mu(\underline{x}', \underline{y}') = (\lambda \underline{x} + \mu \underline{x}', \lambda \underline{y} + \mu \underline{y}')$ belongs to $K_{m,n}$, since

$$|\lambda \underline{x} + \mu \underline{x}'| + |\lambda \underline{y} + \mu \underline{y}'| \leq \lambda(|\underline{x}| + |\underline{y}|) + \mu(|\underline{x}'| + |\underline{y}'|) \leq \lambda + \mu = 1.$$

It remains to compute $V(m,n)$, the volume of $K_{m,n}$. Let $K_{m,n}^*$ denote the set of points $(\underline{x}, \underline{y})$ in $K_{m,n}$ for which $x_i \geq 0$ ($1 \leq i \leq m$), $y_j \geq 0$ ($1 \leq j \leq n$), and $x_1 = |\underline{x}|$. By (3.2), $0 \leq x_1 \leq 1$; moreover, for each fixed x_1 , $0 \leq x_i \leq x_1$ ($2 \leq i \leq m$) and $0 \leq y_j \leq 1 - x_1$ ($1 \leq j \leq n$). If $V^*(m,n)$ denotes the volume of $K^*(m,n)$, it follows that

$$\begin{aligned} V(m,n) &= m 2^{m+n} V^*(m,n) \\ &= m 2^{m+n} \int_0^1 x_1^{m-1} (1-x_1)^n dx_1. \end{aligned}$$

The value of the integral is $\frac{(m-1)!n!}{(m+n)!}$ (this follows by induction on m or by the reader's knowledge of the Beta function), and the proof of the lemma is complete.

Proof of Theorem 3A. Suppose that $t > 0$. Let $K_{m,n}$ be as in

Lemma 3B, and let $\tilde{K}_{m,n}$ be the set of points $(\underline{x}, \underline{y}) = (x_1, \dots, x_m, y_1, \dots, y_n)$ in E^{m+n} satisfying

$$t^{-n} \|\underline{x}\| + t^m \|\mathcal{L}(\underline{x}) - \underline{y}\| \leq C ,$$

where $C = 2V(m,n)^{-1/(m+n)}$. The linear transformation defined by

$$x_i \mapsto Cx_i \quad (1 \leq i \leq m)$$

$$y_j \mapsto C(L_j(\underline{x}) - y_j) \quad (1 \leq j \leq n)$$

maps $K_{m,n}$ onto $\tilde{K}_{m,n}$ and has determinant $(-1)^n C^{m+n}$, so that $\tilde{K}_{m,n}$ has volume $C^{m+n} V(m,n) = 2^{m+n}$. Since compactness, symmetry and convexity are preserved under linear transformations, $\tilde{K}_{m,n}$ has these three properties in view of Lemma 3B. By Minkowski's Convex Body Theorem, it follows that $\tilde{K}_{m,n}$ contains an integer point $(\underline{x}, \underline{y}) \neq (\underline{0}, \underline{0})$.

Now put $M = \|\underline{x}\|$, $L = \|\mathcal{L}(\underline{x}) - \underline{y}\|$: then

$$(3.3) \quad t^{-n} M + t^m L \leq C .$$

For a given integer point $(\underline{x}, \underline{y})$, equality in (3.3) can hold for at most finitely many values of t . Since the number of integer points in E^{m+n} is countable, it follows that

$$(3.4) \quad t^{-n} M + t^m L < C$$

for all but countably many t . In the remainder of the proof, we consider only those t which satisfy (3.4).

Recall the Arithmetic-Geometric Inequality: If z_1, z_2, \dots, z_ℓ are ℓ nonnegative numbers, then

$$(z_1 z_2 \dots z_\ell)^{1/\ell} \leq \frac{z_1 + z_2 + \dots + z_\ell}{\ell} .$$

We apply this with $\ell = m+n$, $z_1 = \dots = z_m = m^{-1} t^{-n} M$, $z_{m+1} = \dots = z_{m+n} = n^{-1} t^m L$, to obtain

$$(m^{-1}t^{-n_M})^m(n^{-1}t^{m_L})^n \leq \left(\frac{t^{-n_M} + t^{m_L}}{t}\right)^t < \left(\frac{C}{t}\right)^t,$$

which yields

$$M_L^m n < \frac{m^m n^m}{(m+n)^{m+n}} C^{m+n} = C_{m,n}.$$

This inequality is equivalent to (3.1), provided $M \neq 0$. If we choose t to satisfy (3.4) and $t \geq C^{1/m}$, then $M \neq 0$, since otherwise (3.4) implies that

$$|\overline{f(\underline{0})} - \underline{y}| = |\underline{y}| < Ct^{-m} \leq 1,$$

hence $\underline{y} = \underline{0}$, a contradiction. This establishes the first assertion of the theorem.

To complete the proof of the theorem, suppose now that, whenever $\underline{x} \neq \underline{0}$ is an integer point in E^m , $f(\underline{x})$ is not an integer point in E^n . Further suppose that t satisfies (3.4) and $t \geq C^{1/m}$. Now $M \neq 0$ by the preceding paragraph, hence $L \neq 0$. It follows that for fixed $(\underline{x}, \underline{y})$, inequality (3.4) can hold only for $t \leq t_0$. Hence as $t \rightarrow \infty$, there will be infinitely many distinct integer points $(\underline{x}, \underline{y})$ with coprime components and with $\underline{x} \neq \underline{0}$ satisfying (3.1).

Remarks. Consider the special case of Theorem 3A in which $m = 1$. Then $C_{1,n} = \left(\frac{n}{n+1}\right)^n$. It was shown by Blichfeldt (1914) that $C_{1,n}$ may be replaced by the still smaller constant

$$\gamma_n = \left(\frac{n}{n+1}\right)^n \left(1 + \left(\frac{n-1}{n+1}\right)^{n+3}\right)^{-1}.$$

When $m = 1$, the inequality (3.1) may be written as

$$\max(|\alpha_1^{x-y_1}|, \dots, |\alpha_n^{x-y_n}|) < \frac{n}{n+1} x^{-1/n}.$$

If $n = 1$, the coefficient of $x^{-1/n}$ here is $\frac{1}{2}$. By Hurwitz' Theorem of Chapter I, a coefficient $1/\sqrt{5}$ is best possible. If $n = 2$, the coefficient in question is $2/3$. The best possible coefficient is unknown. However, if c_0 denotes the infimum of admissible coefficients when $n = 2$, we have shown that $c_0 \leq 2/3$, and it is known that

$$0.53 \approx \sqrt{2/7} \leq c_0 \leq 0.615.$$

See Cassels (1955a), Davenport (1952), Davenport & Mahler (1946), Mack (1977).

In the next section we are going to show that the best coefficient is positive for every n, m .

§4. Badly Approximable Systems of Linear Forms.

Let α_{ij} ($1 \leq i \leq n$, $1 \leq j \leq m$) be nm real numbers, and consider linear forms

$$L_i(\underline{x}) = \alpha_{i1}x_1 + \dots + \alpha_{im}x_m, \quad (1 \leq i \leq n)$$

where $\underline{x} = (x_1, \dots, x_m)$. Put

$$\mathcal{L}(\underline{x}) = (L_1(\underline{x}), \dots, L_n(\underline{x})).$$

Definition. L_1, \dots, L_n is a badly approximable system of linear forms if there is a constant $\gamma = \gamma(L_1, \dots, L_n) > 0$ such that

$$(4.1) \quad \left| \underline{x} \right|^m \left| \mathcal{L}(\underline{x}) - \underline{y} \right|^n > \gamma$$

for every integer point $(\underline{x}, \underline{y}) = (x_1, \dots, x_m, y_1, \dots, y_n)$ with $\underline{x} \neq \underline{0}$.

(In other words, these are linear forms for which $C_{m,n}$ in Theorem 3A cannot be replaced by arbitrarily small constants.)

Remark. Consider, for the moment, the special case $m = n = 1$. Then

(4.1) reduces to $|x| \cdot |\alpha_{11}x - y| > \gamma$, or

$$|\alpha_{11} - \frac{y}{x}| > \frac{\gamma}{x^2}.$$

Thus, the linear form $L(x) = \alpha x$ is badly approximable precisely if the coefficient α is badly approximable as defined in §5 of Chapter I.

In the special case when $n = 1$, a single linear form

$L(\underline{x}) = \alpha_1 x_1 + \dots + \alpha_m x_m$ is badly approximable if

$$|\alpha_1 q_1 + \dots + \alpha_m q_m - p| > \frac{\gamma}{q^m}$$

for every integer point (q_1, \dots, q_m, p) with $q = \max(|q_1|, \dots, |q_m|) > 0$.

The other extreme is when $m = 1$ and we have linear forms $L_1(x) = \alpha_1 x, \dots, L_n(x) = \alpha_n x$, and $\max(|\alpha_1 q - p_1|, \dots, |\alpha_n q - p_n|) > \gamma_2 / q^{1/n}$ for integers $q > 0, p_1, \dots, p_n$, or

$$\max(|\alpha_1 - \frac{p_1}{q}|, \dots, |\alpha_n - \frac{p_n}{q}|) > \frac{\gamma_2}{q^{1+(1/n)}}.$$

In this case we simply call $(\alpha_1, \dots, \alpha_n)$ a badly approximable n -tuple.

It will be shown in Theorem 5B of Chapter IV (Khintchine's transference principle) that the n -tuple $(\alpha_1, \dots, \alpha_n)$ is badly approximable precisely if the linear form $\alpha_1 x_1 + \dots + \alpha_n x_n$ is badly approximable. More generally, it will be shown that the system of linear forms

$L_1(\underline{x}), \dots, L_n(\underline{x})$ is badly approximable if and only if the "dual" system $L_1^*(\underline{y}), \dots, L_m^*(\underline{y})$ is badly approximable, where

$$L_j^*(\underline{y}) = \alpha_{1j} y_1 + \dots + \alpha_{nj} y_n \quad (1 \leq j \leq m).$$

THEOREM 4A. Suppose $1, \alpha_1, \dots, \alpha_m$ is a basis of a real algebraic number field of degree $m + 1$. Then the linear form $\alpha_1 x_1 + \dots + \alpha_m x_m$ is badly approximable.

Proof. We may restrict ourselves to integers q_1, \dots, q_m, p with $|\alpha_1 q_1 + \dots + \alpha_m q_m - p| < 1$, and then $\dagger)$ $|p| \ll q = \max(|q_1|, \dots, |q_m|)$. Each conjugate $\alpha_1^{(i)} q_1 + \dots + \alpha_m^{(i)} q_m - p$ has an absolute value $\ll q$, and therefore the norm satisfies $\dagger)$

$$(4.2) \quad \mathfrak{N}(\alpha_1 q_1 + \dots + \alpha_m q_m - p) \ll q^m |\alpha_1 q_1 + \dots + \alpha_m q_m - p|.$$

On the other hand, there is a rational integer $h > 0$ such that $h\alpha_1, \dots, h\alpha_n$ are algebraic integers. Then $|\mathfrak{N}(h\alpha_1 q_1 + \dots + h\alpha_m q_m - hp)| \geq 1$ and

$$(4.3) \quad |\mathfrak{N}(\alpha_1 q_1 + \dots + \alpha_m q_m - p)| \geq h^{-m-1} \gg 1.$$

Combining (4.2) and (4.3) we obtain the theorem.

In general we have

THEOREM 4B. (Perron (1921b)) For every pair of integers $m \geq 1$, $n \geq 1$, there exist algebraic numbers α_{ij} ($1 \leq i \leq n$, $1 \leq j \leq m$) such that the system of linear forms L_1, \dots, L_n given above is badly approximable.

Proof. Put $\ell = m + n$, and let $\varphi_1, \dots, \varphi_\ell$ be ℓ real conjugate algebraic integers. (The existence of such integers will be discussed below.) Writing $\underline{x} = (x_1, \dots, x_m)$ and $\underline{y} = (y_1, \dots, y_n)$, we define

$$M_k(\underline{x}, \underline{y}) = \sum_{i=1}^n \varphi_k^{i-1} y_i + \sum_{j=1}^m \varphi_k^{n+j-1} x_j \quad (1 \leq k \leq \ell).$$

If $(\underline{x}, \underline{y}) \neq (\underline{0}, \underline{0})$ is an integer point, then $M_k(\underline{x}, \underline{y}) \neq 0$ for each k , since $1, \varphi_k, \varphi_k^2, \dots, \varphi_k^{\ell-1}$ are linearly independent over the rationals. Furthermore, $M_1(\underline{x}, \underline{y}), M_2(\underline{x}, \underline{y}), \dots, M_\ell(\underline{x}, \underline{y})$ are conjugate algebraic integers, whence

$\dagger)$ The "Vinogradov symbol" \ll is explained in the notation section. In the present context the implied constant depends on n, m

$$(4.4) \quad \left| \prod_{k=1}^{\ell} M_k(\underline{x}, \underline{y}) \right| \geq 1.$$

There are linear forms

$$L_i(\underline{x}) = \alpha_{i1}x_1 + \dots + \alpha_{im}x_m \quad (1 \leq i \leq n)$$

such that

$$(4.5) \quad M_k(\underline{x}, \underline{y}) = \sum_{i=1}^n \varphi_k^{i-1}(y_i - L_i(\underline{x})) \quad (1 \leq k \leq n).$$

Namely, we have to choose these forms such that

$$(4.6) \quad \sum_{i=1}^n \varphi_k^{i-1} L_i(\underline{x}) = - \sum_{j=1}^m \varphi_k^{n+j-1} x_j \quad (1 \leq k \leq n).$$

This can be done, because the determinant on the left,

$$\det(\varphi_k^{i-1}), \quad (1 \leq i, k \leq n)$$

is a Van der Monde determinant, and is nonzero, since $\varphi_1, \dots, \varphi_n$ are distinct. It is clear from (4.6) that each α_{ij} is a rational function in $\varphi_1, \dots, \varphi_n$, and hence is an algebraic number. Also, we have

$$(4.7) \quad M_k(\underline{x}, \underline{y}) = \sum_{i=1}^n \varphi_k^{i-1}(y_i - L_i(\underline{x})) + \sum_{j=1}^m \omega_{kj} x_j \quad (n < k \leq \ell)$$

for some coefficients ω_{kj} .

Suppose now that $(\underline{x}, \underline{y})$ is an integer point with $\underline{x} \neq \underline{0}$. Write $M = \overline{|\underline{x}|}$, $L = \overline{|\mathcal{L}(\underline{x}) - \underline{y}|}$, and suppose, for the moment, that $L \leq 1$. Then $L \leq 1 \leq M$. From (4.5) and (4.7) we have, respectively,

$$|M_k(\underline{x}, \underline{y})| \ll L \quad (1 \leq k \leq n)$$

and

$$|M_k(\underline{x}, \underline{y})| \ll L + M \ll M \quad (n < k \leq \ell).$$

It follows that

$$\prod_{k=1}^{\ell} |M_k(\underline{x}, \underline{y})| \ll L^{n_m m},$$

and therefore

$$(4.8) \quad L^{n_m m} \gg 1,$$

from (4.4). Clearly, (4.8) also holds for $L > 1$. This completes the proof of the theorem.

It remains to be shown that there are ℓ conjugate real algebraic integers. The simplest argument is as follows. Consider the polynomial

$$P(X) = (X - q)(X - 2q) \dots (X - \ell q) - 1,$$

where q is a large integer. Then

$$P((\ell + \frac{1}{2})q) > 0, -P((\ell - \frac{1}{2})q) > 0, \dots, (-1)^{\ell-1} P(\frac{3}{2}q) > 0, (-1)^{\ell} P(\frac{1}{2}q) > 0.$$

Thus P has real roots $\alpha_1, \dots, \alpha_{\ell}$ with $-\frac{1}{2} + iq < \alpha_i < \frac{1}{2} + iq$ ($i = 1, \dots, \ell$). These roots are algebraic integers. For $i \neq j$ we have $|\alpha_i - jg| \geq \frac{1}{2}q$, and therefore

$$(4.9) \quad 0 < |\alpha_i - iq| \leq (\frac{1}{2}q)^{1-\ell} \quad (i = 1, \dots, \ell).$$

Suppose a proper subset of $\alpha_1, \dots, \alpha_{\ell}$, say $\alpha_1, \dots, \alpha_m$, is a complete set of conjugates of α_1 . Then $a = (\alpha_1 - q)(\alpha_2 - q) \dots (\alpha_m - q)$ is a non-zero rational integer, hence has $|a| \geq 1$. But from (4.9) we get $|a| \ll q^{1-\ell} q^{m-1} = q^{m-\ell}$, which gives a contradiction if $m < \ell$ and q is large.

Slightly more sophisticated is

LEMMA 4C. For every ℓ there exists a real number field of degree ℓ which is normal (over the rationals).

Proof. Let φ be the Euler φ -function and pick an integer n for which $\varphi(n)$ is a multiple of 2ℓ . Let ζ be a primitive n -th root of unity. The field $\mathbb{Q}(\zeta)$ (obtained by adjoining ζ to the field \mathbb{Q} of rationals) is of degree $\varphi(n)$ and normal, and its Galois group is abelian (Van der Waerden (1955) §55). The subfield $\mathbb{Q}(\zeta + (1/\zeta))$ is real, of degree $\frac{1}{2}\varphi(n)$ and normal, and its Galois group G is abelian. The group G contains a subgroup H of order $(1/2\ell)\varphi(n)$. The fixed field of H is real, normal and of degree ℓ .

COROLLARY 4D. The linear forms of Theorem 4B can be constructed to have their coefficients in a field of degree $m + n$.

Proof. They can be constructed to have their coefficients in any real field F which is normal and of degree $\ell = m + n$. We may write F as $F = \mathbb{Q}(\varphi)$ where φ is an algebraic integer, and in the proof of Theorem 4B let $\varphi_1 = \varphi, \varphi_2, \dots, \varphi_\ell$ be the conjugates of φ .

Furtwängler (1927-28) gave a quantitative version of Theorem 4B[†]).

We close this chapter with a famous conjecture of Littlewood. We have

$$|x|^{\max(\|\alpha_1 x\|, \dots, \|\alpha_n x\|)^n} \geq |x| \|\alpha_1 x\| \dots \|\alpha_n x\|,$$

with equality when $n = 1$. A badly approximable n -tuple has

$$|x|^{\max(\|\alpha_1 x\|, \dots, \|\alpha_n x\|)^n} > \gamma > 0.$$

On the other hand, Littlewood conjectured that for $n \geq 2$, arbitrary

[†]) See also forthcoming work of T.W. Cusick.

$\alpha_1, \dots, \alpha_n$ and arbitrary $\varepsilon > 0$, there is always an $x > 0$ with

$$x \|\alpha_1 x\| \dots \|\alpha_n x\| < \varepsilon.$$

Put differently,

$$\liminf_{x \rightarrow \infty} x \|\alpha_1 x\| \dots \|\alpha_n x\| = 0.$$

If this conjecture is true for some n , then it is true a fortiori for $n' > n$.

III. Games and Measures

First Part: Games

References: Cassels (1956), Schmidt (1966).

§1. The (α, β) -Game.

Suppose that $0 < \alpha < 1$, $0 < \beta < 1$. Consider the following game by players Black and White. First, Black picks a compact real interval B_1 of length $\ell(B_1)$. Next, White picks a compact interval $W_1 \subset B_1$ of length $\ell(W_1) = \alpha \ell(B_1)$. Then Black picks a compact interval $B_2 \subset W_1$ of length $\ell(B_2) = \beta \ell(W_1)$, etc. In this way, a nested sequence of compact intervals

$$B_1 \supset W_1 \supset B_2 \supset W_2 \supset \dots$$

is generated, with lengths

$$\ell(B_k) = (\alpha\beta)^{k-1} \ell(B_1) \quad \text{and} \quad \ell(W_k) = (\alpha\beta)^{k-1} \alpha \ell(B_1) \quad (k = 1, 2, 3, \dots).$$

It is clear that $\bigcap_{k=1}^{\infty} B_k = \bigcap_{k=1}^{\infty} W_k$ consists of a single point.

In E^n , the same game is played with intervals replaced by closed balls and length replaced by radius. Let S be a subset of E^n . We say that White wins the game if $\bigcap_{k=1}^{\infty} B_k \in S$. Furthermore, S is an (α, β) -winning set if White is able to win the game no matter how Black plays.

The game just described is a metrical variation of the Banach-Mazur game as discussed in, for example, J.C. Oxtoby (1957).

THEOREM 1A. Suppose that $2\beta < 1 + \alpha\beta$. Then every (α, β) -winning set has the power of the continuum.

We need

LEMMA 1B. Suppose that $0 < \alpha < 1$, $0 < \beta < 1$ and $\gamma = 1 + \alpha\beta - 2\beta > 0$.
Let t be an integer with $(\alpha\beta)^t < \frac{1}{2}\gamma$, and \underline{u} a vector of length 1 .
Suppose a ball W_k occurs in the (α, β) -game. Then Black can play so
that (no matter how White plays) every point \underline{x} of W_{k+t} satisfies

$$(\underline{x} - \underline{w}_k) \underline{u} > \frac{1}{2} \gamma \rho_k^{\dagger}$$

where \underline{w}_k and ρ_k denote respectively the center and the radius of W_k .

Proof. Black can choose $B_{k+1} \subset W_k$ with center $\underline{b}_{k+1} = \underline{w}_k + (1-\beta)\rho_k \underline{u}$.
 Then

$$(1.1) \quad (\underline{b}_{k+1} - \underline{w}_k) \underline{u} = (1-\beta)\rho_k.$$

If now White picks any ball $W_{k+1} \subset B_{k+1}$ with center \underline{w}_{k+1} , then we must have

$$|\underline{w}_{k+1} - \underline{b}_{k+1}| \leq (\text{radius of } B_{k+1}) - (\text{radius of } W_{k+1}) = \beta\rho_k - \alpha\beta\rho_k.$$

By the Cauchy-Schwarz Inequality, it follows that

$$(\underline{w}_{k+1} - \underline{b}_{k+1}) \underline{u} \geq -|\underline{w}_{k+1} - \underline{b}_{k+1}| \cdot |\underline{u}| \geq (\alpha\beta - \beta)\rho_k,$$

and in conjunction with (1.1) we get

$$(\underline{w}_{k+1} - \underline{w}_k) \underline{u} \geq (1 + \alpha\beta - 2\beta)\rho_k = \gamma\rho_k.$$

Repetition of this argument gives

$$(\underline{w}_{k+j} - \underline{w}_{k+j-1}) \underline{u} \geq \gamma \rho_{k+j-1} \quad (j = 1, \dots, t),$$

where \underline{w}_{k+j} and ρ_{k+j} is the center and the radius of W_{k+j} ($j = 0, \dots, t$).

By adding these t inequalities we obtain

\dagger $\underline{v} \underline{u}$ denotes the inner product of vectors \underline{v} , \underline{u} .

$$(1.2) \quad (\underline{w}_{k+t} - \underline{w}_k)_{\underline{u}} \geq \gamma \rho_k + \dots + \gamma \rho_{k+t-1} \geq \gamma \rho_k .$$

The ball W_{k+t} has radius $(\alpha\beta)^t \rho_k < \frac{1}{2} \gamma \rho_k$, and hence every $\underline{x} \in W_{k+t}$ has $|\underline{x} - \underline{w}_{k+t}| < \frac{1}{2} \gamma \rho_k$, which together with (1.2) yields

$$(\underline{x} - \underline{w}_k)_{\underline{u}} > \frac{1}{2} \gamma \rho_k .$$

Proof of Theorem 1A. Put $\underline{u} = (0, \dots, 0, 1)$, and given a ball W with center \underline{w} let the upper half W^+ of W consist of $\underline{x} \in W$ having $(\underline{x} - \underline{w})_{\underline{u}} > 0$, and the lower half W^- of $\underline{x} \in W$ with $(\underline{x} - \underline{w})_{\underline{u}} < 0$.

Suppose White plays a winning strategy, and suppose B_1, W_1 are already given. Choose t with $(\alpha\beta)^t < \frac{1}{2} \gamma$. By Lemma 1B, Black can play such that W_{1+t} lies in the upper half W_1^+ of W_1 . Or, if he chooses otherwise, he can play such that $W_{1+t} \subset W_1^-$. Let W_{1+t}^+ be the ball W_{1+t} if Black uses his first strategy, and W_{1+t}^- the ball W_{1+t} if Black uses his second strategy. In the next t moves Black can play such that W_{1+2t} lies in the upper half of W_{1+t} , or in the lower half of W_{1+t} . In this way there are four possibilities for W_{1+2t} , which we denote by $W_{1+2t}^{++}, W_{1+2t}^{+-}, W_{1+2t}^{-+}, W_{1+2t}^{--}$. And so on. If i_1, i_2, \dots is a sequence of $+$ and $-$ signs, we get a sequence of balls $W_{1+t}^{i_1}, W_{1+2t}^{i_1 i_2}, W_{1+3t}^{i_1 i_2 i_3}, \dots$ whose intersection must lie in S . Since the number of sequences i_1, i_2, \dots has the power of the continuum, and since for distinct sequences we obtain distinct intersections, it follows that S has the power of the continuum.

LEMMA 1C. Suppose S is (α, β) -winning. Then it is $(\alpha(\beta\alpha)^k, \beta)$ -winning for every positive integer k .

Proof. Suppose in the (α, β) -game, White not only chooses the balls W_n , but also the balls B_n except those where $n \equiv 1 \pmod{(k+1)}$.

Thus Black can pick only the balls $B_1, B_{1+(k+1)}, B_{1+2(k+1)}, \dots$. The balls

$$B_1, W_{k+1}, B_{1+(k+1)}, W_{2(k+1)}, B_{1+2(k+1)}, \dots$$

are balls of an $(\alpha(\beta\alpha)^k, \beta)$ -play. If White can win the (α, β) -game, he certainly can with the $(\alpha(\beta\alpha)^k, \beta)$ -game.

Exercise. Suppose $0 < \alpha' \leq \alpha < 1$, $0 < \beta \leq \beta' < 1$ and $\alpha\beta = \alpha'\beta'$. Prove that every (α, β) -winning set is (α', β') -winning.

The following is an open

Question. Suppose that $0 < \alpha' \leq \alpha < 1$ and $0 < \beta \leq \beta' < 1$. Is it true that every (α, β) -winning set is (α', β') -winning? The author's guess is that in general it is not true.

Definition. Given α in $0 < \alpha < 1$, a set S is called α -winning if it is (α, β) -winning for every β in $0 < \beta < 1$.

It follows from Theorem 1A that an α -winning set has the power of the continuum. In fact it was shown by Schmidt (1966) that such a set has Hausdorff dimension n .

LEMMA 1D. The intersection of countably many α -winning sets is α -winning.

Proof. Suppose that α -winning sets S_1, S_2, S_3 are given. Also suppose that β is given, $0 < \beta < 1$. White plays as follows:

For his 1st, 3rd, 5th, ... move, he uses an $(\alpha, \beta\alpha\beta)$ -winning strategy for S_1 . Hence

$$\bigcap_{k=1}^{\infty} B_k \in S_1.$$

For his 2nd, 6th, 10th, ... move, White uses an $(\alpha, \beta(\alpha\beta)^3)$ -winning strategy for S_2 . Hence

$$\bigcap_{k=1}^{\infty} B_k \in S_2.$$

Similarly, for his 4th, 12th, 20th, ... move, he uses an $(\alpha, \beta(\alpha\beta)^7)$ -winning strategy for S_3 . Hence

$$\bigcap_{k=1}^{\infty} B_k \in S_3.$$

Continuing in this manner, White can play so that

$$\bigcap_{k=1}^{\infty} B_k \in \bigcap_{j=1}^{\infty} S_j,$$

and the lemma is proved.

§2. Badly Approximable n-tuples and (α, β) -Games.

We first give some literature.

(i) Cassels (1956) proved:

Suppose that $\lambda_1, \dots, \lambda_k$ are k real numbers. Then there exist continuum many α such that each of

$$\alpha + \lambda_1, \alpha + \lambda_2, \dots, \alpha + \lambda_k$$

is badly approximable.

(ii) Davenport (1964) extended Cassel's method to show:

Suppose that I is a nondegenerate real interval and that $f_1(x), \dots, f_k(x)$ are k real-valued functions, each of which has a nonvanishing derivative on I . Then there exist continuum many α such that each of

$$f_1(\alpha), f_2(\alpha), \dots, f_k(\alpha)$$

is badly approximable.

(iii) Schmidt (1965a) proved:

There exist continuum many α such that every irrational which is algebraically dependent on α is badly approximable.

(iv) Davenport (1954).

In this paper, Davenport showed the existence of continuum many badly approximable pairs (α_1, α_2) .

(v) Cassels (1955b).

Here Cassels proved the existence of continuum many badly approximable n -tuples.

(vi) Davenport (1962).

Davenport gave a simple construction of continuum many badly approximable n -tuples by extending Cassels' method (i).

(vii) Schmidt (1966).

(α, β) -games are introduced and applied to badly approximable n -tuples.

(viii) Schmidt (1969).

In this paper, Schmidt established the existence of continuum many badly approximable systems of linear forms for any positive integers m, n .

Here we prove

THEOREM 2A. Suppose that $0 < \alpha < 1$, $0 < \beta < 1$, $2\alpha < 1 + \alpha\beta$. Then the set S of badly approximable n -tuples is (α, β) -winning. In fact the image of S under a non-singular linear transformation is (α, β) -winning.

In view of Theorem 1A and Lemma 1D, we get

COROLLARY 2B. Let T_1, T_2, \dots be non-singular linear transformations,
and let a_1, a_2, \dots be n-tuples. The set of n-tuples \underline{x} such that

$$T_i(\underline{x}) + a_i$$

is badly approximable for $i = 1, 2, \dots$ is α -winning for $0 < \alpha < \frac{1}{2}$, and
hence has the power of the continuum.

In particular this includes the results of (i), (iv), (v), (vi).

The general result (viii) on badly approximable systems of linear forms is more difficult.

Proof of the Theorem. For notational convenience, we take $n = 2$.

We are to show that, for any nonsingular linear transformation T , the set of \underline{x} such that $T\underline{x}$ is badly approximable is (α, β) -winning.

We put

$$\|T\| = \sup_{\underline{x} \neq 0} \frac{|T\underline{x}|}{|\underline{x}|}.$$

Let $\gamma = 1 + \alpha\beta - 2\alpha > 0$, and let t be an integer, necessarily positive, with $(\alpha\beta)^t < \frac{\gamma}{2}$. Define R by

$$R = (\alpha\beta)^{-\frac{2}{3}t} > 1.$$

We assume, without loss of generality, that Black starts the game with a ball B_1 of radius ρ_1 so small that

$$6\rho_1^2 R^3 (1 + \|T\|)^2 < 1.$$

Finally, we put

$$\delta = \min \left(\rho_1, \frac{\rho_1 \gamma}{4 \|T^{-1}\|} \right).$$

To prove the theorem, it suffices to show that White can play such that, for each nonnegative integer n ,

$$(2.1) \quad \max \left(\left| y_1 - \frac{p_1}{q} \right|, \left| y_2 - \frac{p_2}{q} \right| \right) > \frac{\delta}{q^{3/2}}$$

for every $(y_1, y_2) = T(x_1, x_2)$ with $(x_1, x_2) \in B_{nt+1}$ and with every ordered pair of rationals $\left(\frac{p_1}{q}, \frac{p_2}{q} \right)$ having $1 \leq q < R^n$. For then if $(x_1, x_2) \in \bigcap_{k=1}^{\infty} B_k$ and if $(y_1, y_2) = T(x_1, x_2)$, it follows that (2.1) holds for every positive integer q , whence (y_1, y_2) is badly approximable.

We proceed now by induction on n . If $n = 0$, the inequality $1 \leq q < R^n$ has no solutions q , so that the statement above is true. Suppose, then, that $B_1, B_{1+t}, \dots, B_{1+(k-1)t}$ already have the desired properties. We shall show that, in his next t moves, White can play so that B_{1+kt} also will have the desired properties.

Inequality (2.1) is certainly satisfied for all $\underline{x} \in B_{1+(k-1)t}$ and all rational pairs $\left(\frac{p_1}{q}, \frac{p_2}{q} \right)$ with $1 \leq q < R^{k-1}$. Hence, White has to "worry" only about pairs with $R^{k-1} \leq q < R^k$. Let us call a rational pair $\left(\frac{p_1}{q}, \frac{p_2}{q} \right)$ with $R^{k-1} \leq q < R^k$ dangerous if (2.1) is violated for some $\underline{y} = T\underline{x}$ with $\underline{x} \in B_{1+(k-1)t}$. If there are no dangerous points, the induction is complete and we are done. If dangerous points exist, we claim that they lie on a line.

To prove this claim, suppose that $\left(\frac{p_1}{q}, \frac{p_2}{q} \right), \left(\frac{p_1'}{q'}, \frac{p_2'}{q'} \right), \left(\frac{p_1''}{q''}, \frac{p_2''}{q''} \right)$ are dangerous points. Since $\left(\frac{p_1}{q}, \frac{p_2}{q} \right)$ is dangerous,

$$(2.2) \quad \max\left(\left|y_1 - \frac{p_1}{q}\right|, \left|y_2 - \frac{p_2}{q}\right|\right) \leq \frac{\delta}{q^{3/2}}$$

for some $\underline{y} = T\underline{x}$ with $\underline{x} \in B_{1+(k-1)t}$. Let $\underline{c} = (c_1, c_2)$ denote the center of $B_{1+(k-1)t}$, and put $\underline{d} = (d_1, d_2) = T\underline{c}$.

$B_{1+(k-1)t}$ has radius $\rho_1(\alpha\beta)^{t(k-1)} = \rho_1 R^{-\frac{3}{2}(k-1)}$, so

$$|\underline{c} - \underline{x}| \leq \rho_1 R^{-\frac{3}{2}(k-1)}, \text{ and therefore}$$

$$(2.3) \quad \begin{aligned} |\underline{d} - \underline{y}| &= |T(\underline{c} - \underline{x})| < \|T\| \cdot |\underline{c} - \underline{x}| \\ &\leq \|T\| \rho_1 R^{-\frac{3}{2}(k-1)}. \end{aligned}$$

It follows that

$$\begin{aligned} |qd_1 - p_1| &\leq |qy_1 - p_1| + q|d_1 - y_1| \\ &\leq \frac{\delta}{q^{1/2}} + q\|T\|\rho_1 R^{-\frac{3}{2}(k-1)} \quad (\text{by (2.2), (2.3)}) \\ &< \frac{\delta R^{1/2}}{R^{k/2}} + \frac{\|T\|\rho_1 R^{3/2}}{R^{k/2}} \quad (\text{using } R^{k-1} \leq q < R^k) \\ &< \frac{\rho_1 R^{3/2}(1 + \|T\|)}{R^{k/2}}; \end{aligned}$$

in the same way, we obtain

$$|qd_2 - p_2| < \frac{\rho_1 R^{3/2}(1 + \|T\|)}{R^{k/2}}.$$

Similar estimates hold for q', p'_1, p'_2 and for q'', p''_1, p''_2 .

Consider the matrix

$$M = \begin{pmatrix} 1 & \frac{p_1}{q} & \frac{p_2}{q} \\ 1 & \frac{p'_1}{q'} & \frac{p'_2}{q'} \\ 1 & \frac{p''_1}{q''} & \frac{p''_2}{q''} \end{pmatrix}.$$

To prove the claim, it is enough to prove that the determinant of M is 0. This will happen if the matrix

$$M_1 = \begin{pmatrix} q & p_1 & p_2 \\ q' & p'_1 & p'_2 \\ q'' & p''_1 & p''_2 \end{pmatrix}$$

has determinant 0. Now $\det M_1 = \det M_2$, where

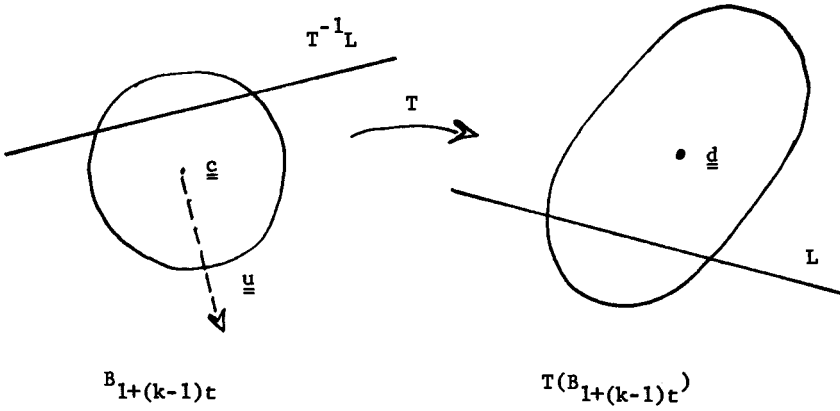
$$M_2 = \begin{pmatrix} q & qd_1 - p_1 & qd_2 - p_2 \\ q' & q'd_1 - p'_1 & q'd_2 - p'_2 \\ q'' & q''d_1 - p''_1 & q''d_2 - p''_2 \end{pmatrix}.$$

Using our estimates, we see that

$$|\det M_2| < 6R^k \left(\frac{\rho_1 R^{3/2} (1 + \|T\|)}{R^{k/2}} \right)^2 = 6\rho_1^2 R^3 (1 + \|T\|)^2 < 1,$$

by our assumption that ρ_1 be small. Since each entry in M_1 is an integer and since $|\det M_1| < 1$, $\det M_1 = 0$. Hence $\det M = 0$, and the claim is established.

We have shown that all dangerous points lie on a line L . Choose a vector \underline{u} of length 1, orthogonal to the line $T^{-1}L$, which when applied to the center of $B_{1+(k-1)t}$, points away from $T^{-1}L$:



Applying Lemma 1B, but with the roles of White and Black interchanged, we see that White can play so that every point $\underline{x} \in B_{1+kt}$ has a distance from $T^{-1}L$ which is (with ρ_L denoting the radius of B_L)

$$\geq \frac{\gamma}{2} \rho_{1+(k-1)t} = \frac{\gamma}{2} (\alpha\beta)^{(k-1)t} \rho_1 = \frac{\gamma}{2} \rho_1^R - \frac{3}{2}(k-1).$$

Suppose now that $\underline{p} = \left(\frac{p_1}{q}, \frac{p_2}{q} \right)$ is dangerous and that $\underline{y} = T\underline{x}$, where $\underline{x} \in B_{1+kt}$. Then $T^{-1}\underline{p} \in T^{-1}L$, so

$$|\underline{x} - T^{-1}\underline{p}| \geq \frac{\gamma}{2} \rho_1^R - \frac{3}{2}(k-1).$$

It follows that

$$\begin{aligned} \max \left(\left| y_1 - \frac{p_1}{q} \right|, \left| y_2 - \frac{p_2}{q} \right| \right) &\geq \frac{1}{\sqrt{2}} |\underline{y} - \underline{p}| = \frac{1}{\sqrt{2}} |T(\underline{x} - T^{-1}\underline{p})| \geq \frac{1}{\sqrt{2}} \cdot \frac{1}{\|T^{-1}\|} |\underline{x} - T^{-1}\underline{p}| \\ &> \frac{1}{2\|T^{-1}\|} \cdot \frac{\gamma}{2} \rho_1^R - \frac{3}{2}(k-1) \geq \frac{\gamma \rho_1}{4\|T^{-1}\|} \cdot \frac{1}{q^{3/2}} \\ &\geq \frac{\delta}{q^{3/2}}. \end{aligned}$$

Thus (2.1) is satisfied for dangerous pairs $\left(\frac{p_1}{q}, \frac{p_2}{q}\right)$ as well, and B_{1+kt} has the desired properties. This completes the induction, and the theorem is proved.

Second Part: Measures

References: Khintchine (1924, 1926a)
 Cassels (1950a, 1950b)
 Schmidt (1960)
 Sprindžuk (1977).

§3. Statement of Results

THEOREM 3A. (Khintchine (1926a)). Let $\psi_1(q), \dots, \psi_n(q)$ be
functions of the positive integer q with $0 < \psi_i(q) \leq 1$ ($i = 1, \dots, n$),
and suppose that

$$\psi(q) = \prod_{i=1}^n \psi_i(q)$$

is non-increasing. If

$$(3.1) \quad \sum_{q=1}^{\infty} \psi(q)$$

is convergent, then for almost all^{†)} n-tuples $(\alpha_1, \dots, \alpha_n)$, there are
only finitely many q with

$$(3.2) \quad \{q\alpha_i\} < \psi_i(q) \quad (i = 1, \dots, n).$$

But if the sum (3.1) is divergent, then for almost all $(\alpha_1, \dots, \alpha_n)$
there are infinitely many q with (3.2).

The sum (3.1) is divergent for $\psi_i(q) = \varepsilon q^{-1/n}$ ($i = 1, \dots, n$).

So for any $\varepsilon > 0$,

$$(3.3) \quad \{q\alpha_i\} < \varepsilon q^{-1/n} \quad (i = 1, \dots, n)$$

^{†)} in the sense of Lebesgue measure.

has infinitely many solutions for almost all $(\alpha_1, \dots, \alpha_n)$, or put differently, almost no n-tuple is badly approximable. In fact even

$$(3.4) \quad \{q\alpha_i\} < q^{-1/n}(\log q)^{-1/n} \quad (i = 1, \dots, n)$$

has infinitely many solutions almost always. On the other hand,

$$\{q\alpha_i\} < q^{-1/n - \varepsilon} \quad (i = 1, \dots, n)$$

has almost always^{†)} only finitely many solutions, and since $\|x\| = \min(\{x\}, \{-x\})$, it follows on using all the n-tuples $(\pm\alpha_1, \dots, \pm\alpha_n)$ that

$$(3.5) \quad \|q\alpha_i\| < q^{-1/n - \varepsilon} \quad (i = 1, \dots, n)$$

almost always has only finitely many solutions. If $\alpha_1, \dots, \alpha_n$ is called very well approximable if for some $\varepsilon > 0$ the system (3.5) has infinitely many solutions, then it follows that almost no n-tuple is very well approximable.

The second (and more difficult) assertion of Theorem 3A follows from

THEOREM 3B. (Schmidt 1960) Make the same assumptions as in Theorem 3A, and suppose that (3.1) is divergent. Write $N(h, \alpha_1, \dots, \alpha_n)$ for the number of solutions of (3.2) with $1 \leq q \leq h$, and put

$$\begin{aligned} \Psi(h) &= \sum_{q=1}^h \psi(q), \\ \Omega(h) &= \sum_{q=1}^h \psi(q)q^{-1}. \end{aligned}$$

Then for $\varepsilon > 0$,

^{†)} For more subtle results involving the Hausdorff dimension see Jarník (1929) and Besicovitch (1934).

$$(3.6) \quad N(h; \alpha_1, \dots, \alpha_n) = \Psi(h) + O(\Psi^{1/2}(h) \Omega^{1/2}(h) (\log \Psi(h))^{2+\varepsilon})$$

for almost^{†)} all $(\alpha_1, \dots, \alpha_n)$.

By Cauchy's inequality we have

$$\Omega(h)^2 = \left(\sum_{q=1}^h \psi(q)^{1/2} \psi(q)^{1/2} q^{-1} \right)^2 \leq \left(\sum_{q=1}^h \psi(q) \right) \left(\sum_{q=1}^h \psi(q) q^{-2} \right) \ll \Psi(h),$$

so that $\Omega(h) \ll \Psi(h)^{1/2}$ and the error term in (3.6) is of smaller order of magnitude than the main term. It follows that almost always the number of solutions $q \leq h$ of (3.3) is $\sim \varepsilon^n \log h$, and the number of solutions $q \leq h$ of (3.4) is $\sim \log \log h$. Again, the number of solutions of

$$\{q\alpha_i\} \leq c_i \quad (i = 1, \dots, n)$$

with $0 < c_i \leq 1$, as q ranges in $1 \leq q \leq h$, is

$$c_1 c_2 \dots c_n h + O(h^{1/2} (\log h)^{(5/2)+\varepsilon})$$

almost always. (In the language of uniform distribution (See Kuipers and Niederreiter (1974)) this means that the sequence of points $(q\alpha_1, \dots, q\alpha_n)$ with $q = 1, 2, \dots$ is uniformly distributed modulo 1 with discrepancy $O(h^{1/2} (\log h)^{(5/2)+\varepsilon})$ for almost every $(\alpha_1, \dots, \alpha_n)$).

Theorem 3B is generalized by Schmidt (1964), and a rather difficult metrical problem concerning n -tuples $(\alpha, \alpha^2, \dots, \alpha^n)$ is settled by Sprindzuk (1965). For a metrical conjecture of Duffin and Schaeffer

^{†)} the constant in 0 may depend on $(\alpha_1, \dots, \alpha_n)$.

concerning approximations $\|q\alpha\| = |q\alpha - p| < \psi(q)$ with $\text{g.c.d.}(p, q) = 1$, see Erdos (1970) and forthcoming work of P. Vaaler. See also Gallagher (1962, 1965).

§4. The convergence part of Theorem 3A. We may suppose that

$(\alpha_1, \dots, \alpha_n)$ lies in the unit cube U^n with $0 \leq \alpha_i < 1$ ($i = 1, \dots, n$).

Let E_q be the set of $\underline{\alpha} = (\alpha_1, \dots, \alpha_n) \in U^n$ with (3.2). For fixed i ,

(3.2) means that $p \leq q\alpha_i < p + \psi_i(q)$ for an integer p with

$0 \leq p \leq q-1$, and this gives a set for α_i of measure $\sum_{p=0}^{q-1} \psi_i(q)q^{-1}$

$= \psi_i(q)$. It follows that E_q has measure

$$(4.1) \quad \mu(E_q) = \psi(q).$$

The convergence part of Theorem 3A is now a consequence of

LEMMA 4A. Let E_q ($q = 1, 2, \dots$) be subsets of U^n with convergent

$$(4.2) \quad \sum_{q=1}^{\infty} \mu(E_q).$$

Then almost no $(\alpha_1, \dots, \alpha_n) \in U^n$ lies in infinitely many of the sets E_q .

Proof. $(\alpha_1, \dots, \alpha_n)$ lies in infinitely many of the sets E_q precisely if it lies in

$$(4.3) \quad \bigcap_{h=1}^{\infty} F_h$$

where $F_h = \bigcup_{q=h}^{\infty} E_q$. But $\mu(F_h)$ tends to zero by the convergence of (4.2), and so (4.3) is of measure zero.

§5. The idea of the proof of Theorem 3B. In probability theory one shows ("law of large numbers") that in an infinite sequence of

throws of a die, the number of "two"s after h throws will be asymptotically equal to $\frac{1}{6}h$ with probability 1. Essential here is that the events "two is thrown at the q -th move", and "two is thrown at the r -th move" where $r \neq q$, are independent. More generally one shows that if E_1, E_2, \dots is a sequence of events, where E_q has probability $\mu(E_q)$ and where (4.2) is divergent, and where E_q, E_r with $q \neq r$ are "independent" in the sense that $\mu(E_q \cap E_r) = \mu(E_q)\mu(E_r)$, then it is almost certain that of the first h events

$$\sim \mu(E_1) + \dots + \mu(E_h)$$

actually take place.

A difficulty in our present situation is that E_q, E_r with $q \neq r$ are not independent. However, it will be shown that if q and r are not too close in a certain sense, then E_q, E_r are nearly independent.

Another point to be observed in the proof will be the following.

If h is large, it will be relatively easy to show that the exceptional set G_h of $\underline{\alpha}$ for which $N(h, \underline{\alpha}) = N(h, \alpha_1, \dots, \alpha_n)$ is far from $\Psi(h)$ has a small measure $\mu(G_h)$. But this is not enough, we need to show that $\mu(\bigcup_{h \geq h_0} G_h)$ is small. To overcome this difficulty one sets

$$(5.1) \quad N(u, v, \underline{x})$$

for the number of q in $u < q \leq v$ with (3.2), and

$$(5.2) \quad \Psi(u, v) = \sum_{q=u+1}^v \psi(q),$$

and one shows that for large u, v , one has $N(u, v, \underline{x}) \sim \Psi(u, v)$ with great probability. If this is true for $(u, v]^\dagger$ and for $(v, w]$, then this is also true for $(u, w]$. The most economical way to obtain every

\dagger) This denotes the half open interval.

interval $(u, v]$ will be to build them up of intervals $(a2^j, (a+1)2^j]$.

We will begin with the case $n = 1$ of Theorem 3B and will indicate the necessary changes for $n > 1$ in §10. We shall follow Schmidt (1960) rather closely.

§6. On certain intervals. Let $\omega(h)$, $h \geq 1$, be a non-negative integral-valued function which increases to infinity. We write $\omega(0) = 0$ and define S' to be the set consisting of 0 and of all integers $h > 0$ such that $\omega(h-1) < \omega(h)$. We define S'' to be the set of integers $h \geq 0$ having $\omega(h) < \omega(h+1)$. Finally, S is the set of values of $\omega(h)$, $h \geq 0$.

Next, we define for fixed $t > 0$ intervals of order t to be the half-open intervals

$$(u2^t + v_1, (u+1)2^t + v_2],$$

where u, v_1, v_2 are non-negative integers such that $v_1 < 2^t$ and v_1, v_2 are the smallest non-negative integers satisfying $u2^t + v_1 \in S$, $(u+1)2^t + v_2 \in S$. (It is possible, of course, that for given u, t there exists no such v_1 .) The intervals of order t cover the positive axis exactly once.

LEMMA 6A. Every interval $(0, x], x \in S$, can be expressed as disjoint union of intervals $\cup I_i$ of the type described above, where no two of the intervals I_i are of the same order.

Proof. Write x in the binary scale,

$$x = \sum_{t=0}^w b_t 2^t,$$

where b_t equals 0 or 1, but $b_w = 1$. There exists an interval $(0, j_1]$ of order w with $j_1 \leq x$. If $j_1 = x$, then we are through. If not, and if

$$j_1 = \sum_{t=0}^w b_t^{(2)} 2^t,$$

then $b_w^{(2)} = b_w = 1$ and there exists a largest integer w_2 having

$$b_{w_2}^{(2)} < b_{w_2}.$$

Hence there exists an interval $(j_1, j_2]$ of order w_2 , $j_2 \leq x$. If $j_2 = x$, then $(0, x] = (0, j_1] \cup (j_1, j_2]$. Otherwise, if

$$j_2 = \sum_{t=0}^w b_t^{(3)} 2^t, \quad b_w^{(3)} = b_w, \dots, b_{w_2}^{(3)} = b_{w_2},$$

then there exists a largest $w_3 < w_2$, having

$$b_{w_3}^{(3)} < b_{w_3}.$$

We proceed as before. Since $j_1 < j_2 < \dots$, we finally arrive at $j_f = x$ and $(0, x] = (0, j_1] \cup \dots \cup (j_{f-1}, j_f]$. The orders of the intervals are $w > w_2 > \dots > w_f > 0$.

§7. Sums involving a function $\varphi(k, q)$. Let k, q be positive and write $\varphi(k, q)$ for the number of integers x , $0 \leq x < q$, so that $\text{g.c.d.}(x, q) \leq k$.

LEMMA 7A.

$$\sum_{q=1}^v \varphi(k, q) q^{-1} = v + O(vk^{-1} + \log v \log k).$$

In this section, the inequality indicated by the O -symbol holds

for all values of all variables involved. Further $\log 1$ is redefined to be equal to 1.

Proof. Clearly,

$$\varphi(k, q) = \sum_{\substack{w \mid q \\ w \leq k}} \varphi\left(\frac{q}{w}\right),$$

where $\varphi(x)$ is the Euler φ -function. Using the well-known relation

$$\varphi(x) = x \sum_{y \mid x} \mu(y) y^{-1},$$

involving the Moebius function we obtain

$$\begin{aligned} \sum_{q=1}^v \varphi(k, q) q^{-1} &= \sum_{q=1}^v q^{-1} \sum_{\substack{w \mid q \\ w \leq k}} q w^{-1} \sum_{y \mid q w^{-1}} \mu(y) y^{-1} \\ &= \sum_{w=1}^{\min(k, v)} w^{-1} \sum_{y=1}^{[(v/w)]} \mu(y) y^{-1} \sum_{q=1}^{[(v/yw)]} 1, \end{aligned}$$

where $[\alpha]$ is the integer part of α . Thus

$$\begin{aligned} \sum_{q=1}^v \varphi(k, q) q^{-1} &= v \sum_{w=1}^{\min(k, v)} w^{-2} \sum_{y=1}^{[(v/w)]} \mu(y) y^{-2} + O(\log v \log k) \\ &= v \sum_{w=1}^{\min(k, v)} w^{-2} \zeta(2)^{-1} + O\left(\sum_{w=1}^{\min(k, v)} w^{-1}\right) + O(\log v \log k) \\ &= v + O(vk^{-1} + \log v \log k). \end{aligned}$$

LEMMA 7B.

$$\sum_{q=1}^v \psi(q) \varphi(k, q) q^{-1} = \Psi(v) + O(\Psi(v) k^{-1} + \Omega(v) \log k).$$

Proof. Put $\Pi(k, 0) = 0$ and

$$\Pi(k, r) = \sum_{q=1}^r \varphi(k, q) q^{-1}$$

for $r \geq 1$. Lemma 7A yields

$$(7.1) \quad \Pi(k, r) = r + O(rk^{-1} + \log r \log k).$$

Using partial summation we obtain

$$\begin{aligned} & \sum_{q=1}^v \psi(q) \varphi(k, q) q^{-1} \\ &= \sum_{q=1}^v \psi(q) (\Pi(k, q) - \Pi(k, q-1)) \\ &= \sum_{q=1}^{v-1} \Pi(k, q) (\psi(q) - \psi(q+1)) + \Pi(k, v) \psi(v) \\ &= \sum_{q=1}^{v-1} q (\psi(q) - \psi(q+1)) + v \psi(v) + R(k, v) \\ &= \Psi(v) + R(k, v), \end{aligned} \quad (7.2)$$

where, according to (7.1),

$$\begin{aligned} & R(k, v) \\ &= O \left(\sum_{q=1}^{v-1} (qk^{-1} + \log q \log k) (\psi(q) - \psi(q+1)) \right) + O(vk^{-1} + \log v \log k) \psi(v) \\ &= O \left(\Psi(v) k^{-1} + \log k \sum_{q=2}^v \psi(q) (\log q - \log(q-1)) + (\log k) \psi(1) \right). \end{aligned} \quad (7.3)$$

Now

$$\begin{aligned}
 (7.4) \quad & \sum_{q=2}^v \psi(q)(\log q - \log(q-1)) \\
 &= 0 \left(\sum_{q=2}^v \psi(q) \log \left(1 + \frac{1}{q-1} \right) \right) \\
 &= O(\Omega(v)) .
 \end{aligned}$$

Lemma 7B is a consequence of (7.2), (7.3), and (7.4).

§8. Bounds for certain integrals. We introduce the following functions and integrals.

$$\beta(q, \alpha) = \begin{cases} 1, & \text{if } 0 \leq \alpha < \psi(q), \\ 0 & \text{otherwise,} \end{cases}$$

$$\gamma(q, \alpha) = \sum_p \beta(q, q\alpha - p) ,$$

$$\gamma(k, q, \alpha) = \sum_{\substack{p \\ \text{g.c.d.}(p, q) \leq k}} \beta(q, q\alpha - p) ,$$

$$I(q) = \int_0^1 \gamma(q, \alpha) d\alpha ,$$

$$I(k; q) = \int_0^1 \gamma(k, q, \alpha) d\alpha ,$$

$$I(k; q, r) = \int_0^1 \gamma(k, q, \alpha) \gamma(k, r, \alpha) d\alpha .$$

We observe that

$$N(v, \alpha) = \sum_{q=1}^v \gamma(q, \alpha) ,$$

and for $0 \leq u < v$ we put

$$N(k; u, v; \alpha) = \sum_{q=u+1}^v \gamma(k, q, \alpha) .$$

LEMMA 8A.

$$(8.1) \quad I(q) = \psi(q) , \quad I(k; q) = \psi(q) \varphi(k, q) q^{-1} ,$$

$$(8.2) \quad I(k; q, r) \leq \psi(q) \psi(r) + A(k; q, r) \psi(q) q^{-1} ,$$

where $A(k; q, r)$ is the number of solutions p, s of

$$(8.3) \quad qs - rp = 0 \quad \text{with} \quad 0 \leq p < q$$

having

$$\text{g.c.d. } (p, q) \leq k , \quad \text{g.c.d. } (s, r) \leq k .$$

We remark that in the notation of §4 and §5,

$$\gamma(q, \alpha) = \begin{cases} 1 & \text{if } \alpha \in E_q , \\ 0 & \text{otherwise.} \end{cases}$$

Hence $\mu(E_q) = I(q) = \psi(q)$, and the independence of E_q , E_r for $q \neq r$ would mean that

$$(8.4) \quad \int_0^1 \gamma(q, \alpha) \gamma(r, \alpha) d\alpha = \psi(q) \psi(r) .$$

However, this is not true, and as a technical device we have to introduce $I(k; q, r)$ to replace the integral on the left of (8.4), and instead of (8.4) we have (8.2)

Proof. $I(q) = \psi(q)$ is rather trivial, while the second half of (8.1) follows from

$$\begin{aligned}
 I(k, q) &= \sum_{\substack{p \\ \text{g.c.d.}(p, q) \leq k}} \int_0^1 \beta(q, \alpha q - p) d\alpha \\
 &= \varphi(k, q) q^{-1} \int_{-\infty}^{\infty} \beta(q, \alpha) d\alpha .
 \end{aligned}$$

As for $I(k; q, r)$, we have

$$I(k; q, r) = \sum_{\substack{p, \text{g.c.d.}(p, q) \leq k \\ s, \text{g.c.d.}(s, r) \leq k}} \int_0^1 \beta(q, \alpha q - p) \beta(r, \alpha r - s) d\alpha .$$

We split this sum into two parts,

$$I(k; q, r) = I_0(k; q, r) + I_1(k; q, r) ,$$

where I_0 consists of the terms with $qs - rp \neq 0$. Then

$$\begin{aligned}
 I_0(k; q, r) &\leq \sum_{\substack{p, s \\ qs - rp \neq 0}} \int_0^1 \beta(q, \alpha q - p) \beta(r, \alpha r - s) d\alpha \\
 (8.5) \quad &= \sum_{\substack{p, s \\ qs - rp \neq 0}} \int_{-(p/q)}^{1-(p/q)} \beta(q, q\alpha') \beta\left(r, r\alpha' - \frac{qs - rp}{q}\right) d\alpha' .
 \end{aligned}$$

To find an estimate for this sum, write $q = q'd, r = r'd$, where $d = \text{g.c.d.}(q, r)$. We have $qs - rp = dh$ where $h = q's - r'p$. For given h , p is determined modulo q' , and has therefore d possibilities modulo q . Thus for given h , the intervals $[-(p/q), 1 - (p/q))$ will cover the real axis d times. Hence we get

$$I_0(k; q, r)$$

$$\begin{aligned} &\cong d \sum_{h \neq 0} \int_{-\infty}^{\infty} \beta(q, q\alpha') \beta\left(r, r\alpha' - \frac{hd}{q}\right) d\alpha' \\ &\cong d \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \beta(q, q\alpha') \beta(r, r\alpha' - \lambda dq^{-1}) d\alpha' d\lambda \\ &= \psi(q) \psi(r) . \end{aligned}$$

In changing from the summation over h to the continuous parameter λ we used the fact that the function

$$\int_{-\infty}^{\infty} \beta(q, q\alpha') \beta(r, r\alpha' - \lambda dq^{-1}) d\alpha'$$

is monotonically decreasing in λ when $\lambda \geq 0$, and monotonically increasing when $\lambda \leq 0$.

To prove Lemma 8A it remains to give an upper bound for $I_1(k; q, r)$.

In analogy to (8.5) we find that

$$\begin{aligned} I_1(k; q, r) &= \sum_{\substack{p, \text{ g.c.d.}(p, q) \leq k \\ s, \text{ g.c.d.}(s, r) \leq k \\ qs - rp = 0}} \int_{-(p/q)}^{1-(p/q)} \beta(q, q\alpha') \beta(r, r\alpha') d\alpha' \\ &\leq A(k; q, r) \psi(q) q^{-1} . \end{aligned}$$

LEMMA 8B.

$$\begin{aligned} \int_0^1 N(v, \alpha) d\alpha &= \Psi(v) , \\ \int_0^1 N(k; u, v; \alpha) d\alpha &= \sum_{q=u+1}^v \psi(q) \varphi(k, q) q^{-1} , \\ \int_0^1 N^2(k; u, v; \alpha) d\alpha &\leq \Psi^2(u, v) + 2 \sum_{q=u+1}^v \psi(q) d_k(q) , \end{aligned}$$

where $\Psi(u,v)$ is defined in (5.2) and where $d_k(q)$ is the number of divisors of q not exceeding k .

Proof. The first two assertions follow from (8.1). As an immediate consequence of (8.2) we have

$$\int_0^1 N^2(k; u, v; \alpha) d\alpha \leq \Psi^2(u, v) + 2 \sum_{u < r \leq q \leq r} A(k; q, r) \psi(q) q^{-1} .$$

Now

$$\sum_{r=1}^q A(k; q, r)$$

is equal to the number of solutions r, p, s of

$$\begin{aligned} qs - rp &= 0, \quad 0 \leq p < q, \quad 1 \leq r \leq q, \\ \text{g.c.d.}(p, q) &\leq k, \quad \text{g.c.d.}(s, r) \leq k. \end{aligned}$$

Define a, b by

$$\frac{a}{b} = \frac{p}{q} = \frac{s}{r}, \quad \text{g.c.d.}(a, b) = 1 .$$

Then b divides q and qb^{-1} divides both p and r , so that $\text{g.c.d.}(p, q) \leq k$ implies $qb^{-1} \leq k$. Thus the number of possible choices for b is $d_k(q)$. Furthermore, there are $\varphi(b) \leq b$ possibilities for a and qb^{-1} possibilities for r , once b is given. Hence

$$\sum_{r=1}^q A(k; q, r) \leq q d_k(q)$$

and

$$\sum_{u < r \leq q \leq v} A(k; q, r) \psi(q) q^{-1} \leq \sum_{q=u+1}^v \psi(q) d_k(q) .$$

§9. Proof of Theorem B. ($n = 1$). Write $\omega(h) = [\Psi(h)\Omega(h)]$ and define S, S', S'' as in §6. For $s \geq 1$ let L_s be the set of pairs $(u, v), u \in S', v \in S''$, so that $(\omega(u), \omega(v))$ is an interval of any order t with respect to ω (see §6), and $\omega(v) \leq 2^s$. From now on, the numbers k, s are always connected by the relation

$$(9.1) \quad k = 2^s.$$

From here on, we make heavy use of the methods developed in Cassels (1950b). Write $h^* = h^*(s)$ for the largest integer h^* having $\omega(h^*) \leq 2^s$.

LEMMA 9A.

$$(9.2) \quad 0 \leq \int_0^1 (N(h^*, \alpha) - N(k; 0, h^*, \alpha)) d\alpha = O(s 2^{s/2})$$

$$(9.3) \quad \sum_{(u,v) \in L_s} \int_0^1 (N(k; u, v; \alpha) - \Psi(u, v))^2 d\alpha = O(s^2 2^s).$$

Proof. The first two equations of Lemma 8B give

$$\begin{aligned} & \int_0^1 (N(h^*, \alpha) - N(k; 0, h^*, \alpha)) d\alpha \\ &= \Psi(h^*) - \sum_{q=1}^{h^*} \psi(q) \varphi(k, q) q^{-1} \\ &= O(\Psi(h^*) k^{-1} + \Omega(h^*) \log k), \end{aligned}$$

according to Lemma 7B. Since

$$\Omega(h^*) = O(2^{s/2}),$$

(9.2) follows.

