Now the application of Lemma 3 gives

$$|d_0(x, D)| \ll b_{16}(\lg x)^{-\frac{1}{32}M} \quad \text{for} \quad x \geqslant D^{3^8+\varepsilon}.$$

Since M can be chosen as an arbitrary positive number, we have thus proved our theorem.

References

[1] Yu. V. Linnik, *On the least prime in an arithmetic progression*, I: Matem. Sbornik 15 (57) (1944), pp. 139-178; II: 15 (57) (1944), pp. 347-368.

[2] K. A. Rodosski, *On the least prime in an arithmetic progression*, Matem. Sbornik 34 (76) (1955), pp. 331-356 (Russian).

[3] Pang Cheng-Tung, *Least prime in an arithmetic progression*, Science Records (China) 1 (1957), pp. 311-313.

[4] A. G. Postnikov, *On the sum of the characters for a prime power modulus*, Izvestia AN SSSR, ser. matem., 19 (1955), pp. 11-16 (Russian).

[5] — *On Dirichlet's L-series with character modulus equal to a power of a prime number*, J. Indian Math., Ser. 20, 1-3 (1956), pp. 217-226.

[6] S. M. Rosin, *On the zeros of Dirichlet's L-series*, Izvestia AN SSSR, ser. matem., 23 (1959), pp. 503-508 (Russian).

[7] M. B. Barban, *The density of the zeros of Dirichlet's L-series and the problem on the addition of prime and "almost prime" numbers*, Matem. Sbornik 61 (4) (1963), pp. 418-425.

[8] N. G. Tschudakov, *Introduction to the theory of L-functions*, Moskow 1947 (Russian).

[9] Karl Prachar, *Primzahlverteilung*, Berlin 1957.

[10] A. Walfisz, *Weylsche Exponentialsummen in der neueren Zahlentheorie*, Berlin 1963.

Reçu par la Rédaction le 21. 1. 1964

ACTA ARITHMETICA
IX (1964)

On n-dimensional additive moduli and Diophantine approximations

by

A. M. Ostrowski (Basel)

Introduction. By the famous Kronecker's Theorem, if $\alpha_1, \ldots, \alpha_n$ are linearly independent, any point in the n-dimensional unity cube

$$C \ (0 \leqslant x_\nu < 1) \quad (\nu = 1, \ldots, n)$$

can be approximated by a point

$$(p\alpha_\nu - p_\nu) \quad (\nu = 1, \ldots, n)$$

for a positive integer p and integers p_ν. The question arises, how large p must be taken if we want to be able to approximate any point of C with the precision δ. The answer depends on the "degree of independence" of the α_ν, defined as the function $\eta(\varepsilon)$ given for any ε, $0 < \varepsilon < 1$, by

$$\eta(\varepsilon) = \operatorname{Inf} |m_1\alpha_1 + \ldots + m_n\alpha_n + m|,$$

where the integers $m_0, m_1, \ldots, m_n$ satisfy the inequality

$$0 < \sqrt{m_1^2 + \ldots + m_n^2} < 1/\varepsilon.$$

The first estimate of a bound for p was given by Landau [3]. A much better estimate was announced (1925) by Thomas [4], whose bound has the order of $\delta^{-n}/\eta(\delta)$; he gave also explicit numerical constants.

However, Thomas' paper written up with unusual carelessness is practically unreadable, as in particular its geometric part contains not only considerable gaps in the argumentation but also evidently erroneous statements (1).

As I needed a corresponding result in another investigation I lost some time trying to prove Thomas' statements in his way and finally decided to take up the geometric investigation of n-dimensional lattices ab

(1) In particular the formula on page 892: $OP_{s-1}^2 = O_sP_s^2 + {}_s\alpha_{s-1}^2 O_{s-1}P_{s-1}^2$ which appears out of the blue and is used in an essential way to obtain the final estimates, is certainly only true in exceptional cases.

ovo in order to clarify the geometric background of the whole situation. This study of metric relations in an n-dimensional lattice turned out to be quite rewarding — this is a more or less new chapter in the affine geometry. The results of my discussion, which could certainly be completed in several respects, are presented in the first sections §§ 1-6. In § 7 I apply this to derive a result corresponding to that of Thomas in which however I make use of the Euclidean length in the n-dimensional space instead of a different vector norm used by Thomas.