Using Lemma 8B again we see that a single integral in (9.3) does not exceed

$$2 \sum_{q=u+1}^v \psi(q) d_k(q) + 2\Psi(u,v) \left(\Psi(u,v) - \sum_{q=u+1}^v \psi(q) \varphi(k,q) q^{-1} \right).$$

We first take the sum over those pairs $(u,v) \in L_s$ where $(\omega(u), \omega(v)]$ is an interval of fixed order t . Since intervals of order t cover the positive axis exactly once, we obtain the upper bound

$$2 \sum_{q=1}^{h^*} \psi(q) d_k(q) + 2\Psi(h^*) \left(\Psi(h^*) - \sum_{q=1}^{h^*} \psi(q) \varphi(k,q) q^{-1} \right).$$

We observe that since $\psi(q)$ is non-increasing

$$\sum_{q=1}^{h^*} \psi(q) d_k(q) = \sum_{t=1}^k \sum_{q_1=1}^{[h^*/t]} \psi(tq_1) \leq \sum_{t=1}^k t^{-1} \sum_{q=1}^{h^*} \psi(q) = O(2^s \log k),$$

and using Lemma 7B we find the upper bound

$$O(2^s \log k) + O(\Psi^2(h^*) k^{-1} + \Psi(h^*) \Omega(h^*) \log k) = O(s 2^s).$$

Summing over t and observing $t \leq s$ we obtain (9.3).

LEMMA 9B. There is a sequence of subsets $\sigma_1, \sigma_2, \dots$ of the unit-interval with measures

$$\mu_s = \int_{\sigma_s} d\alpha = O(s^{-1-\epsilon})$$

such that

$$N(h, \alpha) = \Psi(h) + O(2^{s/2} s^{2+\epsilon})$$

for any h with $\omega(h) \leq 2^s$, $h \in S'$, and any α in $0 \leq \alpha < 1$ which is not in σ_s .

Proof. We define σ_s to be the set of all α in $0 \leq \alpha < 1$, for which not both of the following two inequalities hold:

$$(9.4) \quad 0 \leq N(h^*, \alpha) - N(k; 0, h^*, \alpha) \leq s^{2+\epsilon} 2^{s/2}$$

$$(9.5) \quad \sum_{(u,v) \in L_s} (N(k; u, v; \alpha) - \Psi(u, v))^2 \leq s^{3+\epsilon} 2^s.$$

As a consequence of Lemma 9A,

$$\mu_s = O(s^{-1-\epsilon}).$$

If $h \leq h^*$, $h \in S'$, then the interval $(0, \omega(h)]$ is the union of at most s intervals $(\omega(u), \omega(v)]$, where $(u, v) \in L_s$. So

$$N(k; 0, h; \alpha) - \Psi(h) = \sum (N(k; u, v; \alpha) - \Psi(u, v)),$$

where the sum is over at most s pairs $(u, v) \in L_s$. This fact, together with (9.5) and Cauchy's inequality yields for $0 \leq \alpha < 1$, $\alpha \notin \sigma_s$,

$$(N(k; 0, h; \alpha) - \Psi(h))^2 \leq s^{4+\epsilon} 2^s.$$

This last equation together with (9.4) (which by monotonicity must be true with h^* replaced by $h \leq h^*$), gives Lemma 9B.

Proof of Theorem 3B. ($n = 1$). Since $\sum s^{-1-\epsilon}$ is convergent, we may appeal to Lemma 4A to see that almost every α lies outside σ_s for $s \geq s_0 = s_0(\alpha)$. Let α have this property, and assume h to be so large that $\omega(h) \geq 2^{s_0}$. Choose s so that $2^{s-1} \leq \omega(h) < 2^s$.

Suppose $h \in S'$. Then we have with Lemma 9B,

$$\begin{aligned}
 N(h, \alpha) &= \Psi(h) + O\left(2^{\frac{1}{2}s} s^{2+\epsilon}\right) \\
 &= \Psi(h) + O\left(\Psi^{\frac{1}{2}}(h) \Omega^{\frac{1}{2}}(h) \log^{2+\epsilon} \Psi(h)\right).
 \end{aligned}$$

Hence Theorem 3B holds for $h \in S'$. By the same argument we can prove the Theorem for $h \in S''$.

To any h there exist h' , h'' with $h' \in S'$, $h'' \in S''$ having $h' \leq h \leq h''$ and

$$\omega(h') = \omega(h) = \omega(h'').$$

Then

$$|\Psi(h)\Omega(h) - \Psi(h')\Omega(h')| \leq 1,$$

and

$$|\Psi(h) - \Psi(h')| \leq \Omega(h)^{-1} \leq \Omega(1)^{-1} = \psi(1)^{-1},$$

and similarly for $\Psi(h'')$. Since

$$N(h', \alpha) \leq N(h, \alpha) \leq N(h'', \alpha),$$

the case $n = 1$ of Theorem 3B follows.

§10. The case $n \geq 2$. Using

$$\begin{aligned}
 v - \sum_{q=1}^v \varphi^n(k, q) q^{-n} \\
 &= \sum_{q=1}^v (q^n - \varphi^n(k, q)) q^{-n} \\
 &\leq n \sum_{q=1}^v (q - \varphi(k, q)) q^{n-1} q^{-n}
 \end{aligned}$$

$$= n(v - \sum_{q=1}^v \varphi(k, q)q^{-1}),$$

we easily generalize Lemmas 7A, 7B to

$$\sum_{q=1}^v \varphi^n(k, q)q^{-n} = v + O(vk^{-1} + \log k \log v),$$

$$\sum_{q=1}^v \psi(q) \varphi^n(k, q)q^{-n} = \psi(v) + O(\psi(v)k^{-1} + \Omega(v) \log k).$$

In analogy to $\beta(q, \alpha)$ of §8 we define $\beta(q, \alpha_1, \dots, \alpha_n)$ to be the characteristic function of the box

$$0 \leq \alpha_i < \psi_i(q) \quad (i = 1, \dots, n),$$

and put

$$\gamma(q, \alpha_1, \dots, \alpha_n) = \sum_{p_1, \dots, p_n} \beta(q, q\alpha_1 - p_1, \dots, q\alpha_n - p_n),$$

$$\gamma(k; q, \alpha_1, \dots, \alpha_n) = \sum_{\substack{p_i, \text{g.c.d.}(p_i, q) \leq k \\ (i = 1, \dots, n)}} \beta(q, q\alpha_1 - p_1, \dots, q\alpha_n - p_n).$$

$I(q)$, $I(k, q)$, $I(k; q, r)$ are now n -dimensional integrals. To find an upper bound for

$$I(k; q, r) = \sum_{\substack{p_i, \text{g.c.d.}(p_i, q) \leq k \\ s_i, \text{g.c.d.}(s_i, r) \leq k \\ (i = 1, \dots, n)}} \int_0^1 \dots \int_0^1 \beta(q, q\alpha_1 - p_1, \dots, \beta(r, r\alpha_1 - s_1, \dots) d\alpha_1 \dots d\alpha_n,$$

we split this sum into $n + 1$ parts,

$$I(k; q, r) = I_0 + \dots + I_n,$$

where I_j consists of the terms with exactly j indices i_1, \dots, i_j having $qs_{i_1} - rp_{i_1} = 0$. We find that

$$I_0(k; q, r) \cong \psi(q)\psi(r)$$

and

$$\begin{aligned} I_j(k; q, r) &\cong c^{(j)}_A{}^j(k; q, r)\psi(q)q^{-j} \\ &\cong c^{(j)}_A(k; q, r)\psi(q)q^{-1}. \end{aligned}$$

There are no other modifications of any depth.

IV. Integer Points in Parallelepipeds.

(A chapter in the Geometry of Numbers).

References: Minkowski (1896 & 1910, 1907)

Cassels (1959),

Lekkerkerker (1969),

Mahler (1939, 1955).

§1. Minkowski's Theorem on Successive Minima.

Let K be a compact convex set in E^n , symmetric about $\underline{0}$, and suppose that $\underline{0}$ lies in the interior of K . We call such a set a convex body. Let $V(K)$ denote the volume of K . (By the volume of K we mean the Riemann integral of the characteristic function of K . It can be proved that every convex body has a volume in this sense. Alternatively, the existence of the volume of K may be added as a hypothesis.)

Let $\lambda_1 = \lambda_1(K)$ be the infimum of those numbers $\lambda \geq 0$ such that λK contains an integer point different from $\underline{0}$. By our rather narrow definition of convex bodies, this infimum is in fact a minimum. It is easily seen that $0 < \lambda_1 < \infty$. Put $\tilde{\lambda} = 2V(K)^{-1/n}$. Then $\tilde{\lambda}K$ is a convex body with volume 2^n . By Minkowski's Convex Body Theorem (Theorem 2B of Chapter II), $\tilde{\lambda}K$ contains an integer point $\neq \underline{0}$. Therefore, $\lambda_1 \leq 2V(K)^{-1/n}$, or

$$(1.1) \quad \lambda_1^n V(K) \leq 2^n.$$

For each integer j with $1 \leq j \leq n$, let $\lambda_j = \lambda_j(K)$ be the infimum of all $\lambda \geq 0$ such that λK contains j linearly independent integer points. Clearly each λ_j is actually a minimum, and

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

We call $\lambda_1, \lambda_2, \dots, \lambda_n$ the successive minima of K . λ_1 is the first minimum of K , λ_2 is the second minimum of K , etc.

Examples. Let K denote the closed rectangle with center $(0,0)$, length 4 and width 1, having sides parallel to the coordinate axes.

Then $\lambda_1(K) = \frac{1}{2}$ and $\lambda_2(K) = 2$.

The closed disc L with radius $\frac{1}{2}$ and center $(0,0)$ satisfies $\lambda_1(L) = \lambda_2(L) = 2$.

THEOREM 1A*. (Minkowski's Second Convex Body Theorem. Minkowski (1907)).

Suppose that K is a convex body in E^n . Then

$$(1.2) \quad \frac{2^n}{n!} \leq \lambda_1 \lambda_2 \cdots \lambda_n V(K) \leq 2^n.$$

Note that the right hand inequality here sharpens (1.1). We shall not prove this theorem here but shall prove relatively weak special cases in Lemmas 1C, 1D, 1E and in Corollary 2B.

The cube $|x_i| \leq 1$ ($i = 1, \dots, n$) has $\lambda_1 = \dots = \lambda_n = 1$ and $V(K) = 2^n$, so that the right inequality in (1.2) is best possible. The "octahedron" O consisting of points with $|x_1| + \dots + |x_n| \leq 1$ has $\lambda_1 = \dots = \lambda_n = 1$ and $V(O) = 2^n/n!$, so that the left inequality in (1.2) is best possible.

Remark. By definition, and since K is compact, $\lambda_1 K$ contains an integer point $\underline{g}_1 \neq \underline{0}$. Also $\lambda_2 K$ contains an integer point \underline{g}_2 such that $\underline{g}_1, \underline{g}_2$ are linearly independent, $\lambda_3 K$ contains an integer point \underline{g}_3 such that $\underline{g}_1, \underline{g}_2, \underline{g}_3$ are linearly independent, etc. In this way, we obtain linearly independent integer points $\underline{g}_1, \underline{g}_2, \dots, \underline{g}_n$ such that

$$\underline{g}_i \in \lambda_i K \quad (1 \leq i \leq n).$$

Now \underline{g}_1 may be replaced by $-\underline{g}_1$, hence it is not unique. Occasionally, more precisely when $\lambda_1 = \lambda_2$, there are more than these two possibilities for \underline{g}_1 . Generally, there is only a finite number of choices for $\underline{g}_1, \dots, \underline{g}_n$.

LEMMA 1B. Let K be a convex body in E^n , and let $\underline{g}_1, \dots, \underline{g}_n$ be linearly independent integer points with $\underline{g}_i \in \lambda_i K$ ($1 \leq i \leq n$). Suppose that J is an integer with $2 \leq J \leq n$ and that \underline{g} is an integer point in λK , where $\lambda < \lambda_J$. Then

$$(1.3) \quad \underline{g} = u_1 \underline{g}_1 + \dots + u_{J-1} \underline{g}_{J-1}$$

with real coefficients u_1, \dots, u_{J-1} .

Proof. If $\lambda < \lambda_1$, then $\underline{g} = \underline{0}$ and (1.3) is satisfied with $u_1 = \dots = u_{J-1} = 0$. Otherwise, there is an integer j with $2 \leq j \leq J$ such that $\lambda_{j-1} \leq \lambda < \lambda_j$. Now the points $\underline{g}_1, \dots, \underline{g}_{j-1}, \underline{g}$ all lie in λK . Since $\lambda < \lambda_j$, these points must be linearly dependent. But $\underline{g}_1, \dots, \underline{g}_{j-1}$ are linearly independent, so \underline{g} is necessarily a linear combination of $\underline{g}_1, \dots, \underline{g}_{j-1}$, and therefore a linear combination of $\underline{g}_1, \dots, \underline{g}_{j-1}, \dots, \underline{g}_{J-1}$.

LEMMA 1C. Let K be a convex body in E^n . Then the left inequality (1.2) holds, or

$$\frac{2^n}{n!} \leq \lambda_1 \lambda_2 \dots \lambda_n V(K).$$

Proof. Let $\underline{g}_1, \dots, \underline{g}_n$ be linearly independent integer points with $\underline{g}_i \in \lambda_i K$ ($1 \leq i \leq n$). Then

$$\underline{h}_i = \lambda_i^{-1} \underline{g}_i \in K \quad (1 \leq i \leq n).$$

The points $\pm \underline{h}_1, \dots, \pm \underline{h}_n$ lie in K , hence their convex hull H lies in

K . It is easily seen that H consists of all points $t_1 \underline{h}_1 + \dots + t_n \underline{h}_n$ with $|t_1| + \dots + |t_n| \leq 1$.

If $\underline{h}_1, \dots, \underline{h}_n$ were the basis vectors $\underline{e}_1 = (1, 0, \dots, 0), \dots, \underline{e}_n = (0, \dots, 0, 1)$, then H would be the octahedron O , consisting of the points $(t_1, \dots, t_n) \in E^n$ with $|t_1| + \dots + |t_n| \leq 1$, and having volume

$$V(O) = \frac{2^n}{n!}.$$

In general, H is obtained from the octahedron O by a linear transformation of determinant $\det(\underline{h}_1, \dots, \underline{h}_n)$. Therefore

$$V(H) = V(O) |\det(\underline{h}_1, \dots, \underline{h}_n)| = \frac{2^n}{n!} \frac{|\det(\underline{g}_1, \dots, \underline{g}_n)|}{\lambda_1 \lambda_2 \dots \lambda_n} \geq \frac{2^n}{n! \lambda_1 \lambda_2 \dots \lambda_n};$$

here $|\det(\underline{g}_1, \dots, \underline{g}_n)| \geq 1$ since $\underline{g}_1, \dots, \underline{g}_n$ are linearly independent integer points. Thus

$$V(K) \geq V(H) \geq \frac{2^n}{n! \lambda_1 \lambda_2 \dots \lambda_n},$$

and the lemma follows.

Definitions. The unit ball (in E^n) is the set of all points $\underline{x} = (x_1, x_2, \dots, x_n)$ with $x_1^2 + \dots + x_n^2 \leq 1$. An ellipsoid is the image of the unit ball under a nonsingular linear transformation.

LEMMA 1D. (Minkowski (1896)). If \mathbb{E} is an ellipsoid, then

$$\lambda_1 \dots \lambda_n V(\mathbb{E}) \leq 2^n.$$

Proof. Let $\underline{e}_1 = (1, 0, \dots, 0), \dots, \underline{e}_n = (0, \dots, 0, 1)$ be the standard basis vectors in E^n . The unit ball B consists of all points $t_1 \underline{e}_1 + \dots + t_n \underline{e}_n$ with $t_1^2 + \dots + t_n^2 \leq 1$. Further $\mathbb{E} = T(B)$, where T is some nonsingular linear transformation. Put $\underline{f}_i = T \underline{e}_i$ ($1 \leq i \leq n$); then

\mathbb{E} consists of all points $t_{1=1}f_1 + \dots + t_{n=n}f_n$ with $t_1^2 + \dots + t_n^2 \leq 1$.

We define a new inner product on E^n as follows: given arbitrary points $\underline{x} = x_{1=1}f_1 + \dots + x_{n=n}f_n$ and $\underline{x}' = x'_{1=1}f_1 + \dots + x'_{n=n}f_n$ in E^n (the f_i -s form a basis since T is nonsingular), we put

$$\underline{x} \otimes \underline{x}' = x_1 x'_1 + \dots + x_n x'_n.$$

With this inner product, we obtain a Euclidean space. For notational convenience, we write

$$|\underline{x}|^{\otimes} = \sqrt{\underline{x} \otimes \underline{x}}.$$

Then \mathbb{E} consists of all points $\underline{x} \in E^n$ with $|\underline{x}|^{\otimes} \leq 1$.

Let $\underline{g}_1, \dots, \underline{g}_n$ be linearly independent integer points with $\underline{g}_i \in \lambda_i \mathbb{E}$ ($1 \leq i \leq n$). Applying the Gram-Schmidt orthonormalization process to $\underline{g}_1, \dots, \underline{g}_n$, we obtain vectors $\underline{c}_1, \dots, \underline{c}_n$ in E^n and real numbers t_{ij} ($1 \leq j \leq i \leq n$) with $\underline{c}_i \otimes \underline{c}_j = \delta_{ij}$ ($1 \leq i, j \leq n$) and

$$\underline{g}_i = t_{i1}\underline{c}_1 + \dots + t_{ii}\underline{c}_i \quad (1 \leq i \leq n).$$

Let T^* be the linear transformation with $T^*\underline{c}_i = \lambda_i \underline{c}_i$ ($1 \leq i \leq n$), and put $\mathbb{E}^* = T^*(\mathbb{E})$. To complete the proof of the lemma, it suffices to show that

$$(1.4) \quad \lambda_1(\mathbb{E}^*) \geq 1;$$

for then

$$\begin{aligned} \lambda_1 \lambda_2 \dots \lambda_n V(\mathbb{E}) &= |\det T^*| V(\mathbb{E}) = V(\mathbb{E}^*) \\ &\leq (\lambda_1(\mathbb{E}^*))^n V(\mathbb{E}^*) \\ &\leq 2^n \end{aligned} \quad (\text{by (1.1)}).$$

In order to establish (1.4) , suppose that

$$\underline{x} = u_1 \underline{g}_1 + \dots + u_n \underline{g}_n$$

is an arbitrary point in \mathbb{R}^* . Then

$$(1.5) \quad \begin{aligned} \underline{x} &= \sum_{i=1}^n \left(\sum_{j=1}^n t_{ji} u_j \right) \underline{c}_i = \sum_{i=1}^n \left(\lambda_i^{-1} \sum_{j=1}^n t_{ji} u_j \right) T^* \underline{c}_i \\ &= T^* \left(\sum_{i=1}^n \lambda_i^{-1} \left(\sum_{j=1}^n t_{ji} u_j \right) \underline{c}_i \right) , \end{aligned}$$

and $\underline{y} = \sum_{i=1}^n \lambda_i^{-1} \left(\sum_{j=1}^n t_{ji} u_j \right) \underline{c}_i$ lies in \mathbb{R} . Necessarily, $|\underline{y}|^{\otimes} \leq 1$,

so that

$$(1.6) \quad \sum_{i=1}^n \lambda_i^{-2} \left(\sum_{j=1}^n t_{ji} u_j \right)^2 \leq 1 .$$

We now further suppose that $\underline{x} \neq \underline{0}$ is an integer point in \mathbb{R}^* .

There is an integer J , $1 \leq J \leq N$, such that

$$\underline{x} = u_1 \underline{g}_1 + \dots + u_J \underline{g}_J$$

with $u_J \neq 0$. By Lemma 1B , $\underline{x} \notin \lambda \mathbb{R}$ whenever $\lambda < \lambda_J$. Therefore for such an \underline{x} ,

$$(1.7) \quad |\underline{x}|^{\otimes} \geq \lambda_J .$$

We have

$$\begin{aligned} \sum_{i=1}^n \lambda_i^{-2} \left(\sum_{j=1}^n t_{ji} u_j \right)^2 &= \sum_{i=1}^J \lambda_i^{-2} \left(\sum_{j=1}^n t_{ji} u_j \right)^2 \geq \lambda_J^{-2} \sum_{i=1}^J \left(\sum_{j=1}^n t_{ji} u_j \right)^2 \\ &= \lambda_J^{-2} \sum_{i=1}^n \left(\sum_{j=1}^n t_{ji} u_j \right)^2 = \lambda_J^{-2} (|\underline{x}|^{\otimes})^2 \quad (\text{by (1.5)}) \\ &\geq \lambda_J^{-2} \cdot \lambda_J^2 = 1 \quad (\text{by (1.7)}) . \end{aligned}$$

This last result, together with (1.6), shows that no integer point $\underline{x} \neq \underline{0}$ belongs to the interior of \mathcal{E}^* . Therefore $\lambda_1(\mathcal{E}^*) \geq 1$, and the lemma follows.

LEMMA 1E. Let Π be a closed parallelepiped in E^n , symmetric about $\underline{0}$. Then

$$\frac{2^n}{n!} \leq \lambda_1 \lambda_2 \cdots \lambda_n V(\Pi) \leq 2^n n!$$

Proof. In view of Lemma 1C, only the right-hand inequality requires proof. If C denotes the cube

$$|x_i| \leq 1, \quad (1 \leq i \leq n)$$

then $\Pi = T(C)$ for some nonsingular linear transformation T . Let $\mathcal{E} = T(B)$, where B is the unit ball in E^n . Now $B \subseteq C$, so $\mathcal{E} \subseteq \Pi$, whence

$$\lambda_i(\Pi) \leq \lambda_i(\mathcal{E}) \quad (1 \leq i \leq n).$$

Then

$$\begin{aligned} \lambda_1(\Pi) \cdots \lambda_n(\Pi) V(\Pi) &\leq \lambda_1(\mathcal{E}) \cdots \lambda_n(\mathcal{E}) V(\mathcal{E}) \frac{V(\Pi)}{V(\mathcal{E})} \\ &\leq 2^n \frac{V(\Pi)}{V(\mathcal{E})} \quad (\text{By Lemma 1D}) \\ &= 2^n \frac{V(C)}{V(B)} = \frac{4^n}{V(B)}. \end{aligned}$$

B contains the octahedron $|x_1| + \cdots + |x_n| \leq 1$ of volume $\frac{2^n}{n!}$, so $V(B) \geq 2^n/n!$ and

$$\lambda_1(\Pi) \cdots \lambda_n(\Pi) V(\Pi) \leq 2^n n!$$

Remarks. In our applications, Π will often be defined by inequalities

$$|L_i(\underline{x})| \leq 1, \quad (1 \leq i \leq n)$$

where $L_i(\underline{x}) = \alpha_{i1}x_1 + \dots + \alpha_{in}x_n$ ($1 \leq i \leq n$) are n linear forms of determinant ± 1 . In this case $V(\Pi) = 2^n$, so Lemma 1E becomes

$$\frac{1}{n!} \leq \lambda_1 \lambda_2 \dots \lambda_n \leq n!.$$

Even more briefly, we have

$$(1.8) \quad 1 \ll \lambda_1 \lambda_2 \dots \lambda_n \ll 1.$$

Here and throughout this chapter, the constants in \ll depend only on the dimension, which will usually be n , but will be l in §5.

The following remark will be useful. Any n points $\underline{x}_1, \dots, \underline{x}_n$ in a parallelepiped Π span the octahedron of points $t_1 \underline{x}_1 + \dots + t_n \underline{x}_n$ with $|t_1| + \dots + |t_n| \leq 1$ of volume $(2^n/n!) |\det(\underline{x}_1, \dots, \underline{x}_n)|$. Since this volume cannot exceed $V(\Pi)$, we get $|\det(\underline{x}_1, \dots, \underline{x}_n)| \leq n! 2^{-n} V(\Pi)$. The n integer points $\underline{g}_1, \dots, \underline{g}_n$ of Lemma 1B are of the type $\underline{g}_i = \lambda_{i1} \underline{x}_1$ with \underline{x}_1 in Π ($1 \leq i \leq n$). Hence

$$(1.9) \quad |\det(\underline{g}_1, \dots, \underline{g}_n)| \leq \lambda_1 \dots \lambda_n n! 2^{-n} V(\Pi) \leq (n!)^2 \ll 1.$$

In most of our applications we will need only parallelepipeds. In the next section, which may at first be omitted, we will show that an argument as was given in Lemma 1E for parallelepipeds may in fact be given for any convex body K .

§2. Jordan's Theorem.

THEOREM 2A. Let K be a symmetric convex body. Then there is an ellipsoid \mathcal{E} with

$$(2.1) \quad \mathcal{E} \subset K \subset \sqrt{n} \mathcal{E}.$$

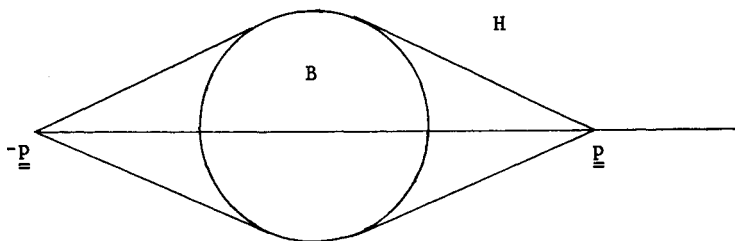
COROLLARY 2B. The successive minima of K satisfy

$$\lambda_1 \dots \lambda_n V(K) \leq 2^n n^{n/2} .$$

Proof of Corollary. $\lambda_1 = \lambda_1(K) \leq \lambda_1(E)$. By Lemma 1D ,

$$\begin{aligned} \lambda_1(K) \dots \lambda_n(K) V(K) &\leq \lambda_1(E) \dots \lambda_n(E) V(E) (V(K)/V(E)) \\ &\leq 2^n V(K)/V(E) \\ &\leq 2^n n^{n/2} . \end{aligned}$$

Proof of Theorem 2A. Compactness arguments show that there is an ellipsoid $E_0 \subseteq K$ of maximum volume. This ellipsoid will be shown to satisfy (2.1) . Without loss of generality we may suppose that $E_0 = B$, the unit ball. It will suffice to show that a point \underline{x} with $|\underline{x}| > \sqrt{n}$ cannot lie in K , so without loss of generality it will suffice to show that the point $\underline{p} = (p, 0, \dots, 0)$ with $p > \sqrt{n}$ cannot lie in K . If we suppose to the contrary that \underline{p} lies in K , then the convex hull H of $B, \underline{p}, -\underline{p}$ is contained in K .



A short computation shows that the ellipsoid

$$E_1 : \left(\frac{x_1}{a}\right)^2 + \left(\frac{x_2}{b}\right)^2 + \dots + \left(\frac{x_n}{b}\right)^2 \leq 1$$

will be contained in H and therefore in K if

$$(2.2) \quad \frac{a^2}{p^2} + b^2 \left(1 - \frac{1}{p^2}\right) \leq 1 \quad \text{and if } a \geq b .$$

If we set

$$a = p/\sqrt{n} \quad \text{and} \quad b = \sqrt{\frac{p^2 - a^2}{p^2 - 1}} ,$$

the first condition (2.2) holds with equality, and the second is true since $p > \sqrt{n}$.

The function $f(x) = x^n(x-1)^{1-n}$ is increasing for $x \geq n$, and therefore $f(p^2) > f(n)$. Now

$$\begin{aligned} (V(E_1)/V(B))^2 &= a^2 b^{2(n-1)} = \frac{p^2}{n} \left(\frac{p^2(1 - (1/n))}{p^2 - 1} \right)^{n-1} \\ &= \frac{p^{2n}}{(p^2 - 1)^{n-1}} \bigg/ \frac{n^n}{(n-1)^{n-1}} = f(p^2)/f(n) > 1 , \end{aligned}$$

contradicting the maximality of B .

For the subject of this section see also John (1948).

§3. Davenport's Lemma.

THEOREM 3A. (Essentially Davenport (1937)). Let $L_1(\underline{x}), \dots, L_n(\underline{x})$ be n linear forms of determinant 1, and let $\lambda_1, \dots, \lambda_n$ be the successive minima of the parallelepiped Π defined by

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

Suppose ρ_1, \dots, ρ_n are real numbers with

$$(3.1) \quad \rho_1 \geq \rho_2 \geq \dots \geq \rho_n > 0 ,$$

$$(3.2) \quad \rho_1 \lambda_1 \leq \rho_2 \lambda_2 \leq \dots \leq \rho_n \lambda_n ,$$

$$(3.3) \quad \rho_1 \rho_2 \dots \rho_n = 1 .$$

Then there is a permutation of the linear forms L_1, \dots, L_n such that

after this permutation the new parallelepiped Π' defined by

$$|\rho_i L_i(\underline{x})| \leq 1 \quad (i = 1, \dots, n)$$

has successive minima $\lambda'_1, \dots, \lambda'_n$ with

$$(3.4) \quad 2^{-n} \lambda_i \rho_i \leq \lambda'_i \leq 2^{n^2} (n!)^2 \lambda_i \rho_i \quad (i = 1, \dots, n) .$$

Moreover, let $\underline{g}_1, \dots, \underline{g}_n$ be linearly independent integer points with $\underline{g}_i \in \lambda_i \Pi$, i.e. with $|L_j(\underline{g}_i)| \leq \lambda_i$ ($1 \leq i, j \leq n$). Then every integer point \underline{x} which is not in the subspace S_{i-1} spanned by $\underline{g}_1, \dots, \underline{g}_{i-1}$ has

$$(3.5) \quad \max(|\rho_1 L_1(\underline{x})|, \dots, |\rho_n L_n(\underline{x})|) \geq 2^{-n} \rho_i \lambda_i .$$

Proof. Put $N(\underline{x}) = \max(|L_1(\underline{x})|, \dots, |L_n(\underline{x})|)$. The n linearly independent integer points $\underline{g}_1, \dots, \underline{g}_n$ with $\underline{g}_j \in \lambda_j \Pi$ satisfy

$$N(\underline{g}_j) = \lambda_j \quad (1 \leq j \leq n) .$$

Put $S_0 = \{0\}$, and let S_i be the subspace spanned by $\underline{g}_1, \dots, \underline{g}_i$ ($1 \leq i \leq n$). In S_i , the linear forms L_1, \dots, L_n satisfy $n-i$ independent linear conditions. In particular, there is a nontrivial relation:

$$\alpha_1 L_1 + \dots + \alpha_n L_n \equiv 0 \text{ in } S_{n-1} .$$

Order the forms in such a way that the coefficient α_n is largest in absolute value. In S_{n-2} , the forms L_1, \dots, L_n satisfy a further nontrivial linear condition, independent of the condition above. In fact, there is such a condition of the type

$$\beta_1 L_1 + \dots + \beta_{n-1} L_{n-1} \equiv 0 \text{ in } S_{n-2} .$$

Reorder the forms L_1, \dots, L_n such that β_{n-1} has the largest absolute

value among $\beta_1, \dots, \beta_{n-1}$. Etc.

If $\underline{x} \in S_{n-1}$, then

$$\alpha_n L_n(\underline{x}) = -\alpha_1 L_1(\underline{x}) - \dots - \alpha_{n-1} L_{n-1}(\underline{x}) ,$$

whence

$$|L_n(\underline{x})| \leq |L_1(\underline{x})| + \dots + |L_{n-1}(\underline{x})| ,$$

since

$$\left| \frac{\alpha_i}{\alpha_n} \right| \leq 1 \quad (1 \leq i \leq n) .$$

Thus

$$|L_1(\underline{x})| + \dots + |L_{n-1}(\underline{x})| \geq \frac{1}{2} (|L_1(\underline{x})| + \dots + |L_n(\underline{x})|)$$

for $\underline{x} \in S_{n-1}$. Similarly,

$$\begin{aligned} |L_1(\underline{x})| + \dots + |L_{n-2}(\underline{x})| &\geq \frac{1}{2} (|L_1(\underline{x})| + \dots + |L_{n-1}(\underline{x})|) \\ &\geq \frac{1}{4} (|L_1(\underline{x})| + \dots + |L_n(\underline{x})|) \end{aligned}$$

for $\underline{x} \in S_{n-2}$. (The last inequality is true since $S_{n-2} \subseteq S_{n-1}$.) Continuing in this manner, we see that

$$(3.6) \quad |L_1(\underline{x})| + \dots + |L_j(\underline{x})| \geq 2^{j-n} (|L_1(\underline{x})| + \dots + |L_n(\underline{x})|)$$

whenever $\underline{x} \in S_j$ ($1 \leq j \leq n$) .

Now suppose that \underline{x} is an integer point and that $\underline{x} \notin S_{i-1}$. Then there is an integer j with $i \leq j \leq n$ such that $\underline{x} \in S_j$ and $\underline{x} \notin S_{j-1}$. Since $\underline{x} \notin S_{j-1}$, $N(\underline{x}) \geq \lambda_j$ by Lemma 1B . Because $\underline{x} \in S_j$, we have

$$\begin{aligned}
\max(\rho_1 |L_1(\underline{x})|, \dots, \rho_n |L_n(\underline{x})|) &\geq \max(\rho_1 |L_1(\underline{x})|, \dots, \rho_j |L_j(\underline{x})|) \\
&\geq \rho_j \max(|L_1(\underline{x})|, \dots, |L_j(\underline{x})|) \quad (\text{by (3.1)}) \\
&\geq \frac{\rho_j}{j} (|L_1(\underline{x})| + \dots + |L_j(\underline{x})|) \\
&\geq \frac{2^{j-n}}{j} \rho_j (|L_1(\underline{x})| + \dots + |L_n(\underline{x})|) \quad (\text{by (3.6)}) \\
&\geq 2^{-n} \rho_j N(\underline{x}) \geq 2^{-n} \rho_j \lambda_j \\
&\geq 2^{-n} \rho_i \lambda_i \quad (\text{by (3.2)}) .
\end{aligned}$$

Thus (3.5) is established. It follows that

$$(3.7) \quad \lambda'_i \geq 2^{-n} \rho_i \lambda_i \quad (1 \leq i \leq n) ,$$

which is the first half of (3.4) .

In view of (3.3) , $V(\Pi') = V(\Pi) = 2^n$, whence

$$\frac{1}{n!} \leq \lambda_1 \lambda_2 \cdots \lambda_n \leq n!$$

and

$$\frac{1}{n!} \leq \lambda'_1 \lambda'_2 \cdots \lambda'_n \leq n!$$

by Lemma 1E . Therefore

$$\begin{aligned}
\lambda'_i &\leq \frac{n!}{\lambda'_1 \cdots \lambda'_{i-1} \lambda'_{i+1} \cdots \lambda'_n} \leq \frac{2^{n(n-1)} n!}{\rho_1 \lambda_1 \cdots \rho_{i-1} \lambda_{i-1} \rho_{i+1} \lambda_{i+1} \cdots \rho_n \lambda_n} \quad (\text{by (3.7)}) \\
&< 2^{n^2} (n!)^2 \rho_i \lambda_i ,
\end{aligned}$$

and the second inequality (3.4) holds.

§4. Reciprocal Parallelepipeds.

Let $\underline{a}_1, \dots, \underline{a}_n$ be a basis of E^n of determinant D . There are

unique vectors a_1^*, \dots, a_n^* such that the inner product

$$a_{i=j}^* = \delta_{ij} \quad (1 \leq i, j \leq n),$$

the Kronecker symbol. Now a_1^*, \dots, a_n^* is called the reciprocal basis to the given basis; its determinant is D^{-1} .

A parallelepiped Π is determined by inequalities

$$(4.1) \quad |a_{i=1}^* x| \leq 1 \quad (i = 1, \dots, n),$$

and conversely a_1, \dots, a_n , except for their ordering and replacement of a_i by $-a_i$, are determined by Π . (The proof of the second statement may be done as an exercise). The reciprocal[†] parallelepiped Π^* is defined by

$$(4.2) \quad |a_{i=1}^* x| \leq 1 \quad (i = 1, \dots, n).$$

THEOREM 4A. (Mahler 1939). Let Π, Π^* be reciprocal parallelepipeds, with successive minima $\lambda_1, \dots, \lambda_n$ and $\lambda_1^*, \dots, \lambda_n^*$, respectively. Then

$$(4.3) \quad \lambda_i^* \ll \lambda_{n+1-i}^{-1} \ll \lambda_i^*,$$

with the constants in \ll depending only on n . Moreover, if Π, Π^* are respectively defined by (4.1), (4.2), if g_1, \dots, g_n are linearly independent integer points with $g_j \in \lambda_j \Pi$ ($j = 1, \dots, n$) and if g_1^*, \dots, g_n^* are reciprocal to g_1, \dots, g_n , then

$$(4.4) \quad |a_{i=j}^* g_j^*| \ll \lambda_j^{-1} \quad (1 \leq i, j \leq n).$$

Proof. The hypothesis $g_j \in \lambda_j \Pi$ means that

$$(4.5) \quad |a_{i=j} g_j| \leq \lambda_j \quad (1 \leq i, j \leq n).$$

[†]) Also called polar parallelepiped.

Setting $E = |\det(\underline{g}_1, \dots, \underline{g}_n)|$ we have

$$(4.6) \quad 1 \leq E < 1$$

by (1.9) .

The identity

$$\sum_{i=1}^n (a_{i\underline{x}})(a_{i\underline{y}}^*) = \underline{xy}$$

is easily proved by the linearity of both sides in \underline{x} and in \underline{y} and by substituting the special values $\underline{x} = a_1^*, \dots, a_n^*$ and $\underline{y} = a_1, \dots, a_n$.

From the identity we obtain

$$\sum_{i=1}^n (a_{i\underline{g}_k})(a_{i\underline{g}_j}^*) = \underline{g}_k \underline{g}_j^* = \delta_{kj} \quad (1 \leq k, j \leq n) .$$

It follows that

$$a_{i\underline{g}_j}^* = A_{ij} / \det A \quad (1 \leq i, j \leq n) ,$$

where A is the matrix with entries $a_{\ell\underline{g}_m}$ ($1 \leq \ell, m \leq n$) , and where A_{ij} is the cofactor of $a_{i\underline{g}_j}$ in this matrix. It is clear that $|\det A| = DE$, and (4.5) yields $|A_{ij}| < \lambda_1 \dots \lambda_{j-1} \lambda_{j+1} \dots \lambda_n$. Thus

$$(4.7) \quad |a_{i\underline{g}_j}^*| < \lambda_1 \lambda_2 \dots \lambda_n D^{-1} E^{-1} \lambda_j^{-1} < \lambda_j^{-1} ,$$

by (4.6) and since $\lambda_1 \dots \lambda_n D^{-1} = 2^{-n} \lambda_1 \dots \lambda_n V(\Pi) < 1$ by Lemma 1E .

Thus (4.4) is true.

The points $\underline{g}_1^*, \dots, \underline{g}_n^*$ are not necessarily integer points, but it is easily seen that $E\underline{g}_1^*, \dots, E\underline{g}_n^*$ are integer points, and these points have

$$|a_{i\underline{g}_j}^*(E\underline{g}_j^*)| < \lambda_j^{-1} \quad (1 \leq i, j \leq n)$$

by (4.6), (4.7) . Considering the integer point $E_{\underline{n}}^*$ we see that $\lambda_1^* \ll \lambda_n^{-1}$. Considering $E_{\underline{n}}^*$, $E_{\underline{n}-1}^*$, we find that $\lambda_2^* \ll \lambda_{n-1}^{-1}$, etc. So generally $\lambda_i^* \ll \lambda_{n+1-i}^{-1}$ ($i = 1, \dots, n$) , or $\lambda_1^* \lambda_{n+1-i} \ll 1$. Since

$$\lambda_1 \dots \lambda_n \lambda_1^* \dots \lambda_n^* = 4^{-n} \lambda_1 \dots \lambda_n V(\Pi) \lambda_1^* \dots \lambda_n^* V(\Pi^*) \gg 1 ,$$

the assertion (4.3) follows.

COROLLARY 4B. Suppose that $V(\Pi) = 2^n$ (i.e. that the determinant
 $D = 1$). Then

$$\lambda_1^* \ll \lambda_1^{1/(n-1)} .$$

Proof. We have $1 \ll \lambda_1 \dots \lambda_n \leq \lambda_1 \lambda_n^{n-1}$ and $\lambda_1^* \ll \lambda_n^{-1} \ll \lambda_1^{1/(n-1)}$.

§5. Khintchine's Transference Principle.

Let α_{ij} ($1 \leq i \leq n, 1 \leq j \leq m$) be nm real numbers. Throughout this section, we consider linear forms

$$L_i(\underline{x}) = \alpha_{i1}x_1 + \dots + \alpha_{im}x_m , \quad (1 \leq i \leq n)$$

and the dual forms

$$M_j(\underline{u}) = \alpha_{1j}u_1 + \dots + \alpha_{nj}u_n , \quad (1 \leq j \leq m)$$

where $\underline{x} = (x_1, \dots, x_m)$ and $\underline{u} = (u_1, \dots, u_n)$. Put

$$\mathcal{L}(\underline{x}) = (L_1(\underline{x}), \dots, L_n(\underline{x})) ,$$

$$\mathcal{M}(\underline{u}) = (M_1(\underline{u}), \dots, M_m(\underline{u})) .$$

Also, put $\ell = m+n$.

THEOREM 5A. Suppose that there is an integer point $(\underline{x}, \underline{y}) =$
 $(x_1, \dots, x_m, y_1, \dots, y_n) \neq (\underline{0}, \underline{0})$ in E^ℓ with

$$\left| \underline{x} \right| \leq X \quad \text{and} \quad \left| f(\underline{x}) - \underline{y} \right| \leq Y .$$

Then there is an integer point $(\underline{u}, \underline{v}) = (u_1, \dots, u_n, v_1, \dots, v_m) \neq (0, 0)$ in
 E^ℓ such that

$$\left| \underline{u} \right| \ll U \quad \text{and} \quad \left| \mathcal{M}(\underline{u}) - \underline{v} \right| \ll V ,$$

where

$$U = X^{m/(\ell-1)} Y^{(1-m)/(\ell-1)} , \quad V = X^{(1-n)/(\ell-1)} Y^{n/(\ell-1)} .$$

(Here and throughout this section, the constants in \ll depend only on ℓ .)

Proof. Consider the following vectors in E^ℓ :

$$\underline{a}_i = (\delta_{i1}, \dots, \delta_{im}, 0, \dots, 0) \quad (1 \leq i \leq m) ,$$

$$\underline{a}_{m+i} = (\alpha_{i1}, \dots, \alpha_{im}, \delta_{i1}, \dots, \delta_{in}) \quad (1 \leq i \leq n) ,$$

$$\underline{a}_j^* = (\delta_{j1}, \dots, \delta_{jm}, -\alpha_{1j}, \dots, -\alpha_{nj}) \quad (1 \leq j \leq m) ,$$

$$\underline{a}_{m+j}^* = (0, \dots, 0, \delta_{j1}, \dots, \delta_{jn}) \quad (1 \leq j \leq n) ,$$

where δ_{ij} is the Kronecker symbol. It is easily shown that

$$\underline{a}_i \underline{a}_j^* = \delta_{ij} \quad (1 \leq i, j \leq \ell) .$$

Now we write

$$\underline{b}_i = X^{-1} (X^m Y^n)^{1/\ell} \underline{a}_i \quad (1 \leq i \leq m) ,$$

$$\underline{b}_i = Y^{-1} (X^m Y^n)^{1/\ell} \underline{a}_i \quad (m < i \leq \ell) ,$$

$$\underline{b}_j^* = X (X^m Y^n)^{-1/\ell} \underline{a}_j^* \quad (1 \leq j \leq m) ,$$

$$\underline{b}_j^* = Y (X^m Y^n)^{-1/\ell} \underline{a}_j^* \quad (m < j \leq \ell) .$$

Then again

$$b_{i=j}^* b_{i=j} = \delta_{ij} \quad (1 \leq i, j \leq \ell) .$$

Thus the bases b_1, \dots, b_ℓ and b_1^*, \dots, b_ℓ^* are reciprocal. We observe that

$$\det(b_1, \dots, b_\ell) = \det(b_1^*, \dots, b_\ell^*) = 1 .$$

If Π denotes the parallelepiped in E^ℓ defined by

$$|b_i z| \leq 1 \quad (1 \leq i \leq \ell) ,$$

then the reciprocal parallelepiped Π^* is defined by

$$|b_j^* z| \leq 1 \quad (1 \leq j \leq \ell) .$$

Put $z = (\underline{x}, -\underline{y})$, where $(\underline{x}, \underline{y})$ is the integer point of the hypothesis.

It is easily checked that

$$|b_i z| \leq (X^m Y^n)^{1/\ell} \quad (1 \leq i \leq \ell) .$$

Thus if λ_1 and λ_1^* is the first minimum of Π and of Π^* , then

$$\lambda_1 \leq (X^m Y^n)^{1/\ell} ,$$

whence

$$\lambda_1^* \ll (X^m Y^n)^{\frac{1}{\ell(\ell-1)}}$$

by Corollary 4B. It follows that there is an integer point

$\underline{w} = (\underline{v}, \underline{u}) = (v_1, \dots, v_m, u_1, \dots, u_n) \neq (0, 0)$ in E^ℓ with

$$|b_j^* \underline{w}| \ll (X^m Y^n)^{\frac{1}{\ell(\ell-1)}} \quad (1 \leq j \leq \ell) ,$$

and \underline{w} consequently satisfies

$$|a_{j=1}^* w| << X^{-1} (X^m Y^n)^{\frac{1}{l} + \frac{1}{l(l-1)}} = V \quad (1 \leq j \leq m)$$

and

$$|a_{j=m+1}^* w| << Y^{-1} (X^m Y^n)^{\frac{1}{l} + \frac{1}{l(l-1)}} = U \quad (m < j \leq l) .$$

Since

$$\left| \mathcal{M}(\underline{u}) - \underline{v} \right| = \max_{1 \leq j \leq m} |a_{j=1}^* w|$$

and

$$\left| \underline{u} \right| = \max_{m < j \leq l} |a_{j=m+1}^* w| ,$$

the proof of the theorem is complete.

Remark. Suppose that the integer point \underline{x} of the theorem is different from $\underline{0}$, and suppose that Y is small in terms of m and n . Then $X \geq 1$, and $V \leq Y^{n/(l-1)}$ is small, and

$$\left| \mathcal{M}(\underline{u}) - \underline{v} \right| < 1 .$$

If we had $\underline{u} = \underline{0}$, then also $\underline{v} = \underline{0}$, a contradiction. Therefore, if $\underline{x} \neq \underline{0}$ and Y is small, then $\underline{u} \neq \underline{0}$.

THEOREM 5B. L_1, \dots, L_n is a badly approximable system of linear forms if and only if M_1, \dots, M_m is a badly approximable system.

Remark. In the special case $m = 1$, sufficiency was proved by Perron (1921b), while necessity was established by Khintchine (1926b). In the case $m = 1$, or $n = 1$, the theorem asserts that an n -tuple is badly approximable if and only if the corresponding linear form $\alpha_1 x_1 + \dots + \alpha_n x_n$

is badly approximable. See Ch. II, §4.

Proof of the Theorem. By symmetry, it is enough to prove the sufficiency. Suppose, then, that M_1, \dots, M_m is a badly approximable system, and that $\underline{x} \neq \underline{0}$, \underline{y} are integer points in E^m, E^n respectively. Put $Y = \begin{bmatrix} f(\underline{x}) - \underline{y} \\ \underline{x} \end{bmatrix}$ and $X = \begin{bmatrix} \underline{x} \end{bmatrix}$. It is required to show that

$$(5.1) \quad X^m Y^n \gg 1.$$

It is no restriction to assume that Y is small. In view of Theorem 5A and the remark following the proof of that theorem, there are integer points $\underline{u} \neq \underline{0}$, \underline{v} in E^n, E^m respectively such that the conclusions of Theorem 5A hold. Since M_1, \dots, M_m is a badly approximable system,

$$U^n V^m \gg 1,$$

and hence

$$U^{n(\ell-1)} V^{m(\ell-1)} \gg 1.$$

But this is equivalent to (5.1), and the theorem is proved.

THEOREM 5C. ("Khinchine's Transference Principle") Let ω be the supremum of the numbers $\eta \geq 0$ such that there are infinitely many integer points $\underline{x} \neq \underline{0}$ with

$$(5.2) \quad \frac{\begin{vmatrix} f(\underline{x}) - \underline{y} \\ \underline{x} \end{vmatrix}}{\begin{vmatrix} \underline{x} \end{vmatrix}} < \frac{1}{\frac{m}{n}(1+\eta)}$$

for some integer point \underline{y} . Similarly, let ω^* be the supremum of the numbers $\eta^* \geq 0$ such that there are infinitely many integer points $\underline{u} \neq \underline{0}$ with

$$(5.3) \quad \left| \overline{m(u) - v} \right| < \frac{1}{\frac{n}{m}(1+\eta^*)} \left| \overline{u} \right|$$

for some integer point \underline{v} . Then

$$(5.4) \quad \omega^* \cong \frac{\omega}{(m-1)\omega + m+n-1} ,$$

and

$$(5.5) \quad \omega \cong \frac{\omega^*}{(n-1)\omega^* + m+n-1} .$$

In particular, $\omega = 0$ precisely if $\omega^* = 0$.

The special case $m = 1$ of the theorem was proved by Khintchine, who established inequality (5.5) in (1925) and inequality (5.4) in (1926b). The general case is due to Dyson (1947b). Jarnik (1959) showed (5.4), (5.5) to be best possible.

Proof of the Theorem. We prove only (5.4) . We may assume that $\omega > 0$, and in the case $m > 1$ that $\omega^* < \frac{1}{m-1}$, since otherwise the result is obvious. Suppose, then, that $0 < \eta < \omega$. Suppose also that $\eta^* > \omega^*$, and that $\eta^* < \frac{1}{m-1}$ if $m > 1$. Since $\eta < \omega$, there are infinitely many integer points $\underline{x} \neq \underline{0}$ satisfying (5.2) .

If we put $X = \left| \overline{x} \right|$ and $Y = \min_{\underline{y}} \left| \overline{f(x) - y} \right|$, where the minimum is taken over all integer points $\underline{y} \in E^n$, then $\underline{x} \neq \underline{0}$ satisfies (5.2) for some integer point \underline{y} if and only if

$$(5.6) \quad YX^{\frac{m}{n}(1+\eta)} < 1 .$$

Here Y will be small if we take X sufficiently large. (This is possible, since (5.2) is satisfied for infinitely many integer points \underline{x} .) By Theorem 5B and the remark following its proof, there exist integer points

$\underline{u} \neq \underline{0}$ and \underline{v} with $\left| \underline{u} \right| \ll U$ and $\left| \eta(\underline{u}) - \underline{v} \right| \ll V$. Since $\eta^* > \omega^*$, (5.3) is satisfied for only finitely many \underline{u} , and therefore

$$(5.7) \quad \frac{\underline{n}}{V U^m} (1 + \eta^*) \gg 1$$

if X , hence U , is large.

From (5.7), we obtain

$$V^{m(\ell-1)} U^{n(\ell-1)(1+\eta^*)} \gg 1,$$

which yields

$$X^{m+mn\eta^*} Y^{n-n(m-1)\eta^*} \gg 1.$$

The exponent of Y here is positive since $\eta^* < \frac{1}{m-1}$, and therefore

$$\frac{\underline{m}}{X^n} \cdot \frac{1+n\eta^*}{1-(m-1)\eta^*} Y \gg 1.$$

Comparing this with (5.6) and taking X sufficiently large, we conclude that

$$1 + \eta \leq \frac{1+n\eta^*}{1-(m-1)\eta^*}.$$

It follows that

$$\eta^* \geq \frac{\eta}{(m-1)\eta + m+n-1}.$$

Since η and η^* may be chosen arbitrarily close to ω and ω^* , respectively, inequality (5.4) is established.

A system of linear forms is called very well approximable if $\omega > 0$.

It follows from Theorem 5C that a system of linear forms is very well approximable if and only if the dual system is. In particular, in the case $m = 1$, an n -tuple $(\alpha_1, \dots, \alpha_n)$ is called very well approximable

if for some $\varepsilon > 0$ there are infinitely many $q > 0$ having

$$\max(\|q\alpha_1\|, \dots, \|q\alpha_n\|) < q^{-(1/n)-\varepsilon},$$

and a form $\alpha_1 q_1 + \dots + \alpha_n q_n$ is very well approximable if for some ω there are infinitely many n -tuples (q_1, \dots, q_n) with

$$|\alpha_1 q_1 + \dots + \alpha_n q_n| < q^{-n-\omega},$$

where $q = \max(|q_1|, \dots, |q_n|) \neq 0$. It now follows from the theorem that

an n -tuple $(\alpha_1, \dots, \alpha_n)$ is very well approximable precisely if the corresponding linear form $\alpha_1 q_1 + \dots + \alpha_n q_n$ is very well approximable.

It was shown in §3,4 of Ch. III that almost no n -tuple is very well approximable, and it now follows that almost no linear form is very well approximable. For further transference theorems see Wang, Yu and Zhu (1979) and Schmidt and Wang (1979).

§6. The Grassmann Algebra.

For a more thorough discussion see Greub (1967).

Let $\underline{e}_1 = (1, 0, \dots, 0), \dots, \underline{e}_n = (0, 0, \dots, 1)$ be the standard basis vectors of \mathbb{R}^n .

Let \mathbb{R}_0^n be the vector space of real numbers. It is a one dimensional Euclidean space, and it has 1 as a basis vector. For $1 \leq p \leq n$, consider all formal expressions

$$(6.1) \quad \underline{e}_{i_1} \wedge \underline{e}_{i_2} \wedge \dots \wedge \underline{e}_{i_p}$$

with integers $1 \leq i_1 < i_2 < \dots < i_p \leq n$. There are

$$(6.2) \quad l = \binom{n}{p}$$

such expressions. Let \mathbb{R}_p^n be the l -dimensional vector space generated by the symbols (6.1). The elements $\underline{x}^{(p)}$ of \mathbb{R}_p^n will be called p -vectors. We shall denote p -vectors by $\underline{x}^{(p)}$, $\underline{y}^{(p)}$ etc. or by \underline{x}_σ , \underline{y}_σ etc. Note that

$\mathbb{R}_1^n = \mathbb{R}^n$, the ordinary n -space, and denote vectors of this space by \underline{x} , \underline{y} , etc., as usual.

Write $C(n, p)$ for the set of p -tuples of integers i_1, \dots, i_p with $1 \leq i_1 < \dots < i_p \leq n$. For $\sigma = \{i_1 < \dots < i_p\}$ in $C(n, p)$ let \underline{E}_σ be the vector (6.1). The ℓ vectors \underline{E}_σ form a basis of \mathbb{R}_p^n . Put

$$\underline{E}_{\sigma=\tau} \underline{E}_\tau = \delta_{\sigma\tau} = \begin{cases} 1 & \text{if } \sigma = \tau \\ 0 & \text{otherwise} \end{cases},$$

for any elements σ, τ of $C(n, p)$, and use linearity to extend this to an inner product on \mathbb{R}_p^n . With this inner product, \mathbb{R}_p^n becomes a Euclidean space.

For convenience of notation we shall allow more general expressions $\underline{e}_{j_1} \wedge \underline{e}_{j_2} \wedge \dots \wedge \underline{e}_{j_p}$, where $1 \leq j_k \leq n$ ($k = 1, \dots, p$) with arbitrary $p \geq 1$, with the convention that such an expression is $\underline{0}$ if $j_h = j_k$ for some $h \neq k$ (so, in particular, it will be $\underline{0}$ if $p > n$), and that it is plus or minus (6.1) if j_1, \dots, j_p is obtained from a p -tuple $i_1 < i_2 < \dots < i_p$ by an even or by an odd permutation.

The direct sum $G_n = \mathbb{R}_0^n \oplus \mathbb{R}_1^n \oplus \dots \oplus \mathbb{R}_n^n$ is a vector space of dimension 2^n . We define a product \wedge in G_n by the formulas $1 \wedge 1 = 1$,

$$(6.3) \quad 1 \wedge (\underline{e}_{i_1} \wedge \dots \wedge \underline{e}_{i_p}) = (\underline{e}_{i_1} \wedge \dots \wedge \underline{e}_{i_p}) \wedge 1 = \underline{e}_{i_1} \wedge \dots \wedge \underline{e}_{i_p}$$

and

$$(6.4) \quad (\underline{e}_{i_1} \wedge \dots \wedge \underline{e}_{i_p}) \wedge (\underline{e}_{j_1} \wedge \dots \wedge \underline{e}_{j_q}) = \underline{e}_{i_1} \wedge \dots \wedge \underline{e}_{i_p} \wedge \underline{e}_{j_1} \wedge \dots \wedge \underline{e}_{j_q},$$

and by extending the definition using linearity to any two vectors of G_n . With this product G_n becomes an associative algebra, the Grassmann or exterior algebra of \mathbb{R}^n .

For any vectors $\underline{x}_1, \dots, \underline{x}_p$ in \mathbb{R}^n with $1 \leq p \leq n$, the vector

$$(6.5) \quad \underline{x}^{(p)} = \underline{x}_1 \wedge \dots \wedge \underline{x}_p$$

is a p -vector. A p -vector of this type is called decomposable.

LEMMA 6A. Suppose that $\underline{x}_i = (\xi_{i1}, \dots, \xi_{in}) = \sum_{j=1}^n \xi_{ij} \underline{e}_j$ ($1 \leq i \leq p$). Then the vector $\underline{X}^{(p)}$ given by (6.5) is

$$(6.6) \quad \underline{X}^{(p)} = \sum_{\sigma \in C(n,p)} \xi_{\sigma} \underline{E}_{\sigma},$$

where the summation is extended over σ in $C(n,p)$, and where the coefficient ξ_{σ} is the $(p \times p)$ -determinant $|\xi_{ij}|$ with $1 \leq i \leq p$ and $j \in \sigma$.

For example, if $n = 4$ and $p = 2$, we have

$$\begin{aligned} \underline{x}_1 \wedge \underline{x}_2 = & \begin{vmatrix} \xi_{11} & \xi_{12} \\ \xi_{21} & \xi_{22} \end{vmatrix} \underline{E}_{12} + \begin{vmatrix} \xi_{11} & \xi_{13} \\ \xi_{21} & \xi_{23} \end{vmatrix} \underline{E}_{13} + \begin{vmatrix} \xi_{11} & \xi_{14} \\ \xi_{21} & \xi_{24} \end{vmatrix} \underline{E}_{14} \\ & + \begin{vmatrix} \xi_{12} & \xi_{13} \\ \xi_{22} & \xi_{23} \end{vmatrix} \underline{E}_{23} + \begin{vmatrix} \xi_{12} & \xi_{14} \\ \xi_{22} & \xi_{24} \end{vmatrix} \underline{E}_{24} + \begin{vmatrix} \xi_{13} & \xi_{14} \\ \xi_{23} & \xi_{24} \end{vmatrix} \underline{E}_{34}. \end{aligned}$$

Proof. Both sides of (6.6) are linear in each \underline{x}_i , and hence we may suppose that $\underline{x}_1, \dots, \underline{x}_p$ are among the basis vectors $\underline{e}_1, \dots, \underline{e}_n$. If two among $\underline{x}_1, \dots, \underline{x}_p$ are the same \underline{e}_i , then both sides of (6.6) vanish. So we may suppose that $\underline{x}_1, \dots, \underline{x}_p$ are distinct vectors among $\underline{e}_1, \dots, \underline{e}_n$. Both sides of (6.6) now change into $+$ or $-$ themselves if we permute $\underline{x}_1, \dots, \underline{x}_p$ by an even or by an odd permutation. Hence it suffices to consider the special case when $\underline{x}_1 = \underline{e}_{i_1}, \dots, \underline{x}_p = \underline{e}_{i_p}$ where $i_1 < i_2 < \dots < i_p$. Then $\underline{X}^{(p)} = \underline{E}_{\tau}$ with $\tau = \{i_1 < \dots < i_p\}$; on the other hand $\xi_{\sigma} = 1$ if $\sigma = \tau$ and $\xi_{\sigma} = 0$ if $\sigma \neq \tau$.

LEMMA 6B. The vector (6.5) is 0 if and only if $\underline{x}_1, \dots, \underline{x}_p$ are linearly dependent.

Proof. This is an obvious consequence of Lemma 6A.

LEMMA 6C. Suppose $\underline{x}_1, \dots, \underline{x}_p$ and $\underline{y}_1, \dots, \underline{y}_p$ are sets of p independent points in \mathbb{R}^n . The points $\underline{x}^{(p)} = \underline{x}_1 \wedge \dots \wedge \underline{x}_p$ and $\underline{y}^{(p)} = \underline{y}_1 \wedge \dots \wedge \underline{y}_p$ of \mathbb{R}_p^n are proportional, i.e.,

$$(6.7) \quad \underline{y}^{(p)} = \lambda \underline{x}^{(p)},$$

if and only if $\underline{x}_1, \dots, \underline{x}_p$ and $\underline{y}_1, \dots, \underline{y}_p$ span the same subspace in \mathbb{R}^n .

Proof. Let S^p be the subspace spanned by $\underline{x}_1, \dots, \underline{x}_p$. If $\underline{y}_1, \dots, \underline{y}_p$ lie in S^p , then clearly (6.7) holds. Conversely, suppose we have (6.7). Since $\underline{y}_1, \dots, \underline{y}_p$ are independent, we have $\lambda \neq 0$.

It is clear that

$$\underline{x} \wedge (\underline{x}_1 \wedge \dots \wedge \underline{x}_p) = \underline{x} \wedge \underline{x}^{(p)} = 0$$

precisely if \underline{x} lies in S^p . Since $\underline{y}_i \wedge \underline{x}^{(p)} = \lambda^{-1} (\underline{y}_i \wedge \underline{y}^{(p)}) = 0$, we see that \underline{y}_i lies in S^p ($1 \leq i \leq p$).

LEMMA 6D. (Laplace identity). For $\underline{x}_1, \dots, \underline{x}_p, \underline{y}_1, \dots, \underline{y}_p$ in \mathbb{R}^n ,

$$(6.8) \quad (\underline{x}_1 \wedge \dots \wedge \underline{x}_p) \cdot (\underline{y}_1 \wedge \dots \wedge \underline{y}_p) = \begin{vmatrix} \underline{x}_1 \underline{y}_1 & \underline{x}_1 \underline{y}_2 & \dots & \underline{x}_1 \underline{y}_p \\ \underline{x}_2 \underline{y}_1 & \underline{x}_2 \underline{y}_2 & \dots & \underline{x}_2 \underline{y}_p \\ \vdots & \vdots & & \vdots \\ \underline{x}_p \underline{y}_1 & \underline{x}_p \underline{y}_2 & \dots & \underline{x}_p \underline{y}_p \end{vmatrix},$$

where the raised dot on the L.H.S. denotes the inner product in \mathbb{R}_p^n , while juxtaposition in the determinant on the right denotes the inner product in $\mathbb{R}^n = \mathbb{R}_1^n$.

Proof. Exercise. Use linearity in $\underline{x}_1, \dots, \underline{y}_p$.

LEMMA 6E. Let $\underline{x}_1, \dots, \underline{x}_n$ be in \mathbb{R}^n . For each $\tau = \{j_1 < \dots < j_p\} \in C(n, p)$ put

$$X_{\tau} = x_{j_1} \wedge x_{j_2} \wedge \cdots \wedge x_{j_p}.$$

Then

$$(6.9) \quad \det_{\tau \in C(n,p)} (X_{\tau}) = (\det(x_1, \dots, x_n))^{lp/n}.$$

(It is understood that the τ 's in (6.9) are ordered lexicographically.

Observe that $lp/n = \binom{n-1}{p-1}$).

Remark. The entries in the matrix (X_{τ}) are of the form

$$X_{\tau\sigma} = \det(x_{ij}) \quad , \quad (i \in \tau, j \in \sigma)$$

where $\sigma \in C(n,p)$. The matrix $(X_{\tau}) = (X_{\tau\sigma})$ is called the p -th compound of the matrix (x_1, \dots, x_n) . For example, if $n = 4$ and $p = 2$, the (6×6) -matrix

$$\begin{pmatrix} \begin{vmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{vmatrix} & \begin{vmatrix} x_{11} & x_{13} \\ x_{21} & x_{23} \end{vmatrix} & \cdots & \begin{vmatrix} x_{13} & x_{14} \\ x_{23} & x_{24} \end{vmatrix} \\ \begin{vmatrix} x_{11} & x_{12} \\ x_{31} & x_{32} \end{vmatrix} & \begin{vmatrix} x_{11} & x_{13} \\ x_{31} & x_{33} \end{vmatrix} & \cdots & \begin{vmatrix} x_{13} & x_{14} \\ x_{33} & x_{34} \end{vmatrix} \\ \vdots & \vdots & & \vdots \\ \begin{vmatrix} x_{31} & x_{32} \\ x_{41} & x_{42} \end{vmatrix} & \begin{vmatrix} x_{31} & x_{33} \\ x_{41} & x_{43} \end{vmatrix} & \cdots & \begin{vmatrix} x_{33} & x_{34} \\ x_{43} & x_{44} \end{vmatrix} \end{pmatrix}$$

is the second compound of the matrix

$$\begin{pmatrix} x_{11} & x_{12} & x_{13} & x_{14} \\ x_{21} & x_{22} & x_{23} & x_{24} \\ x_{31} & x_{32} & x_{33} & x_{34} \\ x_{41} & x_{42} & x_{43} & x_{44} \end{pmatrix}$$

Proof. Consider the following two types of operations on n-tuples of vectors in \mathbb{R}^n :

$$(I) \quad T_i(\lambda): (\underline{a}_1, \dots, \underline{a}_i, \dots, \underline{a}_n) \mapsto (\underline{a}_1, \dots, \lambda \underline{a}_i, \dots, \underline{a}_n),$$

$$(II) \quad U_{ij}(\lambda): (\underline{a}_1, \dots, \underline{a}_i, \dots, \underline{a}_j, \dots, \underline{a}_n) \mapsto (\underline{a}_1, \dots, \underline{a}_i, \dots, \underline{a}_j + \lambda \underline{a}_i, \dots, \underline{a}_n) .$$

Let $\underline{x}_1, \dots, \underline{x}_n$ be arbitrary vectors in \mathbb{R}^n . From elementary linear algebra, there are finitely many operations of types (I) and (II) , say O_1, \dots, O_R , such that

$$O_R O_{R-1} \cdots O_1 : (\underline{e}_1, \dots, \underline{e}_n) \mapsto (\underline{x}_1, \dots, \underline{x}_n) .$$

If $(\underline{x}_1, \dots, \underline{x}_n) = (\underline{e}_1, \dots, \underline{e}_n)$, then both sides of (6.9) are 1 . To prove the lemma, it therefore suffices to show that both sides of (6.9) change in the same way if we apply operations of the type (I) or (II) to $(\underline{x}_1, \dots, \underline{x}_n)$.

If we apply $T_i(\lambda)$ to $(\underline{x}_1, \dots, \underline{x}_n)$, the R.H.S. of (6.9) is multiplied by $\lambda^{\ell p/n}$. A factor of λ is introduced by this operation in each entry of the τ -th row of the matrix $(X_{\tau\sigma})$ precisely when $i \in \tau$. Since there are $\ell = \binom{n}{p}$ elements of $C(n, p)$, there are $\binom{n-1}{p-1} = \frac{\ell p}{n}$ elements of $C(n, p)$ containing i . Thus the L.H.S. of (6.9) is multiplied by $\lambda^{\ell p/n}$ if we apply $T_i(\lambda)$.

In applying operations of type (II) , we see that the R.H.S. of (6.9) remains unchanged. Let us examine the effect of the operation $U_{ij}(\lambda)$ on the τ -th row

$$\underline{X}_{\tau} = \underline{x}_{i_1} \wedge \underline{x}_{i_2} \wedge \cdots \wedge \underline{x}_{i_p} \quad (\tau = \{i_1 < i_2 < \cdots < i_p\})$$

of the matrix in the L.H.S. of (6.9) . If $j \notin \tau$, or if $j \in \tau$ and $i \in \tau$, then it is clear that \underline{X}_{τ} remains unchanged. Suppose, then, that

$j \in \tau$ and $i \notin \tau$. After applying $U_{ij}(\lambda)$, X_{τ} becomes $X_{\tau} \pm \lambda X_{\bar{\tau}}$ where $\bar{\tau}$ is the element of $C(n, p)$ obtained from τ by replacing j with i . Note that X_{τ} is unaffected by the operation (II). It follows that the L.H.S. of (6.9) remains unchanged under operations of the type (II), and the lemma is proved.

LEMMA 6F. Let a_1, \dots, a_n be a basis of \mathbb{R}^n . Suppose $1 \leq p \leq n$, and for $\sigma = \{i_1 < \dots < i_p\}$ in $C(n, p)$ put

$$(6.10) \quad A_{\sigma} = a_{i_1} \wedge \dots \wedge a_{i_p}.$$

Then the $\ell = \binom{n}{p}$ vectors A_{σ} with $\sigma \in C(n, p)$ are a basis of \mathbb{R}_p^n .

If a_1, \dots, a_n have determinant 1, then so do the vectors A_{σ} .

Moreover, let a_1^*, \dots, a_n^* be the basis of \mathbb{R}^n which is reciprocal to a_1, \dots, a_n , and for $\sigma = \{i_1 < \dots < i_p\}$ in $C(n, p)$ put

$$(6.11) \quad (A^*)_{\sigma} = a_{i_1}^* \wedge \dots \wedge a_{i_p}^*.$$

Then

$$(6.12) \quad A_{\sigma} (A^*)_{\tau} = \delta_{\sigma\tau},$$

i.e. the basis $(A^*)_{\sigma}$ with $\sigma \in C(n, p)$ is reciprocal to the basis A_{σ} with $\sigma \in C(n, p)$.

Proof. Only (6.12) requires a proof. But (6.12) is a consequence of the Laplace identity.

§7. Mahler's Theory of Compound Sets.

Let a_1, \dots, a_n be points of \mathbb{R}^n with determinant 1. The inequalities

$$|a_i x| \leq 1 \quad (1 \leq i \leq n)$$

define a parallelepiped Π of volume 2^n . Suppose $1 \leq p \leq n$. For $\sigma = \{i_1 < \dots < i_p\}$ in $C(n, p)$ define \underline{A}_σ by (6.10). By Lemma 6F the vectors \underline{A}_σ have determinant 1, and the inequalities

$$|\underline{A}_\sigma^{(p)}| \leq 1 \quad (\sigma \in C(n, p))$$

define a parallelepiped $\Pi^{(p)}$ in \mathbb{R}_p^n of volume 2^ℓ where $\ell = \binom{n}{p}$. This parallelepiped is closely related to Mahler's p -th compound body of Π and will be called the p -th pseudocompound of Π . (Mahler (1955) had more generally defined compounds of symmetric convex sets).

Let $\lambda_1, \dots, \lambda_n$ be the successive minima of Π , and for τ in $C(n, p)$ put

$$(7.1) \quad \lambda_\tau = \prod_{i \in \tau} \lambda_i.$$

There is an ordering $\tau_1, \tau_2, \dots, \tau_\ell$ of the elements of $C(n, p)$ such that

$$(7.2) \quad \lambda_{\tau_1} \leq \lambda_{\tau_2} \leq \dots \leq \lambda_{\tau_\ell}.$$

Now let $\underline{g}_1, \dots, \underline{g}_n$ be independent integer points such that \underline{g}_j lies in $\lambda_j \Pi$, i.e., that

$$(7.3) \quad |\underline{a}_{i \underline{g}_j}| \leq \lambda_j \quad (1 \leq i, j \leq n).$$

For $\tau = \{j_1 < \dots < j_p\}$ in $C(n, p)$, put $\underline{G}_\tau = \underline{g}_{j_1} \wedge \dots \wedge \underline{g}_{j_p}$. By Laplace's identity (6.8), $\underline{A}_{\sigma=\tau} \underline{G}_\tau$ is equal to the $p \times p$ -determinant $|\underline{a}_{i \underline{g}_j}|$ with $i \in \sigma$, $j \in \tau$. Hence in view of (7.3) we obtain

$$(7.4) \quad |\underline{A}_{\sigma=\tau} \underline{G}_\tau| \leq p! \lambda_\tau \quad (\sigma, \tau \in C(n, p)).$$

Thus we obtain the following result due to Mahler (1955).

THEOREM 7A. Suppose the parallelepiped Π with minima $\lambda_1, \dots, \lambda_n$

and the p -th pseudocompound $\Pi^{(p)}$ with minima v_1, \dots, v_ℓ are defined as above. Also let $\underline{g}_1, \dots, \underline{g}_n$ and \underline{G}_τ be points as above. Then (7.4) holds and we have

$$(7.5) \quad \lambda_{\tau_i} \ll v_i \ll \lambda_{\tau_i} \quad (1 \leq i \leq \ell) .$$

Here the constants in \ll depend only on n .

Proof. The points \underline{G}_τ with $\tau \in C(n, p)$ are ℓ independent integer points, and hence we have $v_i \leq p! \lambda_{\tau_i}$ by (7.2) and (7.4). The lower bound for v_i follows from $1 \ll v_1 \cdots v_\ell \ll 1$ and from $1 \ll \prod_{i=1}^{\ell} \lambda_{\tau_i} = (\lambda_1 \cdots \lambda_n)^t \ll 1$, where $t = \binom{n-1}{p-1}$.

Exercises. (Will not be needed in sequel). Let $\underline{a}_1, \dots, \underline{a}_n$ of \mathbb{R}^n have determinant 1. (1) Given $\sigma = \{i_1 < \cdots < i_p\}$ in $C(n, p)$, let σ' be the element of $C(n, n-p)$ which is the complement of σ in $\{1, \dots, n\}$. Define \underline{A}_σ by (6.10) and write $\underline{A}_\sigma = \{A_{\sigma\tau}\}$ with $\tau \in C(n, p)$, so that each $A_{\sigma\tau}$ is a certain $(p \times p)$ -determinant. Let $\hat{\underline{A}}_\sigma$ be the vector $\hat{\underline{A}}_\sigma = \{\hat{A}_{\sigma\tau}\}$ with

$$\hat{A}_{\sigma\tau} = (-1)^\sigma (-1)^\tau A_{\sigma'\tau'} ,$$

where $(-1)^\sigma = (-1)^{i_1 + \cdots + i_p}$ if $\sigma = \{i_1 < \cdots < i_p\}$. Expand the determinant $\det(\underline{a}_1, \dots, \underline{a}_n)$ according to Laplace^{†)} to show that

$$\underline{A}_{\sigma_1} \hat{\underline{A}}_{\sigma_2} = \begin{cases} 1 & \text{if } \sigma_1 = \sigma_2 , \\ 0 & \text{otherwise} . \end{cases}$$

(2) Show that the map which sends $\underline{X} = \{X_\sigma\}$ in \mathbb{R}_p^n into $\underline{Y} = \{Y_{\sigma'}\}$ in \mathbb{R}_{n-p}^n with $Y_{\sigma'} = (-1)^\sigma X_\sigma$, sends the parallelepiped $(\Pi^{(p)})^*$ (i.e. the reciprocal of $\Pi^{(p)}$) onto $\Pi^{(n-p)}$.

†) Not the Laplace identity of Lemma 6D.

(3) Take $p = 1$ and deduce (4.3) in Theorem 4A from (7.5) in Theorem 7A.

§8. Point Lattices.

Given n linearly independent points $\underline{b}_1, \dots, \underline{b}_n$ in \mathbb{R}^n , the set of points

$$c_1 \underline{b}_1 + \dots + c_n \underline{b}_n$$

with integer coefficients c_1, \dots, c_n is called a lattice, and $\underline{b}_1, \dots, \underline{b}_n$ is called a basis of this lattice. The basis of a lattice is not unique.

If $\underline{b}'_1, \dots, \underline{b}'_n$ is another basis, then $\underline{b}'_i = \sum_{j=1}^n c_{ij} \underline{b}_j$ ($1 \leq i \leq n$) and $\underline{b}_i = \sum_{j=1}^n c'_{ij} \underline{b}'_j$ ($1 \leq i \leq n$) with integers c_{ij} and c'_{ij} . Since the matrices (c_{ij}) and (c'_{ij}) are inverses of each other, they are unimodular, i.e. they have determinant 1 or -1.

Therefore if our space is equipped with a Euclidean metric, if it is a Euclidean space E^n , then the volume of the parallelepiped of points $x_1 \underline{b}_1 + \dots + x_n \underline{b}_n$ with $0 \leq x_i < 1$ is independent of the basis and depends only on the lattice. Given a lattice Λ this volume is denoted by $d(\Lambda)$ and is called the determinant of the lattice. It equals $|\det(b_{ij})|$, if $\underline{b}_i = (b_{i1}, \dots, b_{in})$ in a Cartesian coordinate system.

Given a Cartesian coordinate system, the set of integer points forms a lattice Λ_0 , the fundamental lattice. Every lattice Λ is of the form $\Lambda = T\Lambda_0$ where T is a non-singular linear transformation, and $d(\Lambda)$ is equal to the determinant of T .

If K is a symmetric convex body, we define the successive minima $\lambda_1, \dots, \lambda_n$ of K with respect to a lattice Λ by the property that λ_i is the smallest positive number such that $\lambda_i K$ contains i linearly

independent points of Λ . The results of this chapter so far were about integer points, i.e. about Λ_0 , but they can all be generalized to arbitrary lattices. In the proofs we simply have to transform Λ back to Λ_0 ; then a convex body of volume $V(K)$ will be mapped into a convex body of volume $V(K)d(\Lambda)^{-1}$. Theorem 1A* now becomes

$$(8.1) \quad \frac{2^n}{n!} d(\Lambda) \leq \lambda_1 \cdots \lambda_n V(K) \leq 2^n d(\Lambda) .$$

We have not proved Theorem 1A*, but we did prove Lemma 1C and Corollary 2B, so that in the general case we have proved

$$(8.2) \quad \frac{2^n}{n!} d(\Lambda) \leq \lambda_1 \cdots \lambda_n V(K) \leq 2^n n^{n/2} d(\Lambda) .$$

THEOREM 8A. Necessary and sufficient for a subset Λ of \mathbb{R}^n to be a lattice is that the following three conditions hold.

- (i) Λ forms a group under vector addition,
- (ii) Λ contains n linearly independent points,
- (iii) Λ is discrete, i.e., it contains no limit point.

Proof. The necessity of the conditions follows from the definition. The sufficiency is easy when $n = 1$: By (ii), (iii) there is a point $\underline{b}_1 \in \Lambda$, $\underline{b}_1 \neq \underline{0}$, which is closest to $\underline{0}$. Then it is easily seen that every $\underline{b} \in \Lambda$ is of the form $\underline{b} = c_1 \underline{b}_1$ with integral c_1 , and hence Λ is a lattice with basis \underline{b}_1 . When $n > 1$ we may choose our coordinate system so that Λ contains $n-1$ linearly independent points on the subspace \mathbb{R}^{n-1} of \mathbb{R}^n where $x_n = 0$. Then $\Lambda' = \Lambda \cap \mathbb{R}^{n-1}$ is a lattice by induction. Let $\underline{b}_1, \dots, \underline{b}_{n-1}$ be a basis of Λ' . There is a point $\underline{b}_n = (b_{n1}, \dots, b_{nn})$ in Λ with $b_{nn} > 0$, which is clear, and in fact there is such a point for which $b_{nn} > 0$ is minimal, which is less clear. But it does become clear when we observe that Λ

is discrete and that b_{nn} remains unchanged if we add to b_{nn} an element of Λ' , so that we may restrict ourselves to b_{nn} for which $(b_{n1}, \dots, b_{n,n-1})$ lies in the parallelepiped of points $\lambda_1 b_{n1} + \dots + \lambda_{n-1} b_{n,n-1}$ with $0 \leq \lambda_i < 1$ ($i = 1, \dots, n-1$). Now if $\underline{g} = (g_1, \dots, g_n) \in \Lambda$, then also $\underline{g}' = \underline{g} - [g_n/b_{nn}] b_{nn}$ lies in Λ . The n -th coordinate of \underline{g}' is ≥ 0 and less than b_{nn} , hence must be zero. Thus $\underline{g}' \in \Lambda'$, and \underline{g} is a linear combination of $b_1, \dots, b_{n-1}, b_{nn}$. Hence Λ is a lattice with basis b_1, \dots, b_n .

Suppose S^d is a d -dimensional rational subspace of E^n , i.e. a subspace spanned by vectors with rational coordinates. Then $\Lambda_0 \cap S^d$ satisfies the conditions of Theorem 8A (with d in place of n) and we have

COROLLARY 8B. The integer points on a rational subspace S^d form a lattice on S^d .

V. Roth's Theorem

References: Thue (1908, 1909)

Roth (1955a)

Schmidt (1971d).

§1. Liouville's Theorem

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

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

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

(In considering inequalities of this type here and elsewhere, we tacitly assume that $q > 0$).

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

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

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

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

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).

Proof. (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

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

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

(c) Unless $d = 1$ and $\frac{p}{q} = \alpha$, we have $P\left(\frac{p}{q}\right) \neq 0$, whence

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

§2. Roth's Theorem and its History.

Suppose α is a real algebraic number of degree $d \geq 2$. Liouville's Theorem implies that the inequality

$$(2.1) \quad \left| \alpha - \frac{p}{q} \right| < \frac{1}{q^\mu}$$

has only finitely many rational solutions $\frac{p}{q}$ if $\mu > d$. The great Norwegian mathematician Thue (1909)[†] showed that (2.1) has only finitely many solutions if $\mu > \frac{1}{2}d + 1$. Then Siegel (1921) in his thesis showed that this is already true if $\mu > 2\sqrt{d}$. (Siegel's result was a little better, with a more complicated function in place of $2\sqrt{d}$). A slight improvement to $\mu > \sqrt{2d}$ was made by Dyson (1947a). See also Gelfond (1952). Finally Roth (1955a) proved that (2.1) has only finitely many solutions if $\mu > 2$. For this result Roth was awarded the Field Prize

[†]No. 12 in Thue's selected papers (1977).

in 1958. Roth's Theorem is as follows.

THEOREM 2A. Suppose α is real and algebraic of degree $d \geq 2$.
Then for each $\delta > 0$, the inequality

$$(2.2) \quad \left| \alpha - \frac{p}{q} \right| < \frac{1}{q^{2+\delta}}$$

has only finitely many solutions in rationals $\frac{p}{q}$.

Remarks. (i) The conclusion is trivially true for a complex α which is not real.

(ii) By Dirichlet's Theorem the exponent 2 in (2.2) is best possible. If α is of degree 2, then

$$(2.3) \quad \left| \alpha - \frac{p}{q} \right| > \frac{c(\alpha)}{q^2}$$

by Liouville's Theorem. In this case Liouville's Theorem is stronger than Roth's. Of course (2.3) follows already from Lemma 2E of Chapter I.

(iii) For no single α of degree ≥ 3 do we know whether (2.3) holds. It is very likely (see Ch. IV, §3) that in fact (2.3) is false for every such α , i.e., that no such α is badly approximable, or, put differently, that such α has unbounded partial quotients in its continued fraction.

(iv) An apparently very difficult conjecture (Lang (1965)) is that for α of degree ≥ 3 ,

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

has only finitely many solutions if $\kappa > 1$, or at least if $\kappa > \kappa_0(\alpha)$.

We now make a preliminary remark about the proof. 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 (2.1) holds, 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 $\left|P\left(\frac{p}{q}\right)\right| \leq cq^{-\mu i}$. Finally (c) we have $P\left(\frac{p}{q}\right) \neq 0$ whence $\left|P\left(\frac{p}{q}\right)\right| \geq q^{-r}$ for all but finitely many rationals $\frac{p}{q}$. Hence if (2.1) has infinitely many solutions, then $\mu i \leq r$ or $\mu \leq \left(\frac{r}{i}\right)^{-1}$. Hence one should try to make $\frac{i}{r}$ as large as possible. But it is clear that always $\frac{i}{r} \leq \frac{1}{d}$, 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 $\mu \leq d$, i.e., nothing better than Liouville's result.

In order to improve on this estimate, Thue uses a polynomial $X_2 Q(X_1) - P(X_1)$ in two variables, and Siegel uses a more general polynomial $P(X_1, X_2)$ in two variables, while Schneider (1936) and Roth use a polynomial $P(X_1, \dots, X_m)$ in many variables.

Now if $\frac{p_1}{q_1}, \dots, \frac{p_m}{q_m}$ are very good rational approximations, one substitutes these into $P(X_1, \dots, X_m)$. A major problem is that it is difficult to ascertain that $P\left(\frac{p_1}{q_1}, \dots, \frac{p_m}{q_m}\right) \neq 0$. This difficulty is overcome by "Roth's Lemma", which requires however that $q_1 < q_2 < \dots < q_m$. A consequence is that we reach a contradiction only if we have at least m very good rational approximations. In the case of Thue's and Siegel's result one has $m = 2$ and hence one needs two very good rational approximations. Davenport (1968) used Thue's approach to show that for cubic α and for $\mu > 1 + \sqrt{3}$, solutions to (2.1) have $q \leq c_1(\alpha, \mu)$ with an explicit $c_1(\alpha, \mu)$, with one possible exception $\frac{p}{q}$, and Schinzel (1967) used Siegel's approach to do the same for α

of degree $d \geq 3$ and for $\mu > \sqrt[3]{d/2}$.

But the method of Thue-Siegel-Roth does not give a bound which holds without exception, and hence it is non-effective. It gives no way to find all the solutions of (2.1). Effective results, which however are weaker than the results of Thue (and hence a fortiori of Roth) were given by Feldman (1971) with Baker's method.

THEOREM 2B*. Suppose that α is algebraic of degree $d \geq 3$. Then there are explicit constants $\mu_0(\alpha) < d$ and $c_2(\alpha)$ such that every solution of (2.1) with $\mu \geq \mu_0$ has $q \leq c_2(\alpha)$.

This Theorem will not be proved in these Notes.

We remark that the method of Thue-Siegel-Roth does permit to estimate the number of solutions of (2.1), and this is carried out by Davenport and Roth (1955).

Our proof of Roth's Theorem in §§ 4-11 will follow rather closely the exposition in Cassels (1957). Generalisations will be presented in chapters VI and VIII.

§3. Thue's Equation.

THEOREM 3A. (Thue (1908)). Suppose $F(X,Y)$ is a binary form with rational coefficients, and with at least 3 distinct linear factors (with algebraic coefficients). Then if m is non-zero, the diophantine equation

$$(3.1) \quad F(x,y) = m$$

has only finitely many solutions in rational integers x,y .

An equation of the type (3.1) will henceforth be called a Thue equation.

Proof. We may factor $F(X,Y)$ as

$$(3.2) \quad F(X,Y) = a(\gamma_1 X + \delta_1 Y)^{e_1} \cdots (\gamma_s X + \delta_s Y)^{e_s}$$

with $s \geq 3$ and with real or complex algebraic $\gamma_1, \delta_1, \dots, \gamma_s, \delta_s$ such that forms $\gamma_i X + \delta_i Y$ and $\gamma_j X + \delta_j Y$ with $i \neq j$ are linearly independent, with the further convention that each γ_i is either 1 or 0, and that $\delta_i = 1$ if $\gamma_i = 0$. By rearranging the factors we may suppose that

$$0 < |\gamma_1 x + \delta_1 y| \leq \cdots \leq |\gamma_s x + \delta_s y|.$$

Since $\gamma_1 X + \delta_1 Y$ and $\gamma_2 X + \delta_2 Y$ are linearly independent,

$$\begin{aligned} |\gamma_s x + \delta_s y| &\geq \cdots \geq |\gamma_2 x + \delta_2 y| \geq \frac{1}{2}(|\gamma_1 x + \delta_1 y| + |\gamma_2 x + \delta_2 y|) \\ &\geq c_1 \max(|x|, |y|) = c_1 \left| \frac{x}{y} \right|, \end{aligned}$$

say, and

$$|F(x,y)| \geq c_2 |\gamma_1 x + \delta_1 y|^{e_1} \left| \frac{x}{y} \right|^{d-e_1},$$

where d is the degree of F . If $\gamma_1 X + \delta_1 Y$ has rational coefficients, we have $|\gamma_1 x + \delta_1 y| \geq c_3$, and since $d > e_1$ it follows that $|F(x,y)|$ tends to infinity with $\left| \frac{x}{y} \right|$. If $\gamma_1 = 1$ and δ_1 is algebraic of degree $l \geq 2$ it follows from Thue's result mentioned in §2 (hence a fortiori from Roth's Theorem) that for $\delta > 0$,

$$|\gamma_1 x + \delta_1 y| \geq c_4(\delta_1, \delta) \left| \frac{x}{y} \right|^{-(l/2)-\delta},$$

so that

$$|F(x,y)| \geq c_5 \left| \frac{x}{y} \right|^{d-e_1((l/2)+1+\delta)}.$$

Together with $\gamma_1 X + \delta_1 Y = X + \delta_1 Y$, the form F must have the conjugate factors, each with multiplicity e_1 , so that $d \geq \ell e_1$, and if $\ell = 2$ we even have $d \geq 2e_1 + 1$, since F has at least 3 distinct factors. Since $\delta > 0$ may be picked arbitrarily small, we have $d > e_1((\ell/2) + 1 + \delta)$, and again $|F(x,y)|$ tends to infinity with $\left[\frac{x}{y}\right]$.

By using the full strength of Roth's Theorem one obtains the following result, whose proof is left as an Exercise.

THEOREM 3B. Suppose $F(X,Y)$ is a binary form of degree $d \geq 3$ with rational coefficients and without multiple factors. Then for given $v < d - 2$ there are only finitely many integer points $\underline{x} = (x,y)$ with

$$0 < |F(x,y)| < \left[\frac{x}{y}\right]^v.$$

In particular, if $G(X,Y)$ is a form of degree $< d - 2$, then the diophantine equation

$$F(x,y) = G(x,y)$$

has only finitely many solutions having[†] $F(x,y) \neq 0$.

The method of proof of the above theorems by Thue's method is non-effective, i.e. it provides no bound for the size of the solutions. But Baker (1967/68) gave the bound

$$\left[\frac{x}{y}\right] < \exp((dH)^{10d^5})$$

for the solutions $\underline{x} = (x,y)$ of Thue's equation (3.1), where d is the degree of F , where F has distinct linear factors, and where the coefficients of F as well as m are rational integers of absolute value

[†]) There are infinitely many solutions of $F(x,y) = G(x,y) = 0$ precisely if F, G have a common linear factor with rational coefficients.

at most H .

Thue's method does provide estimates for the number of solutions of his equations, and this was carried out with respect to Theorem 3A by Mahler (1933), and with respect to Theorem 3B by Davenport and Roth (1955). Siegel (1970) in the case $F(X,Y) = (\alpha X + \beta Y)^d + (\gamma X + \delta Y)^d$ with $\alpha\delta - \beta\gamma \neq 0$ gives a bound which depends only on d and on m . Note that for $d = 3$ every non-degenerate form is of this type.

§4. Combinatorial Lemmas.

LEMMA 4A. Suppose that r_1, \dots, r_m are positive integers and that
 $0 < \varepsilon < 1$. Then the number of m -tuples of integers i_1, \dots, i_m with

$$0 \leq i_h \leq r_h \quad (1 \leq h \leq m)$$

and with

$$\left| \left(\sum_{h=1}^m \frac{i_h}{r_h} \right) - \frac{m}{2} \right| \geq \varepsilon m$$

is at most

$$(r_1 + 1) \cdots (r_m + 1) \cdot 2e^{-\varepsilon^2 m/4} .$$

Remarks. (a) The lemma may be given a probabilistic interpretation. Namely, regard the expressions $\frac{i_h}{r_h}$ as random variables with expectation $\frac{1}{2}$. By the Law of Large Numbers, $\sum \frac{i_h}{r_h} \approx \frac{m}{2}$ with high probability. (b) A slightly weaker version of this lemma has long been known: see, e.g., Schneider (1936). (c) Lemma 4A is an immediate consequence of Lemma 4C below. Our proof will be like that given in Mahler (1961),

Appendix A. Mahler attributes the proof given there to G. E. H. Reuter.

LEMMA 4B. Suppose that $n \geq 1$, $r \geq 0$ are integers. The number of
 n -tuples of nonnegative integers i_1, \dots, i_n with $i_1 + \dots + i_n = r$ is

$$(4.1) \quad N(n, r) = \binom{r + n - 1}{r}.$$

Proof. The argument is by induction on n and r . We have

$$N(n, 0) = 1 = \binom{n-1}{0}$$

for any $n \geq 1$, so that (4.1) holds for $r = 0$. Also,

$$N(1, r) = 1 = \binom{r}{r}$$

for any $r \geq 0$, whence the lemma is true for $n = 1$.

Now suppose that $r \geq 1$, $n \geq 2$ are given, and assume that (4.1) holds for all pairs (n', r') with $n' \leq n$, $r' \leq r$, and $(n', r') \neq (n, r)$. We will prove (4.1) for the pair (n, r) . Let $N^*(n, r)$ denote the number of n -tuples of nonnegative integers i_1, \dots, i_n with $i_1 + \dots + i_n \leq r$. Then

$$\begin{aligned} N(n, r) &= N^*(n-1, r) = N^*(n-1, r-1) + N(n-1, r) = N(n, r-1) + N(n-1, r) \\ &= \binom{r + n - 2}{r-1} + \binom{r + n - 2}{r} = \binom{r + n - 1}{r}. \end{aligned}$$

LEMMA 4C. Suppose that r_1, \dots, r_m are positive integers and that
 $0 < \varepsilon < 1$. Further suppose that $n \geq 2$ is an integer. Then the number
of nm -tuples of nonnegative integers

$$i_{11}, \dots, i_{1n}$$

$$i_{21}, \dots, i_{2n}$$

...

$$i_{m1}, \dots, i_{mn}$$

with

$$(4.2) \quad \sum_{k=1}^n i_{hk} = r_h \quad (1 \leq h \leq m)$$

and with

$$\left| \left(\sum_{h=1}^m \frac{i_{h1}}{r_h} \right) - \frac{m}{n} \right| \geq \varepsilon m$$

is at most

$$\binom{r_1 + n - 1}{r_1} \cdots \binom{r_m + n - 1}{r_m} \cdot 2e^{-\varepsilon^2 m/4}.$$

Remarks. (a) Lemma 4C again may be given a probabilistic interpretation: the "random variables" $\frac{i_{hk}}{r_h}$ have expectation $\frac{1}{n}$, so $\sum \frac{i_{h1}}{r_h} \approx \frac{m}{n}$ with high probability by the Law of Large Numbers. (b) Lemma 4A follows from Lemma 4C on taking $n = 2$ and considering $2m$ -tuples

$$\begin{array}{c} i_1, r_1 - i_1 \\ \dots \dots \dots \\ i_m, r_m - i_m \end{array}.$$

Proof of Lemma 4C. Let M_+ denote the number of nm -tuples with (4.2) and with

$$\left(\sum_{h=1}^m \frac{i_{h1}}{r_h} \right) - \frac{m}{n} \geq \varepsilon m ;$$

and let M_- denote the number of nm -tuples with (4.2) and with

$$\left(\sum_{h=1}^m \frac{i_{h1}}{r_h} \right) - \frac{m}{n} \leq -\varepsilon m.$$

To prove the lemma, it clearly suffices to show that

$$(4.3) \quad M_{\pm} \leq \binom{r_1 + n - 1}{r_1} \dots \binom{r_m + n - 1}{r_m} e^{-\varepsilon^2 m/4}.$$

Let integers j and c_j be given, with $1 \leq j \leq m$ and $0 \leq c_j \leq r_j$. Then write $f_j(c_j)$ for the number of $(n-1)$ -tuples of nonnegative integers i_{j2}, \dots, i_{jn} with $i_{j2} + \dots + i_{jn} = r_j - c_j$. It is clear from the definitions of M_+ and M_- that

$$(4.4) \quad M_{\pm} = \sum_{\underline{c}} f_1(c_1) \dots f_m(c_m),$$

where the sum is over all $\underline{c} = (c_1, \dots, c_m)$ with $0 \leq c_j \leq r_j$ ($1 \leq j \leq m$) for which

$$\left(\sum_{h=1}^m \frac{c_h}{r_h} \right) - \frac{m}{n}$$

is $\geq \varepsilon m$ or $\leq -\varepsilon m$, respectively. It follows that

$$(4.5) \quad \begin{aligned} M_{\pm} e^{\varepsilon^2 m/2} &\leq \sum_{c_1=0}^{r_1} \dots \sum_{c_m=0}^{r_m} f_1(c_1) \dots f_m(c_m) \exp \left(\pm \frac{\varepsilon}{2} \left(\sum_{h=1}^m \frac{c_h}{r_h} - \frac{m}{n} \right) \right) \\ &= \prod_{j=1}^m \left(\sum_{c_j=0}^{r_j} f_j(c_j) \exp \left(\pm \frac{\varepsilon}{2} \left(\frac{c_j}{r_j} - \frac{1}{n} \right) \right) \right). \end{aligned}$$

For the moment, we keep j fixed.

It is clear from the definitions that

$$(4.6) \quad \sum_{c=0}^{r_j} f_j(c) = \sum_{c+i_2+\dots+i_n=r_j} 1 = N(n, r_j) = \binom{r_j + n - 1}{r_j}.$$

Recall that $e^x \leq 1 + x + x^2$ whenever $|x| \leq 1$. We obtain

$$(4.7) \quad \begin{aligned} \sum_{c=0}^{r_j} f_j(c) \exp \left(\pm \frac{\varepsilon}{2} \left(\frac{c}{r_j} - \frac{1}{n} \right) \right) &\leq \sum_{c=0}^{r_j} f_j(c) \left(1 \pm \frac{\varepsilon}{2} \left(\frac{c}{r_j} - \frac{1}{n} \right) + \frac{\varepsilon^2}{4} \left(\frac{c}{r_j} - \frac{1}{n} \right)^2 \right) \\ &\leq \sum_{c=0}^{r_j} f_j(c) \left(1 + \frac{\varepsilon^2}{4} \right) \pm \frac{\varepsilon}{2r_j} \left(\sum_{c=0}^{r_j} c f_j(c) - \frac{r_j}{n} \sum_{c=0}^{r_j} f_j(c) \right) \\ &= \binom{r_j + n - 1}{r_j} \left(1 + \frac{\varepsilon^2}{4} \right), \end{aligned}$$

the last equality following from (4.6) and from the fact that the coefficient of $\pm \frac{\varepsilon}{2r_j}$ in the next to the last line displayed above is 0: To show that the coefficient in question is indeed 0, we observe that

$$f_j(c) = \sum_{\substack{i_2, \dots, i_n \\ i_2 + \dots + i_n = r_j - c}} 1 = \sum_{\substack{i_1, \dots, i_n \\ i_1 + \dots + i_n = r_j \\ i_1 = c}} 1,$$

whence

$$cf_j(c) = \sum_{\substack{i_1, \dots, i_n \\ i_1 + \dots + i_n = r_j \\ i_1 = c}} i_1.$$

Then

$$\begin{aligned} \sum_{c=0}^{r_j} cf_j(c) &= \sum_{\substack{i_1, \dots, i_n \\ i_1 + \dots + i_n = r_j}} i_1 = \frac{1}{n} \sum_{\substack{i_1, \dots, i_n \\ i_1 + \dots + i_n = r_j}} (i_1 + \dots + i_n) \\ &= \frac{r_j}{n} \sum_{\substack{i_1, \dots, i_n \\ i_1 + \dots + i_n = r_j}} 1 = \frac{r_j}{n} \sum_{c=0}^{r_j} f_j(c), \end{aligned}$$

as claimed.

It follows from (4.5) and (4.7) that

$$M_{\pm} e^{\varepsilon^2 m/2} \leq \binom{r_1+n-1}{r_1} \dots \binom{r_m+n-1}{r_m} \left(1 + \frac{\varepsilon^2}{4}\right)^m \leq \binom{r_1+n-1}{r_1} \dots \binom{r_m+n-1}{r_m} e^{\varepsilon^2 m/4}.$$

This inequality is equivalent to (4.3), and the lemma is proved.

§5. Further Auxiliary Lemmas.

Beginning with this section, we consider polynomials in m variables with rational integer coefficients. We write

$$P(X_1, \dots, X_m) = \sum C(j_1, \dots, j_m) X_1^{j_1} \dots X_m^{j_m},$$

where the sum is over all m -tuples of nonnegative integers j_1, \dots, j_m ; all but finitely many of the coefficients $C(j_1, \dots, j_m)$ are 0. We define the height of P by

$$[P] = \max |C(j_1, \dots, j_m)|.$$

Finally, if i_1, \dots, i_m are nonnegative integers, we write

$$P_{i_1 \dots i_m} = \frac{1}{i_1! \dots i_m!} \frac{\partial^{i_1 + \dots + i_m}}{\partial X_1^{i_1} \dots \partial X_m^{i_m}} P.$$

It will often be convenient to write $P_{\underline{i}}$ instead of $P_{i_1 \dots i_m}$, where it is understood that $\underline{i} = (i_1, \dots, i_m)$.

LEMMA 5A. If P has rational integer coefficients, then so does

$P_{\underline{i}}$. Furthermore, if P has degree $\leq r_h$ in the variable X_h ($1 \leq h \leq m$), then

$$[P_{\underline{i}}] \leq 2^{r_1 + \dots + r_m} [P].$$

Proof. We may write

$$P(X_1, \dots, X_m) = \sum_{j_1=0}^{r_1} \dots \sum_{j_m=0}^{r_m} C(j_1, \dots, j_m) X_1^{j_1} \dots X_m^{j_m};$$

hence

$$(5.1) \quad P_{\underline{i}}(X_1, \dots, X_m) = \sum_{j_1=0}^{r_1} \dots \sum_{j_m=0}^{r_m} \binom{j_1}{i_1} \dots \binom{j_m}{i_m} C(j_1, \dots, j_m) X_1^{j_1-i_1} \dots X_m^{j_m-i_m}.$$

The new coefficients are again integers, since the binomial coefficients are integers. (We adopt the convention that $\binom{m}{n} = 0$ if $m < n$.) Since

$$\binom{j_k}{i_k} \leq 2^{j_k} \leq 2^{r_k} \quad (1 \leq k \leq m),$$

the second assertion of the lemma follows from (5.1) .

LEMMA 5B. (Siegel's Lemma). Let

$$L_j(\underline{z}) = \sum_{k=1}^N a_{jk} z_k \quad (1 \leq j \leq M)$$

be M linear forms with rational integer coefficients. Suppose that
 $N > M$ and that

$$|a_{jk}| \leq A, \quad (1 \leq j \leq M, 1 \leq k \leq N)$$

where A is a positive integer. Then there exists an integer point

$$\underline{z} = (z_1, \dots, z_N) \neq \underline{0} \text{ with}$$

$$(5.2) \quad L_j(\underline{z}) = 0 \quad (1 \leq j \leq M)$$

and with

$$(5.3) \quad \left[\underline{z} \right] \leq \left[(NA)^{\frac{M}{N-M}} \right] = Z, \quad \text{say.}$$

Proof. Since $N > M$, rational solutions of (5.2) with $\underline{z} \neq \underline{0}$ always exist. But if \underline{z} is a solution of (5.2), then so is $\lambda \underline{z}$ for any real λ , and therefore integer points $\underline{z} \neq \underline{0}$ exist which satisfy (5.2).

It remains to show that (5.3) can be satisfied in addition to (5.2). Our proof of this is much like the proof of Dirichlet's Theorem. First, we have $Z + 1 > (NA)^{\frac{M}{N-M}}$, whence $NA < (Z+1)^{\frac{N-M}{M}}$, and therefore

$$NAZ + 1 \leq NA(Z+1) < (Z+1)^{N/M}.$$

For every integer point $\underline{z} = (z_1, \dots, z_N)$ with

$$(5.4) \quad 0 \leq z_i \leq Z, \quad (1 \leq i \leq N)$$

we have

$$-B_j Z \leq L_j(\underline{z}) \leq C_j Z, \quad (1 \leq j \leq M)$$

where $-B_j$ and C_j are the sums of the negative and the positive coefficients of $L_j(\underline{z})$, respectively. Now $B_j + C_j \leq NA$, so that each $L_j(\underline{z})$ lies in an interval of length $\leq NAZ$. Therefore each $L_j(\underline{z})$ takes at most $NAZ + 1$ distinct values, and hence the M -tuple $L_1(\underline{z}), \dots, L_M(\underline{z})$ takes at most

$$(NAZ + 1)^M < (Z + 1)^N$$

values.

On the other hand, the number of possibilities for \underline{z} with (5.4) is $(Z + 1)^N$. It follows that there are N -tuples $\underline{z}^{(1)} \neq \underline{z}^{(2)}$ with (5.4) and with

$$L_j(\underline{z}^{(1)}) = L_j(\underline{z}^{(2)}) \quad (1 \leq j \leq M).$$

The integer point $\underline{z} = \underline{z}^{(1)} - \underline{z}^{(2)}$ satisfies the conditions of the lemma.

As is well known, an algebraic integer α satisfies an equation $\alpha^d + a_1 \alpha^{d-1} + \dots + a_d = 0$ with rational integer coefficients a_1, \dots, a_d . If α is just algebraic of degree d and satisfies $a_0 \alpha^d + \dots + a_d = 0$, then $\beta = a_0 \alpha$ is algebraic of degree d and has $\beta^d + a_1 \beta^{d-1} + a_2 a_0 \beta^{d-2} + \dots + a_d a_0^{d-1} = 0$, hence is an algebraic integer. Now if $|\alpha - (p/q)| < q^{-2-\delta}$, then $|\beta - (a_0 p/q)| < a_0 q^{-2-\delta}$. Hence if Roth's Theorem is true for β , then it is true for α , and it will suffice to prove the theorem for algebraic integers. We also recall the well known fact (Hardy & Wright (1954), §14.2) that if α is an algebraic integer, then already the defining polynomial of α has its leading coefficient equal to 1.

LEMMA 5C. Let α be an algebraic integer, with defining polynomial $Q(X) = X^d + a_1 X^{d-1} + \dots + a_{d-1} X + a_d$. Then for each integer $\ell \geq 0$, there are rational integers $a_1^{(\ell)}, \dots, a_d^{(\ell)}$ with

$$\alpha^\ell = a_1^{(\ell)} \alpha^{d-1} + \dots + a_{d-1}^{(\ell)} \alpha + a_d^{(\ell)}$$

and with

$$|a_i^{(\ell)}| \leq (\overline{Q} + 1)^\ell \quad (1 \leq i \leq d).$$

Proof. We proceed by induction on ℓ . The lemma is true trivially if $\ell < d$. Suppose, then, that the lemma is true for $\ell - 1$. We have

$$\begin{aligned} \alpha^\ell &= \alpha^{\ell-1} \cdot \alpha = \left(a_1^{(\ell-1)} \alpha^{d-1} + \dots + a_d^{(\ell-1)} \right) \alpha \\ &= a_1^{(\ell-1)} \alpha^d + a_2^{(\ell-1)} \alpha^{d-1} + \dots + a_d^{(\ell-1)} \alpha \\ &= a_1^{(\ell-1)} \left(-a_1 \alpha^{d-1} - \dots - a_{d-1} \alpha - a_d \right) + a_2^{(\ell-1)} \alpha^{d-1} + \dots + a_d^{(\ell-1)} \alpha \\ &= \left(a_2^{(\ell-1)} - a_1 a_1^{(\ell-1)} \right) \alpha^{d-1} + \dots + \left(a_d^{(\ell-1)} - a_{d-1} a_1^{(\ell-1)} \right) \alpha - a_d a_1^{(\ell-1)}, \end{aligned}$$

so

$$\alpha^\ell = a_1^{(\ell)} \alpha^{d-1} + \dots + a_{d-1}^{(\ell)} \alpha + a_d^{(\ell)}$$

in an obvious notation. For each i , $1 \leq i \leq d$, we have the estimate

$$|a_i^{(\ell)}| \leq (\overline{Q} + 1)^{\ell-1} + \overline{Q} (\overline{Q} + 1)^{\ell-1} = (\overline{Q} + 1)^\ell.$$

§6. The Index of a Polynomial.

Let $P(X_1, \dots, X_m)$ be a polynomial with rational integer coefficients, let r_1, \dots, r_m be positive integers, and let $(\alpha_1, \dots, \alpha_m)$ be an arbitrary point of \mathbb{R}^m .

Definition. Suppose at first that $P \neq 0$. The index of P with

respect to $(\alpha_1, \dots, \alpha_m; r_1, \dots, r_m)$ is the least value of

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

for which $P_{i_1 \dots i_m}(\alpha_1, \dots, \alpha_m)$ does not vanish. In particular, the index of P is 0 if $P(\alpha_1, \dots, \alpha_m) \neq 0$.

If $P = 0$, we define the index of P to be $+\infty$. In either case, we denote the index of P by $\text{Ind } P$.

LEMMA 6A. Let $(\alpha_1, \dots, \alpha_m)$ and r_1, \dots, r_m be given as above.

With respect to these parameters, we have:

- (i) $\text{Ind } P_{\underline{i}} \geq \text{Ind } P - \sum_{h=1}^m \frac{i_h}{r_h}$,
- (ii) $\text{Ind}(P^{(1)} + P^{(2)}) \geq \min(\text{Ind } P^{(1)}, \text{Ind } P^{(2)})$,
- (iii) $\text{Ind}(P^{(1)} P^{(2)}) = \text{Ind } P^{(1)} + \text{Ind } P^{(2)}$.

We remark that (ii) together with (iii) show that the index is a "valuation" of the ring of polynomials in m variables.

Proof. (i). Let $T = P_{\underline{i}}$, and suppose that $T_{\underline{j}}(\alpha_1, \dots, \alpha_m) \neq 0$.

Then $P_{\underline{i}+\underline{j}}(\alpha_1, \dots, \alpha_m) \neq 0$, whence

$$\frac{i_1+j_1}{r_1} + \dots + \frac{i_m+j_m}{r_m} \geq \text{Ind } P.$$

Thus

$$\frac{j_1}{r_1} + \dots + \frac{j_m}{r_m} \geq \text{Ind } P - \sum_{h=1}^m \frac{i_h}{r_h},$$

and therefore

$$\text{Ind } T \geq \text{Ind } P - \sum_{h=1}^m \frac{i_h}{r_h}.$$

(ii). Suppose that $(P^{(1)} + P^{(2)})_{\underline{j}}(\alpha_1, \dots, \alpha_m) \neq 0$. Then either $P_{\underline{j}}^{(1)}(\alpha_1, \dots, \alpha_m) \neq 0$ or $P_{\underline{j}}^{(2)}(\alpha_1, \dots, \alpha_m) \neq 0$, hence

$$\text{either } \frac{j_1}{r_1} + \dots + \frac{j_m}{r_m} \geq \text{Ind } P^{(1)} \quad \text{or} \quad \frac{j_1}{r_1} + \dots + \frac{j_m}{r_m} \geq \text{Ind } P^{(2)}.$$

Consequently,

$$\text{Ind}(P^{(1)} + P^{(2)}) \geq \min(\text{Ind } P^{(1)}, \text{Ind } P^{(2)}).$$

(iii). In an obvious notation, we have

$$(6.1) \quad (P^{(1)}P^{(2)})_{\underline{j}} = \sum_{\substack{\underline{i} + \underline{i}' = \underline{j} \\ \underline{i}, \underline{i}' \geq 0}} C(\underline{i}, \underline{i}') P_{\underline{i}}^{(1)} P_{\underline{i}'}^{(2)}$$

for any integer point $\underline{j} = (j_1, \dots, j_m)$ with $j_h \geq 0$ ($1 \leq h \leq m$). (In fact it may be shown that $C(\underline{i}, \underline{i}') = 1$; an obvious advantage of our $P_{\underline{i}}$ over partial derivatives.)

Suppose now that \underline{j} is chosen so that $\sum_{h=1}^m \frac{j_h}{r_h} = \text{Ind}(P^{(1)}P^{(2)})$ and $(P^{(1)}P^{(2)})_{\underline{j}}(\alpha_1, \dots, \alpha_m) \neq 0$. By (6.1), there exist \underline{i} and \underline{i}' with $\underline{i} + \underline{i}' = \underline{j}$ such that $P_{\underline{i}}^{(1)}(\alpha_1, \dots, \alpha_m) \neq 0$ and $P_{\underline{i}'}^{(2)}(\alpha_1, \dots, \alpha_m) \neq 0$.

Then $\sum_{h=1}^m \frac{i_h}{r_h} \geq \text{Ind } P^{(1)}$ and $\sum_{h=1}^m \frac{i'_h}{r_h} \geq \text{Ind } P^{(2)}$, whence

$$(6.2) \quad \text{Ind}(P^{(1)}P^{(2)}) = \sum_{h=1}^m \frac{j_h}{r_h} = \sum_{h=1}^m \frac{i_h}{r_h} + \sum_{h=1}^m \frac{i'_h}{r_h} \geq \text{Ind } P^{(1)} + \text{Ind } P^{(2)}.$$

Conversely, there exist m -tuples \underline{i} with $\sum_{h=1}^m \frac{i_h}{r_h} = \text{Ind } P^{(1)}$ and $P_{\underline{i}}^{(1)}(\alpha_1, \dots, \alpha_m) \neq 0$. Of these \underline{i} , assume that $\underline{i} = (\bar{i}_1, \dots, \bar{i}_m)$ is the first lexicographically. Similarly, let $\underline{i}' = (\bar{i}'_1, \dots, \bar{i}'_m)$ be the first m -tuple \underline{i}' lexicographically with $\sum_{h=1}^m \frac{i'_h}{r_h} = \text{Ind } P^{(2)}$ and with $P_{\underline{i}'}^{(2)}(\alpha_1, \dots, \alpha_m) \neq 0$. Putting $\underline{j} = \underline{i} + \underline{i}'$, we have

$$(P^{(1)}P^{(2)})_{\underline{j}}(\alpha_1, \dots, \alpha_m) = C(\underline{i}, \underline{i}') P_{\underline{i}}^{(1)}(\alpha_1, \dots, \alpha_m) P_{\underline{i}'}^{(2)}(\alpha_1, \dots, \alpha_m) \neq 0$$

by (6.1) . This establishes the inequality reverse to (6.2) .

§7. The Index Theorem.

THEOREM 7A. Suppose that α is an algebraic integer of degree d ,
 $d \geq 2$. Also suppose that $\varepsilon > 0$, and that m is an integer satisfying

$$(7.1) \quad m > 16\varepsilon^{-2} \log 4d .$$

Let r_1, \dots, r_m be positive integers.

Then there is a polynomial $P(X_1, \dots, X_m) \neq 0$ with rational integer coefficients such that

$$(i) \quad P \text{ has degree } \leq r_h \text{ in } X_h , \quad (1 \leq h \leq m)$$

$$(ii) \quad P \text{ has index } \geq \frac{m}{2}(1-\varepsilon) \text{ with respect to } (\alpha, \alpha, \dots, \alpha; r_1, \dots, r_m) ,$$

and

$$(iii) \quad |P| \leq B^{r_1 + \dots + r_m} , \text{ where } B = B(\alpha) .$$

Proof. We seek a polynomial

$$P(X_1, \dots, X_m) = \sum_{j_1=0}^{r_1} \dots \sum_{j_m=0}^{r_m} C(j_1, \dots, j_m) X_1^{j_1} \dots X_m^{j_m}$$

with rational integer coefficients $C(j_1, \dots, j_m)$ such that (ii) and (iii) hold. The coefficients are

$$N = (r_1+1) \dots (r_m+1)$$

integers to be determined. By (ii) , we need

$$(7.2) \quad P_1(\alpha, \alpha, \dots, \alpha) = 0$$

whenever

$$\left(\sum_{h=1}^m \frac{i_h}{r_h} \right) - \frac{m}{2} < -\frac{\varepsilon}{2^m} .$$

In view of Lemma 4A , the number of such m -tuples \underline{i} is at most

$(r_1+1)\dots(r_m+1)\cdot 2e^{-\varepsilon^2 m/16}$. It follows that the number of these conditions (7.2) is at most

$$N \cdot \frac{2}{4d} = \frac{N}{2d}$$

by (7.1) .

Each condition (7.2) is a linear equation in the coefficients $C(j_1, \dots, j_m)$. The coefficients of these equations will be rational integers times powers of α , hence will be algebraic. But each power of α is a linear combination of $1, \alpha, \dots, \alpha^{d-1}$ with rational integer coefficients. Hence each condition (7.2) follows from d linear relations in $C(j_1, \dots, j_m)$ with rational integer coefficients. Altogether, we obtain

$$M \leq d \cdot \frac{N}{2d} = \frac{N}{2}$$

linear equations for the $C(j_1, \dots, j_m)$ with rational integer coefficients.

Let A be the maximum of the absolute values of these rational integer coefficients. For each $C(j_1, \dots, j_m)$ in (7.2) , the coefficients in question have absolute value at most

$$\binom{j_1}{i_1} \dots \binom{j_m}{i_m} (\sqrt{Q} + 1)^\ell \leq 2^{j_1 + \dots + j_m} (\sqrt{Q} + 1)^\ell$$

by Lemma 5C: Here $Q(X)$ is the defining polynomial for α , and $\ell = (j_1 - i_1) + \dots + (j_m - i_m)$. Thus

$$A \leq (2(\sqrt{Q} + 1))^{r_1 + \dots + r_m} .$$

By Lemma 5B , our system of linear equations has a nontrivial integer solution with

$$|C(j_1, \dots, j_m)| \leq Z \leq (NA)^{\frac{M}{N-M}} \leq NA$$

$$\leq 2^{r_1 + \dots + r_m} (2(\sqrt{Q} + 1))^{r_1 + \dots + r_m} = B^{r_1 + \dots + r_m}$$

for each m -tuple (j_1, \dots, j_m) . The polynomial P with these coefficients $C(j_1, \dots, j_m)$ satisfies

$$|P| \leq B^{r_1 + \dots + r_m}$$

with $B = B(\alpha) = 4(\sqrt{Q} + 1)$.

The construction of the polynomial P in this section corresponds to part (a) in the proof of Liouville's Theorem. The next section will correspond to part (b).

§8. The index of $P(X_1, \dots, X_m)$ at rational points near $(\alpha, \alpha, \dots, \alpha)$.

Suppose that α is an algebraic integer of degree d , $d \geq 2$.

For any $\varepsilon > 0$, let $m = m(\alpha, \varepsilon)$ be an integer satisfying

$$m > 16\varepsilon^{-2} \log 4d.$$

Also, let r_1, \dots, r_m be positive integers and let P be a polynomial satisfying the conclusions of Theorem 7A.

THEOREM 8A. Suppose that $0 < \delta < 1$, and that

$$(8.1) \quad 0 < \varepsilon < \frac{\delta}{36}.$$

Let $\frac{p_1}{q_1}, \dots, \frac{p_m}{q_m}$ be rational approximations to α , with

$$(8.2) \quad \left| \alpha - \frac{p_h}{q_h} \right| < q_h^{-2-\delta} \quad (1 \leq h \leq m)$$

and with

$$(8.3) \quad q_h^\delta > D, \quad (1 \leq h \leq m)$$

where $D = D(\alpha) > 0$. Also, suppose that

$$(8.4) \quad r_1 \log q_1 \leq r_h \log q_h \leq (1+\varepsilon) r_1 \log q_1 \quad (1 \leq h \leq m) .$$

Then the index of P with respect to $\left(\frac{p_1}{q_1}, \dots, \frac{p_m}{q_m}; r_1, \dots, r_m\right)$ is $\geq \varepsilon m$.

Proof. Suppose that j_1, \dots, j_m are nonnegative integers with

$$\sum_{h=1}^m \frac{j_h}{r_h} < \varepsilon m .$$

Put $T(X_1, \dots, X_m) = P_{\underline{j}}(X_1, \dots, X_m)$, where $\underline{j} = (j_1, \dots, j_m)$. We have to

show that $T\left(\frac{p_1}{q_1}, \dots, \frac{p_m}{q_m}\right) = 0$.

By the Index Theorem (Theorem 7A),

$$|P| \leq B^{r_1 + \dots + r_m} ,$$

whence

$$|T| \leq (2B)^{r_1 + \dots + r_m} ,$$

from Lemma 5A . Applying Lemma 5A again, we have

$$|T_{\underline{i}}| \leq (4B)^{r_1 + \dots + r_m}$$

for any m -tuple $\underline{i} = (i_1, \dots, i_m)$ of nonnegative integers. Hence in

$T_{\underline{i}}(\alpha, \alpha, \dots, \alpha)$, each monomial has absolute value

$$\leq (4B)^{r_1 + \dots + r_m} (\max(1, |\alpha|))^{r_1 + \dots + r_m} .$$

Since $T_{\underline{i}}$ is a sum of at most

$$(r_1+1) \dots (r_m+1) \leq 2^{r_1 + \dots + r_m}$$

monomials, we get

$$(8.5) \quad |T_{\frac{1}{\alpha}}(\alpha, \alpha, \dots, \alpha)| \leq (8B \max(1, |\alpha|))^{r_1 + \dots + r_m} = C^{r_1 + \dots + r_m},$$

where $C = C(\alpha)$.

By the Index Theorem, the index of P with respect to $(\alpha, \alpha, \dots, \alpha; r_1, \dots, r_m)$ is $\geq \frac{m}{2}(1-\varepsilon)$. It follows from part (i) of Lemma 6A that the index of T with respect to $(\alpha, \alpha, \dots, \alpha; r_1, \dots, r_m)$ is

$$\geq \frac{m}{2}(1-\varepsilon) - \sum_{h=1}^m \frac{j_h}{r_h} > \frac{m}{2}(1-3\varepsilon).$$

By Taylor's formula,

$$T\left(\frac{p_1}{q_1}, \dots, \frac{p_m}{q_m}\right) = \sum_{i_1=0}^{r_1} \dots \sum_{i_m=0}^{r_m} T_{i_1 \dots i_m}(\alpha, \alpha, \dots, \alpha) \cdot \left(\frac{p_1}{q_1} - \alpha\right)^{i_1} \dots \left(\frac{p_m}{q_m} - \alpha\right)^{i_m}.$$

From the paragraph above, the summands vanish unless $\sum_{h=1}^m \frac{i_h}{r_h} > \frac{m}{2}(1-3\varepsilon)$.

In view of (8.2) and (8.5), we obtain

$$(8.6) \quad \left|T\left(\frac{p_1}{q_1}, \dots, \frac{p_m}{q_m}\right)\right| \leq \sum'_{\frac{1}{\alpha}} C^{r_1 + \dots + r_m} \left(\frac{i_1}{q_1} \frac{i_2}{q_2} \dots \frac{i_m}{q_m}\right)^{-2-\delta},$$

where Σ' indicates that the sum is to be taken over all m -tuples of integers i_h in $0 \leq i_h \leq r_h$ for which $\sum_{h=1}^m \frac{i_h}{r_h} > \frac{m}{2}(1-2\varepsilon)$.

For such m -tuples, we have in view of (8.4), (8.1),

$$\begin{aligned} \frac{i_1}{q_1} \frac{i_2}{q_2} \dots \frac{i_m}{q_m} &= \frac{r_1}{q_1} \frac{i_1}{r_1} \frac{r_2}{q_2} \frac{i_2}{r_2} \dots \frac{r_m}{q_m} \frac{i_m}{r_m} \geq \frac{r_1}{q_1} \left(\frac{i_1}{r_1} + \dots + \frac{i_m}{r_m}\right) \\ &> \frac{r_1}{q_1} \frac{m}{2} \left(\frac{1}{2} - 2\varepsilon\right) \geq \left(\frac{r_1}{q_1} \frac{r_2}{q_2} \dots \frac{r_m}{q_m}\right)^{\frac{1}{2} - 2\varepsilon} / (1+\varepsilon) \\ &> \left(\frac{r_1}{q_1} \frac{r_2}{q_2} \dots \frac{r_m}{q_m}\right)^{\frac{1}{2}(1-6\varepsilon)}. \end{aligned}$$

The number of summands in (8.6) is $\leq 2^{r_1 + \dots + r_m}$, so that

$$\left|T\left(\frac{p_1}{q_1}, \dots, \frac{p_m}{q_m}\right)\right| \leq \prod_{h=1}^m \left(2C \frac{r_h}{q_h}\right)^{\frac{1}{2}(1-6\varepsilon)(2+\delta)}.$$

Now

$$\frac{1}{2}(1-6\varepsilon)(2+\delta) > 1 + \frac{\delta}{2} - 9\varepsilon \quad (\text{since } \delta < 1)$$

$$> 1 + \frac{\delta}{4}$$

by (8.1) ; therefore,

$$\frac{1}{2Cq_h} \frac{-\frac{1}{2}(1-6\varepsilon)(2+\delta)}{2Cq_h} < 2Cq_h^{-1-\frac{\delta}{4}} < q_h^{-1}$$

if $q_h^\delta > (2C)^4$. This is true by (8.3) if we put $D = (2C)^4$, say.

It follows that

$$\left| T\left(\frac{p_1}{q_1}, \dots, \frac{p_m}{q_m}\right) \right| < \frac{1}{q_1^{r_1} q_2^{r_2} \dots q_m^{r_m}}.$$

Recall that P , hence T , has degree $\leq r_h$ in X_h ($1 \leq h \leq m$).

Thus

$$T\left(\frac{p_1}{q_1}, \dots, \frac{p_m}{q_m}\right) = \frac{N}{q_1^{r_1} q_2^{r_2} \dots q_m^{r_m}}$$

for some integer N . By the inequality above, N is necessarily 0, whence

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

Part (c) in the proof of Liouville's Theorem was trivial. But in the present context this part is most difficult, and will be resolved in Roth's Lemma in §10.

§9. Generalized Wronskians.

Suppose that $\varphi_1, \dots, \varphi_k$ are rational functions in m variables X_1, \dots, X_m . We consider differential operators

$$\Delta = \frac{\partial^{i_1 + \dots + i_m}}{\partial X_1^{i_1} \dots \partial X_m^{i_m}}.$$

The order of such a differential operator is $i_1 + \dots + i_m$.

Definition. A generalized Wronskian of $\varphi_1, \dots, \varphi_k$ is any determinant of the form

$$(9.1) \quad \det(\Delta_i \varphi_j) \quad (1 \leq i, j \leq k),$$

where $\Delta_1, \dots, \Delta_k$ are operators as above, with Δ_i of order $\leq i-1$ ($1 \leq i \leq k$).

Remark. Take $m = 1$. Then Δ_1 is the identity operator, Δ_2 is the identity operator or $\frac{\partial}{\partial x}$, Δ_3 is the identity operator or $\frac{\partial}{\partial x}$ or $\frac{\partial^2}{\partial x^2}$, etc. A necessary condition for the generalized Wronskian (9.1) not to vanish identically is that $\Delta_1 =$ identity operator, $\Delta_2 = \frac{\partial}{\partial x}$, $\Delta_3 = \frac{\partial^2}{\partial x^2}, \dots$. In this case, (9.1) becomes

$$\det \begin{pmatrix} \varphi_1 & \varphi_2 & \dots & \varphi_k \\ \varphi_1' & \varphi_2' & \dots & \varphi_k' \\ \vdots & \vdots & & \vdots \\ \varphi_1^{(k-1)} & \varphi_2^{(k-1)} & \dots & \varphi_k^{(k-1)} \end{pmatrix},$$

which is the (ordinary) Wronskian of $\varphi_1, \dots, \varphi_k$.

LEMMA 9A. Suppose that $\varphi_1, \dots, \varphi_k$ are rational functions in x_1, \dots, x_m with real coefficients, and linearly independent over the reals. Then at least one generalized Wronskian of $\varphi_1, \dots, \varphi_k$ is not identically zero.

Remark. The converse of Lemma 9A is also true: namely, if $\varphi_1, \dots, \varphi_k$ are linearly dependent over the reals, then all generalized Wronskians of

$\varphi_1, \dots, \varphi_k$ vanish.

Proof. We use induction on k . If $k = 1$, then Δ_1 is necessarily the identity operator, and the generalized Wronskian is φ_1 . But φ_1 is linearly independent over the reals, so $\varphi_1 \neq 0$.

Suppose now that $\varphi_1, \dots, \varphi_k$ are k (≥ 2) rational functions satisfying the hypotheses of the lemma. Let Ω be any rational function in X_1, \dots, X_m with real coefficients, $\Omega \neq 0$. Consider the functions

$$\varphi_i^* = \Omega \varphi_i \quad (1 \leq i \leq k).$$

Then $\varphi_1^*, \dots, \varphi_k^*$ are again linearly independent over the reals. Any generalized Wronskian of $\varphi_1^*, \dots, \varphi_k^*$ is a linear combination of general Wronskians of $\varphi_1, \dots, \varphi_k$. (The coefficients in this linear combination are rational functions involving the partial derivatives of Ω .) To prove the lemma, it therefore will suffice to show that some generalized Wronskian of $\varphi_1^*, \dots, \varphi_k^*$ does not vanish trivially. If we put $\Omega = \varphi_1^{-1}$, then $\varphi_1^* = 1$, $\varphi_2^* = \frac{\varphi_2}{\varphi_1}, \dots, \varphi_k^* = \frac{\varphi_k}{\varphi_1}$. This argument shows that it is no restriction to assume that, in our list of given rational functions $\varphi_1, \dots, \varphi_k$, the function φ_1 is identically 1.

The set of all linear combinations

$$c_1 \varphi_1 + \dots + c_k \varphi_k$$

with real coefficients c_1, \dots, c_k forms a real vector space V of dimension k . Since $k > 1$, and since $\varphi_1 = 1$ and φ_2 are linearly independent, φ_2 is not a constant. Thus $\frac{\partial \varphi_2}{\partial X_j} \neq 0$ for some j . Without loss of generality, we suppose that $\frac{\partial \varphi_2}{\partial X_1} \neq 0$. Let W be the subspace of V consisting of all elements $c_1 \varphi_1 + \dots + c_k \varphi_k$ with

$$\frac{\partial}{\partial X_1} (c_1 \varphi_1 + \dots + c_k \varphi_k) = 0 \quad .$$

W is not the zero subspace, since $\varphi_1 \in W$. Also, $W \neq V$, since $\varphi_2 \notin W$. Accordingly, if we put $t = \dim W$, then $1 \leq t \leq k-1$.

We choose rational functions ψ_1, \dots, ψ_k such that ψ_1, \dots, ψ_t is a basis of W and ψ_1, \dots, ψ_k is a basis for V . By the induction hypothesis, there are operators $\Delta_1^*, \dots, \Delta_t^*$ of orders $\leq 0, 1, \dots, t-1$, respectively, with

$$W_1 = \det(\Delta_i^* \psi_j) \neq 0 \quad (1 \leq i, j \leq t) \quad .$$

If c_{t+1}, \dots, c_k are real numbers, not all of which are 0, then

$$\frac{\partial}{\partial X_1} (c_{t+1} \psi_{t+1} + \dots + c_k \psi_k) \neq 0 \quad :$$

This is true because the subspace spanned by $\psi_{t+1}, \dots, \psi_k$ has intersection $\underline{0}$ with W . In other words, the rational functions

$$\frac{\partial}{\partial X_1} \psi_{t+1}, \dots, \frac{\partial}{\partial X_1} \psi_k$$

are linearly independent over the reals. By induction, there are operators $\Delta_{t+1}^*, \dots, \Delta_k^*$ of orders $\leq 0, 1, \dots, k-t-1$, respectively, with

$$W_2 = \det \left(\Delta_i^* \frac{\partial}{\partial X_1} \psi_j \right) \neq 0 \quad (t < i, j \leq k) \quad .$$

We define operators Δ_i ($1 \leq i \leq k$) as follows:

$$\Delta_i = \begin{cases} \Delta_i^* & , \text{ if } 1 \leq i \leq t \\ \Delta_i^* \frac{\partial}{\partial X_1} & , \text{ if } t < i \leq k \end{cases} \quad .$$

Note that each Δ_i has order $\leq i-1$. We have

$$\det(\Delta_i \psi_j) = \det \begin{pmatrix} \Delta_i^* \psi_j & \Delta_i^* \psi_j \\ 0 & \Delta_i^* \frac{\partial}{\partial X_1} \psi_j \end{pmatrix} \quad \begin{matrix} 1 \leq j \leq t & t < j \leq k \\ 1 \leq i \leq t & t < i \leq k \end{matrix}$$

$$= W_1 W_2 \neq 0 \quad .$$

Because ψ_1, \dots, ψ_k is a basis of the vector space spanned by $\varphi_1, \dots, \varphi_k$, it follows that

$$\det(\Delta_i \varphi_j) \neq 0 \quad .$$

This completes the proof.

§10. Roth's Lemma.

THEOREM 10A. (Roth (1955a)). Suppose that

$$(10.1) \quad 0 < \varepsilon < \frac{1}{12} \quad .$$

Let m be a fixed positive integer. Put

$$(10.2) \quad \omega = \omega(m, \varepsilon) = 24 \cdot 2^{-m} \left(\frac{\varepsilon}{12} \right)^{2^{m-1}} \quad .$$

Let r_1, \dots, r_m be positive integers with

$$(10.3) \quad \omega r_h \geq r_{h+1} \quad (1 \leq h < m) \quad .$$

Suppose that $0 < \gamma \leq 1$, and let $(p_1, q_1), \dots, (p_m, q_m)$ be pairs of coprime integers with $q_h > 0$ ($1 \leq h \leq m$) and with

$$(10.4) \quad q_h^{r_h} \geq q_1^{\gamma r_1} \quad (1 \leq h \leq m) \quad ,$$

$$(10.5) \quad q_h^{\omega \gamma} \geq 2^{3m} \quad (1 \leq h \leq m) \quad .$$

Further, suppose that $P(X_1, \dots, X_m) \neq 0$ is a polynomial of degree $\leq r_h$

in X_h ($1 \leq h \leq m$) with rational integer coefficients and with

$$(10.6) \quad \overline{P} \leq q_1^{\omega \gamma r_1}.$$

Then the index of P with respect to $\left(\frac{p_1}{q_1}, \dots, \frac{p_m}{q_m}; r_1, \dots, r_m\right)$ is $\leq \varepsilon$.

Remark This theorem is usually stated with $\gamma = 1$. The presence of the parameter γ is not necessary for the proof of Roth's Theorem, but will be useful in later generalizations.

Proof. We use induction on m .

$m = 1$. We may write

$$P(X) = \left(X - \frac{p_1}{q_1}\right)^l M(X),$$

where $M(X)$ is a polynomial with rational coefficients having $M\left(\frac{p_1}{q_1}\right) \neq 0$.

Thus

$$(10.7) \quad P(X) = (q_1 X - p_1)^l R(X),$$

where $R(X) = q_1^{-l} M(X)$. It is a consequence of Gauss' Lemma that $R(X)$ has integer coefficients.

It now follows from (10.7) that the leading coefficient of $P(X)$ is divisible by q_1^l . Thus

$$\begin{aligned} q_1^l &\leq \overline{P} \leq q_1^{\omega r_1} && \text{(by (10.6))} \\ &= q_1^{\varepsilon r_1} && \text{(by (10.2) and since } m = 1), \end{aligned}$$

so that $\frac{l}{r_1} \leq \varepsilon$, since $q_1 > 1$ by (10.5). But $\frac{l}{r_1}$ is the index of P with respect to $\left(\frac{p_1}{q_1}; r_1\right)$, and the theorem is true if $m = 1$.

Inductive Step $m-1 \Rightarrow m$. Consider decompositions

$$(10.8) \quad P(X_1, \dots, X_m) = \sum_{j=1}^k \varphi_j(X_1, \dots, X_{m-1}) \psi_j(X_m) ,$$

where $\varphi_1, \dots, \varphi_k$ and ψ_1, \dots, ψ_k are polynomials with rational coefficients.

For example, take $k = r_m + 1$ and take ψ_1, \dots, ψ_k equal to

$1, X_m, X_m^2, \dots, X_m^{r_m}$, respectively. We now choose a decomposition where k is minimal. Then in particular,

$$k \leq r_m + 1 .$$

Since this decomposition is one in which k is minimal, the functions $\varphi_1, \dots, \varphi_k$ are linearly independent over the reals. For otherwise, there would exist real numbers c_1, \dots, c_k , not all 0, such that

$$c_1 \varphi_1 + \dots + c_k \varphi_k = 0 .$$

Since $\varphi_1, \dots, \varphi_k$ have rational coefficients, there would, in fact, exist rational numbers c_1, \dots, c_k with these properties. But then if, say, $c_k \neq 0$, we would have

$$P = \sum_{j=1}^{k-1} \varphi_j \left(\psi_j - \frac{c_j}{c_k} \psi_k \right) ,$$

contradicting the minimality of k . Similarly, ψ_1, \dots, ψ_k are linearly independent over the reals.

Let us write

$$U(X_m) = \det \left(\frac{1}{(i-1)!} \frac{\partial^{i-1}}{\partial X_m^{i-1}} \psi_j(X_m) \right) \quad (1 \leq i, j \leq k) .$$

By Lemma 9A and the remark preceding it,

$$U(X_m) \neq 0 .$$

Also by Lemma 9A, there exist operators

$$\Delta'_i = \frac{1}{i_1! \dots i_{m-1}!} \frac{\partial^{i_1 + \dots + i_{m-1}}}{\partial X_1^{i_1} \dots \partial X_{m-1}^{i_{m-1}}} \quad (1 \leq i \leq k)$$

with orders

$$(10.9) \quad i_1 + \dots + i_{m-1} \leq i-1 \leq k-1 \leq r_m$$

having

$$V(X_1, \dots, X_{m-1}) \stackrel{\text{def}}{=} \det_{1 \leq i, j \leq k} (\Delta'_i \varphi_j) \neq 0.$$

Put

$$W(X_1, \dots, X_m) = \det \left(\frac{1}{(j-1)!} \frac{\partial^{j-1}}{\partial X_m^{j-1}} \Delta'_i \varphi_j \right) \quad (1 \leq i, j \leq k).$$

Then

$$\begin{aligned} W(X_1, \dots, X_m) &= \det \left(\sum_{r=1}^k (\Delta'_i \varphi_r) \left(\frac{1}{(j-1)!} \frac{\partial^{j-1}}{\partial X_m^{j-1}} \varphi_r \right) \right) \quad (\text{by (10.8)}) \\ &= V(X_1, \dots, X_{m-1}) U(X_m) \neq 0. \end{aligned}$$

The entries in the determinant defining W are of the type

$\varphi_{i_1 \dots i_{m-1} j-1}$, and hence these entries have rational integer coefficients.

Thus W is a polynomial with rational integer coefficients. Before proceeding further we need

LEMMA 10B. Let Θ be the index of W with respect to

$\left(\frac{p_1}{q_1}, \dots, \frac{p_m}{q_m}; r_1, \dots, r_m \right)$. Then

$$\Theta \leq \frac{ke^2}{6}.$$

Proof of the Lemma. The polynomials U, V do not necessarily have integer coefficients. But clearly there is a factorization

$$W(X_1, \dots, X_m) = V^*(X_1, \dots, X_{m-1})U^*(X_m),$$

where U^*, V^* have rational integer coefficients.

Next, we obtain estimates for the heights of U^* and V^* . First, we have

$$\begin{aligned} \overline{P_{i_1 \dots i_{m-1} j-1}} &\leq 2^{r_1 + \dots + r_m} \overline{P} && \text{(by Lemma 5a)} \\ &\leq 2^{r_1 + \dots + r_m} \frac{\omega \gamma r_1}{q_1} && \text{(by (10.6)).} \end{aligned}$$

Furthermore, the number of terms in $P_{i_1 \dots i_{m-1} j-1}$ is $\leq 2^{r_1 + \dots + r_m}$, and the number of summands in the determinant expansion for W is $k! \leq k^{k-1} \leq k^m \leq 2^{kr_m}$. It follows that

$$\overline{W} \leq 2^{kr_m} \left(2^{r_1 + \dots + r_m} \frac{\omega \gamma r_1}{q_1} \right)^k \leq \left(2^{3mr_1} \frac{\omega \gamma r_1}{q_1} \right)^k,$$

since $r_1 \geq r_2 \geq \dots \geq r_m$ by (10.2) and (10.3). From (10.5) we obtain

$$\overline{W} \leq \left(\frac{2\omega \gamma r_1}{q_1} \right)^k = \frac{2\omega \gamma r_1^k}{q_1^k}.$$

This yields the estimates

$$(10.10) \quad \overline{U^*} \leq \frac{2\omega \gamma r_1^k}{q_1^k} \leq \frac{2\omega r_m^k}{q_m^k}, \quad \overline{V^*} \leq \frac{2\omega \gamma r_1^k}{q_1^k}.$$

We now apply the inductive hypothesis. More precisely, we are going to apply Theorem 10A with $m-1$ for m , with kr_1, \dots, kr_{m-1} for r_1, \dots, r_m , with $\frac{\varepsilon^2}{12}$ for ε , and with $V^*(X_1, \dots, X_{m-1})$ for $P(X_1, \dots, X_m)$. Namely, (10.3) and (10.5) are true with $\omega = \omega(m, \varepsilon)$ by our assumption, hence are true a fortiori with $\omega(m-1, \frac{\varepsilon^2}{12}) = 2\omega(m, \varepsilon)$. It is clear that (10.4) holds and that (10.1) holds with $\frac{\varepsilon^2}{12}$ for ε . The analogue of (10.6) holds by (10.10), since

$$|V^*| \leq q_1^{\omega\left(m-1, \frac{\varepsilon^2}{12}\right)} \gamma(kr_1)$$

Hence the index of V^* with respect to $\left(\frac{p_1}{q_1}, \dots, \frac{p_{m-1}}{q_{m-1}}; kr_1, \dots, kr_{m-1}\right)$ is $\leq \frac{\varepsilon^2}{12}$, and therefore the index of V^* with respect to

$\left(\frac{p_1}{q_1}, \dots, \frac{p_{m-1}}{q_{m-1}}; r_1, \dots, r_{m-1}\right)$ is $\leq \frac{k\varepsilon^2}{12}$. If we now consider V^* as a polynomial in X_1, \dots, X_m , then V^* still has index $\leq \frac{k\varepsilon^2}{12}$ with respect to $\left(\frac{p_1}{q_1}, \dots, \frac{p_m}{q_m}; r_1, \dots, r_m\right)$.

It is easily checked that the hypotheses of Roth's Lemma are satisfied with $\gamma = 1$, with $m = 1$, with kr_m for r_1, \dots, r_m , with $\frac{\varepsilon^2}{12}$ for ε , and with $U^* = U^*(X_m)$ for $P(X_1, \dots, X_m)$. (Note that $\omega\left(1, \frac{\varepsilon^2}{12}\right) \geq 2\omega(m, \varepsilon)$.) Since the case $m = 1$ of Roth's Lemma has been established, U^* has index $\leq \frac{k\varepsilon^2}{12}$ with respect to $\left(\frac{p_1}{q_1}, \dots, \frac{p_m}{q_m}; r_1, \dots, r_m\right)$. Since $W = U^*V^*$, it follows from part (iii) of Lemma 6A that

$$\theta \leq \frac{k\varepsilon^2}{12} + \frac{k\varepsilon^2}{12} = \frac{k\varepsilon^2}{6}.$$

Lemma 10B is established.

Completion of the proof of Theorem 10A. Let θ denote the index of P with respect to $\left(\frac{p_1}{q_1}, \dots, \frac{p_m}{q_m}; r_1, \dots, r_m\right)$. Then

$$\begin{aligned} \text{Ind } P_{i_1 \dots i_{m-1} j-1} &\geq \theta - \frac{i_1}{r_1} - \dots - \frac{i_{m-1}}{r_{m-1}} - \frac{j-1}{r_m} && \text{(by Lemma 6A)} \\ &\geq \theta - \frac{i_1 + \dots + i_{m-1}}{r_{m-1}} - \frac{j-1}{r_m} && \text{(since } r_1 \geq \dots \geq r_{m-1}) \\ &\geq \theta - \frac{r_m}{r_{m-1}} - \frac{j-1}{r_m} && \text{(by (10.9))} \\ &\geq \theta - \omega - \frac{j-1}{r_m} && \text{(by (10.3))} \\ &\geq \theta - \frac{\varepsilon^2}{24} - \frac{j-1}{r_m} && \text{(by (10.2), since } m \geq 2). \end{aligned}$$

As noted earlier, each entry in column j in the determinant defining W is of the type $P_{i_1 \dots i_{m-1} j-1}$. Recall the identity

$$\text{Ind } P^{(1)} P^{(2)} = \text{Ind } P^{(1)} + \text{Ind } P^{(2)}$$

and the inequality $\text{Ind}(P^{(1)} + P^{(2)}) \geq \min(\text{Ind } P^{(1)}, \text{Ind } P^{(2)})$ from Lemma 6A. Since W is a sum of products of k elements, one from each column, we see that

$$\begin{aligned} \Theta = \text{Ind } W &\geq \sum_{j=1}^k \max \left(\theta - \frac{\varepsilon^2}{24} - \frac{j-1}{r_m}, 0 \right) \\ &\geq -\frac{k\varepsilon^2}{24} + \sum_{i=0}^{k-1} \max \left(\theta - \frac{i}{r_m}, 0 \right). \end{aligned}$$

Thus

$$(10.11) \quad \sum_{i=0}^{k-1} \max \left(\theta - \frac{i}{r_m}, 0 \right) \leq \Theta + \frac{k\varepsilon^2}{24} \leq \frac{k\varepsilon^2}{6} + \frac{k\varepsilon^2}{24} < \frac{k\varepsilon^2}{4}.$$

We consider two cases:

Case I. $\theta > \frac{k-1}{r_m}$. Then (10.11) becomes

$$\frac{1}{2}k \left(\theta + \theta - \frac{k-1}{r_m} \right) < k \frac{\varepsilon^2}{4},$$

which is equivalent to

$$\theta + \left(\theta - \frac{k-1}{r_m} \right) < \frac{\varepsilon^2}{2}.$$

But $\theta - \frac{k-1}{r_m} > 0$, so $\theta < \frac{\varepsilon^2}{2} < \varepsilon$.

Case II. $\theta \leq \frac{k-1}{r_m}$. Then (10.11) becomes

$$\sum_{i=0}^{[\theta r_m]} \left(\theta - \frac{i}{r_m} \right) < \frac{k\varepsilon^2}{4},$$

which yields

$$\frac{1}{2}\theta([\theta r_m]+1) < \frac{k\varepsilon^2}{4},$$

whence

$$\frac{1}{2}\theta^2 r_m < \frac{k\varepsilon^2}{4}.$$

But $k \leq r_m + 1 \leq 2r_m$, so $\frac{1}{2}\theta^2 r_m < \frac{1}{2}\varepsilon^2 r_m$, and therefore $\theta < \varepsilon$.

In either case, we see that the index θ of P with respect to $\left(\frac{p_1}{q_1}, \dots, \frac{p_m}{q_m}; r_1, \dots, r_m\right)$ is $< \varepsilon$. This completes the proof of Theorem 10A.

§11. Conclusion of the proof of Roth's Theorem.

As was pointed out in §5, we may restrict ourselves to algebraic integers. Suppose then that there exists a $\delta > 0$ such that

$$(11.1) \quad \left| \alpha - \frac{p}{q} \right| < q^{-2-\delta}$$

has infinitely many rational solutions $\frac{p}{q}$, where α is an algebraic integer of degree d , $d \geq 2$. We proceed as follows.

(i) Assume, without loss of generality, that $0 < \delta < 1$.

(ii) Choose ε with $0 < \varepsilon < \frac{\delta}{36}$. This is (8.1). It implies that $0 < \varepsilon < \frac{1}{12}$, which is (10.1).

(iii) Choose an integer m with $m > 16\varepsilon^{-2} \log 4d$. Thus (7.1) holds. Define $\omega = \omega(m, \varepsilon)$ by (10.2).

(iv) Let $\frac{p_1}{q_1}$ be a solution of (11.1) with $(p_1, q_1) = 1$, $q_1 > 0$, such that $q_1^\omega > B^m$, where $B = B(\alpha)$ is the quantity of Theorem 7A, and such that (8.3), (10.5) hold with $\gamma = 1$ and $h = 1$.

(v) Choose $\frac{p_2}{q_2}, \dots, \frac{p_m}{q_m}$ successively with (11.1), with $(p_h, q_h) = 1$, $q_h > 0$ ($2 \leq h \leq m$), and so that

$$\omega \log q_{h+1} \geq 2 \log q_h \quad (1 \leq h \leq m-1).$$

This implies that $q_1 < q_2 < \dots < q_m$, and hence (8.3) and (10.5) with $\gamma = 1$ hold for $h = 1, 2, \dots, m$.

(vi) Let r_1 be an integer so large that $\varepsilon r_1 \log q_1 \geq \log q_m$.

(vii) For $2 \leq h \leq m$, put

$$r_h = \left\lfloor \frac{r_1 \log q_1}{\log q_h} \right\rfloor + 1.$$

Then for $2 \leq h \leq m$ we have

$$r_1 \log q_1 < r_h \log q_h \quad (\text{by (vii)})$$

$$\leq r_1 \log q_1 + \log q_h \quad (\text{by (vii)})$$

$$\leq (1+\varepsilon)r_1 \log q_1 \quad (\text{by (vi)}).$$

This gives (8.4), (10.4) with $\gamma = 1$. From this sequence of inequalities, it follows that

$$r_{h+1} \log q_{h+1} \leq (1+\varepsilon)r_h \log q_h \quad (1 \leq h \leq m-1).$$

Consequently,

$$\omega r_h \geq \omega \frac{r_{h+1} \log q_{h+1}}{(1+\varepsilon) \log q_h} \quad (1 \leq h \leq m-1)$$

$$\geq \frac{2}{1+\varepsilon} r_{h+1}$$

by (v), whence

$$\omega r_h \geq r_{h+1} \quad (1 \leq h \leq m-1),$$

which is (10.3).

The conditions of Theorem 7A (Index Theorem) are satisfied, since

(7.1) holds. Let $P(X_1, \dots, X_m)$ be a polynomial satisfying the conclusions of the Index Theorem. The hypotheses of Theorem 8A, (i.e., (8.1), (8.2), (8.3), (8.4)) hold. Hence the index of P with respect to

$$\left(\frac{p_1}{q_1}, \dots, \frac{p_m}{q_m}; r_1, \dots, r_m \right) \text{ is}$$

$$(11.2) \qquad \qquad \qquad \geq \varepsilon m \quad .$$

On the other hand, the hypotheses of Theorem 10A, (i.e., (10.1), (10.2), (10.3), (10.4), (10.5), (10.6)) hold with $\gamma = 1$: Namely, (10.6) holds with $\gamma = 1$ because

$$\begin{aligned} |\overline{P}| &\leq B^{r_1 + \dots + r_m} && \text{(by Theorem 6A)} \\ &\leq B^{mr_1} \leq q_1^{\omega r_1} && \text{(by (iv))} . \end{aligned}$$

The conclusion of Theorem 10A is that the index of P with respect to

$$\left(\frac{p_1}{q_1}, \dots, \frac{p_m}{q_m}; r_1, \dots, r_m \right) \text{ is } \leq \varepsilon \quad . \text{ But this contradicts (11.2) .}$$

This is the desired contradiction, and Roth's Theorem follows.

VI. Simultaneous Approximation to Algebraic Numbers

References: Schmidt (1970, 1971b,c)

§1. Basic Results. The only easy result here is as follows.

Suppose[†]) $1, \alpha_1, \dots, \alpha_v$ are algebraic and linearly independent over the rationals. Let d be the degree of the number field generated by $\alpha_1, \dots, \alpha_v$ and let $1, \alpha_1, \dots, \alpha_v, \dots, \alpha_{d-1}$ be a basis of this field. We saw in Theorem 4A of Chapter II that $\alpha_1, \dots, \alpha_{d-1}$ are badly approximable, so that

$$|\alpha_1 q_1 + \dots + \alpha_{d-1} q_{d-1} - p| > c_1 q^{-d+1}$$

where q_1, \dots, q_{d-1}, p are rational integers and where $q = \max(|q_1|, \dots, |q_{d-1}|) \neq 0$.

Taking $q_{v+1} = \dots = q_{d-1} = 0$, we have

LEMMA 1A. Suppose $1, \alpha_1, \dots, \alpha_v$ are linearly independent over \mathbb{Q} , and they generate an algebraic number field of degree d . Then

$$|\alpha_1 q_1 + \dots + \alpha_v q_v - p| > c_1 q^{-d+1}$$

for arbitrary integers q_1, \dots, q_v, p having $q = \max(|q_1|, \dots, |q_v|) > 0$.

When $v = 1$ we obtain Liouville's result. The exponent is clearly best possible when $d = v + 1$, but will be improved in Corollary 1E below if $d > v + 1$. We now turn to more difficult "Roth-type" results:

THEOREM 1B. Suppose $\alpha_1, \dots, \alpha_u$ are real algebraic numbers such that $1, \alpha_1, \dots, \alpha_u$ are linearly independent over the rationals, and suppose $\delta > 0$. Then there are only finitely many positive integers q with

$$(1.1) \quad q^{1+\delta} \|\alpha_1 q\| \dots \|\alpha_u q\| < 1.$$

[†]) We ask the reader to forgive us for using u or v in this section instead of the more customary letters n or m , which will be needed later.

The inequalities

$$(1.2) \quad \left| \alpha_i - \frac{p_i}{q} \right| < q^{-1-(1/u)-\delta} \quad (i = 1, \dots, u)$$

imply that $\|\alpha_i q\| < q^{-(1/u)-\delta}$, and hence they imply (1.1). Therefore (1.2) has only finitely many solutions, and we obtain

COROLLARY 1C. Suppose $\alpha_1, \dots, \alpha_u$ and δ are as in Theorem 1B. Then there are only finitely many rational u -tuples $(\frac{p_1}{q}, \dots, \frac{p_u}{q})$ with (1.2).

THEOREM 1D. Assume that $1, \alpha_1, \dots, \alpha_v$ are real algebraic and linearly independent over the rationals, and suppose $\delta > 0$. Then there are only finitely many v -tuples of nonzero integers q_1, \dots, q_v with

$$(1.3) \quad |q_1 q_2 \cdots q_v|^{1+\delta} \|\alpha_1 q_1 + \dots + \alpha_v q_v\| < 1.$$

By applying Theorem 1D to all the non-empty subsets of $\alpha_1, \dots, \alpha_v$ one deduces

COROLLARY 1E. If $\alpha_1, \dots, \alpha_v$ are as in Theorem 1B, then there are only finitely many $(v+1)$ -tuples of integers q_1, \dots, q_v, p with $q = \max(|q_1|, \dots, |q_v|) > 0$ and with

$$(1.4) \quad |\alpha_1 q_1 + \dots + \alpha_v q_v - p| < q^{-v-\delta}.$$

The exponents in Corollaries 1C and 1E are essentially best possible by Dirichlet's theorems on simultaneous approximation (see Ch. II). In view of Khintchine's transference principle (see the end of §5 of Chapter IV), these two corollaries are equivalent[†], and they say that an u -tuple such as given in Theorem 1B is not very well approximable. The case $u = 1$ is of course Roth's Theorem. Theorems 1B and 1D were

[†]Also, Theorems 1B and 1D are equivalent. See Wang, Yu and Zhu (1979).

proved by Schmidt (1970), after he had derived weaker results in (1965b, 1967). Baker (1967), Feldman (1970) and Osgood (1970) proved weaker but effective versions of Corollary 1E for special algebraic numbers

$$\alpha_1, \dots, \alpha_v.$$

THEOREM 1F. (Subspace Theorem) (Schmidt (1972)). Suppose
 $L_1(\underline{x}), \dots, L_n(\underline{x})$ are linearly independent forms in $\underline{x} = (x_1, \dots, x_n)$ with
real or complex algebraic coefficients. Given $\delta > 0$, there are finitely
many proper subspaces T_1, \dots, T_w of \mathbb{R}^n such that every integer point
 $\underline{x} \neq \underline{0}$ with

$$(1.5) \quad |L_1(\underline{x}) \cdots L_n(\underline{x})| < |\underline{x}|^{-\delta}$$

lies in one of these subspaces.

The integer points in a subspace T span a rational linear subspace, i.e. one defined by linear equations with rational coefficients, and hence T_1, \dots, T_w may be taken to be rational subspaces.

We now proceed to derive Theorem 1B from the Subspace Theorem.

Given q with (1.1) choose p_1, \dots, p_u with $\|\alpha_i q\| = |\alpha_i q - p_i|$
 $(i = 1, \dots, u)$. Then with $n = u + 1$ and

$$\underline{x} = (x_1, \dots, x_n) = (p_1, \dots, p_u, q)$$

it is clear that $|\underline{x}| \ll q$. Introduce linear forms

$$(1.6) \quad \begin{aligned} L_i(\underline{X}) &= \alpha_i X_n - X_i & (1 \leq i \leq u), \\ L_n(\underline{X}) &= X_n. \end{aligned}$$

Then (1.1) implies that

$$(1.7) \quad |L_1(\underline{x}) \cdots L_n(\underline{x})| < |\underline{x}|^{-\delta/2},$$

at least when q is large. By the Subspace Theorem, the solutions to (1.7) lie in finitely many proper rational subspaces. If a typical such subspace T is defined by

$$(1.8) \quad c_1 x_1 + \dots + c_u x_u + c_n x_n = 0,$$

then for \underline{x} in T we have

$$c_1(\alpha_1 q - p_1) + \dots + c_u(\alpha_u q - p_u) = (c_1 \alpha_1 + \dots + c_u \alpha_u + c_n)q,$$

and thus

$$(1.9) \quad |c_1| \|\alpha_1 q\| + \dots + |c_u| \|\alpha_u q\| \geq \gamma q$$

where $\gamma = |c_1 \alpha_1 + \dots + c_u \alpha_u + c_n| > 0$ in view of our linear independence condition. But (1.9) shows that q is bounded.

Theorem 1D also is a consequence of the Subspace Theorem. Given q_1, \dots, q_v with (1.3) choose p with $\|\alpha_1 q_1 + \dots + \alpha_v q_v\| = |\alpha_1 q_1 + \dots + \alpha_v q_v - p|$. Put $n = v+1$ and write

$$\underline{x} = (x_1, \dots, x_n) = (q_1, \dots, q_v, p);$$

then $|\underline{x}| \ll q$. In view of (1.3) we have (1.7), at least when q is large, with linear forms

$$(1.10) \quad \begin{aligned} L_1(\underline{x}) &= X_i & (1 \leq i \leq v), \\ L_n(\underline{x}) &= \alpha_1 X_1 + \dots + \alpha_v X_v - X_n. \end{aligned}$$

In view of the Subspace Theorem, the solutions lie in a finite number of rational subspaces; let a typical such subspace be given by

$$c_1 x_1 + \dots + c_v x_v + c_n x_n = 0.$$

Now if $c_v \neq 0$, we note that for \underline{x} satisfying this equation,

$$\begin{aligned}
c_v \|\alpha_1 q_1 + \dots + \alpha_v q_v\| &= c_v |\alpha_1 q_1 + \dots + \alpha_v q_v - p| \\
&= |(c_v \alpha_1 - c_1 \alpha_v) q_1 + \dots + (c_v \alpha_{v-1} - c_{v-1} \alpha_v) q_{v-1} - (c_v + c_n \alpha_v) p| \\
&= |c_v + c_n \alpha_v| |\alpha'_1 q_1 + \dots + \alpha'_{v-1} q_{v-1} - p|
\end{aligned}$$

where $\alpha'_i = (c_v \alpha_i - c_i \alpha_v) / (c_v + c_n \alpha_v)$ ($i = 1, \dots, v$). For $v = 1$, we get $c_1 \|\alpha_1 q_1\| \gg |p| \geq 1$ if q is large, so that (1.3) has only finitely many solutions. For $v > 1$ we obtain

$$c_v \|\alpha_1 q_1 + \dots + \alpha_v q_v\| \gg \|\alpha'_1 q_1 + \dots + \alpha'_{v-1} q_{v-1}\|,$$

and hence

$$|q_1 q_2 \dots q_{v-1}|^{1+(\delta/2)} \|\alpha'_1 q_1 + \dots + \alpha'_{v-1} q_{v-1}\| < 1$$

from (1.3). Since $1, \alpha'_1, \dots, \alpha'_{v-1}$ are linearly independent over \mathbb{Q} , this has only finitely many solutions by induction.

The situation is the same if some $c_j \neq 0$ where $1 \leq j \leq v$, and it is similar if $c_1 = \dots = c_v = 0$ but $c_n \neq 0$. For another result on linear forms see §VII.11.

§2. Roth Systems.

This section may be omitted at first reading.

Suppose n, u, v are positive integers with

$$(2.1) \quad n = u + v.$$

Let $L_1(\underline{x}), \dots, L_u(\underline{x})$ be linear forms in $\underline{x} = (x_1, \dots, x_n)$ with real coefficients. The rank of these forms is the maximum number of linearly independent forms among them. Suppose at the moment that the rank is u . By reordering the variables we may assume that the n linear forms

x_1, \dots, x_v , L_1, \dots, L_u are linearly independent. By Minkowski's Theorem on Linear Forms (see Theorem 2C of Chapter II) there is for every $Q > 0$ an integer point $\underline{x} \neq 0$ with

$$|x_i| << Q^u \quad (1 \leq i \leq v) \quad \text{and} \quad |L_j(\underline{x})| \leq Q^{-v} \quad (1 \leq j \leq u) ,$$

and with the constant in $<<$ depending only on L_1, \dots, L_u . Writing

$$[\underline{x}] = \max(|x_1|, \dots, |x_n|) \quad \text{we obtain}$$

$$(2.2) \quad |\underline{x}| << Q^u \quad \text{and} \quad |L_j(\underline{x})| \leq Q^{-v} \quad (1 \leq j \leq u) ,$$

and hence

$$(2.3) \quad |L_j(\underline{x})| << [\underline{x}]^{-v/u} \quad (1 \leq j \leq u) .$$

Since Q in (2.2) may be taken arbitrarily large, the inequalities (2.3) have infinitely many solutions in integer points $\underline{x} \neq 0$. (The discerning reader will note that (2.3) is related to Corollary 1F in Chapter II. If there is no integer point $\underline{x}_0 \neq 0$ with $L_1(\underline{x}_0) = \dots = L_u(\underline{x}_0) = 0$, then we get infinitely many solutions to (2.3) in integer points with coprime coordinates. If there is such an \underline{x}_0 , then the multiples of \underline{x}_0 all satisfy (2.3) .)

It is easy to see that (2.2) and (2.3) still have solutions if L_1, \dots, L_u have rank less than u . More precisely, if L_1, \dots, L_u have rank r , then we may replace (2.3) by

$$(2.4) \quad |L_j(\underline{x})| << |\underline{x}|^{-(n-r)/r} \quad (1 \leq j \leq u) .$$

When $r < u$, this is stronger than (2.3) .

Suppose now that L_1, \dots, L_u are linear forms with real algebraic coefficients. We shall call L_1, \dots, L_u a Roth system if for every $\delta > 0$

the inequalities

$$(2.5) \quad |L_j(\underline{x})| < |\underline{x}|^{-(v/u)-\delta} \quad (1 \leq j \leq u)$$

have only finitely many solutions in integer points \underline{x} . Roth's theorem says precisely that for $u = v = 1$, $n = 2$, the single form

$L(\underline{x}) = \alpha X_1 - X_2$ is a Roth system if α is an algebraic irrationality.

If $v = 1$ and $L_i(\underline{x}) = \alpha_i X_n - X_i$ ($1 \leq i \leq u$), and if $1, \alpha_1, \dots, \alpha_u$ are linearly independent, then L_1, \dots, L_u is a Roth System by Corollary 1C.

If $u = 1$ and $L_1(\underline{x}) = \alpha_1 X_1 + \dots + \alpha_v X_v - X_n$ with linearly independent $1, \alpha_1, \dots, \alpha_v$, then L_1 is a Roth System by Corollary 1E.

We shall say that L_1, \dots, L_u have rank r on a subspace S of \mathbb{R}^n if the restrictions of L_1, \dots, L_u to S have rank r .

THEOREM 2A. Suppose L_1, \dots, L_u have real algebraic coefficients. Necessary and sufficient for L_1, \dots, L_u to be a Roth system is that on every rational subspace S^d of dimensions d with $1 \leq d \leq n$, the forms L_1, \dots, L_u have rank r satisfying

$$(2.6) \quad r \geq du/n.$$

Both Corollary 1C and 1E can be derived from this theorem as an exercise.

The necessity of the condition is easy to see: the integer points in S^d form a point lattice (see Corollary 8B of Ch. IV), and the restrictions of L_1, \dots, L_u to this point lattice behave like linear forms defined on the integer points of d -dimensional space. Hence by (2.4) there are infinitely many integer points $\underline{x} \neq \underline{0}$ in S^d having

$$(2.7) \quad |L_j(\underline{x})| \ll |\underline{x}|^{-(d-r)/r} = |\underline{x}|^{1-(d/r)} \quad (j = 1, \dots, u).$$

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

$$|L_j(\underline{x})| << |\underline{x}|^{1-(n/u)(1+\delta)} << |\underline{x}|^{-(v/u)-\delta} \quad (j = 1, \dots, u),$$

and L_1, \dots, L_u is not a Roth System.

We now turn to a generalization. Let L_1, \dots, L_n be linear forms in $\underline{x} = (X_1, \dots, X_n)$ with real coefficients. Let c_1, \dots, c_n be reals with

$$(2.8) \quad c_1 + \dots + c_n = 0.$$

By Minkowski's theorem on linear forms, there is for every $Q > 0$ an integer point $\underline{x} \neq 0$ with

$$(2.9) \quad |L_j(\underline{x})| << Q^{c_j} \quad (1 \leq j \leq n),$$

with the constant in $<<$ depending only on the determinant of the forms L_1, \dots, L_n .

We shall say that $(L_1, \dots, L_n; c_1, \dots, c_n)$ is a general Roth system if L_1, \dots, L_n have algebraic coefficients and if for every $\delta > 0$ there is a $Q_1 = Q_1(\delta; L_1, \dots, L_n; c_1, \dots, c_n)$ such that the inequalities

$$(2.10) \quad |L_j(\underline{x})| \leq Q^{c_j - \delta} \quad (1 \leq j \leq n)$$

have no integer solution $\underline{x} \neq 0$ if $Q > Q_1$. Again by Roth's Theorem, $(L_1 = X_1, L_2 = \alpha X_1 - X_2; c_1 = 1, c_2 = -1)$ is a general Roth system if α is an algebraic irrationality.

From now on we shall assume that

$$(2.11) \quad c_1 \leq c_2 \leq \dots \leq c_n.$$

Now let S^d be a subspace of R^n of dimension $d > 0$, and suppose that

L_1, \dots, L_n have rank r on S^d . If $r = d$, let t_1 be the smallest integer such that $L_{t_1} \neq 0$ on S^d , i.e., that L_{t_1} has rank 1 on S^d . Let t_2 be the smallest integer such that L_{t_1}, L_{t_2} have rank 2 on S^d ; and so on. In this way we obtain d integers t_1, t_2, \dots, t_d . Put

$$(2.12) \quad c(S^d) = c_{t_1} + \dots + c_{t_d}.$$

All this was under the assumption that $r = d$. When $r < d$, put $c(S^d) = +\infty$.

THEOREM 2B. Suppose L_1, \dots, L_n are linear forms with real algebraic coefficients, and c_1, \dots, c_n are constants subject to (2.8) and (2.11). Necessary and sufficient for $(L_1, \dots, L_n; c_1, \dots, c_n)$ to be a general Roth system is that

$$(2.13) \quad c(S^d) \leq 0$$

for every rational subspace $S^d \neq \underline{0}$.

The necessity of the condition (2.13) is easy to show. Observe again that the restrictions of L_1, \dots, L_n to the integer points of a rational subspace S^d behave like linear forms on the integer points of \mathbb{R}^d . Suppose that $c(S^d) > 0$ for some rational subspace $S^d \neq \underline{0}$. If L_1, \dots, L_n have rank $r < d$ on S^d , then for every $\varepsilon > 0$ there is an integer point $\underline{x} \neq \underline{0}$ on S^d with $|L_j(\underline{x})| < \varepsilon$ ($1 \leq j \leq n$), and hence for arbitrary $Q > 0$, $\delta > 0$ there is an $\underline{x} \neq \underline{0}$ with (2.10). Thus in this case we do not have a general Roth system. Assume, therefore, that L_1, \dots, L_n have rank $r = d$ on S^d . Let δ be the positive number with $2d\delta = c(S^d)$. By Minkowski's theorem on linear forms, applied

to the lattice of integer points on S^d , there is for every $Q > 0$ an integer point $\underline{x} \neq \underline{0}$ in S^d with

$$|L_{t_i}(\underline{x})| << Q^{c_{t_i}-2\delta} \quad (1 \leq i \leq d) .$$

By (2.11), by our choice of t_1, \dots, t_d , and since $\underline{x} \in S^d$, this implies that

$$|L_j(\underline{x})| << Q^{c_j-2\delta} \quad (1 \leq j \leq n) ,$$

and hence it implies (2.10) if Q is sufficiently large. Thus again $(L_1, \dots, L_n; c_1, \dots, c_n)$ is not a general Roth system.

We now shall derive Theorem 2A from Theorem 2B. By what we said above, it will suffice to prove the sufficiency of (2.6). This inequality, with $d = n$, implies that L_1, \dots, L_u are independent. By renaming the variables, if necessary, we may assume that the n forms L_1, \dots, L_u , X_1, \dots, X_v are independent. Let us take a look at the system

$$(2.14) \quad (L_1, \dots, L_u, X_1, \dots, X_v; c_1 = -v, \dots, c_u = -v, c_{u+1} = u, \dots, c_n = u) .$$

Since $L_1, \dots, L_u, X_1, \dots, X_v$ have rank n , they have rank d on every subspace of dimension d . Let S^d be a rational subspace of dimension d , and let r be the rank of L_1, \dots, L_u on S^d . Then $t_1 < t_2 < \dots < t_r \leq u < t_{r+1} < \dots < t_d$ and

$$\begin{aligned} c(S^d) &= c_{t_1} + \dots + c_{t_r} + c_{t_{r+1}} + \dots + c_{t_d} = r(-v) + (d-r)u \\ &= ud - nr . \end{aligned}$$

The condition (2.6) implies that $c(S^d) \leq 0$, and hence that (2.14) is a general Roth system by Theorem 2B. Thus for any $\delta > 0$, the inequalities

$$|L_j(\underline{x})| \leq Q^{-v-\delta} \quad (1 \leq j \leq u), \quad |x_i| \leq Q^{u-\delta} \quad (1 \leq i \leq v)$$

have no integer solution $\underline{x} \neq \underline{0}$ if Q is large. This implies that

L_1, \dots, L_u is a Roth system.

We end this section with a deduction of Theorem 2B from the Subspace Theorem. By what we said above, it will suffice to prove the sufficiency of (2.13).

Suppose we did not have a General Roth System. Then for some $\delta > 0$ and some arbitrarily large values of Q , the relations (2.10) have a nontrivial solution \underline{x} . Choose a subspace S which is minimal in the sense that (2.10) has for some arbitrarily large values of Q a non-zero integer solution $\underline{x} \in S$, but that for every proper subspace S' of S the relations (2.10) have no such solution in S' if $Q > Q_0(S')$.

Suppose $\dim S = d$ and write $S = S^d$. In view of (2.13), the forms L_1, \dots, L_n have rank d on S . Pick t_1, \dots, t_d as in the definition of $c(S^d)$. Then conditions (2.10) for large Q imply that $|\underline{x}| \leq Q^2$, hence that

$$(2.15) \quad |L_{t_1}(\underline{x}) \dots L_{t_d}(\underline{x})| \leq Q^{c(S^d)-d\delta} \leq Q^{-\delta} \leq |\underline{x}|^{-\delta/2}.$$

The integer points \underline{x} of S^d form a point lattice \bigwedge . (See §IV.8). There is a nonsingular linear transformation T from \mathbb{R}^d onto S^d which provides a 1-1 map from the lattice of integer points in \mathbb{R}^d onto \bigwedge . If $\underline{x} = T(\underline{y})$ and if $|\underline{x}|$ is large, then

$$(2.16) \quad |\underline{y}| \leq |\underline{x}|^2.$$

The forms

$$\tilde{L}_i(\underline{y}) = L_{t_i}(T(\underline{y})) \quad (i = 1, \dots, d)$$

are d linearly independent linear forms on \mathbb{R}^d with algebraic

coefficients. If \underline{x} is a nonzero integer point in S^d with (2.10) and if $\underline{x} = T(\underline{y})$, then (2.15), (2.16) yield

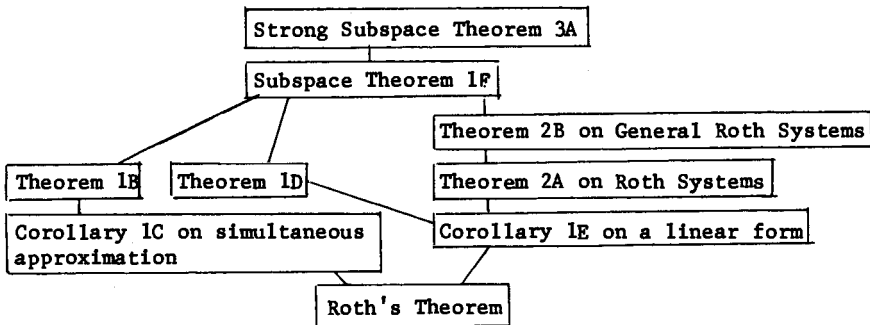
$$(2.17) \quad |\tilde{L}_1(\underline{y}) \dots \tilde{L}_d(\underline{y})| \leq |\underline{y}|^{-\delta/4}.$$

By the Subspace Theorem there are finitely many subspaces V_1, \dots, V_w of \mathbb{R}^d such that each solution \underline{y} of (2.17) lies in one of these subspaces. But then $\underline{x} = T(\underline{y})$ lies in one of the spaces $W_1 = T(V_1), \dots, W_w = T(V_w)$, which are proper subspaces of S^d . So for one of these subspaces W_j , the relations (2.10) have solutions \underline{x} in W_j for arbitrarily large values of Q . This contradicts the minimality of S^d .

Generalizations of the results of this section may be found in Vehmanen (1976).

§3. The Strong Subspace Theorem.

Everything in this chapter will follow from the Strong Subspace Theorem. The interdependence of results is then as follows.



Moreover, by our transference results, the two corollaries are equivalent, and according to Wang, Yu and Zhu (1979), Theorems 1B and 1D are equivalent.

So what is the Strong Subspace Theorem? Here it is:

THEOREM 3A. Let L_1, \dots, L_n be independent linear forms in $\underline{x} = (x_1, \dots, x_n)$ with real algebraic coefficients, and let c_1, \dots, c_n be constants with (2.8). For every $Q > 0$ the inequalities

$$(3.1) \quad |L_j(\underline{x})| \leq Q^{c_j} \quad (1 \leq j \leq n)$$

define a parallelepiped $\Pi = \Pi(Q)$. Denote the successive minima (in the sense of the Geometry of Numbers) of Π by $\lambda_1 = \lambda_1(Q), \dots, \lambda_n = \lambda_n(Q)$. Suppose there is a $\delta > 0$, a number d with $1 \leq d \leq n-1$ and an unbounded set \mathcal{N} of positive numbers such that

$$(3.2) \quad \lambda_d < \lambda_{d+1} Q^{-\delta}$$

for every Q in \mathcal{N} .

Then there is a fixed rational subspace S^d of dimension d and an unbounded subset \mathcal{N}' of \mathcal{N} such that for every Q in \mathcal{N}' the first d successive minima of $\Pi(Q)$ are assumed by points $\underline{g}_1, \dots, \underline{g}_d$ in S^d .

We now proceed to derive the Subspace Theorem.

LEMMA 3B. Let $L_1, \dots, L_n, c_1, \dots, c_n$ be as in Theorem 3A. Then there are finitely many proper rational subspaces T_1, \dots, T_w of \mathbb{R}^n such that every integer point $\underline{x} \neq \underline{0}$ with

$$(3.3) \quad |L_i(\underline{x})| \leq |\underline{x}|^{c_i - \delta} \quad (i = 1, \dots, n)$$

lies in one of these subspaces.

Proof. Suppose this were wrong. We could then construct a sequence $\underline{x}_1, \underline{x}_2, \dots$ of solutions of (3.3) with any n elements of this sequence

linearly independent. For every $Q > 0$ consider the parallelepiped Π defined by (3.1), choose independent integer points $\underline{g}_1, \dots, \underline{g}_n$ with $\underline{g}_i \in \lambda_i \Pi$ ($i = 1, \dots, n$), and let S_i be the subspace spanned by $\underline{g}_1, \dots, \underline{g}_i$. Now it is clear that $\lambda_1 \leq Q^{-\delta}$ if $Q = |\underline{x}_j|$ for some j , and hence (since $\lambda_1 \dots \lambda_n \gg 1$) $\lambda_n > 1$, at least when $Q = |\underline{x}_j|$ is large. Thus \underline{x}_j will certainly lie in S_{n-1} ; let k be the smallest integer such that \underline{x}_j lies in S_k . Then $\lambda_k \leq Q^{-\delta}$, and there is a d in $k \leq d \leq n-1$ with

$$(3.4) \quad \lambda_d < \lambda_{d+1} Q^{-\delta/n}$$

There is a subsequence $\underline{x}_{j_1}, \underline{x}_{j_2}, \dots$ and a fixed d such that (3.4) holds for $Q = |\underline{x}_{j_u}|$. Now if \mathfrak{N} is the set of values $|\underline{x}_{j_u}|$ ($u = 1, 2, \dots$), then by Theorem 3A there is an infinite subsequence whose absolute values make up a set \mathfrak{N}' such that for $Q \in \mathfrak{N}'$, the points $\underline{g}_1, \dots, \underline{g}_d$ lie in a fixed subspace S^d . Then $\underline{x}_{j_u} \in S_k \subset S_d = S^d$ for elements of this subsequence, contradicting the linear independence condition on our sequence.

Now as for Theorem 1F, the points \underline{x} having $L_j(\underline{x}) = 0$ for some j clearly lie in finitely many proper subspaces. We may therefore restrict our attention to points \underline{x} with $L_1(\underline{x}) \dots L_n(\underline{x}) \neq 0$.

Using norms, one sees that for these integer points one has

$$|\underline{x}|^{-A} \ll |L_j(\underline{x})| \ll |\underline{x}| \quad (j = 1, \dots, n) \text{ with some } A > 0, \text{ hence}$$

$$|\underline{x}|^{-2A} \leq |L_j(\underline{x})| < |\underline{x}|^2 \text{ if } |\underline{x}| \text{ is large. We divide the interval}$$

$$-2A \leq \xi < 2 \text{ into finitely many subintervals } c' \leq \xi < c \text{ of length less}$$

than $\delta/2n$. If $c'_1 \leq \xi < c''_1, \dots, c'_n \leq \xi < c''_n$ are any n such intervals, it will suffice to show that the solutions \underline{x} of (1.5) with

$$(3.5) \quad |\underline{x}|^{c'_j} \leq |L_j(\underline{x})| < |\underline{x}|^{c''_j} \quad (j = 1, \dots, n)$$

lie in finitely many subspaces. Now (1.5) is compatible with (3.5) only if $c'_1 + \dots + c'_n < -\delta$, so that $c''_1 + \dots + c''_n < -\delta/2$. Setting $c_j = c''_j - n^{-1}(c''_1 + \dots + c''_n)$, we have (2.8) and $c''_j < c_j - (\delta/2n)$, whence

$$|L_j(\underline{x})| < |\underline{x}|^{c_j - (\delta/2n)} \quad (j = 1, \dots, n),$$

and by Lemma 3B the solutions to this system lie in a finite number of subspaces.

We are finished with the proof of Theorem 1F, except that so far all our forms had real coefficients. To prove the general case, we use induction on the number of linear forms whose coefficients are not real.

Suppose, say, that $L_n(\underline{x})$ has some non-real coefficients, and that

$L_n(\underline{x}) = R(\underline{x}) + i I(\underline{x})$ where R, I have real coefficients. If

L_1, \dots, L_{n-1}, R are linearly independent, put $L'_n = R$; otherwise

L_1, \dots, L_{n-1}, I are linearly independent, and set $L'_n = I$. The forms

$L_1, \dots, L_{n-1}, L'_n$ again satisfy the conditions of Theorem 1F, and fewer of them are non-real. The truth of Theorem 1F is now evident.

The remainder of this chapter will be devoted to a proof of the Strong Subspace Theorem. We will imitate Ch. V as much as possible. The reader should observe that the basic ideas of the (perhaps somewhat intricate) proof are geometrical.

§4. The Index of a Polynomial.

Suppose that n and m are positive integers. Let \mathfrak{R} denote the ring of polynomials

$$P(X_{11}, \dots, X_{1n}; X_{21}, \dots, X_{2n}; \dots; X_{m1}, \dots, X_{mn})$$

with real coefficients. Let L_1, \dots, L_m be linear forms, none identically 0, where L_h is a linear form in the variables X_{h1}, \dots, X_{hn} : thus,

$$L_h = \alpha_{h1} X_{h1} + \dots + \alpha_{hn} X_{hn} \quad (1 \leq h \leq m),$$

where for every h at least one α_{hk} is different from 0. Let r_1, \dots, r_m be positive integers. For each $c \geq 0$, let $I(c)$ be the ideal in \mathfrak{R} generated by the polynomials $L_1^{i_1} L_2^{i_2} \dots L_m^{i_m}$ with

$$\frac{i_1}{r_1} + \frac{i_2}{r_2} + \dots + \frac{i_m}{r_m} \geq c.$$

Clearly, $I(0) = \mathfrak{R}$. Also, as c increases, $I(c)$ remains constant or decreases.

Definition. Suppose that $P \in \mathfrak{R}$, $P \neq 0$. The index of P with respect to $(L_1, \dots, L_m; r_1, \dots, r_m)$ is the largest number c such that $P \in I(c)$. (The existence of such a c is easily seen.) We define the index of the zero polynomial to be $+\infty$. We denote the index of P by $\text{Ind } P$.

LEMMA 4A. Suppose that $L_1, \dots, L_m; r_1, \dots, r_m$ are given as above.

With respect to these parameters, we have:

- (i) $\text{Ind}(P+Q) \geq \min(\text{Ind } P, \text{Ind } Q)$,
- (ii) $\text{Ind}(PQ) = \text{Ind } P + \text{Ind } Q$.

Proof. (i). $P \in I(\text{Ind } P)$ and $Q \in I(\text{Ind } Q)$, so $P, Q \in I(\min(\text{Ind } P, \text{Ind } Q))$. Then $P + Q \in I(\min(\text{Ind } P, \text{Ind } Q))$, so

$$\text{Ind}(P+Q) \geq \min(\text{Ind } P, \text{Ind } Q).$$

(ii). Put

$$L_h = \alpha_{h1} X_{h1} + \dots + \alpha_{hn} X_{hn} \quad (1 \leq h \leq m),$$

and assume without loss of generality that $\alpha_{h1} \neq 0$ for each h . Then every polynomial in X_{h1}, \dots, X_{hn} is a polynomial in $L_h, X_{h2}, \dots, X_{hn}$.

We may write (uniquely)

$$(4.1) \quad P = \sum c(j_1, a_{12}, \dots, a_{1n}; \dots; j_m, a_{m2}, \dots, a_{mn}) \cdot \\ L_1^{j_1} X_{12}^{a_{12}} \dots X_{1n}^{a_{1n}} \dots L_m^{j_m} X_{m2}^{a_{m2}} \dots X_{mn}^{a_{mn}} .$$

Put

$$\mu = \min \left(\frac{j_1}{r_1} + \frac{j_2}{r_2} + \dots + \frac{j_m}{r_m} \right) ,$$

where the minimum is taken over all m -tuples j_1, j_2, \dots, j_m of nonnegative integers such that $c(j_1, a_{12}, \dots) \neq 0$ for some choice of the a_{ij} -s .

We claim that $\mu = \text{Ind } P$. Namely, it is clear that $P \in I(\mu)$, whence that $\text{Ind } P \geq \mu$. But since $P \in I(\text{Ind } P)$, P is a linear combination (with polynomial coefficients) of $L_1^{j_1} L_2^{j_2} \dots L_m^{j_m}$ with

$$(4.2) \quad \frac{j_1}{r_1} + \frac{j_2}{r_2} + \dots + \frac{j_m}{r_m} \geq \text{Ind } P .$$

By the uniqueness of the representation (4.1) for P , each term in (4.1) with nonzero coefficient satisfies (4.2) . Hence $\mu \geq \text{Ind } P$, and the claim is true.

Now let P_1 and P_2 be the sum of those terms in (4.1) for which

$$\sum_{h=1}^m \frac{j_h}{r_h} = \text{Ind } P \quad \text{and} \quad \sum_{h=1}^m \frac{j_h}{r_h} > \text{Ind } P ,$$

respectively. Similarly define Q_1 and Q_2 . Then

$$PQ = P_1 Q_1 + P_1 Q_2 + P_2 Q_1 + P_2 Q_2 .$$

Here $P_1 Q_1$ is non-zero and has index $\text{Ind } P + \text{Ind } Q$, while each of $P_1 Q_2$, $P_2 Q_1$, $P_2 Q_2$ has an index $> \text{Ind } P + \text{Ind } Q$. It follows that $\text{Ind}(PQ) = \text{Ind } P + \text{Ind } Q$.

Remark. Let $P(X_1, \dots, X_m)$ be a polynomial of degree $\leq r_h$ in X_h ($1 \leq h \leq m$). Let

$$\hat{P}(X_{11}, X_{12}; \dots; X_{m1}, X_{m2})$$

denote the polynomial obtained from P by replacing $X_h^{i_h}$ by $X_{h1}^{i_h} X_{h2}^{r_h - i_h}$ ($0 \leq i_h \leq r_h$; $1 \leq h \leq m$). Then \hat{P} is homogeneous in X_{h1}, X_{h2} of degree r_h ($1 \leq h \leq m$).

Now let $(\alpha_1, \dots, \alpha_m) \in \mathbb{R}^m$. By Taylor's formula,

$$(4.3) \quad P(X_1, \dots, X_m) = \sum c(j_1, \dots, j_m) (X_1 - \alpha_1)^{j_1} \dots (X_m - \alpha_m)^{j_m},$$

where the sum is over all m -tuples of nonnegative integers j_1, \dots, j_m with $0 \leq j_h \leq r_h$ ($1 \leq h \leq m$). In the notation of Chapter V, §5,

$$c(j_1, \dots, j_m) = P_{j_1 \dots j_m}(\alpha_1, \dots, \alpha_m).$$

Therefore the Roth index of P with respect to $(\alpha_1, \dots, \alpha_m; r_1, \dots, r_m)$ is

$$\mu = \min \left(\frac{j_1}{r_1} + \dots + \frac{j_m}{r_m} \right),$$

where the minimum is taken over all m -tuples j_1, \dots, j_m for which $c(j_1, \dots, j_m) \neq 0$ in (4.3).

We have

$$\begin{aligned} & \hat{P}(X_{11}, X_{12}; \dots; X_{m1}, X_{m2}) \\ &= \sum c(j_1, \dots, j_m) X_{12}^{r_1 - j_1} \dots X_{m2}^{r_m - j_m} (X_{11} - \alpha_1 X_{12})^{j_1} \dots (X_{m1} - \alpha_m X_{m2})^{j_m} \\ &= \sum c(j_1, \dots, j_m) L_1^{j_1} X_{12}^{r_1 - j_1} L_2^{j_2} X_{22}^{r_2 - j_2} \dots L_m^{j_m} X_{m2}^{r_m - j_m}, \end{aligned}$$

where

$$L_h = X_{h1} - \alpha_h X_{h2} \quad (1 \leq h \leq m).$$

By the characterization of the index of a polynomial established in the proof of part (ii) of Lemma 4A, the index of \hat{P} with respect to $(L_1, \dots, L_m; r_1, \dots, r_m)$ also is μ . Thus, the Roth index of P with respect to $(\alpha_1, \dots, \alpha_m; r_1, \dots, r_m)$ equals the index of \hat{P} with respect to $(L_1, \dots, L_m; r_1, \dots, r_m)$.

We now introduce the following notation. If r_1, \dots, r_m is an m -tuple of positive integers, we write $\underline{r} = (r_1, \dots, r_m)$. \mathfrak{J} will denote an nm -tuple of nonnegative integers

$$(i_{11}, \dots, i_{1n}; i_{21}, \dots, i_{2n}; \dots; i_{m1}, \dots, i_{mn}).$$

Given \underline{r} and \mathfrak{J} , we put

$$(\mathfrak{J}/\underline{r}) = \sum_{h=1}^m \frac{i_{h1} + i_{h2} + \dots + i_{hn}}{r_h}.$$

Finally, if $P \in \mathfrak{R}$, we put

$$P^{\mathfrak{J}} = \frac{1}{i_{11}! \dots i_{mn}!} \frac{\partial^{i_{11} + \dots + i_{mn}}}{\partial x_{11}^{i_{11}} \dots \partial x_{mn}^{i_{mn}}} P.$$

LEMMA 4B. Suppose that L_1, \dots, L_m are linear forms as above and that $\underline{r} = (r_1, \dots, r_m)$ has positive integer components. Then for any $P \in \mathfrak{R}$ and any \mathfrak{J} ,

$$\text{Ind } P^{\mathfrak{J}} \geq \text{Ind } P - (\mathfrak{J}/\underline{r}).$$

Further, if $(\mathfrak{J}/\underline{r}) < \text{Ind } P$, then $P^{\mathfrak{J}} = 0$ on the $(nm-m)$ -dimensional subspace T of \mathbb{R}^{nm} defined by $L_1 = L_2 = \dots = L_m = 0$.

Proof. The first assertion follows from the representation of P in (4.1) and from the characterization of the index in terms of μ .

If $(\mathfrak{J}/r) < \text{Ind } P$, then $\text{Ind } P^{\mathfrak{J}} > 0$ from what has already been shown. Let us write

$$(4.4) \quad P^{\mathfrak{J}} = \sum c(j_1, a_{12}, \dots, a_{1n}; \dots; j_m, a_{m2}, \dots, a_{mn}) L_1^{j_1} X_{12}^{a_{12}} \dots X_{1n}^{a_{1n}} \dots L_m^{j_m} X_{m2}^{a_{m2}} \dots X_{mn}^{a_{mn}}.$$

It was shown in the proof of part (ii) of Lemma 4A that

$$\text{Ind } P^{\mathfrak{J}} = \mu^{\mathfrak{J}} = \min \left(\frac{j_1}{r_1} + \dots + \frac{j_m}{r_m} \right),$$

where the minimum is taken over all m -tuples j_1, \dots, j_m of nonnegative integers such that $c(j_1, a_{12}, \dots) \neq 0$ for some choice of the a_{ij} -s.

Since $\text{Ind } P^{\mathfrak{J}} > 0$, each m -tuple (j_1, \dots, j_m) for which $c(j_1, a_{12}, \dots) \neq 0$ in (4.4) is different from $\underline{0}$, and therefore each term in (4.4) is divisible by at least one of L_1, \dots, L_m . Hence $P^{\mathfrak{J}}$ vanishes on T .

LEMMA 4C. Suppose that L_1, \dots, L_m and r are as in Lemma 4B.

Let T be the subspace of \mathbb{R}^{mn} defined by $L_1 = \dots = L_m = 0$. Suppose that $P \in \mathfrak{R}$, $P \neq 0$. Then there is an \mathfrak{J} with $(\mathfrak{J}/r) = \text{Ind } P$ such that $P^{\mathfrak{J}} \neq 0$ on T . Moreover, if in

$$L_h = \alpha_{h1} X_{h1} + \dots + \alpha_{hn} X_{hn} \quad (1 \leq h \leq m)$$

we have $\alpha_{h1} \neq 0$ for each h , then there is such an \mathfrak{J} of the type

$$\mathfrak{J} = (i_1, 0, \dots, 0; i_2, 0, \dots, 0; \dots; i_m, 0, \dots, 0) \quad .$$

Proof. We may assume that $\alpha_{h1} \neq 0$ ($1 \leq h \leq m$) by relabeling if necessary. Then P may be written (uniquely) as

$$P = \sum P(j_1, \dots, j_m | X_{12}, \dots, X_{1n}; \dots; X_{m2}, \dots, X_{mn}) \cdot L_1^{j_1} L_2^{j_2} \dots L_m^{j_m},$$

where, for each fixed j_1, \dots, j_m , $P(j_1, \dots, j_m | X_{12}, \dots, X_{mn})$ is a polynomial in the $mn-m$ variables $X_{12}, \dots, X_{1n}; \dots; X_{m2}, \dots, X_{mn}$. From the

definition of the index of P with respect to $(L_1, \dots, L_m; r_1, \dots, r_m)$, we have that $P(j_1, \dots, j_m | X_{12}, \dots, X_{mn}) = 0$ if

$$\frac{j_1}{r_1} + \dots + \frac{j_m}{r_m} < \text{Ind } P,$$

but $P(\bar{j}_1, \dots, \bar{j}_m | X_{12}, \dots, X_{mn}) \neq 0$ for some $\bar{j}_1, \dots, \bar{j}_m$ with

$$\frac{\bar{j}_1}{r_1} + \dots + \frac{\bar{j}_m}{r_m} = \text{Ind } P.$$

It is clear that with

$$\mathfrak{J} = (\bar{j}_1, 0, \dots, 0; \bar{j}_2, 0, \dots, 0; \bar{j}_m, 0, \dots, 0),$$

we have $P^{\mathfrak{J}} \neq 0$ on T .

§5. Some Auxiliary Lemmas.

If $P \in \mathfrak{R}$, $P = \sum c(j_{11}, \dots, j_{mn}) X_{11}^{j_{11}} \dots X_{mn}^{j_{mn}}$, put

$$|P| = \max_{j_{11}, \dots, j_{mn}} (|c(j_{11}, \dots, j_{mn})|).$$

LEMMA 5A. Suppose P has rational integer coefficients. Then so does $P^{\mathfrak{J}}$. Further suppose that P is homogeneous in X_{h1}, \dots, X_{hn} of degree r_h ($1 \leq h \leq m$). Then for any \mathfrak{J} ,

$$(5.1) \quad |P^{\mathfrak{J}}| \leq 2^{r_1 + \dots + r_m} |P|.$$

Proof. The first assertion follows as in Chapter V. It will suffice to prove the estimate (5.1) for monomials. Now

$$\left(X_{11}^{j_{11}} \dots X_{mn}^{j_{mn}} \right)^{\mathfrak{J}} = \binom{j_{11}}{i_{11}} \dots \binom{j_{mn}}{i_{mn}} X_{11}^{j_{11} - i_{11}} \dots X_{mn}^{j_{mn} - i_{mn}}.$$

Since $\binom{j_{11}}{i_{11}} \dots \binom{j_{mn}}{i_{mn}} \leq 2^{j_{11} + \dots + j_{mn}} \leq 2^{r_1 + \dots + r_m}$, everything follows.

We now suppose that $\alpha_1, \dots, \alpha_n$ are fixed algebraic integers. Denote the degree of the field $K = \mathbb{Q}(\alpha_1, \dots, \alpha_n)$ by Δ , and let

$$\beta_1, \dots, \beta_\Delta$$

be an integer basis of K . There are relations

$$\beta_i \beta_j = \sum_{k=1}^{\Delta} c_{ijk} \beta_k \quad (1 \leq i, j \leq \Delta)$$

with rational integer coefficients c_{ijk} . Put $A = \max_{i,j,k} (|c_{ijk}|)$. An element $\gamma \in K$ may uniquely be written as

$$\gamma = \sum_{j=1}^{\Delta} r_j \beta_j \quad \text{with rational } r_1, \dots, r_\Delta.$$

Put

$$|\gamma| = A\Delta^2 \max_j |r_j|.$$

Then for $\gamma, \delta \in K$,

$$(5.2) \quad |\gamma + \delta| \leq |\gamma| + |\delta|.$$

If, say $\delta = \sum_{j=1}^{\Delta} s_j \beta_j$ with rational s_1, \dots, s_Δ , then

$$\gamma\delta = \sum_i \sum_j r_i s_j \beta_i \beta_j = \sum_i \sum_j \sum_k r_i s_j c_{ijk} \beta_k.$$

Here the coefficient of β_k is rational and has absolute value

$$\leq A\Delta^2 \max(|r_i|) \max(|s_j|) = (A\Delta^2)^{-1} |\gamma| |\delta|,$$

and this proves that

$$(5.3) \quad |\gamma\delta| \leq |\gamma| |\delta|.$$

Put

$$B = \max(\lceil \alpha_1 \rceil, \dots, \lceil \alpha_n \rceil) ,$$

where $\alpha_1, \dots, \alpha_n$ were the given integers.

LEMMA 5B. Suppose $\alpha_1, \dots, \alpha_n$ are algebraic integers as above.

Consider the polynomials

$$P(X_{11}, \dots, X_{1n}; \dots; X_{m1}, \dots, X_{mn}) = \sum c(j_{11}, \dots, j_{mn}) X_{11}^{j_{11}} \dots X_{mn}^{j_{mn}}$$

in \mathfrak{R} which are homogeneous of degree r_h in the variables X_{h1}, \dots, X_{hn}
($1 \leq h \leq m$) . We construct new polynomials P^* in the $nm - m$ variables

$$X_{12}, \dots, X_{1n}; \dots; X_{m2}, \dots, X_{mn}$$

by putting

$$P^* = P^{\mathfrak{J}}(-\alpha_2 X_{12} - \dots - \alpha_n X_{1n}, \alpha_1 X_{12}, \dots, \alpha_1 X_{1n}; \dots; \\ ; -\alpha_2 X_{m2} - \dots - \alpha_n X_{mn}, \alpha_1 X_{m2}, \dots, \alpha_1 X_{mn}).$$

where \mathfrak{J} is of the type

$$\mathfrak{J} = (j_1, 0, \dots, 0; \dots; j_m, 0, \dots, 0) .$$

Then every coefficient of P^* is a linear form in the coefficients
 $c(j_{11}, \dots, j_{mn})$ of P . The coefficients γ of these linear forms are
integers of the field $K = Q(\alpha_1, \dots, \alpha_n)$ with

$$\lceil \gamma \rceil \leq (4^n B)^{r_1 + \dots + r_m} .$$

Proof. Only the upper bound for $\lceil \gamma \rceil$ is to be proved. By Lemma
 4B of Chapter V , P^* is a sum of at most

$$\binom{r_1 + n - 1}{n - 1} \dots \binom{r_m + n - 1}{n - 1} \leq 2^{(r_1 + n - 1) + \dots + (r_m + n - 1)} \leq 2^{r_1 n + \dots + r_m n} = 2^{n(r_1 + \dots + r_m)}$$

summands S of the type

$$S = \pm c(\dots) \binom{j_{11}}{j_1} \dots \binom{j_{m1}}{j_m} \left(\alpha_2^{X_{12} + \dots + \alpha_n^{X_{1n}}} \right)^{j_{11} - j_1} (\alpha_1^{X_{12}})^{j_{12}} \dots (\alpha_1^{X_{1n}})^{j_{1n}} \\ \dots (\alpha_2^{X_{m2} + \dots + \alpha_n^{X_{mn}}})^{j_{m1} - j_m} (\alpha_1^{X_{m2}})^{j_{m2}} \dots (\alpha_1^{X_{mn}})^{j_{mn}} \\ = \pm c(\dots) \binom{j_{11}}{j_1} \dots \binom{j_{m1}}{j_m} S_1 S_2 \dots S_m, \quad ,$$

where for $1 \leq h \leq m$,

$$S_h = (\alpha_2^{X_{h2} + \dots + \alpha_n^{X_{hn}}})^{j_{h1} - j_h} (\alpha_1^{X_{h2}})^{j_{h2}} \dots (\alpha_1^{X_{hn}})^{j_{hn}} \\ = \sum_{\substack{c_2, \dots, c_n \\ c_2 + \dots + c_n = j_{h1} - j_h}} \frac{(j_{h1} - j_h)!}{c_2! \dots c_n!} \alpha_2^{c_2} \dots \alpha_n^{c_n} \alpha_1^{j_{h2} + \dots + j_{hn}} X_{h2}^{c_2 + j_{h2}} \dots X_{hn}^{c_n + j_{hn}}.$$

Here the sum of the numerical coefficients in front is $(n-1)^{j_{h1} - j_h} < n^{r_h}$,

and we note that

$$\left| \alpha_2^{c_2} \dots \alpha_n^{c_n} \alpha_1^{j_{h2} + \dots + j_{hn}} \right| \leq B^{c_2 + \dots + c_n + j_{h2} + \dots + j_{hn}} = B^{r_h - j_h} \leq B^{r_h}.$$

Thus every coefficient α of S_h is an integer of K with

$$|\alpha| \leq (nB)^{r_h}. \quad \text{Thus every coefficient of } S \text{ is of the type } c(\dots)\beta$$

where $\beta \in K$ and

$$|\beta| \leq \binom{j_{11}}{j_1} \dots \binom{j_{m1}}{j_m} (nB)^{r_1 + \dots + r_m} \leq (2nB)^{r_1 + \dots + r_m}.$$

Now since, as pointed out above, P^* is a sum of at most $2^{(r_1 + \dots + r_m)n}$

polynomials of the type S , each coefficient of P^* is a linear form

in the $c(\dots)$ with coefficients γ having

$$|\gamma| \leq 2^{n(r_1 + \dots + r_m)} (2nB)^{r_1 + \dots + r_m} \leq (2^{2nB})^{r_1 + \dots + r_m}.$$

§6. The Index Theorem.

Let $L = \alpha_1 X_1 + \dots + \alpha_n X_n$ be a non-zero linear form in X_1, \dots, X_n . We make m polynomials out of it by putting $L_h = \alpha_1 X_{h1} + \dots + \alpha_n X_{hn}$ ($1 \leq h \leq m$). When we speak of the index of a polynomial $P \in \mathfrak{R}$ with respect to

$$(L, L, \dots, L; r_1, \dots, r_m),$$

we really mean the index of P with respect to $(L_1, \dots, L_m; r_1, \dots, r_m)$.

[This is like considering in Chapter V the index of a polynomial with respect to $(\alpha, \dots, \alpha; r_1, \dots, r_m)$.]

THEOREM 6A. (Index Theorem). Let n, t be positive integers and $L^{(1)}, \dots, L^{(t)}$ linear forms in X_1, \dots, X_n , none identically zero.
Suppose that

$$L^{(i)} = \alpha_{i1} X_1 + \dots + \alpha_{in} X_n \quad (1 \leq i \leq t),$$

where the α_{ij} are algebraic integers. We put Δ_i for the degree of the field $K_i = \mathbb{Q}(\alpha_{i1}, \dots, \alpha_{in})$, and $\Delta = \max(\Delta_1, \dots, \Delta_t)$.

Suppose $\varepsilon > 0$, and m is an integer so large that

$$(6.1) \quad m \geq 4\varepsilon^{-2} \log(2t\Delta).$$

Let r_1, \dots, r_m be positive integers.

Then there exists a non-zero polynomial $P \in \mathfrak{R}$ with rational integer coefficients such that

- (i) P is homogeneous in X_{h1}, \dots, X_{hn} of degree r_h ($1 \leq h \leq m$),
- (ii) P has index $> (\frac{1}{n} - \varepsilon)m$ with respect to

$$(L^{(i)}, \dots, L^{(i)}; r_1, \dots, r_m) \quad (1 \leq i \leq t),$$

$$(iii) \quad |P| \leq D^{r_1 + \dots + r_m}$$

with a constant D depending only on the α_{ij} -s.

Proof. We write P in the form

$$P = \sum_{\substack{j_{h1} + \dots + j_{hn} = r_h \\ (1 \leq h \leq m)}} c(j_{11}, \dots, j_{mn}) X_{11}^{j_{11}} \dots X_{mn}^{j_{mn}},$$

and we shall check what the conditions (ii) imply for the coefficients $c(j_{11}, \dots, j_{mn})$. What for instance does it mean that the index with respect to $(L^{(1)}, \dots, L^{(1)}; r_1, \dots, r_m)$ should be $\geq (\frac{1}{n} - \varepsilon)m$?

Without loss of generality we may assume that $\alpha_{11} \neq 0$. By Lemma 4C, the index satisfies the desired inequality if each polynomial $P^{\mathfrak{J}}$, where

$$\mathfrak{J} = (i_1, 0, \dots, 0; \dots; i_m, 0, \dots, 0)$$

and where

$$\sum_{h=1}^m \frac{i_h}{r_h} < (\frac{1}{n} - \varepsilon)m \quad \text{and} \quad 0 \leq i_h \leq r_h,$$

vanishes on the subspace T defined by $L_1^{(1)} = L_2^{(1)} = \dots = L_m^{(1)} = 0$.

(For the definition of $L_1^{(1)}, \dots, L_m^{(1)}$ see the beginning of this section.)

Keep i_1, \dots, i_m fixed at the moment. Our condition then is that

$$P^* = P^{\mathfrak{J}}(-\alpha_{12}X_{12} - \dots - \alpha_{1n}X_{1n}, \alpha_{11}X_{12}, \dots, \alpha_{11}X_{1n}; \dots \\ ; -\alpha_{12}X_{m2} - \dots - \alpha_{1n}X_{mn}, \alpha_{11}X_{m2}, \dots, \alpha_{11}X_{mn})$$

is identically zero. In general P^* is homogeneous in X_{h2}, \dots, X_{hn} of degree $r_h - i_h$ ($1 \leq h \leq m$), hence has at most

$$\binom{r_1 - i_1 + n - 2}{n - 2} \dots \binom{r_m - i_m + n - 2}{n - 2} = f_1(i_1) \dots f_m(i_m)$$

coefficients. By Lemma 5B, each coefficient is a linear form in the $c(\dots)$, with the coefficients of these linear forms being integers of $K_1 = \mathbb{Q}(\alpha_{11}, \dots, \alpha_{1n})$. We thus obtain

$$f_1(i_1) \dots f_m(i_m)$$

linear homogeneous equations in the numbers $c(\dots)$, and the coefficients of these linear equations are algebraic integers γ of K with

$|\gamma| \leq (4^{n_{B_1}})^{r_1 + \dots + r_m}$. Here B_1 is the constant B of the last section with respect to $\alpha_{11}, \dots, \alpha_{1n}$. Each of these linear equations follows from Δ_1 linear homogeneous equations with rational integer coefficients k having

$$(6.2) \quad |k| \leq (4^{n_{B_1}})^{r_1 + \dots + r_m}.$$

By taking the sum over all m -tuples i_1, \dots, i_m with $\sum_{h=1}^m \frac{i_h}{r_h} < (\frac{1}{n} - \epsilon)m$, one sees: In order that (ii) is true with $(L^{(1)}, \dots, L^{(1)}; r_1, \dots, r_m)$, the $c(\dots)$ have to satisfy at most

$$(6.3) \quad \Delta_1 \sum f_1(i_1) \dots f_m(i_m)$$

linear homogeneous equations with rational integer coefficients k satisfying (6.2), where the sum in (6.3) is extended over all i_1, \dots, i_m with

$$\sum_{h=1}^m \frac{i_h}{r_h} < (\frac{1}{n} - \epsilon)m \quad \text{and} \quad 0 \leq i_h \leq r_h.$$

Now by the formula (4.4) in Chapter V the expression in (6.3) equals

$$\Delta_1 M_- ,$$

since $f_j(i_j) = \binom{r_j - i_j + n - 2}{n - 2}$ is the number of $(n-1)$ -tuples i_{j2}, \dots, i_{jn} with $i_{j2} + \dots + i_{jn} = r_j - i_j$. By (4.3) of Chapter V we have

$$M_- \leq \binom{r_1 + n - 1}{r_1} \dots \binom{r_m + n - 1}{r_m} e^{-\varepsilon^2 m / 4},$$

and the total number of equations to satisfy (ii) with $i = 1$ is

$$\leq \Delta_1 \binom{r_1 + n - 1}{r_1} \dots \binom{r_m + n - 1}{r_m} \frac{1}{2t\Delta} \quad \text{by (6.1),}$$

and hence is

$$\leq \frac{N}{2t}$$

where

$$N = \binom{r_1 + n - 1}{r_1} \dots \binom{r_m + n - 1}{r_m}.$$

Thus the total number M of homogeneous linear equations to be satisfied by the N unknowns $c(j_{11}, \dots, j_{mn})$ to obtain (ii) for $i = 1, \dots, t$ satisfies

$$M \leq \frac{N}{2}.$$

The absolute values of the rational integer coefficients of these M equations are at most

$$A = (4^{n_B})^{r_1 + \dots + r_m}, \quad \text{where } B = \max(B_1, \dots, B_t).$$

By Lemma 5B of Chapter V, this system of homogeneous linear equations has rational integer solutions $c(\dots)$, not all zero, with

$$\begin{aligned} |c(\dots)| &\leq (NA)^{M/(N-M)} \leq NA = \binom{r_1 + n - 1}{r_1} \dots \binom{r_m + n - 1}{r_m} A \\ &\leq 2^{n(r_1 + \dots + r_m)} A = (8^{n_B})^{r_1 + \dots + r_m} = D^{r_1 + \dots + r_m}, \quad \text{say.} \end{aligned}$$

So we obtain a polynomial $P \neq 0$ with (i), (ii) and (iii).

§7. The Polynomial Theorem.

THEOREM 7A. (The Polynomial Theorem). Let

$$L^{(i)} = \alpha_{i1}X_1 + \dots + \alpha_{in}X_n \quad (1 \leq i \leq n)$$

be n linear forms with algebraic integer coefficients α_{ij} of non-vanishing determinant. Put

$$L_h^{(i)} = \alpha_{i1}X_{h1} + \dots + \alpha_{in}X_{hn} \quad (1 \leq h \leq m, 1 \leq i \leq n),$$

Suppose that $\epsilon > 0$ and that m satisfies

$$(7.1) \quad m > 4\epsilon^{-2} \log(2n\Delta),$$

where Δ is defined as in the Index Theorem (Theorem 6A). Also, let r_1, r_2, \dots, r_m be positive integers.

Then there is a $P \in \mathbb{R}$, $P \neq 0$, with rational integer coefficients, which has the following properties:

- (i) P is homogeneous in X_{h1}, \dots, X_{hn} of degree r_h ($1 \leq h \leq m$).
- (ii) $|P| \leq D^{r_1 + \dots + r_m}$, where D depends only on the α_{ij} -s.
- (iii) If we write (uniquely)

$$(7.2) \quad P^{\mathfrak{J}} = \sum d^{\mathfrak{J}}(j_{11}, \dots, j_{mn}) L_1^{(1)j_{11}} \dots L_1^{(n)j_{1n}} \dots L_m^{(1)j_{m1}} \dots L_m^{(n)j_{mn}},$$

then

$$|d^{\mathfrak{J}}(j_{11}, \dots, j_{mn})| \leq E^{r_1 + \dots + r_m}$$

for any \mathfrak{J} and any (j_{11}, \dots, j_{mn}) , where E depends only on the α_{ij} -s.

- (iv) If $(\mathfrak{J}/r) \leq 2\epsilon m$, then $d^{\mathfrak{J}}(j_{11}, \dots, j_{mn}) = 0$ in (7.2), unless

$$\left| \left(\sum_{h=1}^m \frac{j_{hk}}{r_h} \right) - \frac{m}{n} \right| \leq 3n\epsilon \quad (1 \leq k \leq n) .$$

Proof. We may apply the Index Theorem with $t = n$. Let $P \neq 0$ be a polynomial in \mathfrak{K} with rational integer coefficients satisfying the conditions (i), (ii), (iii) of the Index Theorem. Then the new conditions (i) and (ii) hold. We shall show that also (iii) and (iv) hold.

Verification of (iii). Since $\det(L^{(1)}, \dots, L^{(n)}) \neq 0$, one can go from X_1, \dots, X_n to $L^{(1)}, \dots, L^{(n)}$ by a nonsingular linear transformation. Thus, every polynomial in X_1, \dots, X_n is a polynomial in $L^{(1)}, \dots, L^{(n)}$, and therefore every polynomial in X_{h1}, \dots, X_{hn} is a polynomial in $L_h^{(1)}, \dots, L_h^{(n)}$ ($1 \leq h \leq m$). This makes it clear that $P^{\mathfrak{J}}$ may be uniquely written as (7.2).

If (β_{ij}) is the inverse of the matrix (α_{ij}) , then

$$X_i = \sum_{k=1}^n \beta_{ik} L^{(k)}, \quad (1 \leq i \leq n)$$

whence

$$(7.3) \quad X_{hi} = \sum_{k=1}^n \beta_{ik} L_h^{(k)} \quad (1 \leq h \leq m, 1 \leq i \leq n) .$$

Let us write

$$(7.4) \quad P^{\mathfrak{J}} = \sum c^{\mathfrak{J}}(j_{11}, \dots, j_{mn}) X_{11}^{j_{11}} \dots X_{mn}^{j_{mn}} .$$

Now $\boxed{P^{\mathfrak{J}}} \leq 2^{r_1 + \dots + r_m} \boxed{P} \leq (2D)^{r_1 + \dots + r_m}$ by (ii), so

$$|c^{\mathfrak{J}}(j_{11}, \dots, j_{mn})| \leq (2D)^{r_1 + \dots + r_m}$$

for any \mathfrak{J} and any (j_{11}, \dots, j_{mn}) .

One gets from (7.4) to (7.2) by substituting the right hand side of (7.3) for each X_{hi} . A typical product $X_{11}^{j_{11}} \dots X_{mn}^{j_{mn}}$ in (7.4)

then becomes

$$(7.5) \quad \left(\sum_{k=1}^n \beta_{1k} L_1^{(k)} \right)^{j_{11}} \dots \left(\sum_{k=1}^n \beta_{nk} L_m^{(k)} \right)^{j_{mn}}.$$

We put $G = \max(1, |\beta_{11}|, \dots, |\beta_{mn}|)$. Then (7.5) is a polynomial in $L_h^{(i)}$ ($1 \leq h \leq m$, $1 \leq i \leq n$), whose coefficients have absolute values

$$\leq (nG)^{j_{11} + \dots + j_{mn}} \leq (nG)^{r_1 + \dots + r_m}.$$

So $c^{\mathfrak{J}}(j_{11}, \dots, j_{mn}) X_{11}^{j_{11}} \dots X_{mn}^{j_{mn}}$ as a polynomial in the $L_h^{(i)}$ has coefficients with absolute value

$$\leq (2DnG)^{r_1 + \dots + r_m}.$$

By Lemma 4B of Chapter V, the total number of summands in (7.4) is

$$\begin{aligned} &\leq \binom{r_1 + n - 1}{r_1} \dots \binom{r_m + n - 1}{r_m} \leq 2^{(r_1 + n - 1) + \dots + (r_m + n - 1)} \\ &\leq 2^{r_1 n + \dots + r_m n} = 2^{n(r_1 + \dots + r_m)}. \end{aligned}$$

Hence $p^{\mathfrak{J}}$ as a polynomial in $L_h^{(i)}$ ($1 \leq h \leq m$, $1 \leq i \leq n$) has coefficients of absolute value

$$\leq (2^n \cdot 2DnG)^{r_1 + \dots + r_m} = E^{r_1 + \dots + r_m}, \text{ say.}$$

Verification of (iv). By the Index Theorem, the index of P with respect to

$$(L_1^{(k)}, \dots, L_m^{(k)}; r_1, \dots, r_m) \text{ is } \geq \left(\frac{1}{n} - \varepsilon \right) m \quad (1 \leq k \leq n).$$

Now if $(\mathfrak{Y}/r) \leq 2\varepsilon m$, then it follows from Lemma 4B that $p^{\mathfrak{J}}$ has index $\geq \left(\frac{1}{n} - \varepsilon \right) m - 2\varepsilon m = \left(\frac{1}{n} - 3\varepsilon \right) m$ with respect to

$$(L_1^{(k)}, L_2^{(k)}, \dots, L_m^{(k)}; r_1, r_2, \dots, r_m) \quad (1 \leq k \leq n).$$

Therefore if $d^3(j_{11}, \dots, j_{mm}) \neq 0$ in (7.2), then

$$\sum_{h=1}^m \frac{j_{hk}}{r_h} \geq \left(\frac{1}{n} - 3\varepsilon \right) m \quad (1 \leq k \leq n) .$$

This inequality may be rewritten as

$$(7.6) \quad \left(\sum_{h=1}^m \frac{j_{hk}}{r_h} \right) - \frac{m}{n} \geq -3\varepsilon m \quad (1 \leq k \leq n),$$

which gives half of the desired inequality.

P is homogeneous in X_{h1}, \dots, X_{hn} of degree r_h , hence P^3 is of total degree $\leq r_h$ in the block of variables X_{h1}, \dots, X_{hn} ($1 \leq h \leq m$). Therefore

$$\sum_{k=1}^n \frac{j_{hk}}{r_h} \leq 1 \quad (1 \leq h \leq m),$$

whence

$$\sum_{k=1}^n \sum_{h=1}^m \frac{j_{hk}}{r_h} \leq m ,$$

or

$$\sum_{k=1}^n \left(\left(\sum_{h=1}^m \frac{j_{hk}}{r_h} \right) - \frac{m}{n} \right) \leq 0 .$$

Together with (7.6) this gives

$$\left(\sum_{h=1}^m \frac{j_{hk}}{r_h} \right) - \frac{m}{n} \leq 3m\varepsilon(n-1) \quad (1 \leq k \leq n) .$$

Thus the other half of the desired inequality in (iv) is established, and the theorem is proved.

§8. Grids.

Let n be an integer with $n \geq 2$. Put $q = n - 1$, and let H^q be a linear subspace of \mathbb{R}^n of dimension q .

Definition. Let $\underline{w}_1, \dots, \underline{w}_q$ be q linearly independent points of H^q , and let s be a positive integer. Denote by

$$\Gamma = \Gamma(s; \underline{w}_1, \dots, \underline{w}_q)$$

the set of all points $\underline{w} = h_1 \underline{w}_1 + \dots + h_q \underline{w}_q$ where h_1, \dots, h_q are integers with $1 \leq h_i \leq s$ ($1 \leq i \leq q$). (Thus Γ consists of s^q points.) We shall say that Γ is a grid on H^q of size s .

In what follows, we will study polynomials $P(X_1, \dots, X_n)$ which will be regarded as functions on \mathbb{R}^n .

LEMMA 8A. Let $P(X_1, \dots, X_n)$ be a polynomial with real coefficients, and of total degree $\leq r$. Let s be a positive integer, and let Γ be a grid of size s on some subspace H^q of \mathbb{R}^n . Finally let t be an integer with

$$(8.1) \quad s(t+1) > r,$$

and assume that P and the partial derivatives

$$\frac{\partial^{t_1 + \dots + t_n}}{\partial X_1^{t_1} \dots \partial X_n^{t_n}} P \quad \text{with} \quad t_1 + \dots + t_n \leq t$$

vanish on Γ .

Then P vanishes identically on H^q .

Proof. After a linear transformation, we may assume that the grid Γ is spanned by $\underline{w}_1 = (1, 0, \dots, 0), \dots, \underline{w}_q = (0, \dots, 0, 1, 0)$, and hence that H^q is the coordinate plane $x_n = 0$.

Put $Q(X_1, \dots, X_q) = P(X_1, \dots, X_q, 0)$. We have to show that a polynomial $Q(X_1, \dots, X_q)$ of total degree $\leq r$ is identically zero if Q and its partial derivatives of order $\leq t$ vanish on the s^q integer

points (h_1, \dots, h_q) where $1 \leq h_i \leq s$ ($1 \leq i \leq q$). We shall prove this by induction on q .

Suppose that $q = 1$. We have now a polynomial $Q(X_1)$ in a single variable. If $Q, Q', Q'', \dots, Q^{(t)}$ each vanish at $X_1 = 1, 2, \dots, s$, then the number of zeros of Q , counted with multiplicities, is $\geq s(t+1)$. Since Q is of degree $\leq r < s(t+1)$, we have $Q = 0$.

Assume now that the above statement is true for $q-1$ in place of q , and that $q \geq 2$. To prove the statement with q , it will suffice to show that $(X_1 - h)^{t+1}$ divides Q for $h = 1, 2, \dots, s$. [For then the polynomial $K = (X_1 - 1)^{t+1}(X_1 - 2)^{t+1} \dots (X_1 - s)^{t+1}$ divides Q . But K has degree $s(t+1) > r \geq \deg Q$, whence $Q = 0$.] To this end, let e_h be the largest exponent so that $(X_1 - h)^{e_h}$ divides Q ($1 \leq h \leq s$). To complete the proof it is enough to show that $e \geq t+1$, where $e = \min(e_1, \dots, e_s)$.

Suppose, to the contrary, that $e \leq t$. Suppose, say, that $e = e_1 \leq t$. We write

$$(8.2) \quad Q(X_1, \dots, X_q) = (X_1 - 1)^{e_1} (X_1 - 2)^{e_2} \dots (X_1 - s)^{e_s} R(X_1, \dots, X_q).$$

Here R has total degree $\leq r - e_1 - e_2 - \dots - e_s \leq r - se$. Take the $e = e_1$ -th derivative with respect to X_1 , then put $X_1 = 1$. The right hand side of (8.2) becomes

$$e!(1-2)^{e_2} \dots (1-s)^{e_s} R(1, X_2, \dots, X_q).$$

$R(1, X_2, \dots, X_q)$ is a polynomial in $q-1$ variables of total degree $\leq r - se$. Because Q and its partial derivatives of order $\leq t$ vanish on the s^q integer points (h_1, h_2, \dots, h_q) , where $1 \leq h_i \leq s$ ($1 \leq i \leq q$), the polynomial $R(1, X_2, \dots, X_q)$ and its partial derivatives of order $\leq t-e$ vanish on the s^{q-1} integer points (h_2, \dots, h_q) where $1 \leq h_i \leq s$. Also,

inequality (8.1) gives

$$s(t - e + 1) > r - se .$$

By the inductive hypothesis, we conclude that $R(1, X_2, \dots, X_q) = 0$. But then $(X_1 - 1)$ divides $R(1, X_2, \dots, X_q)$, and therefore $(X_1 - 1)^{e_1+1}$ divides $Q(X_1, X_2, \dots, X_q)$. Since this contradicts the definition of e_1 , we conclude that $e \geq t+1$.

LEMMA 8B. Let P be a polynomial in \mathfrak{R} of total degree $\leq r_h$ in the block of variables X_{h1}, \dots, X_{hn} ($1 \leq h \leq m$). We may write $P = P(\mathfrak{X}_1, \dots, \mathfrak{X}_m)$, where $\mathfrak{X}_h \in \mathbb{R}^n$ ($1 \leq h \leq m$). Thus P may be regarded as a function on the product space

$$\underbrace{\mathbb{R}^n \times \mathbb{R}^n \times \dots \times \mathbb{R}^n}_{m \text{ factors}} .$$

Let H_1^q, \dots, H_m^q be subspaces of \mathbb{R}^n , each of dimension $q = n-1$, and let $T = H_1^q \times H_2^q \times \dots \times H_m^q$ consist of all points $(\mathfrak{X}_1, \dots, \mathfrak{X}_m)$ with $\mathfrak{X}_h \in H_h^q$ ($1 \leq h \leq m$). (Thus, T is a subspace of \mathbb{R}^{nm} of dimension $mn-m$.) For $h = 1, 2, \dots, m$, let Γ_h be a grid on H_h^q of size s_h , and let t_1, t_2, \dots, t_m be integers with

$$(8.3) \quad s_h(t_h+1) > r_h \quad (1 \leq h \leq m) .$$

In an obvious notation, let $\Gamma^* = \Gamma_1 \times \Gamma_2 \times \dots \times \Gamma_m$. Finally, suppose that P and all polynomials $P^{\mathfrak{Z}}$ with $\mathfrak{Z} = (t_{11}, \dots, t_{1n}; \dots; t_{m1}, \dots, t_{mn})$ satisfying

$$t_{h1} + t_{h2} + \dots + t_{hn} \leq t_h \quad (1 \leq h \leq m)$$

vanish on Γ^* .

Then $P = 0$ on T .

Proof. Apply Lemma 8A and induction on m .

§9. The Index of P with Respect to certain Rational Linear Forms.

Suppose that $n \geq 2$ is an integer. Put

$$(9.1) \quad q = n-1 .$$

Suppose that $\underline{w}_1, \dots, \underline{w}_q$ are linearly independent integer points in \mathbb{R}^n .

Then there exists a linear form

$$M(\underline{X}) = m_1 X_1 + \dots + m_n X_n ,$$

not identically zero, with coprime rational integer coefficients and such that

$$M(\underline{w}_t) = 0 \quad (1 \leq t \leq q) .$$

It is clear that M is unique up to a factor ± 1 . We shall put

$$M = M\{\underline{w}_1, \dots, \underline{w}_q\} .$$

THEOREM 9A. (Corresponds to Theorem 8A of Chapter V.) Let

c_1, \dots, c_n be real numbers with

$$(9.2) \quad c_1 + \dots + c_n = 0 \quad \text{and} \quad |c_i| \leq 1 \quad (1 \leq i \leq n) .$$

Suppose that $\varepsilon > 0$, $0 < \delta < 1$, and

$$(9.3) \quad 16n^2 \varepsilon < \delta .$$

Let $L^{(1)}, \dots, L^{(n)}$ be linear forms and $m; r_1, \dots, r_m$ positive integers
satisfying the hypotheses of Theorem 7A . Let E be the constant of

part (iii) of Theorem 7A , and let P be the polynomial described in that theorem.

Let Q_1, \dots, Q_m be real numbers with

$$(9.4) \quad Q_h^\varepsilon > 2^n \varepsilon, \quad Q_h^\varepsilon > n(\varepsilon^{-1} + 1) \quad (1 \leq h \leq m),$$

$$(9.5) \quad r_1 \log Q_1 \leq r_h \log Q_h \leq (1 + \varepsilon) r_1 \log Q_1 \quad (1 \leq h \leq m).$$

For each h , $1 \leq h \leq m$, let g_{h1}, \dots, g_{hq} be q linearly independent integer points in \mathbb{R}^n with

$$(9.6) \quad |L^{(k)}(g_{ht})| \leq Q_h^{c_k - \delta} \quad (1 \leq k \leq n, 1 \leq h \leq m, 1 \leq t \leq q).$$

Then P has index $\geq m\varepsilon$ with respect to $(M_1, \dots, M_m; r_1, \dots, r_m)$, where each M_h is the linear form in X_{h1}, \dots, X_{hn} given by

$$M_h = M_h\{g_{h1}, \dots, g_{hq}\}.$$

Proof. Let T be the subspace of \mathbb{R}^{mn} where $M_1 = M_2 = \dots = M_m = 0$. By Lemma 4C , it will suffice to show that $P^{\mathfrak{J}} = 0$ on T whenever $(\mathfrak{Y}/\mathfrak{r}) < \varepsilon m$. Let Γ_h be the grid

$$\Gamma_h = \Gamma([\varepsilon^{-1}] + 1; g_{h1}, \dots, g_{hq}) \quad (1 \leq h \leq m).$$

By Lemma 8B , it will suffice to show that

$$(P^{\mathfrak{J}})^{\mathfrak{X}}(\underline{v}_1, \dots, \underline{v}_m) = 0$$

for $\underline{v}_h \in \Gamma_h$ ($1 \leq h \leq m$) and $\mathfrak{X} = (t_{11}, \dots, t_{1n}; \dots; t_{m1}, \dots, t_{mn})$ satisfying

$$t_{h1} + \dots + t_{hn} \leq [r_h \varepsilon] \quad (1 \leq h \leq m).$$

Namely, $s_h = [\varepsilon^{-1}] + 1$, $t_h = [r_h \varepsilon]$ satisfy $s_h(t_h + 1) > \varepsilon^{-1} r_h \varepsilon = r_h$, i.e.

(8.3). Now $(\mathfrak{Y}/\underline{r}) < \varepsilon m$ and $(\mathfrak{X}/\underline{r}) \leq \frac{[r_1 \varepsilon]}{r_1} + \dots + \frac{[r_m \varepsilon]}{r_m} \leq \varepsilon m$, so $(P^{\mathfrak{J}})^{\mathfrak{X}}$ (except for a constant factor) is of type $P^{\mathfrak{J}'}$ with $(\mathfrak{Y}'/\underline{r}) < 2\varepsilon m$.

Hence it will suffice to show that

$$(9.7) \quad P^{\mathfrak{J}}(\underline{v}_1, \dots, \underline{v}_m) = 0$$

when $(\mathfrak{Y}/\underline{r}) < 2\varepsilon m$ and $\underline{v}_h \in \Gamma_h$ ($1 \leq h \leq m$).

The left hand side of (9.7) may be written as

$$(9.8) \quad \sum d^{\mathfrak{J}}(j_{11}, \dots, j_{mn})_{L(1)}(\underline{v}_1)^{j_{11} \dots L(n)}(\underline{v}_1)^{j_{1n} \dots L(1)}(\underline{v}_m)^{j_{m1} \dots L(n)}(\underline{v}_m)^{j_{mn}}.$$

By (9.3), (9.4) and (9.6), we have

$$(9.9) \quad |L^{(k)}(\underline{v}_h)| < n\left(\frac{1}{\varepsilon} + 1\right) Q_h^{c_k - \delta} < Q_h^{c_k - \delta + \varepsilon} \leq Q_h^{c_k - 15n^2\varepsilon} \\ (1 \leq k \leq n, 1 \leq h \leq m),$$

whence

$$|L^{(k)}(\underline{v}_1)^{j_{1k} \dots L(k)}(\underline{v}_m)^{j_{mk}}| \leq \exp((c_k - 15n^2\varepsilon) \sum_{h=1}^m j_{hk} \log Q_h) \quad (1 \leq k \leq n).$$

By (9.5) and by (iv) of Theorem 7A, we have $d^{\mathfrak{J}}(j_{11}, \dots, j_{mn}) = 0$

unless

$$\sum_{h=1}^m j_{hk} \log Q_h \geq r_1 \log Q_1 \cdot \sum_{h=1}^m \frac{j_{hk}}{r_h} \geq r_1 \log Q_1 \cdot \left(\frac{1}{n} - 3n\varepsilon\right)m$$

and

$$\begin{aligned} \sum_{h=1}^m j_{hk} \log Q_h &\leq (1+\varepsilon)r_1 \log Q_1 \cdot \sum_{h=1}^m \frac{j_{hk}}{r_h} \\ &\leq r_1 \log Q_1 \cdot (1+\varepsilon)\left(\frac{1}{n} + 3n\varepsilon\right)m \\ &\leq r_1 \log Q_1 \cdot \left(\frac{1}{n} + 7n\varepsilon\right)m. \end{aligned}$$

Both inequalities together give

$$\left| \sum_{h=1}^m j_{hk} \log Q_h - (r_1 \log Q_1) \cdot \frac{m}{n} \right| \leq (r_1 \log Q_1) \cdot 7n\epsilon \quad .$$

This estimate together with $|c_k - 15n^2\epsilon| \leq 2$ (which follows from (9.2), (9.3)) yields

$$\begin{aligned} |L^{(k)}(\underline{v}_1)^{j_{1k}} \dots L^{(k)}(\underline{v}_m)^{j_{mk}}| &\leq Q_1^{r_1 \frac{m}{n}(c_k - 15n^2\epsilon) + 2r_1 \cdot 7n\epsilon} \\ &= Q_1^{r_1 \left(\frac{m}{n} c_k - n\epsilon \right)} . \end{aligned}$$

Altogether each summand of (9.8) has absolute value

$$\leq E^{r_1 + \dots + r_m} Q_1^{r_1 \frac{m}{n}(c_1 + \dots + c_n) - r_1 n^2 m \epsilon} = E^{r_1 + \dots + r_m} Q_1^{-r_1 n^2 m \epsilon} \quad (\text{by (9.2)})$$

$$\leq E^{r_1 + \dots + r_m} \left(\frac{1}{Q_1^{r_1 \epsilon} \dots Q_m^{r_m \epsilon}} \right)^{\frac{n}{1+\epsilon}} \quad (\text{by (9.5)})$$

$$\leq E^{r_1 + \dots + r_m} Q_1^{-r_1 \epsilon} \dots Q_m^{-r_m \epsilon} .$$

The number of summands in (9.8) is $\leq 2^{n(r_1 + \dots + r_m)}$. Therefore (9.8), which is the left hand side of (9.7), has absolute value

$$\leq \prod_{h=1}^m \left(2^{nEQ_h^{-\epsilon}} \right)^{r_h} < 1$$

by (9.4). But the left side of (9.7) is necessarily a rational integer, so it must be zero. This completes the proof of Theorem 9A.

§10. An Analogue of Roth's Lemma.

LEMMA 10A. Suppose that $n \geq 2$ is an integer. Let m_1, \dots, m_n be coprime rational integers, $m_1 \neq 0$. Then there is a j with $2 \leq j \leq n$ such that

$$\text{g.c.d.}(m_1, m_j) \leq |m_1|^{\frac{n-2}{n-1}} .$$

Proof. Put $d_j = \text{g.c.d.}(m_1, m_j)$ for $j = 2, \dots, n$. Then $\text{g.c.d.}(d_2, \dots, d_n) = 1$. Each d_j divides m_1 , hence $d_2 d_3 \dots d_n$ divides m_1^{n-2} , and therefore

$$d_2 d_3 \dots d_n \leq |m_1|^{n-2}.$$

The Lemma follows at once from this inequality.

THEOREM 10B. Suppose that $0 < \varepsilon < \frac{1}{12}$, that m is a positive integer, and

$$(10.1) \quad \omega = 24 \cdot 2^{-m} (\varepsilon/12)^{2^{m-1}}.$$

Let r_1, \dots, r_m be positive integers with

$$(10.2) \quad \omega r_h \geq r_{h+1} \quad (1 \leq h \leq m).$$

Suppose that $n \geq 2$ is an integer, and put $q = n-1$. Let M_1, \dots, M_m be nonzero linear forms in n variables with coprime rational integer coefficients. Suppose that $0 < \Gamma \leq q$ and that

$$(10.3) \quad |M_h|^{r_h} \geq |M_1|^{r_1 \Gamma} \quad (1 \leq h \leq m),$$

$$(10.4) \quad |M_h|^{\omega \Gamma} \geq 2^{3mq^2} \quad (1 \leq h \leq m).$$

Let $P \in \mathfrak{K}$, $P \neq 0$, with rational integer coefficients, homogeneous in X_{h1}, \dots, X_{hn} of degree r_h $(1 \leq h \leq m)$. Further suppose that

$$(10.5) \quad |P|^q \leq |M_1|^{\omega r_1 \Gamma}.$$

Then the index of P with respect to $(M_1, \dots, M_m; r_1, \dots, r_m)$ is

$$\leq \varepsilon.$$

Proof. Without loss of generality, we may suppose that $|M_h| = |m_{h1}|$

for each h , where

$$M_h = m_{h1}X_{h1} + \dots + m_{hn}X_{hn} \quad (1 \leq h \leq m)^{\dagger}.$$

Further, by Lemma 10A, we may assume that

$$(10.6) \quad \text{g.c.d.}(m_{h1}, m_{h2}) \leq |m_{h1}|^{(n-2)/(n-1)} = |m_{h1}|^{(q-1)/q} \quad (1 \leq h \leq m).$$

Let η be the index of P with respect to $(M_1, \dots, M_m; r_1, \dots, r_m)$.

Then P lies in the ideal $I(\eta)$ generated by polynomials $M_1^{i_1} M_2^{i_2} \dots M_m^{i_m}$ with

$$(10.7) \quad \sum_{h=1}^m \frac{i_h}{r_h} \geq \eta.$$

We construct a new polynomial $\hat{P}(X_{11}, X_{12}; X_{21}, X_{22}; \dots; X_{m1}, X_{m2})$ from P as follows: If $n = 2$, put $\hat{P} = P$. Otherwise, if $n > 2$, we get rid of the $p = (n-2)m$ variables $X_{13}, \dots, X_{1n}; \dots; X_{m3}, \dots, X_{mn}$ in the following manner. P may be written (uniquely) as a polynomial in $M_1, X_{12}, X_{13}, \dots, X_{1n}; \dots; M_m, X_{m2}, X_{m3}, \dots, X_{mn}$. Let a be the largest non-negative integer u such that X_{13}^u divides P . Put $\bar{P} = X_{13}^{-a} P$. Then $\bar{P} \in \mathfrak{R}$, $[\bar{P}] = [P]$ and $\bar{P} \in I(\eta)$. Also, \bar{P} is homogeneous in the block of variables X_{h1}, \dots, X_{hn} of some degree $\leq r_h$ ($1 \leq h \leq m$). Put

$$P^{(1)} = \bar{P}(X_{11}, X_{12}, 0, X_{14}, \dots, X_{1n}; \dots; X_{m1}, X_{m2}, X_{m3}, \dots, X_{mn}).$$

Then $P^{(1)} \neq 0$, $[P^{(1)}] \leq [\bar{P}]$, further $P^{(1)}$ is homogeneous in $X_{11}, X_{12}, X_{14}, \dots, X_{1n}$ of degree $\leq r_1$ and homogeneous in $X_{h1}, X_{h2}, \dots, X_{hn}$ of degree r_h ($2 \leq h \leq m$), and $P^{(1)}$ lies in the ring of polynomials in $mn-1$ variables in the ideal $I^{(1)}(\eta)$ generated by polynomials

\dagger) Even though the letter m is used both for the coefficients of M_h and for the number of indices h , there should be no confusion.

$$M_1(X_{11}, X_{12}, 0, X_{14}, \dots, X_{1n})^{i_1} M_2^{i_2} \dots M_m^{i_m}$$

with (10.7) .

Carrying out this procedure p times we obtain a polynomial $P^{(p)}$. Put $\hat{P} = P^{(p)}$. Then \hat{P} is a polynomial in $X_{11}, X_{12}; X_{21}, X_{22}; \dots; X_{m1}, X_{m2}$, homogeneous in X_{h1}, X_{h2} of degree $\leq r_h$ ($1 \leq h \leq m$) , with rational integer coefficients, $\hat{P} \neq 0$, $|\hat{P}| \leq |P|$, and \hat{P} lies in the ideal $\hat{I}(\eta)$ generated by polynomials

$$(m_{11}X_{11} + m_{12}X_{12})^{i_1} \dots (m_{m1}X_{m1} + m_{m2}X_{m2})^{i_m}$$

with (10.7) . Put

$$\tilde{P}(X_1, \dots, X_m) = \hat{P}(X_1, 1; X_2, 1; \dots; X_m, 1) .$$

Then $\tilde{P} \neq 0$ since $\hat{P} \neq 0$ and since \hat{P} was homogeneous in each block of variables. Also, $|\tilde{P}| = |\hat{P}| \leq |P|$. Further \tilde{P} lies in the ideal $\tilde{I}(\eta)$ generated by polynomials

$$(10.8) \quad \left(X_1 + \frac{m_{12}}{m_{11}}\right)^{i_1} \dots \left(X_m + \frac{m_{m2}}{m_{m1}}\right)^{i_m}$$

with (10.7) .

Now put

$$q_h = \frac{|m_{h1}|}{(m_{h1}, m_{h2})} , \quad p_h = -\frac{m_{h2}}{(m_{h1}, m_{h2})} \cdot \frac{m_{h1}}{|m_{h1}|} \quad (1 \leq h \leq m) .$$

Clearly $q_h > 0$ and $(p_h, q_h) = 1$. The polynomials (10.8) may be written as

$$\left(X_1 - \frac{p_1}{q_1}\right)^{i_1} \dots \left(X_m - \frac{p_m}{q_m}\right)^{i_m}$$

with (10.7). For j_1, \dots, j_m with $\sum_{h=1}^m \frac{j_h}{r_h} < \eta$, we have

$$\tilde{P}_{j_1 \dots j_m} \left(\frac{p_1}{q_1}, \dots, \frac{p_m}{q_m} \right) = 0.$$

Therefore \tilde{P} has Roth index $\geq \eta$ with respect to $\left(\frac{p_1}{q_1}, \dots, \frac{p_m}{q_m}; r_1, \dots, r_m \right)$.

This will give the desired relation $\eta \leq \varepsilon$ if we can show that \tilde{P} and $\frac{p_1}{q_1}, \dots, \frac{p_m}{q_m}$ satisfy the hypotheses of Roth's Lemma (Theorem 10A of Chapter V).

First of all, \tilde{P} has degree $\leq r_h$ in X_h . It is clear that conditions (10.1), (10.2) and (10.3) of Roth's Lemma hold. Put $\gamma = \Gamma/q$ (so that $0 < \gamma \leq 1$) and recall that $(p_h, q_h) = 1$ ($1 \leq h \leq m$). We have

$$\begin{aligned} \left| M_h \right|^{\frac{1}{q}} &= |m_{h1}|^{\frac{1}{q}} \leq \frac{|m_{h1}|}{(m_{h1}, m_{h2})} && \text{(by (10.6))} \\ &= q_h \leq \left| M_h \right| && (1 \leq h \leq m). \end{aligned}$$

Thus

$$\begin{aligned} r_h &\geq \left| M_h \right|^{r_h/q} \geq \left| M_1 \right|^{r_1 \Gamma/q} && \text{(by (10.3))} \\ &= \left| M_1 \right|^{r_1 \gamma} \geq q_1^{r_1 \gamma}. \end{aligned}$$

But this is inequality (10.4) of Chapter V. Also,

$$q_h^{\omega \gamma} = q_h^{\omega \Gamma/q} \geq \left| M_h \right|^{\omega \Gamma/q^2} \geq 2^{3m} \quad (1 \leq h \leq m)$$

by our (10.4), which gives (10.5) of Chapter V. Finally, our (10.5) yields

$$\left| \tilde{P} \right| \leq \left| P \right| \leq \left| M_1 \right|^{\omega r_1 \Gamma/q^2} = \left| M_1 \right|^{\omega r_1 \gamma/q} \leq q_1^{\omega \gamma r_1}.$$

This is (10.6) of Chapter V, and the desired conclusion follows.

§11. The Size of \underline{g}_n^* .

Let again $L_i(\underline{X}) = \underline{a}_i \underline{X}$ ($i = 1, \dots, n$) be independent linear forms with real algebraic coefficients, and let $\underline{a}_1^*, \dots, \underline{a}_n^*$ be reciprocal to $\underline{a}_1, \dots, \underline{a}_n$. Let c_1, \dots, c_n be constants with (9.2), let $\delta > 0$, and let \mathcal{C} be the subset of $\{1, \dots, n\}$ of numbers i with

$$(11.1) \quad c_i + (\delta/2) \geq 0 \quad .$$

Clearly \mathcal{C} is nonempty.

LEMMA 11A. Suppose $Q > 1$, and let Π be the parallelepiped

$$(11.2) \quad |L_i(\underline{x})| = |\underline{a}_i \underline{x}| \leq Q^{c_i} \quad (1 \leq i \leq n) \quad .$$

Let $\lambda_1, \dots, \lambda_n$ be the successive minima of Π , let $\underline{g}_1, \dots, \underline{g}_n$ be independent integer points with $\underline{g}_j \in \lambda_j \Pi$ ($1 \leq j \leq n$), and let $\underline{g}_1^*, \dots, \underline{g}_n^*$ be the reciprocal points. Suppose we have

$$(11.3) \quad \lambda_{n-1} < Q^{-\delta} \quad ,$$

and suppose there is an $i \in \mathcal{C}$ having

$$(11.4) \quad \underline{a}_i^* \underline{g}_n^* \neq 0 \quad .$$

Then

$$(11.5) \quad Q^{C_1} \leq |\underline{g}_n^*| \leq Q^{C_2}$$

if $Q \geq C_3$. Here $C_i = C_i(\delta; L_1, \dots, L_n; c_1, \dots, c_n) > 0$.

Proof. By Mähler's theory of reciprocal parallelepipeds (Theorem 4A of Chapter IV), applied to the reciprocal bases

$$Q^{-c_i} \underline{a}_i \quad (1 \leq i \leq n) \quad \text{and} \quad Q^{c_i} \underline{a}_i^* \quad (1 \leq i \leq n) \quad ,$$

we have

$$(11.6) \quad |a_{i\bar{n}}^* g_n^*| << \lambda_n^{-1} Q^{-c_i} << Q^{-\delta(n-1)-c_i} \quad (1 \leq i \leq n),$$

since $\lambda_n^{-1} << \lambda_1 \dots \lambda_{n-1} < Q^{-\delta(n-1)}$ by (11.3). (The constants in $<<$ here and in the rest of this section depend only on L_1, \dots, L_n). Since $|c_i| \leq 1$ and since $a_{\bar{1}}^*, \dots, a_{\bar{n}}^*$ are independent, it follows that $|g_n^*| << Q^{1-\delta(n-1)}$. In particular we have $|g_n^*| \leq Q^{C_2}$ if Q is large. As for the lower bound, we note that in view of (11.1) and (11.6) we have

$$(11.7) \quad |a_{i\bar{n}}^* g_n^*| << Q^{-\delta(n-1)+(\delta/2)} \leq Q^{-\delta/2}$$

for $i \in \mathcal{C}$. In particular, this is true for the integer i in (11.4). The components of $a_{\bar{i}}^*$ are algebraic and generate a number field, of degree Δ_i^* say. Since g_n^* has rational coordinates with denominator $E \leq (n!)^2$ (an easy consequence of (1.9) of Chapter IV), the norm of $a_{i\bar{n}}^* g_n^*$ satisfies $|\mathcal{N}(a_{i\bar{n}}^* g_n^*)| \gg 1$. The conjugates of $a_{i\bar{n}}^* g_n^*$ have absolute values $<< |g_n^*|$, and we obtain

$$|a_{i\bar{n}}^* g_n^*| \gg |g_n^*|^{1-\Delta_i^*} \geq |g_n^*|^{1-\Delta^*},$$

where Δ^* is the maximum of $\Delta_1^*, \dots, \Delta_n^*$. In conjunction with (11.7) this shows that $\Delta^* > 1$ and that $|g_n^*|^{\Delta^*-1} \gg Q^{\delta/2}$. Thus $|g_n^*| \geq Q^{C_1}$ with $C_1 = \frac{1}{3\delta}(\Delta^*-1)$ if Q is large, and the proof of Lemma 11A is complete.

Let M be the linear form $M = M\{g_1, \dots, g_{n-1}\}$. Then $M(\underline{X}) = \underline{m}\underline{X}$ where \underline{m} has coprime integer coordinates, and $g_n^* = (t/E)\underline{m}$ with integral t . In fact, t is a divisor of E since $g_n g_n^* = (t/E)g_n \underline{m} = 1$. Putting $F = E/t$ we have

$$(11.8) \quad \underline{g}_n^* = (1/F) \underline{m} \quad \text{with} \quad 1 \leq |F| \leq (n!)^2 .$$

It follows that $\underline{m} \ll \underline{g}_n^* \leq \underline{m}$, and $\underline{m} \gg \ll \underline{g}_n^*$, so that

$$(11.9) \quad Q^{C_4} \leq \underline{m} \leq Q^{C_5}$$

provided $Q \geq C_6$, where $C_i = C_i(\delta; L_1, \dots, L_n; c_1, \dots, c_n) > 0$ ($i = 4, 5, 6$) .

§12. The Next to Last Minimum.

THEOREM 12A. (Theorem on the Next to Last Minimum). Suppose

$L_1, \dots, L_n, a_1, \dots, a_n, a_1^*, \dots, a_n^*, c_1, \dots, c_n$ and δ and \mathfrak{S} are as in §11.

Again let $Q > 1$ and Π the parallelepiped (11.2) , with minima

$\lambda_1, \dots, \lambda_n$ having (11.3) . Let $\underline{g}_1, \dots, \underline{g}_n$ and $\underline{g}_1^*, \dots, \underline{g}_n^*$ be as above.

Then if $Q > Q_1(\delta; L_1, \dots, L_n; c_1, \dots, c_n)$, we have

$$(12.1) \quad a_{i \in \mathfrak{S}}^* \underline{g}_n^* = 0 \quad \text{for} \quad i \in \mathfrak{S} .$$

This theorem appears to be weak. Firstly, a very good rational approximation makes the first minimum λ_1 of some parallelepiped small, so that the hypothesis (11.3) seems to be far too strong. Secondly, the conclusion (12.1) appears to be weak, since ideally we would want to say that "for $Q > Q_1$ this is impossible." However, it will turn out that Theorem 12A gives us everything we need. So here is a situation where a result which looks temporary gives in fact the final answer.

Let us briefly discuss a special case pertaining to Corollary 1C.

Set $n = u+1$ and let L_1, \dots, L_n be given by (1.6) , and $c_1 = \dots = c_u = -1/u$, but $c_n = 1$. Here $\mathfrak{S} = \{n\}$ and $\underline{a}_n^* = (\alpha_1, \dots, \alpha_u, 1)$. By the linear independence of $1, \alpha_1, \dots, \alpha_u$, the relation (12.1) is impossible (since \underline{g}_n^* has rational components), so that in fact we do get the desired conclusion.

Proof of Theorem 12A. Let Ω be the set of $Q > 1$ such that the parallelepiped $\Pi = \Pi(Q)$ given by (11.2) has (11.3) but not (12.1). We have to show that Ω is bounded. We shall assume indirectly that Ω is unbounded, and we shall derive a contradiction.

A short reflection shows that we may suppose without loss of generality that the coefficients of L_1, \dots, L_n are real algebraic integers. Again write

$$q = n-1.$$

(i) Pick δ_1 with $0 < \delta_1 < 1$ and $\delta_1 \leq \delta$.

(ii) Pick $\varepsilon > 0$ small enough to satisfy

$$(12.2) \quad 16n^2 \varepsilon < \delta_1.$$

(iii) Define Δ as in §7 and choose an integer $m > 4\varepsilon^{-2} \log(2n\Delta)$, so that (7.1) holds. Define $w = w(m, \varepsilon)$ by (10.1).

(iv) Let $Q_1 \in \Omega$ be so large that (9.4) holds for $h=1$, and that $Q_1 \geq C_6$ and $Q_1^{wC_4^2/C_5} \geq 2^{3mq^2}$ and $Q_1^{wC_4^2/C_5} \geq D^{mq^2}$, where C_4, C_4, C_6 are the constants of §11, and D is described in §7.

(v) Successively choose Q_2, \dots, Q_m in Ω with

$$w \log Q_{h+1} \geq 2 \log Q_h \quad (h = 1, \dots, m-1).$$

Then $Q_1 < \dots < Q_m$, so that (9.4) holds for each h , and also

$$(12.3) \quad Q_h \geq C_6 \quad \text{and} \quad Q_h^{wC_4^2/C_5} \geq 2^{3mq^2} \quad (h = 1, \dots, m).$$

(vi) Let r_1 be an integer so large that $\varepsilon r_1 \log Q_1 \geq \log Q_m$.

(viii) For $h = 2, \dots, m$, put

$$r_h = \left[\frac{r_1 \log Q_1}{\log Q_h} \right] + 1.$$

We then have for $h = 2, \dots, m$,

$$\begin{aligned} r_1 \log Q_1 &\leq r_h \log Q_h \\ &\leq r_1 \log Q_1 + \log Q_h \\ &\leq (1+\varepsilon)r_1 \log Q_1, \end{aligned}$$

which is (9.5).

The linear forms L_1, \dots, L_n as well as ε, m and r_1, \dots, r_m satisfy the hypotheses of the Polynomial Theorem (Theorem 7A). Let $P(X_{11}, \dots, X_{1n}; \dots; X_{m1}, \dots, X_{mn})$ be the polynomial described there. Now also the hypotheses of Theorem 9A are satisfied, with δ_1 in place of δ . In particular, if for $1 \leq h \leq m$ the integer points g_{h1}, \dots, g_{hn} are linearly independent with $g_{hi} \in \lambda_i(Q_h) \Pi(Q_h)$, then (9.6) holds since $q = n-1$ and since $\lambda_q(Q_h) = \lambda_{n-1}(Q_h) < Q_h^{-\delta}$. Let M_1, \dots, M_m be the linear forms described in Theorem 9A. The conclusion is that

P has index at least $m\varepsilon$ with respect to $(M_1, \dots, M_m; r_1, \dots, r_m)$.

From (9.5) we have $r_{h+1} \log Q_{h+1} \leq (1+\varepsilon)r_h \log Q_h$, so that

$$\begin{aligned} \omega r_h &\geq \omega \frac{r_{h+1} \log Q_{h+1}}{(1+\varepsilon) \log Q_h} \quad (1 \leq h \leq m-1) \\ &\geq \frac{2}{1+\varepsilon} r_{h+1} \end{aligned}$$

from (v), and therefore $\omega r_h \geq r_{h+1}$, which is (10.2).

Since (12.1) is not true for Q_1, \dots, Q_m , the inequality (11.9) applies:

$$Q_h^{C_4} \leq \overline{M_h} \leq Q_h^{C_5} \quad (1 \leq h \leq m).$$

Therefore

$$\overline{M_h}^{r_h} \geq Q_h^{r_h C_4} \geq Q_1^{r_1 C_4} \geq \overline{M_1}^{r_1 C_4 / C_5},$$

whence

$$\left[M_h \right]^{r_h} \cong \left[M_1 \right]^{r_1 \Gamma} \quad (1 \leq h \leq m)$$

with $\Gamma = C_4/C_5$, which is (10.3). Furthermore, by (12.3),

$$\left[M_h \right]^{\omega \Gamma} \cong Q_h^{\omega \Gamma C_4} \cong 2^{3mq^2} \quad (1 \leq h \leq m),$$

which is (10.4). By Theorem 7A we have

$$\left[P \right] \leq_D r_1 + \dots + r_m \leq_D m r_1,$$

which by (iv) gives

$$\left[P \right]^q \leq_D m q^2 r_1 \leq_{Q_1} \omega r_1 C_4^2 / C_5 \leq \left[M_1 \right]^{\omega r_1 C_4 / C_5} = \left[M_1 \right]^{\omega r_1 \Gamma},$$

i.e. (10.5). All the hypotheses of Theorem 10B are satisfied, and we may conclude that

P has index at most ε with respect to $(M_1, \dots, M_m, r_1, \dots, r_m)$.

This contradicts the lower bound for the index given above.

§13. The Constancy of \underline{g}_n^* .

LEMMA 13A. Suppose $L_1, \dots, L_n, \underline{a}_1, \dots, \underline{a}_n, c_1, \dots, c_n$ and δ are as in §11, 12. For $Q > 1$ let Π be the parallelepiped (11.2), with minima $\lambda_1, \dots, \lambda_n$. Let $\underline{g}_1, \dots, \underline{g}_n$ and $\underline{g}_1^*, \dots, \underline{g}_n^*$ be as above. Let Ω be an unbounded set of reals > 1 such that (11.3) holds for $Q \in \Omega$. Then there is a fixed point \underline{h} and an unbounded subset Ω' of Ω such that

$$\underline{g}_n^* = \underline{h}$$

for every $Q \in \Omega'$.

Proof. Let \bar{Q} be a fixed large element of Ω . By Theorem 12A we have $\underline{a}_{i=1}^* \underline{g}_n^* = 0$ for $i \in \mathfrak{S}$, where $\underline{g}_n^* = \underline{g}_n^*(\bar{Q})$. Thus there is an integer point $\underline{h} \neq \underline{0}$ with coprime components such that

$$(13.1) \quad \underline{a}_{i=1}^* \underline{h} = 0 \quad \text{for } i \in \mathfrak{S}.$$

This point \underline{h} will be fixed in the sequel. Since \underline{h} is fixed we have $|\underline{a}_{i=1}^* \underline{h}| \ll 1$ ($1 \leq i \leq n$). For $i \notin \mathfrak{S}$ we have $c_i + (\delta/2) < 0$, so that

$$|\underline{a}_{i=1}^* \underline{h}| \ll 1 \leq Q^{-c_i - (\delta/2)} \quad (i \notin \mathfrak{S}).$$

So \underline{h} lies in the parallelepiped $\Pi^* = \Pi^*(Q)$ reciprocal to $\Pi = \Pi(Q)$. This holds for every large Q , in particular for $Q \in \Omega$. For $Q \in \Omega$ we have $\lambda_{n-1} < Q^{-\delta}$, whence $\lambda_2^* \gg Q^\delta$ by Theorem 4A of Chapter IV. So $\lambda_2^* > 1$ if $Q \in \Omega$ is large, and therefore every integer point in Π^* is a multiple of \underline{h} .

Again let $Q \in \Omega$ be large, and let $\underline{g}_n^* = \underline{g}_n^*(Q)$. By the argument leading to (11.8) we have $\underline{g}_n^* = (1/F)\underline{m}$ where $1 \leq |F| \leq (n!)^2$ and where \underline{m} has coprime integer coordinates. A priori F and \underline{m} may vary with Q . But now by (11.6), $F\underline{g}_n^*$ lies in Π^* , and so is proportional to \underline{h} . It follows that $\underline{m} = \pm \underline{h}$ and $\underline{g}_n^* = \pm (1/F)\underline{h}$. The number of possibilities for F and hence for \underline{g}_n^* is finite.

A slightly more general result will be needed:

LEMMA 13B. Suppose $L_i(\underline{X}) = \underline{a}_i \underline{X}$ ($i = 1, \dots, n$) are independent linear forms with real algebraic coefficients, and let $\delta > 0$. Let \mathfrak{P} be an unbounded set of reals $Q > 1$, and suppose that with each $Q \in \mathfrak{P}$ we are given positive A_1, \dots, A_n satisfying

$$(13.1) \quad A_1 A_2 \dots A_n = 1 \quad \text{and} \quad \max(A_1, \dots, A_n, A_1^{-1}, \dots, A_n^{-1}) \leq Q \quad .$$

Let $\Pi(Q)$ be the parallelepiped given by

$$|L_i(x)| \leq A_i \quad (i = 1, \dots, n) \quad .$$

Let $\lambda_1, \dots, \lambda_n$ be the minima of $\Pi = \Pi(Q)$, and define $\underline{g}_1, \dots, \underline{g}_n$, $\underline{g}_1^*, \dots, \underline{g}_n^*$ as before. Suppose that (11.3) holds for $Q \in \mathcal{P}$. Then there is a fixed point \underline{h} and an unbounded subset \mathcal{P}' of \mathcal{P} such that $\underline{g}_n^* = \underline{h}$ for $Q \in \mathcal{P}'$.

Proof. Set $A_i = Q^{c_i}$ ($i = 1, \dots, n$) and use a compactness argument, such as was used in the deduction of Theorem 1F from Lemma 3B. The reader is urged to do the details as an exercise.

§14. The Last Two Minima. Since the last minimum λ_n has $\lambda_n \gg 1$, it is clear that the hypothesis

$$(14.1) \quad \lambda_{n-1} < \lambda_n Q^{-\delta}$$

is weaker than (11.3). In general, an inequality (14.1) does not imply an inequality of the type (11.3). But we will see that if (14.1) holds for a parallelepiped Π , then something like (11.3) holds for a related parallelepiped Π' .

LEMMA 14A. The conclusion of Lemma 13A remains valid if (11.3) is replaced by the weaker hypothesis (14.1).

Proof. Given $Q \in \mathcal{Q}$ put $p_0 = (\lambda_1 \lambda_2 \dots \lambda_{n-2} \lambda_{n-1}^2)^{1/n}$ and

$$p_1 = p_0 / \lambda_1, \dots, p_{n-1} = p_0 / \lambda_{n-1}, \text{ but } p_n = p_0 / \lambda_{n-1} \quad .$$

Then equations (3.1), (3.2), (3.3) of Chapter IV hold. By Davenport's

Lemma (Theorem 3A of Chapter IV) there is a permutation t_1, \dots, t_n of $1, \dots, n$ such that the successive minima $\lambda'_1, \dots, \lambda'_n$ of the parallelepiped Π' given by

$$|L_i(\underline{x})| \leq Q^{c_i} p_{t_i}^{-1} \quad (= A_i, \text{ say}) \quad (1 \leq i \leq n)$$

satisfy

$$\lambda_j^{p_j} \ll \lambda'_j \ll \lambda_j^{p_j} \quad (1 \leq j \leq n) .$$

In particular, we have

$$(14.2) \quad \lambda'_{n-1} \ll p_{n-1} \lambda_{n-1} = p_0 \ll (\lambda_{n-1}/\lambda_n)^{1/n} \ll Q^{-\delta/n}$$

by (14.1) .

We observe that

$$\max(Q^{-c_1} |L_1(\underline{x})|, \dots, Q^{-c_n} |L_n(\underline{x})|) \geq Q^{-1} \max(|L_1(\underline{x})|, \dots, |L_n(\underline{x})|) \gg Q^{-1} |\underline{x}| \geq Q^{-1}$$

for every integer point $\underline{x} \neq \underline{0}$, and therefore $\lambda_1 \gg Q^{-1}$. On the other hand, the basis vectors $\underline{e}_1, \dots, \underline{e}_n$ satisfy $|L_1(\underline{e}_j)| \ll 1 \ll Q^{c_1+1}$ ($1 \leq i, j \leq n$) , so that $\lambda_n \ll Q$. We have

$$p_1 = \lambda_1^{-1} p_0 \ll \lambda_1^{-1} (\lambda_{n-1}/\lambda_n)^{1/n} \leq \lambda_1^{-1} \ll Q$$

and

$$p_n = \lambda_{n-1}^{-1} p_0 \gg \lambda_{n-1}^{-1} (\lambda_{n-1}/\lambda_n)^{1/n} \geq \lambda_n^{-1} \gg Q^{-1} .$$

Hence we obtain

$$Q^{-1} \ll p_n \leq p_{n-1} \leq \dots \leq p_1 \ll Q ,$$

and the quotients $A_i = Q^{c_i} p_{t_i}^{-1}$ satisfy $Q^{-2} \ll A_i \ll Q^2$. Since

$c_1 + \dots + c_n = 0$, and in view of $p_1 p_2 \dots p_n = 1$, we have for large Q ,

$$A_1 \dots A_n = 1 \text{ and } \max(A_1, \dots, A_n, A_1^{-1}, \dots, A_n^{-1}) \leq Q^3 .$$

Further, for every large Q in Ω we have

$$\lambda'_{n-1} < Q^{-(3\delta/4n)} = (Q^3)^{-\delta_2}$$

by (14.2) , with $\delta_2 = \delta/(4n)$. The hypotheses of Lemma 13B are thus satisfied with δ_2 in place of δ , and with \mathfrak{P} consisting of large Q^3 with $Q \in \Omega$. Let $\underline{g}'_1, \dots, \underline{g}'_n$ be independent integer points with $\underline{g}'_j \in \lambda'_j \Pi'$ ($1 \leq j \leq n$) , and let $\underline{g}'_1^*, \dots, \underline{g}'_n^*$ be their reciprocals. By Lemma 13B there is an unbounded subset \mathfrak{P}' of \mathfrak{P} such that for $Q \in \mathfrak{P}'$,

$$(14.3) \quad \underline{g}'_n^* = \underline{h}' ,$$

where \underline{h}' is fixed.

Let T^{n-1} be the subspace spanned by $\underline{g}_1, \dots, \underline{g}_{n-1}$. By the last assertion of Theorem 3A of Chapter IV (Davenport's Lemma), we see that for every integer point $\underline{g} \notin T^{n-1}$,

$$\max_{1 \leq i \leq n} (|L_i(\underline{g})| p_i^{-c_i} Q^{-c_i}) = \max_{1 \leq i \leq n} (|L_i(\underline{g})| A_i^{-1}) \gg \lambda'_n \gg 1 .$$

On the other hand, we have $\lambda'_{n-1} \ll Q^{-\delta/n}$ by (14.2) , so that $\underline{g}'_1, \dots, \underline{g}'_{n-1}$ must lie in T^{n-1} . Thus $\underline{g}_1, \dots, \underline{g}_{n-1}$ and $\underline{g}'_1, \dots, \underline{g}'_{n-1}$ span the same space, and \underline{g}_n^* and \underline{g}'_n^* are proportional. By the argument leading to (11.8) we see that $\underline{g}_n^* = F^{-1} \underline{m}$, $\underline{g}'_n^* = F'^{-1} \underline{m}$ where \underline{m} has coprime integer components and where $1 \leq |F|, |F'| \leq (n!)^2$. Since for $Q^3 \in \mathfrak{P}'$ the vector \underline{g}'_n^* is fixed by (14.3) , we get only finitely many possibilities for \underline{g}_n^* , so that indeed \underline{g}_n^* is fixed for an unbounded

subset Ω' of Ω .

§15. Proof of the Strong Subspace Theorem. Our arguments will depend heavily on Mähler's Theory of compound parallelepipeds developed in §7 of Chapter IV. We clearly may suppose that $L_1(\underline{X}) = \underline{a}_1 \underline{X}, \dots, L_n(\underline{X}) = \underline{a}_n \underline{X}$ have determinant 1. Further, for reasons of homogeneity, we may suppose that c_1, \dots, c_n satisfy not only (2.8), i.e., $c_1 + \dots + c_n = 0$, but also $|c_i| \leq 1/n$ ($i = 1, \dots, n$). Suppose that (3.2) holds for some arbitrarily large values of Q , say for Q in some unbounded set \mathfrak{N} . Put

$$(15.1) \quad p = n - d.$$

For $\sigma = \{i_1 < \dots < i_p\}$ in $C(n, p)$, define \underline{A}_σ by (6.10) of Chapter IV, and put

$$(15.2) \quad c_\sigma = \sum_{i \in \sigma} c_i.$$

Then

$$(15.3) \quad \sum_{\sigma \in C(n, p)} c_\sigma = 0 \quad \text{and} \quad |c_\sigma| \leq 1 \quad \text{for } \sigma \in C(n, p).$$

We now apply the theory of §7 of Chapter IV to the points $Q^{-c_i} \underline{a}_1, \dots, Q^{-c_n} \underline{a}_n$. The parallelepiped Π given by $|\underline{a}_i \underline{x}| \leq Q^{c_i}$ ($i = 1, \dots, n$) has a p -th pseudocompound $\Pi^{(p)}$ given by

$$(15.4) \quad |\underline{A}_\sigma \underline{X}^{(p)}| \leq Q^{c_\sigma} \quad (\sigma \in C(n, p)).$$

Denote the successive minima of $\Pi^{(p)}$ by v_1, \dots, v_ℓ where $\ell = \binom{n}{d} = \binom{n}{p}$.

It is clear that in (7.2) of Chapter IV we may take

$$\tau_\ell = \{n - p + 1, n - p + 2, \dots, n\} = \{d + 1, d + 2, \dots, n\} \quad \text{and}$$

$\tau_{l-1} = \{n-p, n-p+2, \dots, n\} = \{d, d+2, \dots, n\}$. By Theorem 7A of Chapter IV we have

$$v_l \ll \lambda_{d+1} \lambda_{d+2} \cdots \lambda_n \ll v_l$$

and

$$v_{l-1} \ll \lambda_d \lambda_{d+2} \cdots \lambda_n \ll v_{l-1} ,$$

and (3.2) yields $v_{l-1} \ll v_l Q^{-\delta}$. In particular, we have

$$(15.5) \quad v_{l-1} < v_l Q^{-\delta/2}$$

if Q in \mathfrak{N} is large.

Next we shall apply Lemma 14A to the vectors \underline{A}_σ and constants c_σ and to the parallelepiped $\Pi^{(p)}$. The conditions (15.3) replace (2.8) . The inequality (15.5) replaces (14.1) . Let $v_1^{(p)}, \dots, v_l^{(p)}$ be independent integer points in \mathbb{R}_p^n with $v_j^{(p)} \in v_j \Pi^{(p)}$ ($1 \leq j \leq l$) , and let $v_1^{(p)*}, \dots, v_l^{(p)*}$ be their reciprocals. The conclusion is that there is a fixed $\underline{H}^{(p)}$ such that

$$(15.6) \quad v_l^{(p)*} = v_l^{(p)*}(Q) = \underline{H}^{(p)}$$

for every Q in some unbounded subset \mathfrak{N}' of \mathfrak{N} .

Let $\underline{g}_1, \dots, \underline{g}_n$ be independent integer points of \mathbb{R}^n with $\underline{g}_j \in \lambda_j \Pi$ ($1 \leq j \leq n$) , and for $\tau = \{j_1 < \dots < j_p\}$ in $C(n, p)$ put $\underline{G}_\tau = \underline{g}_{j_1} \wedge \dots \wedge \underline{g}_{j_p}$. By Theorem 7A of Chapter IV we see that

$$|\underline{A}_{\sigma=\tau}| \ll v_{l-1} Q^{c_\sigma}$$

for σ, τ in $C(n, p)$ with $\tau \neq \tau_l = \{d+1, d+2, \dots, n\}$. Since

$v_{l-1} \ll v_l Q^{-\delta}$, this implies that the $l-1$ vectors \underline{G}_τ with $\tau \neq \tau_l$

span the same subspace as $v_{\underline{1}}^{(p)}, \dots, v_{\underline{l-1}}^{(p)}$. Hence the vectors $(G_{\underline{\tau}_l})^*$ and $v_{\underline{l}}^{(p)*}$ are proportional, and in view of (6.12) of Chapter IV the vectors $(G_{\underline{\tau}_l}^*)_{\underline{l}}$ and $v_{\underline{l}}^{(p)*}$ are proportional. Thus for every Q in \mathfrak{N}' we have

$$(G_{\underline{\tau}_l}^*)_{\underline{l}} = g_{\underline{d+1}}^* \wedge \dots \wedge g_{\underline{n}}^* = \lambda H_{\underline{\tau}}^{(p)}.$$

By Lemma 6C of Chapter IV this implies that there is a fixed subspace S^* of dimension $p = n-d$ such that $g_{\underline{d+1}}^*, \dots, g_{\underline{n}}^*$ lies in S^* when $Q \in \mathfrak{N}'$. Let S^d be the orthogonal complement to S^* . Whenever $g_{\underline{d+1}}^*, \dots, g_{\underline{n}}^*$ lie in S^* , the points $g_{\underline{1}}, \dots, g_{\underline{d}}$ lie in S^d . Hence for every Q in \mathfrak{N}' , the points $g_{\underline{1}}, \dots, g_{\underline{d}}$ lie in S^d .

VII. Norm Form Equations.

References: Borevich and Shafarevich (1966),
Schmidt (1971c, 1972).

§1. Norm Form Equations. In this chapter we shall require some knowledge about algebraic number fields. Throughout, K will be an algebraic number field of degree k , and \mathfrak{N} will denote the norm from K to the field \mathbb{Q} of rationals. There are k isomorphic embeddings $\varphi_1, \dots, \varphi_k$ of K into the complex numbers; denote the image of an element α of K under φ_i by $\alpha^{(i)}$. We will always tacitly assume that φ_1 is the identity map, so that $\alpha^{(1)} = \alpha$, and we shall say that $\alpha^{(1)} = \alpha, \alpha^{(2)}, \dots, \alpha^{(k)}$ are the conjugates of α . In this notation $\mathfrak{N}(\alpha) = \alpha^{(1)} \dots \alpha^{(k)}$. Given a linear form

$$(1.1) \quad M(\underline{X}) = M(X_1, \dots, X_n) = \alpha_1 X_1 + \dots + \alpha_n X_n,$$

with coefficients in K we write

$$\mathfrak{N}(M(\underline{X})) = \prod_{i=1}^k M^{(i)}(X_1, \dots, X_n) = \prod_{i=1}^k (\alpha_1^{(i)} X_1 + \dots + \alpha_n^{(i)} X_n).$$

A form $N(\underline{X})$ of the type $N(\underline{X}) = \mathfrak{N}(M(\underline{X}))$ will be called a norm form.

For example, if $\vartheta = \sqrt[4]{2}$ and $K = \mathbb{Q}(\vartheta)$, and if $M(X_1, X_2, X_3) = X_1 + \vartheta X_2 + \vartheta^2 X_3$, then

$$\mathfrak{N}(M(\underline{X})) = X_1^4 - 2X_2^4 + 4X_3^4 - 4X_1^2 X_3^2 + 8X_1 X_2^2 X_3.$$

A norm form is a form of degree k with rational coefficients. These coefficients will be rational integers if $\alpha_1, \dots, \alpha_n$ are integers in K .

A norm form equation is an equation of the type

$$(1.2) \quad \mathfrak{N}(M(\underline{x})) = a,$$

where $a \neq 0$ is a given rational number, and one is interested in solutions $\underline{x} = (x_1, \dots, x_n)$ with rational integer components. As \underline{x} runs through integer points, $M(\underline{x})$ will run through a subset \mathfrak{M} of K . This subset \mathfrak{M} is clearly a module in K , i.e. it is a nonempty subset which is closed under addition and subtraction, it is closed under multiplication by elements of \mathbb{Z} , the rational integers, and there is a fixed rational integer $z \neq 0$ such that $z\mu$ is an algebraic integer for every $\mu \in \mathfrak{M}$. The equation (1.2) may be written in the form

$$(1.3) \quad \mathfrak{N}(\mu) = a, \quad$$

and we are seeking solutions $\mu \in \mathfrak{M}$.

LEMMA 1A. A module \mathfrak{M} in K has a basis. That is, there is an n in $1 \leq n \leq k$ and there are elements β_1, \dots, β_n in \mathfrak{M} , linearly independent over \mathbb{Q} , such that every $\mu \in \mathfrak{M}$ may uniquely be written as

$$\mu = \beta_1 y_1 + \dots + \beta_n y_n$$

with y_1, \dots, y_n in \mathbb{Z} .

For a proof, see e.g. the Corollary on p. 85 of Borevich and Shafarevich (1966). The number n depends only on \mathfrak{M} and is called the rank of \mathfrak{M} . In view of the lemma we may restrict ourselves to forms M whose coefficients $\alpha_1, \dots, \alpha_n$ are linearly independent over \mathbb{Q} .

LEMMA 1B. The norm form

$$\mathfrak{N}(\alpha_1 X_1 + \dots + \alpha_n X_n)$$

is irreducible over the rationals if and only if $K = \mathbb{Q}(\alpha_2/\alpha_1, \dots, \alpha_n/\alpha_1)$.

Proof. Since the form equals $\mathfrak{N}(\alpha_1)\mathfrak{N}(X_1 + (\alpha_2/\alpha_1)X_2 + \dots + (\alpha_n/\alpha_1)X_n)$,

we may suppose without loss of generality that $\alpha_1 = 1$. Now if $L = \mathbb{Q}(\alpha_2, \dots, \alpha_n)$, then

$$\mathfrak{N}(X_1 + \alpha_2 X_2 + \dots + \alpha_n X_n) = \mathfrak{N}_L(X_1 + \alpha_2 X_2 + \dots + \alpha_n X_n)^{[K:L]},$$

where \mathfrak{N}_L is the norm from L to \mathbb{Q} and $[K:L]$ is the degree of K over L . Hence $K = L$ is necessary for the form to be irreducible. It is also sufficient, for if $K = L$, then by the Theorem on the Primitive Element (Van der Waerden (1955), §43) we have $K = \mathbb{Q}(\beta)$ where $\beta = c_2 \alpha_2 + \dots + c_n \alpha_n$ for certain rational c_2, \dots, c_n . Since β is of degree k , the form $\mathfrak{N}(X + \beta Y)$ is irreducible, and hence, since $\mathfrak{N}(X + \beta Y) = \mathfrak{N}(X + \alpha_2 c_2 Y + \dots + \alpha_n c_n Y)$, the given form is a fortiori irreducible.

An irreducible form $F(X, Y)$ of degree k may be written as

$$F(X, Y) = c(X - \alpha^{(1)} Y) \dots (X - \alpha^{(k)} Y) = c \mathfrak{N}(X - \alpha Y)$$

where $\alpha^{(1)} = \alpha$, $\alpha^{(2)}, \dots, \alpha^{(k)}$ are the roots of $F(X, 1)$ and where \mathfrak{N} is the norm in $K = \mathbb{Q}(\alpha)$. Hence irreducible forms and irreducible norm forms in two variables are essentially the same. It follows that a "Thue equation" $F(x, y) = a$ can be written as a norm form equation $\mathfrak{N}(\alpha_1 x + \alpha_2 y) = a'$, and vice versa. By Thue's Theorem (Theorem 3A of Ch. V) an irreducible norm form equation in two variables and of degree $k \geq 3$ has only a finite number of solutions.

When $n \geq 3$, an irreducible form is in general not a norm form.

A module \mathfrak{M} is called full if it has rank $n = k$. A full module which contains the number 1 and which is a ring is called an order of the field K . In particular the algebraic integers in K form an order. If \mathfrak{O} is an order and $\mu \in \mathfrak{O}$, then $\mu^h \in \mathfrak{O}$ for $h = 1, 2, \dots$. Hence with z the rational integer mentioned above, $z\mu^h$ is an algebraic

integer for $h = 1, 2, \dots$, so that μ itself must be an algebraic integer[†]). Hence every order \mathfrak{O} is contained in the order of algebraic integers, which is therefore called the maximal order of K .

The units in an order \mathfrak{O} are the divisors of 1, and it is easily seen that $\varepsilon \in \mathfrak{O}$ is a unit precisely if $N(\varepsilon) = \pm 1$. The units form a group.

LEMMA 1C. Let \mathfrak{O} be an order in the field K . Then its group of units is infinite except when $K = \mathbb{Q}$ or K is imaginary quadratic.

This is an immediate consequence of Dirichlet's Unit Theorem when extended to orders. (See Borevich and Shafarevich, Theorem 5 on p. 112).

Given a full module \mathfrak{M} , write $\mathfrak{O}_{\mathfrak{M}}$ for the set of $\lambda \in K$ such that $\lambda\mathfrak{M} \subseteq \mathfrak{M}$, i.e. that $\lambda\mu \in \mathfrak{M}$ for every $\mu \in \mathfrak{M}$. Then $\mathfrak{O}_{\mathfrak{M}}$ is easily seen to be an order in K ; it is called the coefficient ring of \mathfrak{M} . Write $\mathfrak{G}_{\mathfrak{M}}$ for the group of the units in $\mathfrak{O}_{\mathfrak{M}}$ which have norm 1. It is a subgroup of index 1 or 2 in the group of all the units of $\mathfrak{O}_{\mathfrak{M}}$. Hence it is infinite except when K is rational or imaginary quadratic. Now if $\mu \in \mathfrak{M}$ is a solution of (1.3), then also $\varepsilon\mu$ is a solution for $\varepsilon \in \mathfrak{G}_{\mathfrak{M}}$. Thus if \mathfrak{M} is a full module, and if K is not exceptional as indicated above, then given a solution of (1.3), we have infinitely many solutions.

But full modules are not the only ones for which (1.3) may have infinitely many solutions. Suppose L is a subfield of K and \mathfrak{M}_0 is a submodule of \mathfrak{M} which is proportional to a full module in L , i.e. $\mathfrak{M}_0 = \sigma\mathfrak{B}$ where $\sigma \neq 0$ is fixed and \mathfrak{B} is a full L -module. Unless L is exceptional, there will be infinitely many $\lambda \in \mathfrak{B}$ which satisfy

[†]) This follows from ideal theory. A more direct argument is that the ring $\mathbb{Z}[\mu]$ is a finite \mathbb{Z} -module (Zariski-Samuel (1958), §V.1).

$$\mathfrak{N}_L(\lambda) = b$$

for a certain b . But then $\sigma\lambda$ lies in \mathfrak{M} and satisfies

$$\mathfrak{N}(\sigma\lambda) = \mathfrak{N}(\sigma)\mathfrak{N}_L(\lambda)^{[K:L]} = \mathfrak{N}(\sigma)b^{[K:L]} = a,$$

say. We therefore define a module \mathfrak{M} in K to be degenerate if it contains a submodule which is proportional to a full module in some subfield L of K , where L is neither rational nor imaginary quadratic. If \mathfrak{M} is degenerate, then for certain values of a we have infinitely many solutions of (1.3) with $\mu \in \mathfrak{M}$.

We have proved (the easy) half of

THEOREM 1D. Let \mathfrak{M} be a module in K . There exists an a such that (1.3), i.e.

$$\mathfrak{N}(\mu) = a,$$

has infinitely many solutions $\mu \in \mathfrak{M}$, if and only if \mathfrak{M} is degenerate.

In the next sections we will formulate and prove a result which is stronger than Theorem 1D. In sections 10-14 we will discuss equations of the type

$$\mathfrak{N}(L(\underline{x})) = P(\underline{x}).$$

§2. Full Modules.

Much stronger than Lemma 1C is

LEMMA 2A. Let \mathfrak{O} be an order in K and \mathfrak{G} the group of units of \mathfrak{O} which have norm 1. Then

- (i) \mathfrak{G} is a subgroup of finite index in the group \mathfrak{G}^K of units of K , i.e. the units of the ring of integers in K .

(ii) There is a constant $\gamma_1 = \gamma_1(\mathfrak{D})$ such that for $\mu \in K$ there is an $\varepsilon \in \mathfrak{E}$ with

$$|(\varepsilon\mu)^{(i)}| \leq \gamma_1 |\mu|^{1/k} \quad (i = 1, \dots, k) .$$

Proof. Both assertions are easy consequences of Dirichlet's Unit Theorem for orders.

Now let \mathfrak{M} be a full module, $\mathfrak{O}_{\mathfrak{M}}$ its coefficient ring, $\mathfrak{E}_{\mathfrak{M}}$ the group of units of $\mathfrak{O}_{\mathfrak{M}}$ with norm 1. Write

$$(2.1) \quad \mu \mathfrak{E}_{\mathfrak{M}}$$

for the set of elements $\mu\varepsilon$ with $\varepsilon \in \mathfrak{E}_{\mathfrak{M}}$. If μ is a solution of (1.3), then so is every element of (2.1). Hence we call (2.1) a family of solutions of (1.3). By the second assertion of the lemma, every family of solutions contains a solution μ having

$$|\mu^{(i)}| \leq \gamma_1 |a|^{1/k} \quad (i = 1, \dots, k) .$$

Since there are only finitely many μ in \mathfrak{M} satisfying these inequalities, there are only finitely many families of solutions. Moreover, a representative of each family may be found in a finite number of steps. Thus the equation (1.3) can be completely solved if \mathfrak{M} is a full module.

§3. An Example.

It will be illuminating to study an example of a degenerate module.

Let K be the field $K = \mathbb{Q}(\sqrt{2}, \sqrt{3})$ and \mathfrak{M} the module with basis $\sqrt{2}, \sqrt{3}, \sqrt{6}$. We shall express this by writing $\mathfrak{M} = \{x\sqrt{2} + y\sqrt{3} + z\sqrt{6}\}$.

This module is not full in K . Let L_1, L_2, L_3 be the following subfields of K :

$$L_1 = \mathbb{Q}(\sqrt{2}) \quad , \quad L_2 = \mathbb{Q}(\sqrt{3}) \quad , \quad L_3 = \mathbb{Q}(\sqrt{6}) \quad ,$$

and let $\mathfrak{M}_1, \mathfrak{M}_2, \mathfrak{M}_3$ be the submodules of \mathfrak{M} given by

$$\mathfrak{M}_1 = \{y\sqrt{3} + z\sqrt{6}\} \quad , \quad \mathfrak{M}_2 = \{x\sqrt{2} + z\sqrt{6}\} \quad , \quad \mathfrak{M}_3 = \{x\sqrt{2} + y\sqrt{3}\} \quad .$$

The module \mathfrak{M}_1 is proportional to a full module in L_1 : We have

$\mathfrak{M}_i = \sigma_i \mathfrak{M}'_i$ ($i = 1, 2, 3$) where $\sigma_1 = \sqrt{3}$, $\sigma_2 = \sqrt{2}$, $\sigma_3 = 1/\sqrt{2}$ and where $\mathfrak{M}'_1 = \{y + z\sqrt{2}\}$, $\mathfrak{M}'_2 = \{x + z\sqrt{3}\}$, $\mathfrak{M}'_3 = \{2x + y\sqrt{6}\}$, so that \mathfrak{M}'_i is full in L_i ($i = 1, 2, 3$).

We now keep i fixed for the moment and study solutions of

$$(3.1i) \quad \mathfrak{N}(\mu_i) = a \quad \text{with} \quad \mu_i \in \mathfrak{M}_i \quad .$$

Clearly μ_i is a solution precisely if $\mu_i = \sigma_i \mu'_i$ where μ'_i is a solution of

$$\mathfrak{N}(\mu'_i) = a \mathfrak{N}(\sigma_i)^{-1} \quad \text{with} \quad \mu'_i \in \mathfrak{M}'_i \quad .$$

We have $\mathfrak{N}(\mu'_i) = (\mathfrak{N}_i(\mu'_i))^2$ where \mathfrak{N}_i is the norm from L_i to \mathbb{Q} , and our equation becomes

$$(3.2i) \quad \mathfrak{N}_i(\mu'_i) = \pm \sqrt{a} \mathfrak{N}(\sigma_i)^{-1} \quad \text{with} \quad \mu'_i \in \mathfrak{M}'_i \quad .$$

Since $\mathfrak{N}(\sigma_i)$ is easily seen to be a rational square, a necessary condition for the existence of a solution μ'_i is that a is a rational square.

Since \mathfrak{M}'_i is full in L_i , all the solutions of (3.2i) lie in finitely many families $v_{i1}\mathfrak{E}_i, \dots, v_{ir_i}\mathfrak{E}_i$, where \mathfrak{E}_i is the full[†] group of units in $\mathfrak{O}_{\mathfrak{M}'_i} = \mathfrak{O}_{L_i}^{(*)}$ and is infinite. Hence all the solutions μ_i of (3.1i) lie in the families

$$(3.3i) \quad \mu_{i1}\mathfrak{E}_i, \dots, \mu_{ir_i}\mathfrak{E}_i$$

[†]) We here use the full group of units because of the \pm sign in (3.2i).

^{*}) I.e. the coefficient ring of \mathfrak{M}'_i in L_i .

with $\mu_{ij} = \sigma_1 v_{ij}$ ($j = 1, \dots, r_1$) .

It will follow from Theorem 4B below that all but finitely many solutions of (1.3) lie in one of the submodules $\mathfrak{M}_1, \mathfrak{M}_2, \mathfrak{M}_3$. Hence all but finitely many solutions lie in one of the families (3.3i) with $i = 1, 2$, or 3 . In particular, (1.3) has only finitely many solutions unless a is a rational square. All but finitely many solutions of the diophantine equation

$$x\sqrt{2} + y\sqrt{3} + z\sqrt{6} = a$$

have $xyz = 0$.

In the special case when $a = 36$, the solutions of (3.2i) are easily seen to consist of the single family $v_1 \mathfrak{E}_1$ where $v_1 = \sqrt{2}$, $v_2 = \sqrt{3}$, $v_3 = 6 + 2\sqrt{6}$. Furthermore, \mathfrak{E}_1 consists of $\pm \varepsilon_1^n$ ($n = 0, \pm 1, \dots$) where $\varepsilon_1 = 1 + \sqrt{2}$, $\varepsilon_2 = 2 + \sqrt{3}$, $\varepsilon_3 = 5 + 2\sqrt{6}$. Hence all but finitely many solutions of

$$(3.4) \quad \mathfrak{N}(\mu) = 36 \quad \text{with } \mu \in \mathfrak{M}$$

lie in one of the families

$$\pm \sqrt{6}(1 + \sqrt{2})^n , \quad \pm \sqrt{6}(2 + \sqrt{3})^n , \quad \pm (3\sqrt{2} + 2\sqrt{3})(5 + 2\sqrt{6})^n$$

where $n = 0, \pm 1, \dots$.

The verification of these assertions concerning (3.4) is left as an Exercise.

§4. The General Case.

Let \mathfrak{M} be a module in K . Write \mathfrak{QM} for the set of products $q\mu$ with $q \in \mathbb{Q}$, $\mu \in \mathfrak{M}$; it is a vector space over \mathbb{Q} . Given a subfield L of K , let

$$\mathfrak{M}^L$$

consist of the elements μ of \mathfrak{M} with $L\mu \subseteq \mathbb{Q}\mathfrak{M}$, i.e. with $\lambda\mu \in \mathbb{Q}\mathfrak{M}$ for every $\lambda \in L$. It is easily seen that \mathfrak{M}^L is a submodule of \mathfrak{M} . We always have $\mathfrak{M}^{\mathbb{Q}} = \mathfrak{M}$, and $\mathfrak{M}^L \subseteq \mathfrak{M}^{L'}$ if $L \supseteq L'$. We have $\mathfrak{M}^K = \mathfrak{M}$ if \mathfrak{M} is full and $\mathfrak{M}^K = \{0\}$ otherwise.

Remarks. (i) If \mathfrak{M} has the basis μ_1, \dots, μ_n , if $\mu_1, \dots, \mu_n, \dots, \mu_k$ is a basis of K , if $\lambda_1, \dots, \lambda_\ell$ is a basis of L , and if

$$\lambda_i \mu_j = \sum_{t=1}^k c_{ijt} \mu_t \quad (1 \leq i \leq \ell, 1 \leq j \leq k),$$

then an element $z_1 \mu_1 + \dots + z_n \mu_n$ of \mathfrak{M} is in \mathfrak{M}^L precisely if

$$\sum_{j=1}^n z_j c_{ijt} = 0 \quad (1 \leq i \leq \ell, n < t \leq k).$$

Hence a basis for \mathfrak{M}^L can easily be constructed.

(ii) If $\lambda \in L$, $\mu \in \mathfrak{M}^L$, then $\lambda\mu \in \mathbb{Q}\mathfrak{M}$, so $z\lambda\mu \in \mathfrak{M}$ for a rational $z \neq 0$. If $\lambda' \in L$ then $\lambda'(z\lambda\mu) = (z\lambda'\lambda)\mu \in L\mathfrak{M}^L \subseteq \mathbb{Q}\mathfrak{M}$, and therefore $z\lambda\mu \in \mathfrak{M}^L$ and $\lambda\mu \in \mathbb{Q}\mathfrak{M}^L$. Thus

$$(4.1) \quad L\mathfrak{M}^L = \mathbb{Q}\mathfrak{M}^L,$$

and $\mathbb{Q}\mathfrak{M}^L$ is a vector space over L .

(iii) We have

$$(4.2) \quad L\mathfrak{M}^L \cap \mathfrak{M} = \mathfrak{M}^L,$$

because $\mu \in L\mathfrak{M}^L \cap \mathfrak{M}$ has $L\mu \subseteq L\mathfrak{M}^L = \mathbb{Q}\mathfrak{M}^L \subseteq \mathbb{Q}\mathfrak{M}$.

(iv) In the example of §3, K has the subfields $\mathbb{Q}, L_1, L_2, L_3, K$. We have $\mathfrak{M}^{\mathbb{Q}} = \mathfrak{M}$, $\mathfrak{M}^{L_1} = \mathfrak{M}_1$, $\mathfrak{M}^{L_2} = \mathfrak{M}_2$, $\mathfrak{M}^{L_3} = \mathfrak{M}_3$, and $\mathfrak{M}^K = \{0\}$.

Now let \mathbb{Q}_L^L be the ring of coefficients of \mathfrak{M}^L , i.e. the set of numbers λ in L such that $\lambda\mu \in \mathfrak{M}^L$ for every $\mu \in \mathfrak{M}^L$.

LEMMA 4A. Suppose $\mathbb{R}^L \neq \{0\}$. Then $\Omega_{\mathbb{R}}^L$ is an order in L .

Proof. $\Omega_{\mathbb{R}}^L$ clearly is a subring of L containing 1. Now let μ_0 be an element of \mathbb{R}^L distinct from 0. Then $\mu_0 \Omega_{\mathbb{R}}^L = \mathbb{R}^L$, whence $\Omega_{\mathbb{R}}^L \subseteq \mu_0^{-1} \mathbb{R}^L$. Now \mathbb{R}^L and therefore also $\mu_0^{-1} \mathbb{R}^L$ is a module in K , and hence there is a rational integer $z \neq 0$ such that $z \mu_0^{-1} \mathbb{R}^L$ and hence a fortiori $z \Omega_{\mathbb{R}}^L$ contains only algebraic integers. This shows that $\Omega_{\mathbb{R}}^L$ is a module in L .

It remains to be shown that $\Omega_{\mathbb{R}}^L$ is full in L . For every $\mu \in \mathbb{R}^L$ and every $\lambda \in L$ there is by Remark (ii) above a rational integer $z \neq 0$ with $z\lambda\mu \in \mathbb{R}^L$. Let μ_1, \dots, μ_r be a basis of \mathbb{R}^L . For every $\lambda \in L$ there are non-zero rational integers z_1, \dots, z_r with $z_i \lambda \mu_i \in \mathbb{R}^L$ ($i = 1, \dots, r$). Thus $z = z_1 \cdots z_r$ has $z\lambda\mu_i \in \mathbb{R}^L$ ($i = 1, \dots, r$), and this implies that $z\lambda\mu \in \mathbb{R}^L$ for every μ of \mathbb{R}^L . Therefore $z\lambda \in \Omega_{\mathbb{R}}^L$, and since λ was arbitrary in L and since $z = z(\lambda) \neq 0$, the module $\Omega_{\mathbb{R}}^L$ is full in L .

Let \mathcal{G}^L be the group of units of the ring of integers in L . Let $\mathcal{G}_{\mathbb{R}}^L$ be the group of elements ε in $\Omega_{\mathbb{R}}^L$ with $\mathfrak{N}(\varepsilon) = 1$. (Here as always \mathfrak{N} denotes the norm from K to \mathbb{Q} , and hence $\mathcal{G}_{\mathbb{R}}^L = \mathcal{G}_{\mathbb{R}}^{K,L}$ depends on K as well as on L and \mathbb{R} .) By Lemma 2A the group $\mathcal{G}_{\mathbb{R}}^L$ is a subgroup of finite index in \mathcal{G}^L .

If the equation $\mathfrak{N}(\mu) = a$ has a solution μ in \mathbb{R}^L , then every member of $\mu \mathcal{G}_{\mathbb{R}}^L$ is a solution. We shall call $\mu \mathcal{G}_{\mathbb{R}}^L$ a family of solutions, or more precisely an (\mathbb{R}, L) -family of solutions.

THEOREM 4B. The solutions of (1.3) are contained in finitely many families of solutions.

Thus for every subfield L of K there are finitely many (\mathbb{R}, L) -

families of solutions, such that μ is a solution of (1.3) precisely if it lies in one of these families for some L .

Let us again look at the example of §3. The subfields of $K = \mathbb{Q}(\sqrt{2}, \sqrt{3})$ are $K, L_1, L_2, L_3, \mathbb{Q}$. Since $\mathfrak{M}^K = \{0\}$, the solutions of (1.3) are contained in a finite number of (\mathfrak{M}, L_1) -, (\mathfrak{M}, L_2) -, (\mathfrak{M}, L_3) -, and $(\mathfrak{M}, \mathbb{Q})$ -families. Since each $(\mathfrak{M}, \mathbb{Q})$ -family consists of only two elements $\mu, -\mu$, all but finitely many solutions belong to an (\mathfrak{M}, L_i) -family, and in particular they lie in one of the submodules $\mathfrak{M}_1, \mathfrak{M}_2$, or \mathfrak{M}_3 .

If (1.3) has infinitely many solutions, then by Theorem 4B we must have $\mathfrak{M}^L \neq \{0\}$ for some field L which is not rational or imaginary quadratic. But then if $\sigma \neq 0$ is in \mathfrak{M}^L and if $\lambda_1, \dots, \lambda_\ell$ is a basis of the field L , we have $z_1 \lambda_1 \sigma, \dots, z_\ell \lambda_\ell \sigma$ in \mathfrak{M}^L for suitable nonzero rationals z_1, \dots, z_ℓ , so that \mathfrak{M}^L contains the module $\sigma \mathfrak{Q}$, where \mathfrak{Q} is the full module in L with basis $z_1 \lambda_1, \dots, z_\ell \lambda_\ell$. Thus \mathfrak{M} contains a submodule (namely $\sigma \mathfrak{Q}$) which is proportional to a full module in L , and hence \mathfrak{M} is degenerate. Therefore Theorem 4B implies Theorem 1D.

It is possible to construct in a finite number of steps a basis for \mathfrak{M}^L and a set of generators of $\mathfrak{S}_{\mathfrak{M}}^L$ for every subfield L of K . But Theorem 4B is ineffective, since in general we have no method to construct a finite number of families of solutions which contain all the solutions. This is a consequence of the fact that our method of proof depends on the preceding chapter, which in turn is a development of the non-effective method of Thue, Siegel, and Roth. In particular, if \mathfrak{M} is degenerate, we have no effective method to decide whether a particular equation (1.3) (with particular a) has infinitely many solutions. However, by combining the present method with Baker's method one can show that if the rank of \mathfrak{M} is at most 5, then one can determine in a finite

number of steps whether there are infinitely many solutions. For this, as well as the concept of maximal families of solutions (which are shown to be finite in number), see Schmidt (1972).

§5. Induction on the rank of \mathfrak{M} .

We shall prove Theorem 4B by induction on the rank of \mathfrak{M} . If \mathfrak{M} has rank 1, then its elements are $x\mu_1$, where $\mu_1 \neq 0$ is fixed and x runs through the rational integers. The equation $\mathfrak{R}(\mu) = a$ with $\mu \in \mathfrak{M}$ becomes $\mathfrak{R}(x\mu_1) = x^k \mathfrak{R}(\mu_1) = a$ where k is the degree of K . If k is odd and if there is a solution $\mu = x\mu_1$ at all, then it is the only solution and forms an (\mathbb{R}, \mathbb{Q}) -family by itself. If k is even and if there is a solution $\mu = x\mu_1$, then there are altogether two solutions μ and $-\mu$, which form a single (\mathbb{R}, \mathbb{Q}) -family, since now $\mathbb{Q}_{\mathfrak{M}}^{\mathbb{Q}}$ consists of $+1, -1$.

In what follows we shall assume that rank $\mathfrak{M} = n > 1$, and that Theorem 4B is true for modules of rank less than n . The following lemma will be useful.

LEMMA 5A. Suppose L is a subfield of K and \mathbb{G}^L is the group of units of the ring of integers of L . Let $\mathbb{G}_1^L, \mathbb{G}_2^L$ be subgroups of finite index in \mathbb{G}^L . Then for every $\kappa \in K$ there are finitely many elements $\kappa_1, \dots, \kappa_r$ in \mathbb{G}_1^L such that $\kappa \mathbb{G}_1^L$ is contained in the union of the sets

$$(5.1) \quad \kappa_1 \mathbb{G}_2^L, \dots, \kappa_r \mathbb{G}_2^L.$$

Proof. The intersection $\mathbb{G}_0^L = \mathbb{G}_1^L \cap \mathbb{G}_2^L$ is a subgroup of finite index in \mathbb{G}_1^L . Hence \mathbb{G}_1^L is the union of finitely many cosets $\varepsilon_1 \mathbb{G}_0^L, \dots, \varepsilon_r \mathbb{G}_0^L$ with $\varepsilon_i \in \mathbb{G}_1^L$ ($i = 1, \dots, r$), and hence it is contained in the union of the sets $\varepsilon_1 \mathbb{G}_2^L, \dots, \varepsilon_r \mathbb{G}_2^L$. This implies that $\kappa \mathbb{G}_1^L$ is contained in the

union of the sets (5.1) with $\kappa_i = \kappa \varepsilon_i \in \mathbb{K}_1^L$ ($i = 1, \dots, r$).

COROLLARY 5B. Suppose \mathbb{M} is a module in K and \mathbb{M}' is a submodule of \mathbb{M} . Let L be a subfield of K . Then every (\mathbb{M}', L) -family of solutions is contained in the union of finitely many (\mathbb{M}, L) -families of solutions.

Proof. Let $\mu_{\mathbb{M}}'^L$ be an (\mathbb{M}', L) -family of solutions. It is contained in a union $\mu_1^L \cup \dots \cup \mu_r^L$ where $\mu_1 \in \mu_{\mathbb{M}}'^L \subseteq \mathbb{M}'^L \subseteq \mathbb{M}^L$.

Given subfields A, B of K , write AB for their compositum, i.e. the smallest subfield of K containing both A and B .

LEMMA 5C. For any two subfields A, B of K , we have

$$(\mathbb{M}^A)^B = \mathbb{M}^{AB}.$$

Moreover, if $A \subseteq B$, then

$$(5.2) \quad \mathbb{M}_A^A \subseteq \mathbb{M}_A^B \quad \text{and} \quad \mathbb{M}_A^A \subseteq \mathbb{M}_A^B.$$

Proof. If $\mu \in \mathbb{M}^{AB}$, then $\mu \in \mathbb{M}^A$ and

$$B\mu \subseteq A B \mathbb{M}^{AB} = \mathbb{M}^{AB} \subseteq \mathbb{M}^A,$$

so that $\mu \in (\mathbb{M}^A)^B$. Conversely, if $\mu \in (\mathbb{M}^A)^B$ and if $\alpha \in A$, $\beta \in B$, then

$$\alpha \beta \mu \in \alpha B \mu \subseteq \alpha \mathbb{M}^A \subseteq A \mathbb{M}^A = \mathbb{M}^A \subseteq \mathbb{M}.$$

Since every element of AB is of the type $\alpha_1 \beta_1 + \dots + \alpha_t \beta_t$ with $\alpha_i \in A$, $\beta_i \in B$, we get $AB\mu \subseteq \mathbb{M}$, and in conjunction with $\mu \in \mathbb{M}$ this yields $\mu \in \mathbb{M}^{AB}$.

As for the second assertion of the lemma, suppose that $A \subseteq B$ and

$\alpha \in \mathfrak{S}_{\mathfrak{M}}^A$. Now if $\mu \in \mathfrak{M}^B$, then $\mu \in \mathfrak{M}^A$, and $\alpha\mu \in \mathfrak{M}^A \subseteq \mathfrak{M}$ and $\alpha\mu \in \mathfrak{M}^B$, so that $\alpha\mu \in \mathfrak{M}^B \cap \mathfrak{M} = \mathfrak{M}^B$. This proves that $\alpha \in \mathfrak{S}_{\mathfrak{M}}^B$, so that we have the first relation in (5.2). The second relation is an immediate consequence of the first.

§6. Linear Inequalities in a Simplex.

LEMMA 6A. Suppose a_1, \dots, a_n are positive reals and $a \geq 0$, $b > 0$.

For $r = 1, \dots, n$, put

$$f(r) = (a_1 + \dots + a_r)/r$$

and

$$H(r) = (a_1 + \dots + a_n)b - (a + nb)f(r).$$

Denote the maximum of $f(r)$ by f_1 , and set $H_1 = (a_1 + \dots + a_n)b - (a + nb)f_1$.

Then the minimum of

$$H(y_1, \dots, y_n) = a_1 y_1 + \dots + a_n y_n$$

in the set of points (y_1, \dots, y_n) with

$$(6.1) \quad y_1 + \dots + y_n + a \geq 0 \quad \text{and} \quad y_1 \leq \dots \leq y_n \leq b$$

is H_1 .

Moreover, if $r_1 < \dots < r_g$ are the values of r with $f(r) = f_1$, and if $g < n$, let f_2 be the maximum of $f(r)$ for r distinct from r_1, \dots, r_g . Put $H_2 = (a_1 + \dots + a_n)b - (a + nb)f_2$. Then every (y_1, \dots, y_n) with (6.1) has

$$(6.2) \quad \begin{aligned} & H(y_1, \dots, y_n) - H_1 \\ & \geq (a + nb)^{-1} (H_2 - H_1) ((y_{r_1} - y_1) + (y_{r_2} - y_{r_1+1}) + \dots \\ & + (y_{r_g} - y_{r_{g-1}+1}) + (b - y_{r_g+1})) , \end{aligned}$$

where we put $y_{n+1} = b$ if $r_g = n$.

Proof. The lemma may be considered as an exercise in Linear Programming. Since $H(y_1, \dots, y_n)$ decreases and the right hand side of (6.2) does not decrease if we replace (y_1, \dots, y_n) by $(y_1 - \delta, \dots, y_n - \delta)$ for some $\delta > 0$, we may restrict ourselves to (y_1, \dots, y_n) with

$$(6.3) \quad y_1 + \dots + y_n + a = 0 \quad \text{and} \quad y_1 \leq \dots \leq y_n \leq b.$$

This set of n -tuples constitutes a simplex in the hyperplane

$$y_1 + \dots + y_n + a = 0:$$

The points p_1, \dots, p_n given by

$$p_r = (b, \dots, b) - \frac{a+nb}{r} \underbrace{(1, 1, \dots, 1, 0, \dots, 0)}_{\leftarrow r \rightarrow}$$

lie on the hyperplane $y_1 + \dots + y_n + a = 0$. Given a point

$\underline{y} = (y_1, \dots, y_n)$ on this hyperplane put $y_{n+1} = b$ and observe that

$$\sum_{r=1}^n (y_{r+1} - y_r) r = -y_1 - \dots - y_n + nb = a + nb$$

and that

$$\sum_{r=1}^n (y_{r+1} - y_r) \underbrace{(1, \dots, 1, 0, \dots, 0)}_{\leftarrow r \rightarrow} = (b - y_1, \dots, b - y_n).$$

Thus

$$\begin{aligned} \sum_{r=1}^n (y_{r+1} - y_r) r p_r &= (a+nb)(b, \dots, b) + (a+nb)(y_1 - b, \dots, y_n - b) \\ &= (a+nb)\underline{y}, \end{aligned}$$

so that

$$\underline{y} = \sum_{r=1}^n r \frac{y_{r+1} - y_r}{a + nb} p_r.$$

The sum of the coefficients here is 1, and the coefficients are non-

negative precisely if \underline{y} satisfies (6.3). Hence (6.3) defines a simplex on the hyperplane $y_1 + \dots + y_n + a = 0$ with vertices

$\underline{p}_1, \dots, \underline{p}_n$. Since

$$H(\underline{p}_r) = (a_1 + \dots + a_n)b - (a + nb)f(r) = H(r) \quad (1 \leq r \leq n),$$

it is clear that $H_1 = \min(H(1), \dots, H(n))$ is the minimum of $H(\underline{y})$ with \underline{y} in (6.3), hence is the minimum with \underline{y} satisfying (6.1).

Now suppose that $f(r) = f_1$ for $r = r_1, \dots, r_g$, and that $f(r) \leq f_2 < f_1$ otherwise. Then $H(\underline{p}_r) \geq H_2$ for r distinct from r_1, \dots, r_g , and

$$\begin{aligned} H(\underline{y}) &\geq \sum_{r=1}^n r \frac{y_{r+1} - y_r}{a + nb} H_1 + \sum_{\substack{r=1 \\ r \neq r_1, \dots, r_g}}^n r \frac{y_{r+1} - y_r}{a + nb} (H_2 - H_1) \\ &\geq H_1 + \frac{H_2 - H_1}{a + nb} ((y_{r_1} - y_1) + (y_{r_2} - y_{r_1+1}) + \dots + (y_{r_g} - y_{r_{g-1}+1}) + (b - y_{r_g+1})). \end{aligned}$$

§7. Construction of a field L .

Let \mathfrak{M} be a module of rank n in K . The elements of \mathfrak{M} are the values of a linear form $M(\underline{X}) = \mu_1 X_1 + \dots + \mu_n X_n$ as the variables run through the integers. The coefficients μ_1, \dots, μ_n lie in K and are linearly independent over \mathbb{Q} .

There are k isomorphisms of K into the complex numbers. As in §1 the images of an element α of K under these isomorphisms are denoted by $\alpha^{(1)}, \dots, \alpha^{(k)}$, and as in §1 we suppose that $\alpha^{(1)} = \alpha$. We shall need the conjugate linear forms

$$M^{(i)}(\underline{X}) = \mu_1^{(i)} X_1 + \dots + \mu_n^{(i)} X_n \quad (1 \leq i \leq k).$$

LEMMA 7A. The forms $M^{(1)}, \dots, M^{(k)}$ have rank n , i.e. there are n but not more than n linearly independent forms among them.

Proof. Complete μ_1, \dots, μ_n to a basis of K , i.e. choose μ_{n+1}, \dots, μ_k such that μ_1, \dots, μ_k is a basis of K . It is then well known that the $(k \times k)$ -matrix $\mu_j^{(i)}$ ($1 \leq i, j \leq k$) has rank k , and hence the $(k \times n)$ -submatrix $\mu_j^{(i)}$ ($1 \leq i \leq k, 1 \leq j \leq n$) has rank n . Hence $M^{(1)}, \dots, M^{(k)}$ have rank n .

Given a subset S of $\{1, \dots, k\}$, we shall write $|S|$ for its cardinality, $r(S)$ for the rank of the set of linear forms $M^{(i)}$ with $i \in S$, and we shall put

$$q(S) = \begin{cases} r(S)/|S| & \text{if } S \text{ is non-empty,} \\ n/k & \text{if } S = \emptyset, \text{ i.e. if } S \text{ is empty.} \end{cases}$$

With this definition we have

$$(7.1) \quad q(\{1, 2, \dots, k\}) = n/k.$$

Now let ω be a primitive element of K over \mathbb{Q} and put $K^* = \mathbb{Q}(\omega^{(1)}, \dots, \omega^{(k)})$, so that K^* is the least normal extension of \mathbb{Q} containing K . Let G be the Galois group of K^* over \mathbb{Q} . An element φ of G permutes $\omega^{(1)}, \dots, \omega^{(k)}$, and hence induces a permutation $\hat{\varphi}$ of $\{1, 2, \dots, k\}$. Given a subset S of $\{1, 2, \dots, k\}$, the image $\hat{\varphi}(S)$ of S under $\hat{\varphi}$ is again such a subset, and we have $|\hat{\varphi}(S)| = |S|$, $r(\hat{\varphi}(S)) = r(S)$ and $q(\hat{\varphi}(S)) = q(S)$.

LEMMA 7B. Let q_0 be the minimum value of $q(S)$ for all subsets S of $\{1, 2, \dots, k\}$. Then

$$q_0 = n/k.$$

Moreover, there are integers ℓ, m with $\ell m = k$, a subset T with $|T| = m$ and elements $\varphi_1, \dots, \varphi_\ell$ of G such that $\{1, 2, \dots, k\}$ is the

disjoint union of the sets

$$(7.2) \quad \hat{\varphi}_1(T), \dots, \hat{\varphi}_\ell(T) ,$$

and a set S has $q(S) = q_0$ if and only if it is the union of some of the sets in (7.2) .

Proof. We have $q_0 \leq n/k$, and since $q(\emptyset) = q(\{1, 2, \dots, k\}) = n/k$, there is a non-empty set T with $q(T) = q_0$. Let T be a minimal non-empty set with this property, i.e. suppose that $q(T) = q_0$ and that $q(T') > q_0$ if $0 < |T'| < |T|$. Now let S be an arbitrary non-empty set with $q(S) = q_0$, and suppose $\varphi \in G$. We have

$$|S \cap \hat{\varphi}(T)| + |S \cup \hat{\varphi}(T)| = |S| + |\hat{\varphi}(T)| = |S| + |T| ,$$

$$r(S \cap \hat{\varphi}(T)) + r(S \cup \hat{\varphi}(T)) \leq r(S) + r(\hat{\varphi}(T)) = r(S) + r(T) .$$

Since $r(S) = q_0|S|$, $r(T) = q_0|T|$, and $r(S \cup \hat{\varphi}(T)) \geq q_0|S \cup \hat{\varphi}(T)|$, we have $r(S \cap \hat{\varphi}(T)) \leq q_0|S \cap \hat{\varphi}(T)|$. Hence either $S \cap \hat{\varphi}(T) = \emptyset$ or $S \cap \hat{\varphi}(T) \neq \emptyset$ and $q(S \cap \hat{\varphi}(T)) = q_0$. Since $|S \cap \hat{\varphi}(T)| \leq |T|$ and since T was minimal with $q(T) = q_0$, we must have $|S \cap \hat{\varphi}(T)| = |T|$ in this case. Hence either S and $\hat{\varphi}(T)$ are disjoint or S contains $\hat{\varphi}(T)$.

In particular, if φ, ψ are elements of G and if we set $S = \hat{\psi}(T)$, we see that $\hat{\varphi}(T), \hat{\psi}(T)$ are either disjoint or equal. It follows that $|T|$ is a divisor of k , and if we put

$$m = |T| , \quad \ell = k/m ,$$

then there are elements $\varphi_1, \dots, \varphi_\ell$ of G such that $\{1, 2, \dots, k\}$ is the disjoint union of the sets (7.2). Every non-empty set S with $q(S) = q_0$ is a union of sets from (7.2). Conversely, if S is a union of $s \geq 1$ of the sets (7.2), then $|S| = sm$, $r(S) \leq sr(T)$, whence

$q(S) \leq r(T)/m = q(T) = q_0$, so that $q(S) = q_0$. In particular, since $\{1, 2, \dots, k\}$ is the union of the sets (7.2), we have $q_0 = q(\{1, 2, \dots, k\}) = n/k$ by (7.1). Finally, since $q_0 = n/k$, all the statements made above remain true if S is the empty set, and the proof of Lemma 7B is complete.

It is clear that the set T in the lemma may be replaced by any of the sets in (7.2). One of these sets contains 1, so that we may assume that $1 \in T$. One of the sets in (7.2) is T itself, and we may suppose that $\hat{\varphi}_1(T) = T$. We may now replace φ_1 by the identity mapping. Then we have

$$(7.3) \quad 1 \in T, \text{ and } \varphi_1 \text{ is the identity mapping.}$$

Let G_L be the subgroup of G consisting of elements φ with $\hat{\varphi}(T) = T$. The left cosets of G_L in G are $\varphi_1 G_L = G_L, \varphi_2 G_L, \dots, \varphi_\ell G_L$, so that G_L has index ℓ in G . Let L be the fixed field of G_L , i.e. the subfield of K^* consisting of elements λ having $\varphi(\lambda) = \lambda$ for every $\varphi \in G_L$. Then L is a field of degree ℓ . Now let G_K be the subgroup of G consisting of elements φ with $\hat{\varphi}(1) = 1$, i.e. with $\varphi(\omega^{(1)}) = \omega^{(1)}$. It is clear that K is the fixed field of G_K . Now if $\varphi \in G_K$, then $\hat{\varphi}(1) = 1$, and since $1 \in T$, and $\hat{\varphi}$ permutes the sets (7.2), this implies that $\hat{\varphi}(T) = T$, whence that $\varphi \in G_L$. Hence $G_K \subseteq G_L$ and $L \subseteq K$.

The conjugates of an element $\lambda \in L$ over \mathbb{Q} are $\lambda = \varphi_1(\lambda), \dots, \varphi_\ell(\lambda)$. The conjugates of an element $\kappa \in K$ over \mathbb{Q} are $\varphi_i(\kappa^{(j)})$ with $1 \leq i \leq \ell$ and $j \in T$. The conjugates of an element $\kappa \in K$ over L are $\kappa^{(j)}$ with $j \in T$.

Suppose $T = \{t_1, \dots, t_m\}$. It will sometimes be convenient to write

$$(7.4) \quad \begin{aligned} \lambda^{[i]} &= \varphi_i(\lambda) & (i = 1, \dots, l) \text{ for } \lambda \in L, \\ \kappa^{[i,j]} &= \varphi_i(\kappa^{(t_j)}) & (i = 1, \dots, l; j = 1, \dots, m) \text{ for } \kappa \in K. \end{aligned}$$

Then every $\lambda \in L$ has

$$\lambda^{[i,j]} = \lambda^{[i]} \quad (i = 1, \dots, l; j = 1, \dots, m).$$

Now let P be a subfield of K . Let p be the degree of P (over Q) and s the degree of K over P , so that $ps = k$. Let $\omega = \omega^{(1)}, \omega^{(i_2)}, \dots, \omega^{(i_s)}$ be the conjugates of ω over P , and put $S = S(P) = \{1, i_2, \dots, i_s\}$.

LEMMA 7C. $q(S(P)) = q_0$ if and only if $P \subseteq L$.

Proof. We have $q(S(P)) = q_0$ precisely if $S(P)$ is a union of sets (7.2). Now if $S(P)$ is such a union, then we have $T \subseteq S(P)$ since $1 \in T$ and $1 \in S(P)$. For every $\pi \in P$ we have $\pi^{(i)} = \pi$ for $i \in S(P)$, whence $\pi^{(i)} = \pi$ for $i \in T$, whence $\varphi(\pi) = \pi^{\hat{\varphi}(1)} = \pi$ if $\hat{\varphi}(1) \in T$, whence $\varphi(\pi) = \pi$ if $\varphi \in G_L$, and π must lie in L . Thus $P \subseteq L$. Conversely, suppose that $P \subseteq L$. Let λ be a primitive element of L (over Q) and let $\lambda = \varphi_1(\lambda) = \varphi_{u_1}(\lambda), \varphi_{u_2}(\lambda), \dots, \varphi_{u_v}(\lambda)$ be the conjugates of λ over P . Then the conjugates of ω over P are $\varphi_{u_i}(\omega^{(j)})$ with $1 \leq i \leq v$ and $j \in T$, and $S(P) = \hat{\varphi}_{u_1}^{(T)} \cup \dots \cup \hat{\varphi}_{u_v}^{(T)}$, so that $q(S(P)) = q_0$ by Lemma 7B.

LEMMA 7D. A subfield P of K has $\mathfrak{M}^P = \mathfrak{M}$ if and only if $P \subseteq L$.

Proof. Suppose \mathfrak{M} contains r but not more than r linearly independent elements over P . Let $\delta_1, \dots, \delta_r$ be such elements. Then

$$M(X) = \delta_1 M_1(X) + \dots + \delta_r M_r(X),$$

where $M_1(\underline{X}), \dots, M_r(\underline{X})$ are linearly independent linear forms with coefficients in P . The elements $\delta_1, \dots, \delta_r$ can be extended to a basis $\delta_1, \dots, \delta_r, \dots, \delta_s$ of K over P , and since the matrix $(\delta_j^{(i)})$ ($i \in S(P)$, $1 \leq j \leq s$) is nonsingular, the matrix $(\delta_j^{(i)})$ ($i \in S(P)$, $1 \leq j \leq r$) has rank r . Hence $r(S(P)) = r$ and $q(S(P)) = r/s$.

Now if $\mathfrak{M} = \mathfrak{M}^P$, then since $\delta_i \in \mathfrak{M}$ ($i = 1, \dots, r$) it follows that $P\delta_i \in \mathfrak{QM}$, and hence there is a basis π_1, \dots, π_p of P over Q such that \mathfrak{M} contains the rp elements $\delta_i \pi_j$ ($1 \leq i \leq r$, $1 \leq j \leq p$), and $n = rp$. Conversely, if $n = rp$, then each of the forms $M_i(\underline{X})$ above represents a basis of P over Q , and hence there is a basis π_1, \dots, π_p of P such that \mathfrak{M} contains $\delta_i \pi_j$ ($1 \leq i \leq r$, $1 \leq j \leq p$), that in fact \mathfrak{QM} is spanned by these rp elements over Q , so that $\mathfrak{QM} = P\mathfrak{M}$ and $\mathfrak{M} = \mathfrak{M}^P$. Thus $\mathfrak{M} = \mathfrak{M}^P$ precisely if $n = rp = r(k/s) = kq(S(P))$, i.e. $q(S(P)) = n/k = q_0$. By Lemma 7C this holds if and only if $P \subseteq L$.

We note that the existence of a maximal subfield L of K with $\mathfrak{M}^L = \mathfrak{M}$ follows from Lemma 5C. But we now have more precise information on L .

§8. The Main Lemma.

For an element κ of our field K put

$$(8.1) \quad |\overline{\kappa}| = \max(|\kappa^{(1)}|, \dots, |\kappa^{(k)}|) \quad .$$

Suppose we are given the equation (1.3), i.e.

$$(8.2) \quad \mathfrak{M}(\mu) = \mu^{(1)} \dots \mu^{(k)} = a \quad \text{with } \mu \in \mathfrak{M} \quad .$$

LEMMA 8A. Suppose $\varepsilon > 0$. There are finitely many submodules $\mathfrak{M}_1, \dots, \mathfrak{M}_w$ of \mathfrak{M} , each of rank less than $n = \text{rank } \mathfrak{M}$, such that every
†) i.e. the value set of $M_i(\underline{x})$ with $\underline{x} \in \mathbb{Z}^n$ contains a basis.

solution μ of (8.2) either lies in one of these submodules, or satisfies

$$(8.3) \quad |\mu^{[i,h]}| < |\mu|^\varepsilon |\mu^{[i,j]}| \quad (1 \leq i \leq \ell; 1 \leq h, j \leq m).$$

Proof. We shall at first deal only with solutions μ having

$$(8.4) \quad |\mu^{(1)}| \leq \dots \leq |\mu^{(k)}|.$$

We may suppose that $\varepsilon < 1$. Pick $\delta > 0$ so small that

$$(8.5) \quad \delta < \varepsilon / (6kn^2).$$

The elements of \mathfrak{M} are the values of a linear form $M(\underline{x})$ as \underline{x} runs through the integer points.

There is a form of this type whose coefficients are linearly independent over the rationals. Hence the forms $M^{(1)}, \dots, M^{(k)}$ have rank n by Lemma 7A, and we have $|\underline{x}| \leq \gamma_1 |M(\underline{x})|$ with a constant γ_1 independent of \underline{x} . We are concerned with solutions \underline{x} of

$$(8.6) \quad M^{(1)}(\underline{x}) \dots M^{(k)}(\underline{x}) = a$$

with

$$(8.7) \quad |M^{(1)}(\underline{x})| \leq \dots \leq |M^{(k)}(\underline{x})|.$$

Put $i_0 = 0$. Let i_1 be the largest integer such that $M^{(1)}, M^{(2)}, \dots, M^{(i_1)}$ have rank 1. Let i_2 be the largest integer such that $M^{(1)}, M^{(2)}, \dots, M^{(i_2)}$ have rank 2; and so on. We obtain numbers $0 = i_0 < i_1 < \dots < i_n = k$. In view of (8.7) and since $M^{(i_j)}_{j-1+1}$ is a linear combination of $M^{(1)}, \dots, M^{(i_{j-1}+1)}$, we have

$$|M^{(i_j)}_{j-1+1}(\underline{x})| \leq \gamma_2 |M^{(i_{j-1}+1)}_{j-1+1}(\underline{x})| \quad (j = 1, \dots, n)$$

with constant γ_2 . Hence all but finitely many integer points \underline{x} with (8.7) have

$$(8.8) \quad |M^{(i_j)}(\underline{x})| |\underline{x}|^{-\delta} \leq |M^{(i_{j-1}+1)}(\underline{x})| \leq |M^{(i_{j-1}+2)}(\underline{x})| \leq \cdots \leq |M^{(i_j)}(\underline{x})| \\ (j = 1, \dots, n).$$

Also, all but finitely many integer points have

$$(8.9) \quad |M^{(k)}(\underline{x})| = |M^{(i_n)}(\underline{x})| \leq |\underline{x}|^{1+\delta} \quad \text{and} \quad 1 < |\underline{x}| \leq \overline{M(\underline{x})}^2,$$

and all but finitely many solutions of (8.6) have

$$(8.10) \quad |M^{(1)}(\underline{x}) \cdots M^{(k)}(\underline{x})| \leq |\underline{x}|^{\delta}.$$

Clearly $M^{(1)}, M^{(i_1+1)}, \dots, M^{(i_{n-1}+1)}$ are n linearly independent linear forms, so that by the Subspace Theorem (Th. 1F of Ch. VI),

$$(8.11) \quad |M^{(1)}(\underline{x}) M^{(i_1+1)}(\underline{x}) \cdots M^{(i_{n-1}+1)}(\underline{x})| \geq |\underline{x}|^{-\delta},$$

unless \underline{x} lies in one of a finite number of proper rational subspaces determined by the linear form M . As \underline{x} runs through the integer points of a proper rational subspace, $M(\underline{x})$ runs through a submodule of \mathfrak{M} of rank less than n . Hence if we disregard solutions $\mu = M(\underline{x})$ of (8.2), (8.4) which lie in a finite number of submodules $\mathfrak{M}_1, \dots, \mathfrak{M}_2$ of rank less than n , then (8.8), (8.9), (8.10) and (8.11) hold.

Assume now, until the last paragraph of this section, that there is a solution $\mu = M(\underline{x})$ of (8.2), (8.4) (and hence with \underline{x} satisfying (8.6), (8.7)), which is not in $\mathfrak{M}_1, \dots, \mathfrak{M}_2$. In what follows, \underline{x} will be a solution of (8.6), (8.7) with $M(\underline{x})$ outside of $\mathfrak{M}_1, \dots, \mathfrak{M}_2$, so that (8.8), (8.9), (8.10) and (8.11) will hold. Define $c_1 = c_1(\underline{x}), \dots, c_n = c_n(\underline{x})$ by

$$|M^{(i_j)}_{\underline{x}}| = |\underline{x}|^{c_j} \quad (j = 1, \dots, n) .$$

Then

$$(8.12) \quad c_1 \leq \dots \leq c_n \leq 1 + \delta$$

by (8.7) and (8.9), and

$$(8.13) \quad c_1 + \dots + c_n + \delta \geq 0$$

by (8.7) and (8.11). By (8.8) and (8.10) we have

$$\begin{aligned} & i_1 c_1 + (i_2 - i_1) c_2 + \dots + (i_n - i_{n-1}) c_n \\ (8.14) & \leq \delta + (i_1 - 1) \delta + (i_2 - i_1 - 1) \delta + \dots + (i_n - i_{n-1} - 1) \delta \\ & = \delta + (k - n) \delta \leq k \delta . \end{aligned}$$

We are going to apply Lemma 6A with $a = \delta$, $b = 1 + \delta$ and with

$a_1 = i_1 = i_1 - i_0$, $a_2 = i_2 - i_1, \dots, a_n = i_n - i_{n-1} = k - i_{n-1}$. Then $f(r) = i_r / r$ ($r = 1, \dots, n$). Let S_r be the subset of $\{1, 2, \dots, k\}$ given by

$$(8.15) \quad S_r = \{1, 2, \dots, i_r\} \quad (r = 1, \dots, n).$$

Then $|S_r| = i_r$, $r(S_r) = r$, $q(S_r) = 1/f(r)$, whence $f(r) \leq 1/q_0 = k/n$ by Lemma 7B, and in particular we have $f_1 \leq k/n$. By Lemma 6A, the minimum of $H(c_1, \dots, c_n)$ with (c_1, \dots, c_n) subject to (8.12) and (8.13) is $H_1 = k(1 + \delta) - (\delta + n(1 + \delta))f_1$. On the other hand, the minimum is $\leq k\delta$ by (8.14), so that $H_1 \leq k\delta$ and $f_1 n(1 + 2\delta) \geq f_1 (\delta + n(1 + \delta)) \geq k(1 + \delta) - k\delta = k$, whence

$$f_1 \geq (k/n)(1 + 2\delta)^{-1} > (k/n)(1 - 2\delta) > (k/n)(1 - (1/kn)) = (k/n) - (1/n^2)$$

by (8.5). Now f_1 is a rational number with denominator at most n ,

and the rational numbers with denominator at most n have mutual distances greater than $1/n^2$. Hence $f_1 \geq k/n$, and by what we said above we have

$$(8.16) \quad f_1 = k/n \quad \text{and} \quad H_1 = k(1+\delta) - (\delta + n(1+\delta))f_1 = -\delta k/n.$$

Suppose that $f(r) = f_1 = k/n$ for $r = r_1, \dots, r_g$ with $1 \leq r_1 < r_2 < \dots < r_g \leq n$, and for no other values of r . Since $f(n) = k/n$, we have $r_g = n$.

Now suppose that $g < n$. Since f_2 (as defined in Lemma 6A) is another rational number with denominator at most n , we have $f_2 < f_1 = (1/n^2)$. By (8.14) and (8.16) we have

$$H(c_1, \dots, c_n) - H_1 \leq k\delta + (\delta k/n) < 2k\delta,$$

and hence (6.2) yields (on noting that $r_g = n$),

$$(8.17) \quad (c_{r_1} - c_1) + (c_{r_2} - c_{r_1+1}) + \dots + (c_{r_g} - c_{r_{g-1}+1}) \\ \leq 2k\delta(a+nb)(H_2 - H_1)^{-1} = 2k\delta(f_1 - f_2)^{-1} < 2k\delta n^2 < \varepsilon/3$$

by (8.5). This inequality is trivially true if $g = n$, i.e. if $r_1 = 1, \dots, r_g = n$, and hence is generally true.

Putting $r_0 = 0$, we have $0 \leq c_{r_e} - c_{r_{e-1}+1} < \varepsilon/3$ ($e = 1, \dots, g$) and hence $|c_h - c_j| < \varepsilon/3$ if $r_{e-1} + 1 \leq h, j \leq r_e$. Put differently,

$$|M^{(i_h)}(\underline{x})| < \overline{|x|}^{\varepsilon/3} |M^{(i_j)}(\underline{x})| \quad \text{if} \quad r_{e-1} + 1 \leq h, j \leq r_e.$$

Combining this with (8.8) and observing that $\delta + (\varepsilon/3) < \varepsilon/2$, we obtain

$$|M^{(u)}(\underline{x})| < \overline{|x|}^{\varepsilon/2} |M^{(v)}(\underline{x})| \quad \text{if} \quad i_{r_{e-1}} + 1 \leq u, v \leq i_{r_e}$$

for some e in $1 \leq e \leq g$, which in turn by (8.9) yields

$$|M^{(u)}(\underline{x})| < \overline{M(\underline{x})}^\varepsilon |M^{(v)}(\underline{x})| \quad \text{if } i_{r_e-1} + 1 \leq u, v \leq i_{r_e}.$$

With S_r defined by (8.15), put

$$V_1 = S_{r_1}, V_2 = S_{r_2} \sim S_{r_1}, \dots, V_g = S_{r_g} \sim S_{r_{g-1}} = \{1, \dots, k\} \sim S_{r_{g-1}},$$

where \sim denotes the set theoretic difference. We have

$$|M^{(u)}(\underline{x})| < \overline{M(\underline{x})}^\varepsilon |M^{(v)}(\underline{x})| \quad \text{if } u, v \in V_e$$

for some e in $1 \leq e \leq g$. Thus $\mu = M(\underline{x})$ satisfies

$$|\mu^{(u)}| < \overline{\mu}^\varepsilon |\mu^{(v)}| \quad \text{if } u, v \in V_e.$$

Now $q(S_{r_e}) = 1/f(r_e) = 1/f_1 = q_0$ ($e = 1, \dots, g$), and hence each set S_{r_e} is the disjoint union of some of the sets (7.2). Hence also each set V_e ($e = 1, \dots, g$) is the union of some sets (7.2). Thus each set $\hat{\varphi}_i(T)$ of (7.2) ($i = 1, \dots, l$) is contained in one of the sets V_e , and we have

$$|\mu^{(u)}| < \overline{\mu}^\varepsilon |\mu^{(v)}| \quad \text{if } u, v \in \hat{\varphi}_i(T)$$

for some i . Now if again $T = \{t_1, \dots, t_m\}$ as in §7, we have

$$|\mu^{(\hat{\varphi}_i(t_h))}| < \overline{\mu}^\varepsilon |\mu^{(\hat{\varphi}_i(t_j))}| \quad \text{if } 1 \leq i \leq l \text{ and } 1 \leq h, j \leq m.$$

In the notation of (7.4) this becomes (8.3).

All this holds if μ is a solution of (8.2) with (8.4) and if μ does not lie in certain submodules $\mathfrak{M}_1, \dots, \mathfrak{M}_2$ of rank less than n . The same is true if (8.4) is replaced by any other ordering. Hence there is a finite number of submodules $\mathfrak{M}_1, \dots, \mathfrak{M}_w$, each of rank less than n , such that (8.3) holds for every solution μ of (8.2) which is outside of these submodules.

§9. Proof of the Main Theorem.

We shall apply Lemma 8A with some fixed ε in $0 < \varepsilon < 1$; say with $\varepsilon = 1/2$. By our inductive assumption (see §5), and since each of the modules $\mathfrak{M}_1, \dots, \mathfrak{M}_w$ of Lemma 8A has rank less than n , the solutions μ of $\mathfrak{M}(\mu) = a$ with $\mu \in \mathfrak{M}_i$ ($i = 1, \dots, w$) lie in finitely many families of solutions. By Corollary 5B, every (\mathfrak{M}_i, P) -family of solutions where P is a subfield of K , is contained in the union of finitely many (\mathfrak{M}, P) -families of solutions. Hence there are finitely many families of solutions

$$(9.1) \quad \mu_1 \begin{smallmatrix} P_1 \\ \mathfrak{G}_1 \\ \mathfrak{M} \end{smallmatrix}, \dots, \mu_z \begin{smallmatrix} P_z \\ \mathfrak{G}_z \\ \mathfrak{M} \end{smallmatrix}$$

with subfields P_1, \dots, P_z of K and with $\mu_i \in \mathfrak{M}^{P_i}$ ($i = 1, \dots, z$), such that every solution μ of (1.3) satisfies (8.3) unless it lies in one of these families.

Since $\mathfrak{M} = \mathfrak{M}^L$ by Lemma 7D, we have $\mathfrak{M}^{P_i L} = (\mathfrak{M}^L)^{P_i} = \mathfrak{M}^{P_i}$ by Lemma 5C. Thus $\mu_i \in \mathfrak{M}^{P_i L}$ ($i = 1, \dots, z$), and since $\mathfrak{G}_i^{P_i} \subseteq \mathfrak{G}_i^{P_i L}$ ($i = 1, \dots, z$), by Lemma 5C again, the family $\mu_i \begin{smallmatrix} P_i \\ \mathfrak{G}_i \\ \mathfrak{M} \end{smallmatrix}$ is contained in the family $\mu_i \begin{smallmatrix} P_i L \\ \mathfrak{G}_i \\ \mathfrak{M} \end{smallmatrix}$. Hence if we replace $\mu_i \begin{smallmatrix} P_i \\ \mathfrak{G}_i \\ \mathfrak{M} \end{smallmatrix}$ in (9.1) by $\mu_i \begin{smallmatrix} P_i L \\ \mathfrak{G}_i \\ \mathfrak{M} \end{smallmatrix}$ and change the notation back from $P_i L$ to P_i , we obtain families (9.1) where the fields P_i satisfy

$$(9.2) \quad L \subseteq P_i \subseteq K \quad (i = 1, \dots, z).$$

Now suppose that an (\mathfrak{M}, L) -family and an (\mathfrak{M}, P_i) -family have a common element μ . Then the (\mathfrak{M}, L) -family is $\mu \begin{smallmatrix} L \\ \mathfrak{G} \\ \mathfrak{M} \end{smallmatrix}$ and the (\mathfrak{M}, P_i) -family is $\mu \begin{smallmatrix} P_i \\ \mathfrak{G} \\ \mathfrak{M} \end{smallmatrix}$. Now in view of (9.2) and by Lemma 5C, we have $\mathfrak{G}^L \subseteq \mathfrak{G}^{P_i}$, so that the (\mathfrak{M}, L) -family is contained in the (\mathfrak{M}, P_i) -family.

Since $\mathfrak{M} = \mathfrak{M}^L$, every solution of (1.3) is contained in an (\mathfrak{M}, L) -family of solutions. If such an (\mathfrak{M}, L) -family intersects one of the

families (9.1), then it is contained in it. Therefore in order to prove Theorem 4B, it will suffice to show the following

LEMMA 9A. Every $\mathbb{Q}(L)$ -family of solutions of (1.3) which is disjoint from the families (9.1) has an element μ with

$$(9.3) \quad |\bar{\mu}| \leq (\gamma_1 |a|^{1/k})^{1/(1-\varepsilon)},$$

where $\gamma_1 = \gamma_1(K, \mathbb{Q})$.

Theorem 4B follows, since there are only finitely many elements μ in \mathbb{Q} with $|\bar{\mu}|$ below a given bound.

Proof of Lemma 9A. Suppose we are given an $\mathbb{Q}(L)$ -family of solutions which is disjoint from the families (9.1), and let μ be a member of the family for which $|\bar{\mu}|$ is least possible. Since it lies outside the families (9.1), the element μ satisfies (8.3).

We saw in §7 that $\varphi_1, \dots, \varphi_\ell$ are the isomorphisms of L into the complex numbers, and for $\lambda \in L$ we put $\lambda^{[i]} = \varphi_i(\lambda)$ ($i = 1, \dots, \ell$), so that $\lambda^{[1]}, \dots, \lambda^{[\ell]}$ are the conjugates of λ . We also introduced the notation $\mu^{[i,j]}$ for the conjugates of elements $\mu \in K$. Now given the μ of the last paragraph, write $\lambda = \mu^{[1,1]} \mu^{[1,2]} \dots \mu^{[1,m]}$. Then λ lies in L and

$$\lambda^{[i]} = \mu^{[i,1]} \mu^{[i,2]} \dots \mu^{[i,m]} \quad (i = 1, \dots, \ell).$$

We have

$$\mathfrak{N}(\mu) = \mathfrak{N}_L(\lambda) = \lambda^{[1]} \dots \lambda^{[\ell]}$$

where \mathfrak{N}_L is the norm from L to \mathbb{Q} .

By Dirichlet's unit theorem (see Lemma 2A), there is for every

$v \in L$ an element ε in the group \mathbb{G}^L of units of the ring of integers in L such that

$$|(\varepsilon v)^{[i]}| \leq \gamma_2 |\mathfrak{N}_L(v)|^{1/l} \quad (i = 1, \dots, l),$$

with a constant γ_2 depending on L only. Since the group $(\mathbb{G}_{\mathfrak{M}}^L)^m$ of elements $\eta = \varepsilon^m$ with $\varepsilon \in \mathbb{G}_{\mathfrak{M}}^L$ is of finite index in \mathbb{G}^L , there is also an element $\eta \in (\mathbb{G}_{\mathfrak{M}}^L)^m$ with $|(\eta v)^{[i]}| \leq \gamma_3 |\mathfrak{N}_L(v)|^{1/l} \quad (i = 1, \dots, l)$.

Here γ_3 depends only on K, L, \mathfrak{M} and hence (since K has only a finite number of subfields) only on K and \mathfrak{M} . In particular, we may apply this to $v = \lambda$. There is an $\varepsilon \in \mathbb{G}_{\mathfrak{M}}^L$ with

$$(9.4) \quad |(\varepsilon^m \lambda)^{[i]}| \leq \gamma_3 |\mathfrak{N}_L(\lambda)|^{1/l} = \gamma_3 |\mathfrak{N}(\mu)|^{1/l} = \gamma_3 |a|^{1/l} \quad (i = 1, \dots, l).$$

On the other hand, (8.3) remains valid if the right hand side is replaced by $|\overline{\mu}|^\varepsilon |\mu^{[i,1]} \dots \mu^{[i,m]}|^{1/m}$, and substituting $\varepsilon \mu$ for μ we obtain

$$|(\varepsilon \mu)^{[i,j]}| \leq |\overline{\varepsilon \mu}|^\varepsilon |(\varepsilon^m \lambda)^{[i]}|^{1/m} \quad \begin{matrix} (i = 1, \dots, l; \\ j = 1, \dots, m)^\dagger \end{matrix}.$$

By this and by (9.4), the element $\mu' = \varepsilon \mu$ satisfies

$$|\overline{\mu'}| = \max(|\mu'^{[1,1]}|, \dots, |\mu'^{[l,m]}|) \leq |\overline{\mu'}|^\varepsilon \gamma_3^{1/m} |a|^{1/(ml)}.$$

Since μ had the minimum value of $|\overline{\mu}|$ in our family, we get

$$|\overline{\mu}| \leq |\overline{\mu}|^\varepsilon \gamma_1 |a|^{1/k}$$

with $\gamma_1 = \gamma_3^{1/m}$, and (9.3) follows.

This completes the proof of Lemma 9A and hence of Theorem 4B.

§10. Equations $\mathfrak{N}(M(\underline{x})) = P(\underline{x})$

We saw in Theorem 3B of Ch. V that if $F(X, Y)$ is a form of

[†]) There should be no confusion of the unit ε and the exponent ε .

degree $k \geq 3$ without multiple factors and if $v < k - 2$, then there are only finitely many integer points $\underline{x} = (x, y)$ with

$$(10.1) \quad 0 < |F(x, y)| < |\underline{x}|^v.$$

Our aim here is to generalize this as far as possible to equations in n variables.

Let K be a number field of degree k . Suppose $K = \mathbb{Q}(\omega)$, and set $K^* = \mathbb{Q}(\omega^{(1)}, \dots, \omega^{(k)})$, so that K^* is a normal extension of the rationals. Further let G be the Galois group of K^* . We shall say that K is h times transitive if G is h times transitive on the set $\{\omega^{(1)}, \dots, \omega^{(k)}\}$, i.e. if $1 \leq h \leq k$ and if given distinct integers i_1, \dots, i_h between 1 and k , there is a $\varphi \in G$ with $\varphi(\omega^{(1)}) = \omega^{(i_1)}, \dots, \varphi(\omega^{(h)}) = \omega^{(i_h)}$. It is easily seen that this definition is independent of the choice of the primitive element ω of K .

THEOREM 10A. Suppose $n < k$ and $M(\underline{X}) = \alpha_1 X_1 + \dots + \alpha_n X_n$ has coefficients in K where K is $n - 1$ times transitive. Further assume that any n conjugates $M^{(i_1)}, \dots, M^{(i_n)}$ of M are linearly independent. Then for

$$(10.2) \quad v < k - n$$

there are only finitely many integer points \underline{x} with

$$|\mathfrak{N}(M(\underline{x}))| \leq |\underline{x}|^v.$$

COROLLARY 10B. Suppose $M(\underline{X})$ is as above, and $P(\underline{X})$ is a polynomial of total degree less than $k - n$. Then the diophantine equation

$$\mathfrak{N}(M(\underline{x})) = P(\underline{x})$$

has only finitely many solutions.

Remarks. (i) The theorem contains the assertion concerning (10.1)

in the case when $F(X, Y)$ is irreducible: For in that case, $F(X, Y)$ is, except for a constant factor, a norm form $\Re(M(\underline{X}))$ in 2 variables, as was seen in §1. In fact the form $M(\underline{X})$ may be taken to be of the type $M(\underline{X}) = X + \alpha Y$ where $K = Q(\alpha)$. It is always true that K is $1 = 2 - 1$ times transitive, and any two conjugates $M^{(i_1)}_1, M^{(i_2)}_2$ of M are linearly independent.

(ii) The hypotheses of Theorem 10A hold if K is n times transitive and if the coefficients $\alpha_1, \dots, \alpha_n$ of $M(\underline{X})$ are linearly independent over Q . For if $M^{(i_1)}_1, \dots, M^{(i_n)}_n$ were linearly dependent for some n -tuple i_1, \dots, i_n , then by the transitivity this would be true for any n -tuple, and $M^{(1)}, \dots, M^{(k)}$ would have rank less than n , in contradiction to Lemma 7A.

(iii) The condition (10.2) is best possible unless $n = 1$, or $n = 2$, and $M^{(1)}, \dots, M^{(k)}$ are complex conjugate in pairs.[†]

For either $n \geq 2$ and one of the forms, say $M^{(1)}$, is real. Then by Minkowski's Theorem on linear forms, there is for $Q > 0$ an integer point $\underline{x} \neq \underline{0}$ with

$$|M^{(1)}(\underline{x})| \leq \gamma Q^{1-n}, \quad |x_i| \leq Q \quad (i = 2, \dots, n),$$

where the coordinates were arranged so that $M^{(1)}, x_2, \dots, x_n$ are linearly independent. It follows that $|\underline{x}| \ll Q$, whence $|M^{(j)}(\underline{x})| \ll Q$ ($j = 2, \dots, k$), and $|\Re(M(\underline{x}))| \ll Q^{1-n} |\underline{x}|^{k-1}$, and

$$(10.3) \quad |\Re(M(\underline{x}))| \ll |\underline{x}|^{k-n}.$$

Since Q may be taken arbitrarily large, one gets infinitely many solutions of (10.3).

Or $n \geq 3$ and two of the forms, say $M^{(1)}, M^{(2)}$, are complex conjugate. Let R, I be the real part and the imaginary part of $M^{(1)}$.

[†] In the exceptional case $|\Re(M(\underline{x}))| \gg |\underline{x}|^k$, as is easily seen. Compare with the end of §12.

Again by Minkowski's Theorem, there is for $Q > 0$ an integer point $\underline{x} \neq \underline{0}$ with

$$|R(\underline{x})| \ll Q^{1-(n/2)}, \quad |I(\underline{x})| \ll Q^{1-(n/2)}, \quad |x_i| \leq Q \quad (i=3, \dots, n),$$

where the coordinates were arranged so that R, I, X_3, \dots, X_n are independent. It follows that $|M^{(1)}(\underline{x})| = |M^{(2)}(\underline{x})| \ll Q^{1-(n/2)}$ and $|\underline{x}| \ll Q$, whence $|M^{(j)}(\underline{x})| \ll Q$ ($j = 3, \dots, k$), and hence again (10.3) holds.

A form $F(\underline{X}) = F(X_1, \dots, X_n)$ is called decomposable if it is a product of linear forms with algebraic coefficients. In particular, a norm form or a product of norm forms is decomposable. Conversely it is easily seen that a decomposable form with rational coefficients is a product of a constant and one or several norm forms. Of course, a form in two variables is always decomposable, while forms in $n \geq 3$ variables are in general not.

THEOREM 10C. Suppose $1 < n \leq \ell$ and $2(\ell-1) < t$. Suppose $F(\underline{X}) = F(X_1, \dots, X_n)$ is a decomposable form of degree t with rational coefficients. Suppose that $F(\underline{X})$ is not divisible by a rational form of degree less than ℓ , and that any ℓ linear factors in the product $F(\underline{X}) = M_1(\underline{X}) \dots M_t(\underline{X})$ have rank n . Then for any v with

$$(10.4) \quad v < t - 2(\ell - 1),$$

there are only finitely many integer points \underline{x} having

$$(10.5) \quad |F(\underline{x})| < |\underline{x}|^v.$$

One obtains an obvious corollary concerning equations $F(\underline{x}) = P(\underline{x})$, where P is a polynomial of degree less than $t - 2(\ell - 1)$. When $n = \ell = 2$, our result is essentially the same as Theorem 3B of Ch. V.

THEOREM 10D*. Suppose the field K of degree k is primitive,
i.e. it has no non-trivial subfields. Suppose that

$$N(\underline{X}) = \mathfrak{N}(\alpha_1 X_1 + \dots + \alpha_n X_n)$$

is an irreducible norm form with respect to K . Here $\alpha_1, \dots, \alpha_n$
lie in K , and the norm refers to K). Now if $\alpha_1, \dots, \alpha_n$ are linearly
independent over Q , and if

$$(10.6) \quad v < k - \frac{n}{n-1} 2^{(n-2)/2} k^{(n-2)/(n-1)},$$

then there are only finitely many integer points \underline{x} with

$$|N(\underline{x})| < |\underline{x}|^v.$$

A proof of this theorem will not be given here. See Schmidt (1973).
 When $n = 2$, then (10.6) reduces to $v < k - 2$ again. When $n = 3$,
 the condition (10.6) becomes $v < k - (3/\sqrt{2})\sqrt{k}$.

In order to prove Theorems 10A and 10C we shall need

§11. Another Theorem on Linear Forms.

Consider linear forms $M(\underline{X})$ in $\underline{X} = (X_1, \dots, X_n)$ with real or
 complex coefficients. Call the form real if all its coefficients are
 real, otherwise call it complex. A system of linear forms M_1, \dots, M_t
 will be called symmetric if every complex form among them occurs as
 often as its complex conjugate.

THEOREM 11A. Suppose M_1, \dots, M_t is a symmetric system of linear
forms with algebraic coefficients. Given $\eta > 0$, the following two
conditions are equivalent.

(a) There are infinitely many integer points \underline{x} with

$$(11.1) \quad |M_1(\underline{x}) \dots M_t(\underline{x})| \leq \gamma_1 |\underline{x}|^{t-\eta}$$

where $\gamma_1 = \gamma_1(M_1, \dots, M_t, \eta)$.

(b) There is a d-dimensional rational subspace $S^d \neq \{0\}$ of \mathbb{R}^n and there is a symmetric (sub-)system of forms M_{i_1}, \dots, M_{i_m} with $1 \leq m \leq t$ and $i_1 < \dots < i_m$ whose restrictions to S^d have rank r with

$$(11.2) \quad r \leq dm/\eta \quad \text{and} \quad r < d .$$

This Theorem could equally well have been included in Chapter VI . The preceding sections of the present chapter, with the exception of §6, are not needed for the proof.

The part $(b) \Rightarrow (a)$ of the proof is[†]) relatively easy: We may suppose that M_1, \dots, M_m are a symmetric system whose rank r on S^d satisfies (11.2) . The real parts and imaginary parts of these forms will then also have rank r on S^d , hence on S^d will be linear combinations of certain real forms R_1, \dots, R_r . Suppose, say, that the linear forms $R_1, \dots, R_r, X_{r+1}, \dots, X_d$ are linearly independent on S^d . By applying Minkowski's linear forms theorem to the lattice of integer points in S^d , we obtain for given $Q > 0$ an integer point $\underline{x} \neq \underline{0}$ in S^d having

$$(11.3) \quad |R_i(\underline{x})| \ll Q^{r-d} \quad (1 \leq i \leq r) \quad \text{and} \quad |x_j| \ll Q^r \quad (r+1 \leq j \leq d) .$$

It follows that $|\underline{x}| \ll Q^r$ and $|M_i(\underline{x})| \ll Q^{r-d}$ for $1 \leq i \leq m$. Thus $|M_i(\underline{x})| \ll |\underline{x}|^{(r-d)/r}$ ($1 \leq i \leq m$) and

$$|M_1(\underline{x}) \dots M_t(\underline{x})| \ll |\underline{x}|^{(m/r)(r-d)+t-m} = |\underline{x}|^{t-(md/r)} \leq |\underline{x}|^{t-\eta} .$$

Since Q in (11.3) may be chosen arbitrarily large, we get infinitely many solutions of (11.1) , i.e. (a) holds.

†) The trivial case $r=0$ is excluded in what follows.

If there is an integer point $\underline{x} \neq \underline{0}$ having $M_j(\underline{x}) = 0$ for one of our forms M_j , then (b) holds with $S^d = S^1$ the subspace spanned by \underline{x} , with $r = 0$, and with the symmetric subsystem consisting of M_j alone if M_j is real, and of M_j and its complex conjugate if M_j is complex. Hence we may suppose that

$$(11.4) \quad M_1(\underline{x}) \dots M_t(\underline{x}) \neq 0$$

for nonzero integer points \underline{x} .

§12. Proof of the Theorem on Linear Forms.

It remains to do the implication (a) \Rightarrow (b). We start with the case when our linear forms are real. There is a smallest d in $1 \leq d \leq n$ and a rational subspace S^d such that for a certain constant γ_2 the relation

$$(12.1) \quad |M_1(\underline{x}) \dots M_t(\underline{x})| \leq \gamma_2 |\underline{x}|^{t-\eta}$$

has infinitely many solutions in integer points in S^d . Now $d = 1$ is possible only if $M_1(\underline{x}) \dots M_t(\underline{x}) = 0$ for some $\underline{x} \neq \underline{0}$, which is ruled out by (11.4). So $2 \leq d \leq n$. After rearrangement of the linear forms, we have infinitely many integer points \underline{x} in S^d with both
(12.1) and

$$(12.2) \quad 0 < |M_1(\underline{x})| \leq \dots \leq |M_t(\underline{x})|.$$

Suppose M_1, \dots, M_t have rank u on S^d . Let i_1 be the smallest number so that $M_{i_1} \neq 0$ on S^d ; in view of (11.4) we have $i_1 = 1$. Let i_2 be the smallest number so that M_{i_1}, M_{i_2} have rank 2 on S^d , and so forth. We obtain^{†)} numbers $1 = i_1 < i_2 < \dots < i_u$.

^{†)} The choice of the i_j is similar to but not the same as the one in §8.

The points \underline{x} of S^d with (12.2) have

$$(12.3) \quad |M_1(\underline{x}) \dots M_t(\underline{x})| \gg \ll |M_1(\underline{x})|^{i_2-1} |M_{i_2}(\underline{x})|^{i_3-i_2} \dots |M_{i_{u-1}}(\underline{x})|^{i_u-i_{u-1}} |M_{i_u}(\underline{x})|^{t+1-i_u}.$$

If $u = d$, then the Subspace Theorem (Th. 1F of Ch. VI) may be applied to the d linear forms M_{i_1}, \dots, M_{i_u} on S^d . Since S^d was chosen minimal, it follows that given $\varepsilon > 0$ we have $|M_{i_1} \dots M_{i_u}| \geq |\underline{x}|^{-\varepsilon}$ for all but finitely many of our points \underline{x} . If $u < d$, we may augment M_{i_1}, \dots, M_{i_u} by $d-u$ of the variables to get d linearly independent forms on S^d , and we may conclude that

$$(12.4) \quad |M_{i_1}(\underline{x}) \dots M_{i_u}(\underline{x})| |\underline{x}|^{d-u} \geq |\underline{x}|^{-\varepsilon}.$$

Thus (12.4) holds for every $u \leq d$. Put

$$v = \begin{cases} u & \text{if } u < d, \\ u-1 = d-1 & \text{if } u = d. \end{cases}$$

It is clear that $|M_{i_u}(\underline{x})| \geq |\underline{x}|^{1+\varepsilon}$ if $|\underline{x}|$ is large. Thus no matter whether $u < d$ or $u = d$, we have

$$(12.5) \quad |M_{i_1}(\underline{x}) \dots M_{i_v}(\underline{x})| |\underline{x}|^{d-v} \geq |\underline{x}|^{-2\varepsilon}.$$

Determine c_1, \dots, c_v by

$$|M_{i_j}(\underline{x})| = |\underline{x}|^{c_j} \quad (1 \leq j \leq v),$$

so that c_1, \dots, c_v are functions of \underline{x} . If $|\underline{x}|$ is large, then $c_j \leq 1+\varepsilon$, and from (12.2), (12.5) we obtain

$$(12.6) \quad c_1 + \dots + c_v + d - v + 2\varepsilon \geq 0 \quad \text{and} \quad c_1 \leq \dots \leq c_v \leq 1 + \varepsilon.$$

On the other hand, (12.1) and (12.3) yield

$$(i_2 - 1)c_1 + (i_3 - i_2)c_2 + \dots + (i_v - i_{v-1})c_{v-1} + (t+1 - i_v)c_v \leq t - \eta + \varepsilon$$

if $u < d$, and

$$(i_2 - 1)c_1 + (i_3 - i_2)c_2 + \dots + (i_{v+1} - i_v)c_v + (t+1 - i_{v+1})(1 - \varepsilon) \leq t - \eta + \varepsilon$$

if $u = d$. Setting

$$i_{v+1} = t + 1 \quad \text{if } u < d,$$

we get in both cases

$$(12.7) \quad H \leq i_{v+1} - 1 - \eta + 2t\varepsilon,$$

where $H = H(c_1, \dots, c_v) = (i_2 - 1)c_1 + \dots + (i_{v+1} - i_v)c_v$.

By Lemma 6A, applied with $a = d - v + 2\varepsilon$, $b = 1 + \varepsilon$, we see that the minimum of H with c_1, \dots, c_v subject to (12.6) is

$$H_1 = (a_1 + \dots + a_v)b - (a + vb)f_1 = (i_{v+1} - 1)(1 + \varepsilon) - (d + v\varepsilon + 2\varepsilon)f_1,$$

where $f_1 = f(r)$ for some r in $1 \leq r \leq v$ and where

$$f(r) = (a_1 + \dots + a_r)/r = (i_{r+1} - 1)/r.$$

Combining this with (12.7) we get

$$(i_{v+1} - 1)(1 + \varepsilon) - (d + v\varepsilon + 2\varepsilon)(i_{r+1} - 1)/r \leq i_{v+1} - 1 - \eta + 2t\varepsilon,$$

and therefore

$$(d/r)(i_{r+1} - 1) \geq \eta - 3dt\varepsilon - 2t\varepsilon.$$

We have $r \leq v < d$. Put $m = i_{r+1} - 1$ and consider the m linear

forms $M_1, \dots, M_{i_{r+1}-1}$. These forms have rank r on S^d , with $md/r \geq \eta - 5d\epsilon$. So for every $\epsilon > 0$ there is a number m and there are m linear forms among M_1, \dots, M_t which have a rank r satisfying $r < d$ and $md/r \geq \eta - 5d\epsilon$. The number of possibilities for m, r and the m forms among M_1, \dots, M_t is finite, and hence there is a number m and there are m forms among M_1, \dots, M_t whose rank r on S^d satisfies (11.2). Theorem 1B is proved for real linear forms.

In the general case, for complex forms, note that a complex form may be written as $M = M^R + iM^I$ where M^R, M^I are real forms.

The proof in the complex case up to (12.2) is just as the real case. In particular, S^d is a minimal subspace of \mathbb{R}^n such that an inequality like (12.1) has infinitely many solutions in integer points in S^d . We note that the forms can be ordered in such a way that complex conjugate forms stand next to each other in (12.2).

We are going to choose real forms R_1, \dots, R_t as follows. If M_j is real, set $R_j = M_j$. If M_j and M_{j+1} are complex conjugate, then infinitely many of our points $\underline{x} \in S^d$ with (12.1), (12.2) have $|M_j^R(\underline{x})| \geq |M_j^I(\underline{x})|$, or there are infinitely many points with the reversed inequality. Without loss of generality suppose, say, that we have $|M_j^R(\underline{x})| \geq |M_j^I(\underline{x})|$. We further may suppose that $M_j^R(\underline{x})M_j^I(\underline{x}) \geq 0$. Set

$$R_j = M_j^R \quad \text{and} \quad R_{j+1} = M_j^R + M_j^I,$$

except if the restriction of M_j^R to S^d is a linear combination of M_1, \dots, M_{j-1} , but the restriction of M_j^I is not such a linear combination. In the exceptional case set

$$R_j = M_j^R + M_j^I \quad \text{and} \quad R_{j+1} = M_j^R.$$

This choice of R_j, R_{j+1} guarantees that if the restriction of R_{j+1} to S^d is independent of R_1, \dots, R_j , then the restriction of R_j is independent of R_1, \dots, R_{j-1} . In either case we have

$$|M_j(\underline{x})| = |M_{j+1}(\underline{x})| \leq 2|R_j(\underline{x})| \leq 4|R_{j+1}(\underline{x})| \leq 8|M_j(\underline{x})|,$$

so that

$$(12.8) \quad |R_j(\underline{x})| \leq 2|R_{j+1}(\underline{x})| \quad (1 \leq j < t).$$

If (12.1) has infinitely many solutions, then so does

$$|R_1(\underline{x}) \dots R_t(\underline{x})| \leq \gamma_3 |\underline{x}|^{t-\eta}.$$

Since R_1, \dots, R_t are real, and since the theorem has been proved for real forms, there will be a subspace S^d of dimension d and m forms among R_1, \dots, R_t , whose rank satisfies (11.2). In fact in the real case we saw that if (12.2) holds, then we can take the first m forms. Now (12.8) is of course just as good as (12.2), and we get forms R_1, \dots, R_m whose rank r on S^d satisfies (11.2). We distinguish three cases:

(i) The system M_1, \dots, M_m is symmetric. Its rank r on S^d satisfies (11.2) and we are finished.

(ii) The system M_1, \dots, M_m is not symmetric and R_{m+1} on S^d is a linear combination of R_1, \dots, R_m . Then R_1, \dots, R_{m+1} have rank r on S^d , hence so does the symmetric system M_1, \dots, M_{m+1} , and we have $r < d$ and $r \leq md/\eta < (m+1)d/\eta$.

(iii) The system is not symmetric and R_{m+1} on S^d is not a linear combination of R_1, \dots, R_m . Then R_m is not a linear combination of R_1, \dots, R_{m-1} , so that R_1, \dots, R_{m-1} have rank $r-1$ on S^d , and in particular $r \geq 1$. Now if $\eta > n$, then $r \leq md/\eta < md/n \leq m$ and

$m \equiv 2$. The $m-1$ forms M_1, \dots, M_{m-1} have rank $r-1 < r \leq \frac{m-1}{m} \leq (m-1)d/\eta$ (since $r < m$) , and we are finished.

There remains the case when $\eta \leq n$. This case is completely elementary. We distinguish two subcases.

(I) We say that we have sufficiently many variables if either $n \equiv 3$, or $n = 2$ and at least one of the forms is proportional to a real form. In this case there is a symmetric system M_{i_1}, \dots, M_{i_m} where either $m < n$, or $m = n = 2$ and both M_{i_1}, M_{i_2} are proportional to the same real form. The system M_{i_1}, \dots, M_{i_m} has rank r with $r < n$ and $r \leq m = nm/n \leq nm/\eta$. So (11.2) is true with $S^d = \mathbb{R}^n$ and $d = n$.

(II) If we don't have sufficiently many variables, then either $n = 1$ and $|M_1(\underline{x}) \dots M_t(\underline{x})| \gg |\underline{x}|^t$ (by (11.4)) . Or $n = 2$ and none of the given linear forms is proportional to a real form. Then M_1, \dots, M_t occur in pairs of complex conjugate forms, with each pair of rank 2 . Therefore again $|M_1(\underline{x}) \dots M_t(\underline{x})| \gg |\underline{x}|^t$, and (a) can in fact not hold for any $\eta > 0$.

§13. Proof of Theorem 10A.

Now that we have Theorem 11A , the proof is quite easy.

By the hypothesis of our theorem, linear forms $M^{(i_1)}_1, \dots, M^{(i_m)}_m$ with $i_1 < \dots < i_m$ and with $m \geq n$ have rank n , and hence forms $M^{(i_1)}_1, \dots, M^{(i_m)}_m$ with $i_1 < \dots < i_m$ and with $m \leq n$ have rank m . Thus the rank r equals $\min(n, m)$. Let $r(M^{(i_1)}_1, \dots, M^{(i_m)}_m; S^d)$ be the rank of the restrictions of $M^{(i_1)}_1, \dots, M^{(i_m)}_m$ to a rational subspace S^d of dimension d . We claim that

$$(13.1) \quad r(M^{(i_1)}_1, \dots, M^{(i_m)}_m; S^d) = \min(d, m) .$$

It is clear that $r(M^{(i_1)}, \dots, M^{(i_m)}; S^d) \leq \min(d, m)$. Equality certainly holds if $m \geq n$, for then $r(M^{(i_1)}, \dots, M^{(i_m)}; R^n) = n$, whence $r(M^{(i_1)}, \dots, M^{(i_m)}; S^d) = d$. On the other hand, when $m < n$, the field K is m times transitive, so that $r(M^{(i_1)}, \dots, M^{(i_m)}; S^d) = r(M^{(e_1)}, \dots, M^{(e_m)}; S^d)$ for any $e_1 < \dots < e_m$. If this rank is less than m , then it follows that $r(M^{(i_1)}, \dots, M^{(i_m)}; S^d) = r(M^{(1)}, \dots, M^{(k)}; S^d) = d = \min(d, m)$.

Suppose the inequality of Theorem 10A had infinitely many solutions. We may apply Theorem 11A to the forms $M^{(1)}, \dots, M^{(k)}$ with $\eta = k - v$. By this theorem there is a rational subspace S^d of R^n with dimension d in $1 \leq d \leq n$, and there are linear forms $M^{(i_1)}, \dots, M^{(i_m)}$ such that $r = r(M^{(i_1)}, \dots, M^{(i_m)}; S^d)$ satisfies

$$r \leq dm/(k - v) \quad \text{and} \quad r < d.$$

In particular, since $v < k - n \leq k - d$, we have $r < m$ and $r < d$, which contradicts (13.1).

§14. Proof of Theorem 10C.

We have to show that under the hypotheses of Theorem 10C, (11.1) cannot hold with $\eta > 2(\ell - 1)$. In view of Theorem 11A, it will be enough to verify that we cannot have $r < dm/(2(\ell - 1))$ and $r < d$, i.e. to show that

$$(14.1) \quad r < d \quad \text{implies} \quad dm \leq 2(\ell - 1)r.$$

If $M(X)$ is a linear form with algebraic coefficients and with its first non-zero coefficient equal to 1, write $\mathfrak{F}(M)$ for the number field generated by the coefficients of M . We shall say that two such forms $M(X), M'(X)$ are conjugate, if there is an isomorphism from $\mathfrak{F}(M)$ onto

$\mathfrak{U}(M')$ which maps the coefficients of $M(\underline{x})$ into the respective coefficients of $M'(\underline{x})$. A form $M(\underline{x})$ has precisely f conjugates where f is the degree of $\mathfrak{U}(M)$, and no two conjugates are proportional.

In proving Theorem 10C, we may assume without loss of generality that all the linear forms M_i in $F(\underline{x}) = M_1(\underline{x}) \dots M_t(\underline{x})$ have their first non-zero coefficient equal to 1. With a form $M_1(\underline{x})$, also all its conjugates divide $F(\underline{x})$. If we block together conjugate linear forms, we may write

$$F(\underline{x}) = F_1(\underline{x}) \dots F_p(\underline{x}),$$

where each F_j is the product of conjugate linear forms. Each form F_j , being a product of conjugate linear forms, has rational coefficients, and hence has degree $\geq \ell$ by the hypothesis of Theorem 10C. Hence every linear form $M_i(\underline{x})$ has at least ℓ conjugates.

Now if $M_i(\underline{x}) = 0$ for some integer point \underline{x} , then also $M_{i_j}(\underline{x}) = 0$ for the conjugates $M_{i_1} = M_i, M_{i_2}, \dots, M_{i_f}$ of M_i . We have $f \geq \ell$, so that by our hypothesis the forms M_{i_1}, \dots, M_{i_f} have rank n , and \underline{x} must be $\underline{0}$. Therefore $M_i(\underline{x}) \neq 0$ ($1 \leq i \leq d$) for every integer point $\underline{x} \neq \underline{0}$. Hence we have rank $r > 0$ on every rational subspace S^d of dimension $d \geq 1$. We therefore have to show (14.1) only for $r > 0$, i.e. we have to show that

$$(14.2) \quad 0 < r < d \text{ implies } dm \leq 2(\ell - 1)r.$$

Let S^d be an arbitrary fixed rational subspace of \mathbb{R}^n of dimension $d \geq 2$; we shall prove (14.2) for S^d . Let $m(r)$ be the maximum number of linear forms M_{i_1}, \dots, M_{i_m} whose restrictions to S^d have rank r , and put $\pi(r) = m(r)/r$. We have to show that

$$(14.3) \quad \pi(r) \leq 2(\ell - 1)/d \quad (0 < r < d) .$$

Let π_0 be the maximum value of $\pi(r)$ in $0 < r < d$. Let r_0 be the largest number in $0 < r < d$ with $\pi(r) = \pi_0$. Put $m = m(r_0)$; there exists an m -tuple M_{i_1}, \dots, M_{i_m} of linear forms whose restrictions to S^d have rank r_0 , hence rank $< d$. It follows that the forms M_{i_1}, \dots, M_{i_m} have rank $< n$ on the full n -dimensional space R^n . By the hypothesis of Theorem 10C, we may conclude that

$$(14.4) \quad m < \ell .$$

Now if M_{i_1} occurs in an m -tuple of linear forms whose restrictions to S^d have rank r_0 , then so does every conjugate of M_{i_1} . Since M_{i_1} has at least ℓ conjugates and since $m < \ell$, there must be two distinct m -tuples M_{i_1}, \dots, M_{i_m} and M_{j_1}, \dots, M_{j_m} whose restrictions to S^d have rank r_0 . Let the intersection of these two m -tuples consist of m^- linear forms with rank r^- on S^d , and let the union of these two m -tuples consist of m^+ linear forms with rank r^+ on S^d . Then $m^+ > m^-$,

$$(14.5) \quad m^- + m^+ = 2m \text{ and } r^- + r^+ \leq 2r_0 .$$

Now if $m^- = 0$, then we have $r^- = 0$ and $\pi(r^+) \geq m^+/r^+ \geq m/r_0 = \pi_0$. If $m^- > 0$, then $0 < r^- \leq r_0$ and $m^-/r^- \leq \pi(r^-) \leq \pi(r_0) = \pi_0$, whence again $\pi(r^+) \geq m^+/r^+ \geq \pi_0$. Now clearly $r^+ \geq r_0$, and if we had $r^+ = r_0$, then $\pi(r^+) = \pi_0 = m/r_0$ and $m^+ = m$, which is impossible. Hence $r^+ > r_0$, and since r_0 was maximal in $0 < r < d$ with $\pi(r) = \pi_0$, we must have $r^+ = d$. Now (14.4), (14.5) yield $d \leq 2r_0$ and

$$\pi_0 = m/r_0 \leq (\ell - 1)/r_0 \leq 2(\ell - 1)/d ,$$

whence (14.3).

VIII. Approximation By Algebraic Numbers.

References: Mahler (1932), Koksma (1939), Wirsing (1961), Davenport and Schmidt (1967).

§1. The Setting. In the first chapters we studied approximation to real numbers by rationals. We now take up approximation to real numbers by algebraic numbers. This is quite different from the questions e.g. considered in Chapter V on approximation to algebraic numbers by rationals.

There are two basic questions here. Firstly, we may suppose a real algebraic number field K is given, and we study approximations to a real number α by elements of K . (A variant of this is when a complex (non-real) algebraic number field is given, and one studies approximations to a complex number α by elements of K . But this variant will not be taken up in these Notes.) Secondly, we may suppose a natural number k is given, and we are interested in approximations to a real number α by real algebraic numbers of degree $\leq k$. (Again there is a complex variant, which will not be developed here.)

With each of these two basic questions, one may strive for Dirichlet type theorems, i.e. theorems which assert that every real α may be approximated to a certain extent. Or one may be looking, e.g., for Roth type theorems, which assert that a given algebraic ξ cannot be approximated too well. Finally, there are questions which have no counterparts in the theory developed so far, and which I want to call questions of comparisons. Here one compares the approximation properties of a number α , either with regard to approximations in different fields, or with regard to approximations of various degrees.

The presentation of these questions will be as follows.

	Dirichlet type results	Roth type results	Questions of comparison
Given a field K	§2	§7,8	§10
Given a bound on the degree	§3,4,5,6.	§9	

In §11 a different viewpoint will be pursued.

Questions about simultaneous approximation, etc., will not be taken up in the context of this chapter.

§2. Field Height and Approximation by Elements of a Given Number

Field. As usual, if $P = P(X) = a_0 X^n + \dots + a_n$ is a polynomial, we set $\overline{P} = \max(|a_0|, \dots, |a_n|)$. If P is a nonzero polynomial with rational coefficients, then there is a polynomial P_0 proportional to P whose coefficients are coprime rational integers, and P_0 is unique up to a factor ± 1 . Define the height of P by

$$H(P) = \overline{P_0}.$$

Then if P has rational integer coefficients, $H(P) \leq \overline{P}$.

Now let K be a real algebraic number field of degree k . There are k isomorphisms σ_1 (= identity map), $\sigma_2, \dots, \sigma_k$ of K into the complex numbers. Write $\sigma_i(\beta) = \beta^{(i)}$ ($i = 1, \dots, k$) for $\beta \in K$; then $\beta^{(1)} = \beta, \beta^{(2)}, \dots, \beta^{(k)}$ are the conjugates of β . The field polynomial of β is $P_0(X)$ where

$$P(X) = (X - \beta^{(1)}) \dots (X - \beta^{(k)}),$$

and its field height is

$$H_K(\beta) = \overline{P_0} = H(P) .$$

Given a bound C , there are only finitely many β in K with

$$H_K(\beta) \leq C .$$

THEOREM 2A. Let K be a real algebraic number field. Then there exists a constant c_K such that for every real α not in K there are infinitely many $\beta \in K$ having

$$|\alpha - \beta| < c_K \max(1, |\alpha|^2) H_K(\beta)^{-2} .$$

It will follow from Theorem 8A that this result is essentially best possible. For a generalization, see Schmidt (1975) .

The connection between the cases $|\alpha| < 1$ and $|\alpha| > 1$ is as follows. Suppose the theorem is true for $|\alpha| < 1$ with some constant c'_K , i.e. there are infinitely many $\beta \in K$ with $|\alpha - \beta| < c'_K H_K(\beta)^{-2}$. Then if $|\alpha| > 1$, there are infinitely many $\gamma \in K$ with $|(1/\alpha) - \gamma| < c'_K H_K(\gamma)^{-2}$. Clearly $|\gamma| > 1/(|\alpha|(1+\epsilon))$ if $\epsilon > 0$ and if $H_K(\gamma)$ is large. Setting $\beta = 1/\gamma$ we have $H_K(\beta) = H_K(\gamma)$ and

$$|\alpha - \beta| = |\alpha\beta| |(1/\alpha) - \gamma| < c'_K |\alpha|^2 (1+\epsilon) H_K(\beta)^{-2} .$$

Hence Theorem 2A is true for any given constant $c_K > c'_K$.

In the case when $K = \mathbb{Q}$, the rationals, the height $H_{\mathbb{Q}}(p/q) = \max(|p|, |q|)$ if p/q is in its lowest terms. If $\alpha \notin \mathbb{Q}$ and $|\alpha| < 1$, Dirichlet's Theorem yields infinitely many p/q with $|\alpha - (p/q)| < 1/q^2$, and since here $|p| < |\alpha q| + 1/q < q$ for large q , we then have $H_{\mathbb{Q}}(p/q) = q$ and $|\alpha - (p/q)| < H_{\mathbb{Q}}(p/q)^{-2}$. Hence we may take $c'_{\mathbb{Q}} = 1$, and $c_{\mathbb{Q}}$ any constant > 1 . If we use Hurwitz' Theorem instead of Dirichlet's, then we may take any constant $> 1/\sqrt{5}$.

Proof of Theorem 2A. We may suppose that $|\alpha| < 1$. Let $\theta_1, \dots, \theta_k$ be an integer basis of K . Given large Q we apply Minkowski's Linear Forms Theorem (Theorem 2C of Ch. II) to find rational integers $q_1, \dots, q_k, p_1, \dots, p_k$, not all zero, with

$$(2.1) \quad |\alpha \theta_1 q_1 + \dots + \alpha \theta_k q_k - \theta p_1 - \dots - \theta_k p_k| |\theta_k|^{-1} < Q^{-2k+1},$$

$$(2.2) \quad |q_1|, \dots, |q_k|, |p_1|, \dots, |p_{k-1}| \leq Q.$$

Then $|p_k| \ll Q$; here and below, the constant in \ll depends only on $\theta_1, \dots, \theta_k$, hence ultimately only on K . If we had $q_1 = \dots = q_k = 0$, then by (2.1), (2.2), the algebraic integer $\theta_1 p_1 + \dots + \theta_k p_k$ would have norm $\ll Q^{-2k+1} Q^{k-1} = Q^{-k}$, whence we would get the contradiction $p_1 = \dots = p_k = 0$. Thus not all of q_1, \dots, q_k are zero, and $\theta_1 q_1 + \dots + \theta_k q_k \neq 0$. Putting

$$\beta = \frac{\theta_1 p_1 + \dots + \theta_k p_k}{\theta_1 q_1 + \dots + \theta_k q_k},$$

we have

$$(2.3) \quad |\alpha - \beta| < Q^{-2k+1} |\theta_1 q_1 + \dots + \theta_k q_k|^{-1}.$$

The field polynomial of β is proportional to the polynomial

$$P(X) = \prod_{i=1}^k ((\theta_1^{(i)} q_1 + \dots + \theta_k^{(i)} q_k) X - (\theta_1^{(i)} p_1 + \dots + \theta_k^{(i)} p_k)).$$

Since P has rational integer coefficients,

$$(2.4) \quad H_K(\beta) \leq |\overline{P}|.$$

To estimate $|\overline{P}|$ we observe that $|\theta_1 p_1 + \dots + \theta_k p_k| \ll |\theta_1 q_1 + \dots + \theta_k q_k|$ by (2.1)[†], and that both $|\theta_1^{(i)} p_1 + \dots + \theta_k^{(i)} p_k|$ and $|\theta_1^{(i)} q_1 + \dots + \theta_k^{(i)} q_k|$ are $\ll Q$ for $i = 2, \dots, k$ by (2.2). It follows that

†) Note that $|\theta_1 p_1 + \dots + \theta_k p_k| \gg Q^{-k+1}$ by Lemma 1A of Ch. VI.

$$H_K(\beta) \leq |\overline{P}| \ll |\theta_1 q_1 + \dots + \theta_k q_k| Q^{k-1} \ll Q^k.$$

Combining this with (2.3) , (2.4) we get

$$|\alpha - \beta| \ll Q^{-k} H_K(\beta)^{-1} \ll H_K(\beta)^{-2}.$$

Since $\alpha \notin K$, the left hand side is not zero, and hence as $Q \rightarrow \infty$, we obtain infinitely many distinct β with $|\alpha - \beta| \ll H_K(\beta)^{-2}$.

Exercise. Show that if K is a complex algebraic number field and if α is complex but not in K , then there are infinitely many $\beta \in K$ with

$$|\alpha - \beta| < \frac{c_K \max(1, |\alpha|^2)}{H(\beta)}.$$

§3. Absolute Height and Approximation by Algebraic Numbers of

Bounded Degree. Let

$$P(X) = a_0 X^n + \dots + a_n = a_0 (X - \alpha^{(1)}) \dots (X - \alpha^{(n)})$$

be a polynomial. There are many quantities besides $|\overline{P}|$ to measure the "size" of P . Among them are

$$L(P) = |a_0| + \dots + |a_n|$$

and Mahler's

$$M(P) = |a_0| \prod_{i=1}^n \max(1, |\alpha^{(i)}|).$$

We have $|\overline{P}| \leq L(P) \leq (n+1) |\overline{P}|$ and

$$(3.1) \quad L(PQ) \leq L(P)L(Q) \quad , \quad |\overline{PQ}| \leq (n+1) |\overline{P}| |\overline{Q}| \quad ,$$

while M satisfies the very satisfying relation

$$(3.2) \quad M(PQ) = M(P)M(Q) \quad .$$

Also,

$$(3.3) \quad |\overline{P}| \leq L(P) \leq |a_0| \prod_{i=1}^n (1 + |\alpha_i|) \leq 2^n M(P) \quad .$$

In the opposite direction we wish to show that

$$(3.4) \quad M(P) \ll_n |\overline{P}| \quad .$$

Suppose that

$$|\alpha_1| \leq \dots \leq |\alpha_n| \quad .$$

If

$$(3.5) \quad |\alpha_1| \leq 4^n, |\alpha_2| \leq 4^n |\alpha_1|, \dots, |\alpha_n| \leq 4^n |\alpha_{n-1}| \quad ,$$

then $M(P) \ll |a_0| \leq |\overline{P}|$. Otherwise, suppose the i -th inequality

(3.5) is the first one which fails to hold. Then

$$(3.6) \quad M(P) \ll |a_0| |\alpha_i| \dots |\alpha_n| \quad .$$

On the other hand

$$\begin{aligned} a_{n-i+1}/a_0 &= \alpha_i \alpha_{i+1} \dots \alpha_n + \alpha_{i-1} \alpha_{i+1} \dots \alpha_n + \dots + \alpha_1 \dots \alpha_{n-i+1} \\ &= \pi_1 + \dots + \pi_t \quad , \end{aligned}$$

say, where $t = 1$ if $i = 1$, and where $t = \binom{n}{i-1} \leq 2^n$ and

$|\pi_j|/|\pi_1| \leq 4^{-n}$ ($j = 2, \dots, t$) if $i > 1$. Thus

$$\begin{aligned} |\overline{P}| &\geq |a_{n-i+1}| \geq |a_0| |\pi_1| (1 - (2^n/4^n)) \\ &\geq \frac{1}{2} |a_0| |\alpha_i| |\alpha_{i+1}| \dots |\alpha_n| \quad , \end{aligned}$$

which in conjunction with (3.6) yields (3.4) . From (3.1) , (3.2) , (3.3) , (3.4) we see that

$$(3.7) \quad L(PQ) \cong L(P)L(Q) \ll L(PQ)$$

and

$$(3.8) \quad [PQ] \ll [P][Q] \ll [PQ] ,$$

with constants in \ll which depend only on the degrees of P and of Q .

More explicit estimates than these were given by Mahler (1960) and others. $L(P)$ and $[P]$ may be defined for polynomials in any number of variables, and (3.7) , (3.8) remain true with constants which depend on the degrees and on the number of variables. It was discovered by Enflo and reproved more simply by Montgomery (1975/76) that (3.7) is true with constants which depend only on the total degrees of P and of Q , but not on the number of variables.

Now if α is algebraic of precise degree n , it satisfies an equation $P_0(\alpha) = 0$ where $P_0(X)$ is a nonzero polynomial of degree n with coprime integer coefficients. The absolute height, or briefly height, of α is given by

$$H(\alpha) = H(P_0) = [P_0] .$$

Also define

$$M(\alpha) = M(P_0) .$$

Further if α lies in a number field K , then the degree $[K:\mathbb{Q}] = k$, say, is a multiple of n . If F_0 is the field polynomial of α , then it is seen that

$$F_0 = P_0^{k/n},$$

so that by (3.8) ,

$$(3.9) \quad H_K(\beta) \gg_k \ll (H(\beta))^{k/n}.$$

CONJECTURE. Suppose α is real, but is not algebraic of degree $\leq k$. Suppose $\varepsilon > 0$. Then there are infinitely many real algebraic β of degree $\leq k$ with

$$(3.10) \quad |\alpha - \beta| < \frac{c_1}{H(\beta)^{k+1-\varepsilon}},$$

where $c_1 = c_1(k, \alpha, \varepsilon)$.

Perhaps even a stronger form of the conjecture is true, without the ε , and with (3.10) replaced by

$$(3.11) \quad |\alpha - \beta| < \frac{c_2}{H(\beta)^{k+1}},$$

where $c_2 = c_2(k, \alpha)$. Such a strong form is true for $k = 1$, by Dirichlet's Theorem. The strong form is also true for $k = 2$:

THEOREM 3A. (Davenport and Schmidt (1967)). For any real α which is not rational or quadratic over \mathbb{Q} , there exists infinitely many rational or real quadratic β such that

$$|\alpha - \beta| < c_3 H(\beta)^{-3},$$

where $c_3 = c_3(\alpha)$.

In fact Davenport and Schmidt proved a more explicit result with $c_3(\alpha) = c_4 \max(1, |\alpha|^2)$, where c_4 is any fixed constant greater than $160/9$. For $k > 2$ the conjectured exponent $k + 1$ or $k + 1 - \varepsilon$

has not yet been obtained, but we have

THEOREM 3B. (Wirsing (1961)). For any real α which is not algebraic of degree $\leq k$, then there are infinitely many real algebraic β of degree $\leq k$ having

$$|\alpha - \beta| < c_5(k, \alpha) H(\beta)^{-(k+3)/2}.$$

Wirsing in fact obtained the slightly better exponent

$$\frac{k+6}{4} + \frac{1}{4}\sqrt{k^2 + 4k - 4}.$$

One could also try to approximate by algebraic integers. A natural conjecture is that if α is not algebraic of degree $\leq k$, then there are infinitely many real algebraic integers β of degree $\leq k$ with

$$|\alpha - \beta| < \frac{c_6(k, \alpha, \varepsilon)}{H(\beta)^{k+\varepsilon}}.$$

The following will not be proved in these Notes.

THEOREM 3C*. (Davenport and Schmidt (1969)). Suppose $k \geq 3$ and suppose α is real, but not algebraic of degree $\leq (k-1)/2$. Then there are infinitely many algebraic integers β of degree $\leq k$ with

$$|\alpha - \beta| < c_7(k, \alpha) H(\beta)^{-[(k+1)/2]}.$$

Incidentally, (3.11), if true, would be best possible. For if α is algebraic of precise degree $k+1$, the resultant $R(\alpha, \beta)$ is a nonvanishing integer, and if

$$a_0(X - \alpha^{(1)}) \dots (X - \alpha^{(k+1)}) \quad \text{and} \quad b_0(X - \beta^{(1)}) \dots (X - \beta^{(d)})$$

are the minimal polynomials of α, β , then (Van der Waerden (1955), §30)

$$\begin{aligned}
 1 \leq |R(\alpha, \beta)| &= \left| a_0^d b_0^{k+1} \prod_{i=1}^{k+1} \prod_{j=1}^d \alpha^{(i)} - \beta^{(j)} \right| \\
 &= |\alpha - \beta| |b_0^{k+1} Q(\beta^{(1)}, \dots, \beta^{(d)})|,
 \end{aligned}$$

where Q is a polynomial with coefficients depending on α , and of degree at most $k+1$ in each $\beta^{(i)}$. Thus with Mahler's $M(\beta)$,

$$1 \ll |\alpha - \beta| M(\beta)^{k+1} \ll |\alpha - \beta| H(\beta)^{k+1},$$

or

$$|\alpha - \beta| > c_8(\alpha) H(\beta)^{-k-1}.$$

For algebraic α of degree $> k+1$, see §9. It can be shown that for almost all α and for $\varepsilon > 0$,

$$(3.12) \quad |\alpha - \beta| > c_9(k, \alpha, \varepsilon) H(\beta)^{-k-1-\varepsilon},$$

for β of degree $\leq k$, and

$$|\alpha - \beta| > c_{10}(k, \alpha, \varepsilon) H(\beta)^{-k-\varepsilon}$$

for algebraic integers β of degree $\leq k$. For more information (involving Hausdorff dimension) on (3.12) see Baker and Schmidt (1970).

§4. Approximation by Quadratic Irrationals. A Theorem on Two Linear Forms. In this section write $\underline{X} = (X, Y, Z)$,

$$\underline{x} = (x, y, z) \text{ and } |\underline{x}| = \max(|x|, |y|, |z|).$$

Let $L(\underline{X}) = \alpha X + \beta Y + \gamma Z$ be a linear form. It follows, e.g. from Minkowski's Linear Forms Theorem, that there are infinitely many nonzero integer points \underline{x} having

$$|L(\underline{x})| \ll |\underline{x}|^{-2},$$

with the constant in \ll depending on L . We will prove the stronger

THEOREM 4A. Suppose L, P are independent linear forms. Then there are infinitely many integer points $\underline{x} \neq \underline{0}$ with

$$(4.1) \quad |L(\underline{x})| \ll |\underline{x}|^{-4} |P(\underline{x})|^2,$$

and with a constant in \ll depending on L and on P .

A more general result was given by Davenport and Schmidt (1970).

To deduce Theorem 3A, put

$$L(\underline{X}) = \alpha^2 X + \alpha Y + Z \quad \text{and} \quad P(\underline{X}) = 2\alpha X + Y.$$

Given \underline{x} with (4.1), the polynomial

$$B(T) = xT^2 + yT + z$$

has

$$(4.2) \quad |B(\alpha)| \ll |\underline{x}|^{-4} |P(\underline{x})|^2 \ll |\underline{x}|^{-3} |P(\underline{x})| \ll |\underline{x}|^{-3} |B'(\alpha)|.$$

Now either $\deg B = 1$. Then $B(T)$ has the rational root β with

$$0 = B(\beta) = B(\alpha) + (\beta - \alpha)B'(\alpha),$$

so that

$$|\alpha - \beta| = |B(\alpha)/B'(\alpha)| \ll |\underline{x}|^{-3} \ll |\underline{B}|^{-3} \ll H(\beta)^{-3}.$$

Or $\deg B = 2$. Then the roots β of B satisfy

$$0 = B(\beta) = B(\alpha) + (\beta - \alpha)B'(\alpha) + \frac{1}{2}(\beta - \alpha)^2 B''(\alpha).$$

Solving this by the quadratic formula for $\beta - \alpha$, we see from (4.2) that the roots β are real, and that one of the roots β has

$$\begin{aligned} |\beta - \alpha| &= |-B'(\alpha) + (B'^2(\alpha) - 2B(\alpha)B''(\alpha))^{1/2}| |B''(\alpha)|^{-1} \\ &<< |B(\alpha)/B'(\alpha)| \\ &<< \underline{|x|}^{-3} << \overline{|B|}^{-3} << H(\beta)^{-3}. \end{aligned}$$

To prove Theorem 4A we may suppose that $L(\underline{x}) \neq 0$ for integer points $\underline{x} \neq \underline{0}$. Choose a linear form F such that F, L, P are independent. Putting

$$(4.3) \quad \langle \underline{x} \rangle = \max(|F(\underline{x})|, |L(\underline{x})|, |P(\underline{x})|),$$

we have

$$(4.4) \quad \langle \underline{x} \rangle >> << \underline{|x|}.$$

For each real $X > 0$ we consider the finite set of integer points $\underline{x} \neq \underline{0}$ satisfying

$$\langle \underline{x} \rangle \leq X.$$

For large X this set is not empty, and by our hypothesis on L , the values of L at the points of this set are distinct. We choose the unique point \underline{x} in this set for which $|L(\underline{x})|$ is minimal and the first nonzero element of the triple $F(\underline{x}), L(\underline{x}), P(\underline{x})$ is positive, and we call this the minimal point corresponding to X .

It is obvious that if \underline{x} is the minimal point corresponding to X^* , then it is the minimal point corresponding to any X in some interval $X^* \leq X < X^{**}$. There is a sequence of numbers

$$(4.5) \quad X_1 < X_2 < \dots$$

which tend to infinity, and a sequence of points

$$(4.6) \quad \underline{x}_1, \underline{x}_2, \dots$$

such that \underline{x}_i is the minimal point corresponding to all X in the range $X_i \leq X < X_{i+1}$, but to no other X . It is clear that

$$(4.7) \quad \langle \underline{x}_i \rangle = X_i.$$

We introduce the notations

$$(4.8) \quad F_i = F(\underline{x}_i), \quad L_i = L(\underline{x}_i), \quad P_i = P(\underline{x}_i).$$

Plainly

$$(4.9) \quad |L_1| > |L_2| > \dots$$

By our construction of the sequences (4.5), (4.6), there is no integer point $\underline{x} \neq 0$ with

$$(4.10) \quad \langle \underline{x} \rangle < X_{i+1}, \quad |L(\underline{x})| < |L_i|.$$

The inequalities (4.10) define a symmetric convex set of volume $\gg X_{i+1}^2 |L_i|$, so that by Minkowski's Theorem (Th. 2B of Ch. II) we have $X_{i+1}^2 |L_i| \ll 1$, or

$$(4.11) \quad |L_i| \ll X_{i+1}^{-2}.$$

Our assertion concerning (4.1) will be proved if we can show that $|L_i| \ll X_i^{-4} P_i^2$ for infinitely many i . We shall assume indirectly that

$$(4.12) \quad P_i^2 = o(|L_i| X_i^4),$$

and eventually we shall reach a contradiction.

§5. Approximation by Quadratic Irrationals, Continued. As a consequence of (4.11) and (4.12) we obtain

$$(5.1) \quad |P_i| = o(X_i^2 X_{i+1}^{-1}) = o(X_i) .$$

Since $L_i \rightarrow 0$ and since $\langle \underline{x}_i \rangle = \max(|F_i|, |L_i|, |P_i|)$, it follows from (4.7) that

$$(5.2) \quad F_i = X_i .$$

LEMMA 5A. The signs of the L_i alternate when i is large.

Proof. The point

$$\underline{y} = \underline{x}_{i+1} - \underline{x}_i$$

has $0 < F(\underline{y}) < X_{i+1}$ by (4.5) and (5.2). Since $|L(\underline{y})| \ll 1$ and since $|P(\underline{y})| \leq |P_{i+1}| + |P_i| = o(X_{i+1})$ by (5.1), we have

$$\langle \underline{y} \rangle < X_{i+1}$$

if i is large. Since (4.10) has no nonzero solution, we must have

$$|L_{i+1} - L_i| = |L(\underline{y})| \geq |L_i| .$$

In conjunction with (4.9) this yields $L_i L_{i+1} < 0$.

LEMMA 5B. Suppose i is large and

$$(5.3) \quad |P_{i+1}| \leq \frac{1}{2} X_i .$$

Then

$$\underline{x}_{i+1} = t\underline{x}_i + \underline{x}_{i-1} ,$$

where t is a positive integer.

Proof. Define integers t, u by

$$t = [X_{i+1}/X_i] , \quad u = [|L_{i-1}| / |L_i|] ,$$

and integer points $\underline{y}, \underline{z}$ by

$$\underline{y} = \underline{x}_{i+1} - t\underline{x}_i , \quad \underline{z} = \underline{x}_{i-1} + u\underline{x}_i .$$

Clearly \underline{x}_i and \underline{x}_{i+1} are independent, and therefore $\underline{y} \neq \underline{0}$.

Similarly $\underline{z} \neq \underline{0}$.

We have $0 \leq F(\underline{y}) < X_i$ from (4.5) and (5.2). It follows from (4.11) that

$$|L(\underline{y})| < \frac{3}{4} X_i$$

if i is large. Finally from (5.1), (5.3),

$$|P(\underline{y})| \leq |P_{i+1}| + X_{i+1} X_i^{-1} |P_i| \leq \frac{1}{2} X_i + o(X_{i+1} X_i^{-1} X_i^2 X_{i+1}^{-1}) < \frac{3}{4} X_i$$

if i is large. Thus $\langle \underline{y} \rangle < X_i$, but since (4.10) has no nonzero solution, we must have

$$|L(\underline{y})| \geq |L_{i-1}| .$$

Thus $|L_{i-1}| \leq |L_{i+1}| + t|L_i|$, whence

$$u \leq t + |L_{i+1}| / |L_i| < t + 1 ,$$

whence $u \leq t$.

Turning to \underline{z} , we have

$$(5.4) \quad |L(\underline{z})| = |L_{i-1}| - u|L_i| < |L_i|$$

since L_{i-1}, L_i are of opposite sign. Since (4.10) has no nonzero solution, $\langle \underline{z} \rangle \cong X_{i+1}$. Since $|L(\underline{z})| = o(X_{i+1})$ and since

$$|P(\underline{z})| \leq |P_{i-1}| + u|P_i| \leq |P_{i-1}| + t|P_i| = o(X_{i+1})$$

by (5.1), we must have $|F(\underline{z})| \geq X_{i+1}$. Thus

$$X_{i-1} + uX_i \geq X_{i+1},$$

whence $t < u + 1$, whence $t \leq u$. Combining this with the opposite inequality shown above, we get

$$u = t.$$

Finally consider the point

$$(5.5) \quad \underline{w} = \underline{x}_{i+1} - t\underline{x}_i - \underline{x}_{i-1} = \underline{y} - \underline{x}_{i-1} = \underline{x}_{i+1} - \underline{z}.$$

From the form $\underline{y} - \underline{x}_{i-1}$, from what we already know about \underline{y} , and from (5.1), we get $|L(\underline{w})|, |P(\underline{w})| < X_i$ if i is large. From $0 \leq F(\underline{y}) < X_i$ we get $|F(\underline{w})| < X_i$, so that

$$\langle \underline{w} \rangle < X_i < X_{i+1}.$$

Also, from the form $\underline{x}_{i+1} - \underline{z}$, on noting that $L(\underline{z})$ has the same sign as L_{i-1} , and so the same sign as L_{i+1} , we obtain

$$|L(\underline{w})| \leq \max(|L_{i+1}|, |L(\underline{z})|) < |L_i|.$$

Thus \underline{w} satisfies (4.10), which has the consequence that $\underline{w} = \underline{0}$.

The proof of Theorem 4A is now completed as follows. Suppose (5.3) is satisfied for some large i . Then by (5.2) and by Lemma 5B,

$$|X_i L_{i+1} - X_{i+1} L_i| = |X_{i-1} L_i - X_i L_{i-1}|$$

and

$$|P_i L_{i+1} - P_{i+1} L_i| = |P_{i-1} L_i - P_i L_{i-1}|.$$

Now suppose that (5.3) holds for every i in $h < i < k$. Then

$$(5.6) \quad |X_h L_{h+1} - X_{h+1} L_h| = |X_{k-1} L_k - X_k L_{k-1}|$$

and

$$|P_h L_{h+1} - P_{h+1} L_h| = |P_{k-1} L_k - P_k L_{k-1}|.$$

The latter gives

$$(5.7) \quad |P_{h+1} L_h| \leq |P_h L_{h+1}| + |P_{k-1} L_k| + |P_k L_{k-1}|.$$

These relations remain trivially true if $k = h + 1$.

The left hand side of (5.6) is $|L(X_{h=h+1} - X_{h+1} x_h)|$, and this is not 0. But by (4.11), the right hand side of (5.6) tends to 0 as $k \rightarrow \infty$. Hence (5.3) cannot be satisfied for all large i , and so there are infinitely many i for which (5.3) fails to hold.

Suppose that (5.3) is satisfied for every i in $h < i < k$, but not for $i = h$ or $i = k$. Then $|P_{h+1}| > \frac{1}{2} X_h$, and by (5.7) and by (4.12), (5.1) we get

$$\begin{aligned} \frac{1}{2} X_h |L_h| &\leq |P_h L_{h+1}| + |P_{k-1} L_k| + |P_k L_{k-1}| \\ &= o(X_h |L_{h+1}| + |L_{k-1}|^{1/2} X_{k-1}^2 |L_k| + |L_k|^{1/2} X_k^2 |L_{k-1}|). \end{aligned}$$

Since $|L_{h+1}| < |L_h|$ and $|L_k| < |L_{k-1}|$, this yields

$$X_h |L_h| = o(X_k^2 |L_k|^{1/2} |L_{k-1}|),$$

which by $|L_{k-1}| \leq |L_h|$ further implies

$$x_h |L_h|^{1/2} = o(x_k^2 |L_{k-1}|^{1/2} |L_k|^{1/2}) = o(x_k |L_k|^{1/2})$$

by (4.11). In particular, if h , and hence k , is large, then

$$x_k |L_k|^{1/2} > 2x_h |L_h|^{1/2}.$$

But this is impossible, since it leads to an infinite sequence of values of j for which $x_j |L_j|^{1/2}$ increases to infinity, while this expression is bounded by (4.11).

§6. Proof of Wirsing's Theorem. Throughout, α will be a fixed real which is not algebraic of degree $\leq k$, and the constants in \ll will depend on α and on k .

Given a polynomial $P(X) = b_0(X - \beta_1) \dots (X - \beta_d)$ with $b_0 > 0$ and of degree $d \leq k$, put

$$p_i = |\alpha - \beta_i| \quad (i = 1, \dots, d)$$

and

$$M_\alpha(P) = b_0 \prod_{i=1}^d \max(1, p_i).$$

Then $M_\alpha(P)$ is equal to Mahler's $M(P_\alpha)$ where $P_\alpha(X) = P(X + \alpha)$, so that

$$(6.1) \quad M_\alpha(P) = M(P_\alpha) \gg \ll \overline{P_\alpha} \gg \ll \overline{P}$$

by (3.3), (3.4). Since

$$M_\alpha(P) \prod_{i=1}^d \min(1, p_i) = b_0 \prod_{i=1}^d p_i = |P(\alpha)|,$$

we get

$$(6.2) \quad |P(\alpha)| \gg \ll |P| \prod_{i=1}^d \min(1, p_i) \quad .$$

LEMMA 6A. Suppose

$$P(X) = b_0(X - \beta_1) \dots (X - \beta_d) \quad , \quad Q(X) = c_0(X - \gamma_1) \dots (X - \gamma_e)$$

are polynomials with integral coefficients and without common factor, of respective degrees d, e where $1 \leq d, e \leq k$. Put

$$p_i = |\alpha - \beta_i| \quad (i = 1, \dots, d) \quad , \quad q_j = |\alpha - \gamma_j| \quad (j = 1, \dots, e) \quad ,$$

and suppose that

$$p_1 \leq \dots \leq p_d \quad \text{and} \quad q_1 \leq \dots \leq q_e \quad .$$

Further suppose that

$$(6.3) \quad p_1 \leq q_1 < 1 \quad .$$

Then one of the following relations holds.

- (i) $p_1 \ll |P(\alpha)| |P|^{-1}$, and β_1 is real,
- (ii) $q_1 \ll |Q(\alpha)| |Q|^{-1}$, and γ_1 is real,
- (iii) $1 \ll Q(\alpha)^2 |P|^k |Q|^{k-2}$,
- (iv) $p_1^2 \ll |P(\alpha)| Q(\alpha)^2 |P|^{k-1} |Q|^{k-2}$, and β_1 is real,
- (v) $p_1^2 \ll P(\alpha)^2 |Q(\alpha)| |P|^{k-2} |Q|^{k-1}$, and β_1 is real.

Proof. If $d = 1$ or if $d > 1$ and $p_2 \geq 1$, then (6.2) yields

(i) . In particular β_1 is real, since $p_1 < p_2$ in the case $d > 1$. Similarly, if $e = 1$ or if $e > 1$ and $q_2 \geq 1$, then (ii) follows.

We may thus suppose that

$$(6.4) \quad 2 \leq d, e \leq k, \text{ and that } p_2 < 1, q_2 < 1.$$

Writing R for the resultant of P, Q , we have

$$(6.5) \quad 1 \leq |R| = b_0^e c_0^d \prod_{i,j} |\beta_i - \gamma_j| \ll b_0^k c_0^k \prod_{i,j} \max(p_i, q_j) = AB,$$

say, where

$$A = \prod_{p_i < 1} \prod_{q_j < 1} \max(p_i, q_j)$$

and where

$$B \leq b_0^k c_0^k \prod_{i,j} \max(1, p_i) \max(1, q_j) \leq M_\alpha(P)^k M_\alpha(Q)^k \ll |P|^k |Q|^k$$

by (6.1).

Now if $p_2 \leq q_1$, then $p_1 \leq p_2 \leq q_1 < 1$ and

$$A \ll \prod_{q_j < 1} q_j^2 \ll Q(\alpha)^2 |Q|^{-2}$$

by (6.2), and (iii) follows.

We may thus suppose that $q_1 < p_2$. We have the two subcases

$$(a) \quad p_1 \leq q_1 < p_2 \leq q_2,$$

$$(b) \quad p_1 \leq q_1 \leq q_2 < p_2.$$

In the case (a) we have

$$A \ll \left(\prod_{\substack{1 \leq i \\ p_i < 1}} p_i \right) q_1 \left(\prod_{\substack{j \geq 2 \\ q_j < 1}} q_j \right)^2.$$

Multiplying (6.5) by $p_1 q_1$ we get

$$\begin{aligned} p_1^2 \leq p_1 q_1 &\ll \left(\prod_{p_i < 1} p_i \right) \left(\prod_{q_j < 1} q_j \right)^2 |P|^k |Q|^k \\ &\ll |P(\alpha)| Q(\alpha)^2 |P|^{k-1} |Q|^{k-2}, \end{aligned}$$

from (6.2) . Thus (iv) holds in this case. (Note that β_1 is real since $p_1 < p_2$.) Finally in the case (b) we multiply (6.5) by p_1^2 to get (v) .

The proof of Wirsing's Theorem 3B is now completed as follows. It is an immediate consequence of Minkowski's Linear Forms Theorem that there are infinitely many nonzero polynomials $P(X)$ of degree $\leq k$ and with rational integer coefficients having

$$|P(\alpha)| << |P|^{-k} .$$

If $P = P_1 \dots P_t$ is the factorization of P into irreducible factors, then

$$|P_1(\alpha) \dots P_t(\alpha)| << (|P_1| \dots |P_t|)^{-k}$$

by (3.8) . Since $P(\alpha) \neq 0$ by hypothesis, it is not very difficult to deduce that

there are infinitely many distinct polynomials P of degree $\leq k$ with coprime rational integer coefficients, which are irreducible over the rationals and which have

$$(6.6) \quad |P(\alpha)| << |P|^{-k} .$$

If P is such a polynomial, and if Q is a polynomial of degree $\leq k$ with integral coefficients which is a multiple of P , then

$|Q| \gg |P|$ by (3.8)[†]. Let's say $|Q| \geq c_1 |P|$. The inequalities

$$|q_0|, \dots, |q_k| \leq \frac{1}{2} c_1 |P| ,$$

$$|q_0 \alpha^k + \dots + q_k| \leq c_2 |P|^{-k}$$

define a symmetric convex set in $(k+1)$ -dimensional space of $(k+1)$ -

[†]) but with a different meaning of Q .

tuples (q_0, \dots, q_k) , of volume $\gg c_1^k c_2$. Hence if c_2 is chosen sufficiently large, then our inequalities have a nontrivial integral solution. The polynomial

$$Q(X) = q_0 X^k + \dots + q_k$$

has

$$(6.7) \quad |Q| \ll |P| \quad \text{and} \quad |Q(\alpha)| \ll |P|^{-k}.$$

Q is not a multiple of P , and by the irreducibility of P , the two polynomials P, Q have no common factor.

Introducing the quantities p_i, q_j as in Lemma 6A, we have $p_1 < 1, q_1 < 1$ from (6.6), (6.7). Suppose that $p_1 \leq q_1$, i.e. (6.3) holds. In the cases (i) and (ii) we have

$$|\alpha - \beta_1| \ll |P(\alpha)| |P|^{-1} \ll |P|^{-k-1} \ll H(\beta_1)^{-k-1}$$

and

$$|\alpha - \gamma_1| \ll |Q(\alpha)| |Q|^{-1} \ll |P|^{-k} |Q|^{-1} \ll |Q|^{-k-1} \ll H(\gamma_1)^{-k-1},$$

respectively. The case (iii) leads to

$$1 \ll |P|^{-2k} |P|^{2k-2},$$

hence is impossible for large $|P|$. The case (iv) yields

$$|\alpha - \beta_1| \ll |P|^{-k/2} |P|^{-k} |P|^{(k-1)/2} |P|^{(k-2)/2} = |P|^{-(k+3)/2} \ll H(\beta_1)^{-(k+3)/2},$$

and the case (v) is similar. The situation is analogous when $q_1 < p_1$.

§7. A Subspace Theorem for Number Fields. Before turning to a Roth type theorem for number fields, we need to obtain a generalized

Subspace Theorem.

Let K be a number field of degree k and I the ring of integers of K . Denote the images of an element $\beta \in K$ under the isomorphisms of K into \mathbb{C} by $\beta^{(1)} = \beta$, $\beta^{(2)}, \dots, \beta^{(k)}$. Given

$$\underline{\beta} = (\beta_1, \dots, \beta_n)$$

with components in K , put $\underline{\beta}^{(j)} = (\beta_1^{(j)}, \dots, \beta_n^{(j)})$ and

$$|\underline{\beta}| = \max(|\beta_1^{(1)}|, \dots, |\beta_n^{(1)}|, \dots, |\beta_1^{(k)}|, \dots, |\beta_n^{(k)}|).$$

THEOREM 7A. Suppose that for each j in $1 \leq j \leq k$ we are given n linearly independent linear forms M_{j1}, \dots, M_{jn} in n variables, with real or complex algebraic coefficients. Suppose $\varepsilon > 0$. Then there are finitely many proper subspaces T_1, \dots, T_w of n -dimensional space K^n such that every nonzero solution $\underline{\beta} \in I^n$ of

$$(7.1) \quad \prod_{j=1}^k |\bar{M}_{j1}(\underline{\beta}^{(j)}) \dots M_{jn}(\underline{\beta}^{(j)})| < |\underline{\beta}|^{-\varepsilon}$$

lies in one of these subspaces.

Proof. Let $\theta_1, \dots, \theta_k$ be an integer basis of K . We have

$$\beta_i = x_{i1}\theta_1 + \dots + x_{ik}\theta_k \quad (i = 1, \dots, n)$$

with rational integers x_{11}, \dots, x_{nk} , and hence $\underline{x} = (x_{11}, \dots, x_{nk})$ has

$$(7.2) \quad |\underline{x}| \gg \ll |\underline{\beta}|.$$

Each

$$M_{ji}(\underline{\beta}^{(j)}) = L_{ji}(\underline{x}),$$

i.e. it is a linear form in \underline{x} . If there were a nontrivial relation

of linear dependence, say $\sum_{i,j} c_{ji} L_{ji} = 0$, then

$$\sum_{i=1}^n c_{ji} M_{ji} = 0 \quad (j = 1, \dots, k),$$

contradicting our hypothesis on M_{j1}, \dots, M_{jn} . So L_{11}, \dots, L_{kn} are linearly independent.

Now suppose that a vector $\underline{\beta} \in I^n$ satisfies (7.1) with $\epsilon/2$ in place of ϵ . Then by (7.2),

$$\prod_{j=1}^k \prod_{i=1}^n |L_{ji}(\underline{x})| < [\underline{x}]^{-\epsilon/3}$$

if $[\underline{x}]$ is large. By the rational Subspace Theorem (Ch. VI, Theorem 1F), applied with kn in place of n , there are finitely many rational subspaces S_1, \dots, S_h of \mathbb{Q}^{kn} , each of dimension $kn - 1$, such that \underline{x} lies in one of these subspaces. Each subspace S_t is defined by a linear relation in x_{11}, \dots, x_{nk} . It is also defined by a linear relation in the $\beta_i^{(j)}$:

$$(7.3) \quad c_{t11}\beta_1^{(1)} + \dots + c_{t1n}\beta_n^{(1)} + \dots + c_{tk1}\beta_1^{(k)} + \dots + c_{tkn}\beta_n^{(k)} = 0.$$

Pick $u = kh$ elements $\lambda_1, \dots, \lambda_u$ in I such that any k among them are linearly independent over \mathbb{Q} . Suppose $\underline{\beta} \in I^n$ satisfies (7.1). If $[\underline{\beta}]$ is large, then each $\underline{\beta}_s = \lambda_s \underline{\beta}$ ($s = 1, \dots, u$) satisfies (7.1) with ϵ replaced by $\epsilon/2$. Hence each $\underline{\beta}_s$ satisfies an equation (7.3). That is, we have

$$(7.4) \quad \lambda_s^{(1)} (c_{t11}\beta_1^{(1)} + \dots + c_{t1n}\beta_n^{(1)}) + \dots + \lambda_s^{(k)} (c_{tk1}\beta_1^{(k)} + \dots + c_{tkn}\beta_n^{(k)}) = 0$$

for some t in $1 \leq t \leq h$. Since $u = kh$, there is a t in $1 \leq t \leq h$ for which (7.4) is satisfied for at least k values of s in $1 \leq s \leq u$. Say it is satisfied for $s = 1, \dots, k$. The linear

independence of $\lambda_1, \dots, \lambda_k$ over Q implies the nonvanishing of the determinant $|\lambda_s^{(i)}|$ ($1 \leq i, s \leq k$), and it follows that

$$(7.5) \quad c_{tj1} \beta_1^{(j)} + \dots + c_{tjn} \beta_n^{(j)} = 0 \quad (j = 1, \dots, k).$$

There is a $j_0 = j_0(t)$ for which the equation (7.5) is nontrivial.

Since $\beta_1^{(j_0)}, \dots, \beta_n^{(j_0)}$ lie in the field $K^{(j_0)}$, the nontrivial relation

(7.5) with $j = j_0$ implies a nontrivial relation

$$d_{t1} \beta_1^{(j_0)} + \dots + d_{tn} \beta_n^{(j_0)} = 0$$

with coefficients d_{t1}, \dots, d_{tn} in $K^{(j_0)}$. From this we get a non-trivial relation

$$(7.6) \quad e_{t1} \beta_1 + \dots + e_{tn} \beta_n = 0$$

with coefficients e_{t1}, \dots, e_{tn} in K .

Thus every solution $\underline{\beta}$ of (7.1) with large $|\underline{\beta}|$ satisfies (7.6) for some t in $1 \leq t \leq h$, hence lies in one of h proper subspaces T_1, \dots, T_h of K^n . By adding a finite number of additional subspaces T_{h+1}, \dots, T_w , we may take care of $\underline{\beta}$ with small $|\underline{\beta}|$.

§8. Approximation to Algebraic Numbers by Elements of Number Fields.

THEOREM 8A. Let α be a real algebraic number and K a real algebraic number field. Given $\delta > 0$, there are at most finitely many elements $\beta \in K$ having

$$(8.1) \quad |\alpha - \beta| < H_K(\beta)^{-2-\delta}.$$

This was first proved by LeVeque (1955) who carried over Roth's arguments to number fields. Here we will deduce it from the Subspace

Theorem. A generalization to simultaneous approximation was carried out by Schmidt (1975). By applying the theorem to each subfield of K , we see that in fact

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

has only finitely many solutions $\beta \in K$, where $H(\beta)$ is the absolute height.

Theorem 8A may be used to generalize the theory of Thue Equations (Theorem 3A of Ch. V) to number fields.

If $\underline{\beta} = (\beta_0, \beta_1) \in K^2$, write

$$(8.2) \quad \overline{|\underline{\beta}|} = \max(|\beta_0^{(1)}|, |\beta_1^{(1)}|, \dots, |\beta_0^{(k)}|, |\beta_1^{(k)}|) ,$$

where again $k = \deg K$ and the upper indices denote conjugation.

LEMMA 8B. Given β in K , there is a vector $\underline{\beta} = (\beta_0, \beta_1)$ with components in the ring I of integers of K , with $\beta_0 \neq 0$, with

$$\beta = \beta_1 / \beta_0 ,$$

and with

$$(8.3) \quad H_K(\beta) \gg \ll \overline{|\underline{\beta}|}^k .$$

Proof. Let \mathfrak{B} be the (possibly fractional) ideal generated by $1, \beta$. In the ideal class of \mathfrak{B} there is an integral ideal \mathfrak{C} with norm $N(\mathfrak{C}) \ll 1$. We have $\mathfrak{C} = \mathfrak{B}(\theta)$ with a principal ideal (θ) , and \mathfrak{C} is generated by γ_0, γ_1 where $\gamma_0 = \theta$, $\gamma_1 = \beta\theta$. By Dirichlet's unit theorem there is a unit ε (of the ring I) such that the quantities $\beta_0 = \gamma_0 \varepsilon$, $\beta_1 = \gamma_1 \varepsilon$ have

$$\max(|\beta_0^{(i)}|, |\beta_1^{(i)}|) \gg \ll \left(\prod_{j=1}^k \max(|\gamma_0^{(j)}|, |\gamma_1^{(j)}|) \right)^{1/k} \quad (i = 1, \dots, k) .$$

Thus $\underline{\beta} = (\beta_0, \beta_1)$ satisfies

$$(8.4) \quad |\underline{\beta}| \gg \ll \max(|\beta_0^{(i)}|, |\beta_1^{(i)}|) \quad (i = 1, \dots, k) .$$

The polynomial

$$B(X) = \prod_{i=1}^k (\beta_0^{(i)} X - \beta_1^{(i)})$$

is an integral multiple of the field polynomial P_0 of β (see §2).

In fact from the analogue in K of Gauss' Lemma,

$$B(X) = \mathfrak{N}(\mathbb{C}) P_0(X) ,$$

since β_0, β_1 generate \mathbb{C} . We obtain

$$H_K(\beta) = |\overline{P_0}| \gg \ll |\overline{B}| \gg \ll M(B) = \prod_{i=1}^k \max(|\beta_0^{(i)}|, |\beta_1^{(i)}|) \gg \ll |\underline{\beta}|^k .$$

Proof of Theorem 8A. Given β with (8.1), construct $\underline{\beta}$ as in Lemma 8B. After multiplying (8.1) by

$$|(\beta_0^{(1)})^2 \beta_0^{(2)} \beta_1^{(2)} \dots \beta_0^{(k)} \beta_1^{(k)}| ,$$

we obtain

$$|(\alpha \beta_0^{(1)} - \beta_1^{(1)}) \beta_0^{(1)} \beta_0^{(2)} \beta_1^{(2)} \dots \beta_0^{(k)} \beta_1^{(k)}| \ll |\underline{\beta}|^{2k-2k-\delta k} = |\underline{\beta}|^{-\delta k} ,$$

from (8.3).

By Theorem 7A the solutions $\underline{\beta} \in I^2$ of this inequality lie in finitely many one dimensional subspaces. But $\beta = \beta_1/\beta_0$ is constant in each such subspace.

It is possible to give results in which both α and β are allowed to vary in a given number field K . One can construct a constant

$c_1 = c_1(d, \delta)$ such that in any number field K of degree d there are only finitely many pairs α, β with

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

§9. Approximation to Algebraic Numbers by Algebraic Numbers of Bounded Degree.

THEOREM 9A. Suppose α is a real algebraic number, and suppose $k \geq 1$, $\delta > 0$. Then there are only finitely many algebraic numbers β of degree $\leq k$ with

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

It may be shown that the $-k - 1$ in the exponent is best possible. Roth (1955a,b) had conjectured and Wirsing (1971) had proved a result with $-2k$ in place of $-k - 1$. For $k = 1$ the theorem becomes Roth's Theorem.

Mahler (1932) introduced for arbitrary real α the quantity $\omega_k = \omega_k(\alpha)$ which is the least upper bound of the numbers ω such that

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

has infinitely many solutions in polynomials P of degree $\leq k$ with integer coefficients. Clearly

$$\omega_1 \leq \omega_2 \leq \dots$$

If α is not algebraic of degree $\leq k$, then $\alpha^k, \dots, \alpha, 1$ are linearly independent over \mathbb{Q} , and it follows from Corollary 1D of Ch. II that

$$(9.1) \quad \omega_k \geq k.$$

If α is algebraic of degree d , then an obvious norm argument (like Lemma 1A of Ch. VI) shows that

$$\omega_k \leq d - 1.$$

On the other hand it follows from Corollary 1E in Ch. VI that

$0 < |P(\alpha)| < |P|^{-k-\delta}$ has only finitely many solutions in polynomials of degree $\leq k$ if α is algebraic, so that

$$(9.2) \quad \omega_k \leq k.$$

Hence

$$(9.3) \quad \omega_k = \min(k, d - 1)$$

if α is of degree d .

Koksma (1939) introduced the quantity $\omega_k^* = \omega_k^*(\alpha)$ which is the least upper bound of the numbers ω^* such that

$$(9.4) \quad |\alpha - \beta| < H(\beta)^{-\omega^*-1}$$

has infinitely many solutions in real algebraic numbers β of degree $\leq k$. Again

$$\omega_1^* \leq \omega_2^* \leq \dots$$

Now if $\beta \neq \alpha$ and if P is the defining polynomial of β , then

(9.4) implies

$$0 < |P(\alpha)| \ll |\alpha - \beta| |P| \ll |P|^{-\omega^*}.$$

Hence always

$$(9.5) \quad \omega_k \leq \omega_k^*.$$

Thus from (9.2) , $\omega_k^* \leq k$ is α is algebraic, and Theorem 9A follows.

We make a few additional remarks. The conjecture of §3 is that
(in analogy to (9.1))

$$\omega_k^* \geq k ,$$

unless α is algebraic of degree $\leq k$, but Theorem 3B gives only

$$(9.6) \quad \omega_k^* \geq (k+1)/2 .$$

Wirsing (1961) proved for α which are not algebraic of degree $\leq k$
that

$$(9.7) \quad \omega_k^* \geq (\omega_k + 1)/2 ,$$

which complements (9.5) , and which in conjunction with (9.1) yields
(9.6) .

Sprindzuk (1965) (see also (1977)) proved a conjecture of Mahler
(1932), to the effect that

$$(9.8) \quad \omega_k(\alpha) = \omega_k^*(\alpha) = k$$

for almost all α , in the sense of Lebesgue measure.

§10. Mahler's Classification of Transcendental Numbers. We give
a report without proofs. It follows from the last section that α is
algebraic if and only if the sequence $\omega_1, \omega_2, \dots$ is bounded, i.e. if

$$\omega_k \ll 1 .$$

Mahler (1932) calls these numbers A - numbers. He divides the remaining
(i.e. the transcendental) real numbers into three classes as follows.

The S - numbers are characterized by

(S) $\omega_k < 1$ but $\omega_k \ll k$,

the T - numbers by

(T) $\omega_k < k$, but each $\omega_k < \infty$,

the U - numbers by

(U) some $\omega_k = \infty$.

If (following Koksma (1939)) one defines an analogous classification in terms of the ω_k^* , then this classification in fact coincides with Mahler's classification by (9.5) , (9.7) . Almost every number is an S - number, which may be seen e.g. from Sprindzuk's result concerning (9.8) . Mahler proved that α, β belong to two distinct classes among S, T, U , then α, β are algebraically independent.

The existence of T - numbers had been left open by Mahler, but was proved by Schmidt (1968) (see also Schmidt (1971a) and Baker (1975)). Schmidt (1971a) constructs T - numbers with $\omega_k \ll k^3$, and Baker asserts the existence of T - numbers with $\omega_k \ll k^2$, but so far the existence of T - numbers with $\omega_k = o(k^2)$ has not been established.

An U - number is of degree k if $\omega_k = \infty$ but $\omega_j < \infty$ for $j < k$. The number α constructed in Chapter V below Liouville's Theorem is a U - number of degree 1 . LeVeque (1953) establishes the existence of U - numbers of prescribed degrees.

§11. A Theorem of Mignotte. Suppose α, β are algebraic but not conjugates[†]). Suppose they are of respective degrees d, e , with defining polynomials

[†]) For differences of conjugate algebraic numbers see Mignotte and Payafar (1979) .

$$A(X) = a_0(X - \alpha^{(1)}) \dots (X - \alpha^{(d)}) \quad \text{and} \quad B(X) = b_0(X - \beta^{(1)}) \dots (X - \beta^{(e)}) .$$

The resultant R of A and B is not zero. It follows that

$$\begin{aligned} 1 \leq |R| &= |\alpha - \beta| a_0^e b_0^d \prod_{\substack{i=j \\ (i,j) \neq (1,1)}}^d \prod_{j=1}^e |\alpha^{(i)} - \beta^{(j)}| \\ &\leq |\alpha - \beta| 2^{de} M(A)^e M(B)^d , \end{aligned}$$

where $M(A)$, $M(B)$ are Mahler's constants' of §3. Thus if we put $M(\alpha) = M(A)$ and $M(\beta) = M(B)$, then

$$|\alpha - \beta| \geq 2^{-de} M(\alpha)^{-e} M(\beta)^{-d} .$$

In particular, if $\alpha = 1$, then $d = 1$, $M(\alpha) = 1$, and

$$(11.1) \quad |1 - \beta| \geq 2^{-e} M(\beta)^{-1} .$$

THEOREM 11A. (Mignotte (to appear)). Suppose α, β are as above,
and put $\alpha^* = \max(1, |\alpha|)$. Suppose further that

$$(11.2) \quad 4d \log(8\alpha^* e M(\beta)) \leq e \log 2e .$$

Then

$$(11.3) \quad |\alpha - \beta| \geq (2e)^{-d} M(\alpha)^{-2e} \exp\left\{-3(de \log 2e)^{1/2} (\log(8\alpha^* e M(\beta)))^{1/2}\right\} .$$

In the estimates given earlier, we had assumed that α was fixed and that β was of bounded degree. The present estimate is different. It is most useful when $M(\beta)$ is fixed and the degree e of β varies. For example, when $\alpha = 1$ and when e is large compared to $M(\beta)$, then

$$(11.4) \quad |1 - \beta| \geq \exp(-4e^{1/2} \log 8e M(\beta)) ,$$

which is (for bounded $M(\beta)$ and large e) better than (11.1) .

The proof has some points in common with a recent result of Dobrowolski (1979) that for $\varepsilon > 0$ and for an algebraic integer β of degree $e > c_1(\varepsilon)$,

$$M(\beta) > 1 + (1 - \varepsilon)(\log \log e / \log e)^3,$$

unless β is a root of unity.

LEMMA 11B*. Suppose $A(X) = a_0(X - \alpha^{(1)}) \dots (X - \alpha^{(d)})$ is the defining polynomial of α . Then if $\{i_1, \dots, i_s\}$ is a subset of $\{1, \dots, d\}$,

$$a_0 \alpha^{(i_1)} \dots \alpha^{(i_s)}$$

is an algebraic integer.

This auxiliary lemma is well known and will not be proved here. See e.g. Schneider (1957, Hilfssatz 17).

LEMMA 11C. Let $P_{ij}(X)$ ($1 \leq i \leq m$, $1 \leq j \leq n$) be polynomials of degree $\leq p$ with rational integer coefficients. Write $L(P_{ij})$ for the sum of the absolute values of the coefficients of P_{ij} , and put

$$L_i = L(P_{i1}) + \dots + L(P_{in}) \quad (i = 1, \dots, m).$$

Let α be algebraic of degree d . If $L_i \neq 0$ ($i = 1, \dots, m$) and if

$$(11.5) \quad n > md,$$

then there is a nonzero integer point $\underline{x} = (x_1, \dots, x_n)$ with

$$(11.6) \quad P_{i1}(\alpha)x_1 + \dots + P_{in}(\alpha)x_n = 0 \quad (i = 1, \dots, m)$$

having

$$(11.7) \quad \left[\frac{x}{\underline{x}} \right] \leq [(\langle L_1 \dots L_m \rangle^{d_{M(\alpha)} mp})^{1/(n-md)}] + 1 .$$

This is a variant on Siegel's Lemma (Lemma 5B of Ch. V). See Mignotte and Waldschmidt (1978) .

Proof. Construct linear forms

$$\Omega_{ij}(\underline{x}) = P_{i1}(\alpha^{(j)})X_1 + \dots + P_{in}(\alpha^{(j)})X_n \quad (1 \leq i \leq m, 1 \leq j \leq d) .$$

Denote the right hand side of (11.7) by B . The $n + md$ inequalities in $\underline{x} = (x_1, \dots, x_n)$,

$$(11.8) \quad |x_i| \leq B \quad (i = 1, \dots, n) ,$$

$$(11.9) \quad |\Omega_{ij}(\underline{x})| < a_0^{-p/d} \quad (1 \leq i \leq m, 1 \leq j \leq d) ,$$

where a_0 is the leading coefficient of A^{\dagger} , define a symmetric convex set \mathfrak{R} in \mathbb{R}^n . The inequalities (11.8) alone define a cube \mathfrak{C} of volume $(2B)^n$. Now consider a fixed inequality (11.9) , and suppose that $\alpha^{(j)}$ is real. Then the inequality (11.9) defines a "strip", i.e. a symmetric set bounded by two parallel hyperplanes. For $\underline{x} \in \mathfrak{C}$,

$$|\Omega_{ij}(\underline{x})| \leq BL(\Omega_{ij}) ,$$

where $L(\Omega_{ij})$ is the sum of the absolute values of the coefficients of Ω_{ij} . Hence the intersection of \mathfrak{C} with the set defined by this particular inequality (11.9) will cut down the volume by a factor

$$\cong \min(1, a_0^{-p/d} B^{-1} L(\Omega_{ij})^{-1}) .$$

Or suppose that $\alpha^{(i)}, \alpha^{(j+1)}$ are conjugate complex, and consider a pair of inequalities (11.9) involving Ω_{ij} and $\Omega_{i,j+1}$. These inequalities

†) A is the defining polynomial of α .

define a circular cylinder. For $\underline{x} \in \mathbb{C}$,

$$|\mathfrak{Q}_{ij}(\underline{x})| = |\mathfrak{Q}_{i,j+1}(\underline{x})| \leq BL(\mathfrak{Q}_{ij}) = BL(\mathfrak{Q}_{i,j+1}) .$$

Hence the intersection with the set defined by this pair of inequalities will cut down the volume by a factor

$$\cong \min(1, a_0^{-2p/d} B^{-2} L(\mathfrak{Q}_{ij})^{-1} L(\mathfrak{Q}_{i,j+1})^{-1}) .$$

It follows that \mathcal{R} has volume

$$\begin{aligned} V(\mathcal{R}) &\cong (2B)^n \prod_{i=1}^m \prod_{j=1}^d \min(1, a_0^{-p/d} B^{-1} L(\mathfrak{Q}_{ij})^{-1}) \\ &\cong 2^{n_B n - md} \prod_{i=1}^m \prod_{j=1}^d (\max(1, a_0^{p/d} L(\mathfrak{Q}_{ij})))^{-1} . \end{aligned}$$

Now $L(\mathfrak{Q}_{ij}) \leq L_i(\max(1, |\alpha^{(j)}|))^p$, whence we get

$$V(\mathcal{R}) \geq 2^{n_B n - md} (L_1 \dots L_m)^{-d} M(\alpha)^{-mp} .$$

By our choice of B , $V(\mathcal{R}) > 2^n$, so that according to Minkowski there is a nontrivial solution \underline{x} of (11.8) and (11.9). Now by Lemma 11B,

$$a_0^p \mathfrak{Q}_{i1}(\underline{x}) \dots \mathfrak{Q}_{id}(\underline{x}) \quad (i = 1, \dots, m)$$

is an algebraic integer, and in fact it must be a rational integer. On the other hand by (11.9) it has absolute value < 1 , so that it must be zero. Thus (11.6) is satisfied.

LEMMA 11D. Suppose α is algebraic of degree d . For $n > md$, there exists a nonzero polynomial $P(X)$ of degree $< n$ with rational integer coefficients which has a zero at $x = \alpha$ of order $\geq m$, and which has

$$(11.10) \quad |\overline{P}| \leq \left[\left(\prod_{i=1}^m \binom{n}{i} \right)^d M(\alpha)^{mn} \right]^{1/(n-md)} + 1.$$

Proof. Set $P(X) = p_0 X^{n-1} + \dots + p_{n-2} X + p_{n-1}$. After dividing a typical condition $P^{(i)}(\alpha) = 0$ by $i!$, we get the conditions

$$p_0 \binom{n-1}{i} \alpha^{n-1-i} + p_1 \binom{n-2}{i} \alpha^{n-2-i} + \dots + a_i = 0 \quad (i = 0, 1, \dots, m-1).$$

We may apply Lemma 11C with $p = n - 1$ and with

$$L_i = \binom{n-1}{i} + \binom{n-2}{i} + \dots + 1 = \binom{n}{i+1} \quad (i = 0, 1, \dots, m-1).$$

Lemma 11D follows.

Henceforth denote the right hand side of (11.10) by C . It is clear that

$$(11.11) \quad |P(z)| \leq nC \max(1, |z|^n)$$

for every complex number z . On the other hand, by the maximum principle, the function $P(z)/(z - \alpha)^m$ takes its maximum modulus in the disk $|z| \leq 2\alpha^*$ on the boundary, so that for $|z| \leq 2\alpha^*$,

$$(11.12) \quad |P(z)| \leq |z - \alpha|^{nC} (2\alpha^*)^n |2\alpha^* - \alpha|^{-m} \leq |z - \alpha|^m nC 2^n \alpha^{*n-m}$$

Suppose now that $\alpha, \beta, A, B, d, e$ are as in Theorem 11A, and that

$$(11.13) \quad n = md + e.$$

Then P , since it is of degree $< n$, and since it is divisible by A^m , cannot be divisible by B , and hence $P(\beta) \neq 0$. Suppose also that

$$(11.14) \quad |\beta| \leq 2\alpha^*.$$

Estimating $P(\beta)$ by (11.12), and the conjugates $P(\beta^{(j)})$ ($j = 2, \dots, e$)

by (11.11) , we get

$$|P(\beta)| \leq |\alpha - \beta|^m n C 2^n \alpha^{*n-m},$$

$$|P(\beta^{(j)})| \leq n C \max(1, |\beta^{(j)}|^n) \quad (j = 2, \dots, e).$$

Now if b_0 is the leading coefficient of the defining polynomial of β , then $b_0^n P(\beta^{(1)}) \dots P(\beta^{(e)})$ is an algebraic integer by Lemma 11B. In fact it is a nonzero rational integer. It follows that

$$(11.15) \quad 1 \leq |\alpha - \beta|^m (nC)^e 2^n \alpha^{*n-m} b_0^n \prod_{j=2}^e \max(1, |\beta^{(j)}|^n) \\ \leq |\alpha - \beta|^m (nC)^e 2^n \alpha^{*n-m} M(\beta)^n.$$

Now suppose that

$$(11.16) \quad m \leq \frac{1}{2} n.$$

Then $\binom{n}{i} \leq \binom{n}{m} \leq n^m$ for $i \leq m$, and by (11.13) and the definition of C ,

$$C^e \leq 2^e n^{m^2 d} M(\alpha)^{mn}.$$

Substituting this into (11.15) we get

$$1 \leq |\alpha - \beta|^m M(\alpha)^{mn} M(\beta)^n n^{m^2 d + e} 2^{n+e} \alpha^{*n},$$

whence (since $e < n$),

$$(11.17) \quad 1 \leq |\alpha - \beta| M(\alpha)^n (4\alpha^{*n} M(\beta))^{n/m} n^{md}.$$

We now choose m to make the right hand side small. We put

$$m = \lceil (e \log(8\alpha^{*e} M(\beta)) / (d \log 2e))^{1/2} \rceil + 1.$$

Then

$$m d \leq (d e \log(8 \alpha^* e_M(\beta)) / \log 2e)^{1/2} + d \leq \frac{1}{2}e + \frac{1}{2}e = e$$

by (11.2) . In conjunction with (11.13) we get (11.16) . We also have $n \leq 2e$, and therefore

$$(4 \alpha^* n_M(\beta))^{n/m} \leq \exp \left(2 (d e (\log 2e) (\log(8 \alpha^* e_M(\beta))))^{1/2} \right) .$$

Finally

$$n^{m d} \leq (2e)^{m d} \leq (2e)^d \exp \left((d e (\log 2e) (\log(8 \alpha^* e_M(\beta))))^{1/2} \right) .$$

Substituting our estimates into (11.17) we get the desired (11.3) .

If (11.14) fails to hold, then $|\alpha - \beta| \geq 1$ and (11.3) is trivially true.

REFERENCES

- A. Baker (1967). Simultaneous approximation to certain algebraic numbers.
Proc. Camb. Phil. Soc. 63, 693-702.
- _____ (1967/68). Contributions to the theory of Diophantine equations.
I. On the representation of integers by binary forms. Philos. Trans.
Roy. Soc. London Ser A 263, 173-191.
- _____ (1975). Transcendental Number Theory. Cambridge University Press.
- A. Baker and W.M. Schmidt (1970). Diophantine Approximation and Hausdorff
Dimension. Proc. London Math. Soc. (3) 21, 1-11.
- A.S. Besicovitch (1934). Sets of fractional dimension (IV): On rational
approximations to real numbers. J. London Math. Soc. 9, 126-131.
- H.F. Blichfeldt (1914). A new principle in the geometry of numbers with
some applications. Trans. A.M.S. 15, 227-235.
- E. Borel (1903a). Sur l'approximation des nombres par des nombres
rationnels. C. R. Acad. Sci. Paris 136, 1054-1055.
- _____ (1903b). Contribution à l'analyse arithmétique du continu.
J. Math. pures appl. (5) 9, 329-375.
- Z.I. Borevich and I.R. Shafarevich (1966). Number Theory. (Translated
from the Russian). Academic Press.
- J.W.S. Cassels (1950a). Some metrical theorems in diophantine approximation
I. Proc. Camb. Phil. Soc. 46, 209-218.
- _____ (1950b). Some metrical theorems in diophantine approximation
III. Ibid. 219-225.
- _____ (1955a). Simultaneous diophantine approximation.

Journ. London Math. Soc. 30, 119-121.

_____ (1955b). Simultaneous diophantine approximation II.

Proc. Lon. Math. Soc. (3) 5, 435-448.

_____ (1956). On a result of Marshall Hall. *Mathematika* 3
109-110.

_____ (1957). An introduction to diophantine approximation.
Cambridge Tracts 45, Cambridge Univ. Press.

_____ (1959). An introduction to the Geometry of Numbers.
Springer Grundlehren 99. Berlin - Göttingen-Heidelberg.

T.W. Cusick (1975). The connection between the Lagrange and Markoff
Spectra. *Duke Math. J.* 42, 507-517.

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

_____ (1952). Simultaneous Diophantine approximation. *Proc. London
Math. Soc.* (3) 2, 406-416.

_____ (1954). Simultaneous Diophantine approximation. *Mathematika*
1, 51-72.

_____ (1962). A note on Diophantine approximation. *Studies in
Math. Analysis and Related Topics*, Stanford Univ. Press, 77-81.

_____ (1964). A note on Diophantine approximation II. *Mathematika*
11, 50-58.

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

H. Davenport and K. Mahler (1946). Simultaneous diophantine approximation.
Duke Math. J. 43, 105-111.

H. Davenport and K.F. Roth (1955). Rational approximation to algebraic

numbers. *Mathematika* 2, 160-167.

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

_____ (1968). Dirichlet's Theorem on diophantine approximation. *Rendic. Convegno Teoria dei numeri, Rome.* 113-132.

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

_____ (1970). Supplement to a theorem on linear forms. *Colloquia Math. Soc. János Bolyai (Number Th., Debrecen, Hung. 1968),* 15-25.

L.G.P. Dirichlet (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.

E. Dobrowolski (1979). On a question of Lehmer and the number of irreducible factors of a polynomial. *Acta. Arith.* 34, 391-401.

F.J. Dyson (1947a). The approximation to algebraic numbers by rationals. *Acta Math. Acad. Sci. Hung.* 79, 225-240.

_____ (1947b). On simultaneous Diophantine approximations. *Proc. Lond. Math. Soc. (2)* 49, 409-420.

P. Erdős (1970). On the Distribution of the Convergents of Almost All Real Numbers. *J. of Number Theory* 2, 425-441.

L. Euler (1737). De fractionibus continuis. *Commentarii Acad. Sci. Imperiali Petropolitanae* 9.

_____ (1748). *Introductio in analysin infinitorum* I.

N.I. Feldman (1951). The approximation of certain transcendental numbers. I. (Russian). *Izv. Akad. Nauk SSSR Ser Mat.* 15, 53-74.

- _____ (1970). Bounds for linear forms of certain algebraic numbers (Russian). *Mat. Zametki* 7, 569-580. English Transl. *Math. Notes* 7 (1970), 343-349.
- _____ (1971). An effective power sharpening of a theorem of Liouville (Russian). *Izv. Akad. Nauk SSSR Ser. Mat.* 35, 973-990.
- P. Furtwängler (1927-28). Über die Simultane Approximation von Irrationalzahlen I, II. *Math. Ann.* 96, 169-175 and 99, 71-83.
- P.X. Gallagher (1962). Metric simultaneous diophantine approximation. *J. Lon. Math. Soc.* 37, 387-390.
- _____ (1965). Metric simultaneous diophantine approximation. II *Mathematika* 12, 123-127.
- A.O. Gelfond (1952). *Transcendental and Algebraic Numbers.* (Russian). Moscow. English transl. (1969), Dover Publications, New York.
- W.H. Greub (1967). *Multilinear Algebra.* Springer Grundlehren 136.
- M. Hall (1947). On the sum and product of continued fractions. *Ann. of Math.* (2) 48, 966-993.
- G.H. Hardy and E.M. Wright (1954). *An introduction to the theory of numbers.* (3rd ed.) Oxford, Clarendon Press.
- C.J. Hightower (1970). The minima of indefinite binary quadratic forms. *J. Number Theory* 2, 364-378.
- A. Hurwitz (1891). Über die angenäherte Darstellung der Irrationalzahlen durch rationale Brüche. *Math. Ann.* 39, 279-284.
- _____ (1906). Über eine Aufgabe der unbestimmten Analysis. *Arch. Math. Phys.* 11 (1906), 185-196.

- V. Jarnik (1929). Diophantische Approximationen und Hausdorffsches Mass. *Recueil math. Moscow* 36, 371-382.
- _____ (1959). Eine Bemerkung zum Übertragungssatz. *Bulgar. Akad. Nauk Izv. Mat. Inst.* 3, No 2, 169-175.
- F. John (1948). Extremum problems with inequalities as subsidiary conditions. *Studies and Essays presented to R. Courant. Interscience*, New York, 187-204.
- A. Khintchine (1924). Einige Sätze über Kettenbrüche, mit Anwendungen auf die Theorie der Diophantischen Approximationen. *Math. Ann.* 92, 115-125.
- _____ (1925). Zwei Bemerkungen zu einer Arbeit des Herrn Perron. *Math. Z.* 22, 274-284.
- _____ (1926a). Zur metrischen Theorie der diophantischen Approximationen. *Math. Z.* 24, 706-714.
- _____ (1926b). Über eine Klasse linearer Diophantischer Approximationen. *Rend. Circ. Math. Palermo* 50, 170-195.
- J.F. Koksma (1939). Über die Mahlersche Klasseneinteilung der transzendenten Zahlen und die Approximation komplexer Zahlen durch algebraische Zahlen. *Mh. Math. Physik* 48, 176-189.
- L. Kuipers and H. Niederreiter (1974). *Uniform distribution of sequences.* John Wiley & Sons. New York-London-Sydney-Toronto.
- J.L. Lagrange (1770a). *Additions aux éléments d'algebra d'Euler.*
- _____ (1770b). *Additions au mémoire sur la résolution des équations numériques. Mém. Berl.* 24.

- S. Lang (1965). Report on diophantine approximations. Bull. de la Soc. Math. de France 93, 117-192.
- C.G. Lekkerkerker (1969). Geometry of Numbers. Wolters-Noordhoff Publishing and American Elsevier Publishing.
- W.J. LeVeque (1953). On Mahler's U-numbers. J. London Math. Soc. 28, 220-229.
- _____ (1955). Topics in Number Theory. Addison-Wesley Publ. Co.: Reading, Mass.
- J. Liouville (1844). Sur des classes très-étendues de quantités dont la irrationnelles algébriques. C.R. Acad. Sci. Paris 18, 883-885 and 910-911.
- J.M. Mack (1977). Simultaneous diophantine approximation. J. Austral. Math. Soc. 24 (Series A), 266-285.
- K. Mahler (1932). Zur Approximation der Exponentialfunktion und des Logarithmus I. J. reine ang. Math. 166, 118-136.
- _____ (1933). Zur Approximation algebraischer Zahlen (II). Über die Anzahl der Darstellungen ganzer Zahlen durch Binärformen. Math. Ann. 108, 37-55.
- _____ (1939). Ein Übertragungsprinzip für konvexe Körper. Časopis Pěst. Mat. Fys. 68, 93-102.
- _____ (1955). On compound convex bodies I. Proc. Lon. Math. Soc. (3), 5, 358-379.
- _____ (1960). An application of Jensen's Formula to Polynomials. Mathematika 7, 98-100.

- _____ (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.
- A.A. Markoff (1879). Sur les formes quadratiques binaires indéfinies. I. Math. Ann. 15, 381-409.
- M. Mignotte. Approximation des nombres algébriques par des nombres algébriques de grand degré. Ann. Fac. Sci. Toulouse, 1 (1979), 165-170.
- M. Mignotte and M. Payafar (1979). Distance entre les racines d'un polynôme. R.A.I.R.O. Analyse numérique 13, 181-192.
- M. Mignotte and M. Waldschmidt (1978). Linear forms and Schneider's method. Math. Ann. 231, 241-267.
- H. Minkowski (1896 & 1910). Geometrie der Zahlen. Teubner: Leipzig u. Berlin (The 1910 ed. prepared posthumously by Hilbert and Speiser).
- _____ (1907). Diophantische Approximationen Teubner: Leipzig u. Berlin.
- H. Montgomery (1975/76). Polynomials in Many Variables. Séminaire Delange-Pisot-Poitou (Théorie des nombres), no 7, 6 p.
- L.J. Mordell (1934). On some arithmetical results in the geometry of numbers. Compositio Math. 1, 248-253.
- I. Niven (1962). On asymmetric Diophantine approximations. Michigan Math. J. 9, 121-123.
- C.F. Osgood (1970). The simultaneous approximation of certain k -th roots. Proc. Camb. Phil. Soc. 67, 75-86.

- _____ (1977). Concerning a possible "Thue-Siegel-Roth Theorem" for algebraic differential equations. Number theory and algebra. Collected papers ded. to H.B. Mann, A.E. Ross and O. Taussky-Todd. New York-London: Academic Press.
- J.C. Oxtoby (1957). The Banach-Mazur game and Banach category theorem. Contrib. to the theory of games, vol. 3, 159-163. Annals of Math. Studies No. 39, Princeton Univ. Press.
- O. Perron (1921a). Über die Approximation irrationaler Zahlen durch rationale I, II. S.B. Heidelberg Akad. W.
- _____ (1921b). Über diophantische Approximationen. Math. Ann. 83, 77-84.
- _____ (1954). Die Lehre von den Kettenbrüchen. 3. Aufl. B.G. Teubner: Stuttgart.
- M. Ratliff (1978). The Thue-Siegel-Roth-Schmidt Theorem for Algebraic Functions. J. of Number Th. 10, 99-126.
- K.F. Roth (1955a). Rational approximations to algebraic numbers. Mathematika 2, 1-20.
- _____ (1955b). Corrigendum. Ibid, 168.
- A. Schinzel (1967). Review of a paper by Hyvär. Zentralblatt Math. 137, 257-258.
- H.P. Schlickewei (1976). Die p -adische Verallgemeinerung des Satzes von Thue-Siegel-Roth-Schmidt. J. reine ang. Math. 288, 86-105.
- _____ (1977). On Norm Form Equations. J. of Number Th. 9, 370-380.

W.M. Schmidt (1960). A metrical theorem in diophantine approximation .

Canad. J. Math. 12, 619-631.

_____ (1964). Metrical theorems on fractional parts of sequences.

Trans. Am. Math. Soc. 110, 493-518.

_____ (1965a). On badly approximable numbers. Mathematika 12, 10-20.

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

_____ (1966). On badly approximable numbers and certain games.

Trans. A.M.S. 123, 178-199.

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

_____ (1968). T-numbers do exist. Rendiconti convegno di Teoria
dei numeri, Roma, 3-26.

_____ (1969). Badly approximable systems of linear forms. J.

Number Theory 1, 139-154.

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

_____ (1971a). Mahler's T-numbers. Proc. of Symp. of Pure
Math. 20 (1969 Inst. of Number Theory) 275-286.

_____ (1971b). Linear forms with algebraic coefficients. I.

J. of Number Th. 3, 253-277.

_____ (1971c). Linearformen mit algebraischen Koeffizienten. II.

Math. Ann. 191, 1-20.

_____ (1971d). Approximation to algebraic numbers. L'Enseignement

Math. 17, 187-253.

_____ (1972). Norm form equations. Ann. of Math. 96, 526-551.

_____ (1973). Inequalities for Resultants and for Decomposable Forms. Proc. Conf. Dioph. Approximation and its Applications. (Ed. C.F. Osgood). Academic Press: New York and London.

_____ (1975). Simultaneous Approximation to Algebraic Numbers by Elements of a Number Field. Monatsh. Math. 79, 55-66.

W.M. Schmidt and Wang Y. A note on a transference theorem of linear forms. Scientia Sinica 22, 276-280.

Th. Schneider (1936). Über die Approximation algebraischer Zahlen. J. reine ang. Math. 175, 182-192.

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

B. Segre (1945). Lattice points in infinite domains and asymmetric Diophantine approximations. Duke Math. J. 12, 337-365.

J.A. Serret (1878). Handbuch der höheren Algebra, vol. I. (German translation).

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

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

V.G. Sprindzuk (1965). A proof of Mahler's conjecture on the measure of the set of S -numbers (Russian). Izv. Akad. Nauk SSSR Ser. Mat. 29, 379-436.

- _____ (1977). *Metrical Theory of Diophantine Approximation*.
(Russian). Moscow.
- A. Thue (1908). Om en generel i store hele tal uløstbar ligning. *Kra.*
Vidensk. Selsk. Skrifter. I. Mat. Nat. Kl. No. 7. Kra.
- _____ (1909). Über Annäherungswerte algebraischer Zahlen. *J. reine ang.*
Math. 135, 284-305.
- _____ (1977). *Selected Mathematical Papers*. Edited by T. Nagell,
A. Selberg, S. Selberg, K. Thalberg. Universitetsforlaget Oslo-
Bergen-Tromsø.
- K. Th. Vahlen (1895). Über Näherungswerte und Kettenbrüche. *J. reine*
angew. Math. 115, 221-233.
- B.L. Van der Waerden (1955). *Algebra* (4th ed.) Springer Grundlehren 33.
Berlin-Göttingen-Heidelberg.
- L. Vehmanen (1976). On linear forms with algebraic coefficients. *Acta*
Univ. Tamperensis, ser A, 71, 36 p.
- Wang Yuan, Yu Kunrui and Zhu Yao-cheng (1979). A note on a transference
theorem concerning the linear forms. *Acta Math. Sinica* 22.
- E. Wirsing (1961). Approximation mit algebraischen Zahlen beschränkten
Grades. *J. reine ang. Math.* 206, 67-77.
- _____ (1971). On approximation of algebraic numbers by algebraic
numbers of bounded degree. *Proc. of Symp. in Pure Math.* XX (1969
Number Theory Institute) 213-247.