In this way, the numerical constants could be rather kept down. What numerical constants come out if Thomas' formulation and line of proof is completely carried out, that I do not know and, in my opinion, it must be left to Mr. Thomas to find this out.

The reader desiring to compare this paper with Mr. Thomas' article will easily see to what extent I made use of the ideas and lineas of argumentation contained or hidden implicitly or explicitly in Mr. Thomas' paper.

It may be finally observed that only a part of our results concerning lattices is really essential for the proof of the Theorem 3 in § 7. As a matter of fact, only the Lemma 1 and the content of §§ 5, 6 need be studied in order to get together the geometric results necessary for the § 7.

§ 1. Two lemmata on matrices.

1. LEMMA 1. *Consider an* $(n\times n)$*-matrix* $A=(a_{\mu\nu})$ *with* $a_{\mu\nu}=0$ $(\nu\geqslant\mu)$ *(a left triangular matrix with zeros along the diagonal). Then we can find two matrices* $G=(g_{\mu\nu})$, $E=(\varepsilon_{\mu\nu})$ *with integers* $g_{\mu\nu}$ *and* $g_{\mu\nu}=0$ $(\nu\geqslant\mu)$, *with*

$$-\tfrac{1}{2}<\varepsilon_{\mu\nu}\leqslant\tfrac{1}{2}\quad(\nu<\mu) \tag{1.1}$$

and $\varepsilon_{\mu\nu}=0$ $(\nu\geqslant\mu)$, *so that*

$$I+A=(I+G)(I+E), \tag{1.2}$$

and the $\varepsilon_{\mu\nu}$ *and* $g_{\mu\nu}$ *are here uniquely determined.*

2. Proof. For $n=1$ the assertion is obvious. We can therefore assume that the assertion has been already proved for matrices of the order $<n$.

Denote the matrices formed by the first $n-1$ rows and columns of A, G and E by $A^{(n-1)}$, $G^{(n-1)}$, $E^{(n-1)}$; then we have from (1.2)

$$I+A^{(n-1)}=(I+G^{(n-1)})(I+E^{(n-1)}) \tag{1.3}$$

and by our assumption $g_{\mu\nu}$ and $\varepsilon_{\mu\nu}$ from $G^{(n-1)}$ and $E^{(n-1)}$ can be chosen in a unique way conformally to the assertion of the Lemma, so that (1.3) holds.

3. Since the elements of the last column in (1.2) come out automatically right, we have only to show, that the elements of the last row in G and E:

$$\begin{matrix} g_{n1} & g_{n2} & \cdots & g_{n,n-1}, \\ \varepsilon_{n1} & \varepsilon_{n2} & \cdots & \varepsilon_{n,n-1} \end{matrix} \tag{1.4}$$

can be uniquely chosen so, that we have

$$a_{n\nu}=g_{n\nu}+\varepsilon_{n\nu}+\sum_{\nu<\sigma<n} g_{n\sigma}\varepsilon_{\sigma\nu}\quad(\nu=n-1,n-2,\ldots,1). \tag{1.5}$$

Now, for $\nu=n-1$ we have the condition

$$a_{n,n-1}=g_{n,n-1}+\varepsilon_{n,n-1}$$

and the integer $g_{n,n-1}$ can obviously be chosen so that

$$-\tfrac{1}{2}<a_{n,n-1}-g_{n,n-1}=\varepsilon_{n,n-1}\leqslant\tfrac{1}{2}.$$

4. Assume now that we have already dealt with (1.5) for $\nu=n-1$, $n-2,\ldots,k+1$ and obtained already in a unique way

$$g_{n,n-1},\ldots,g_{n,k+1};\ \varepsilon_{n,n-1},\ldots,\varepsilon_{n,k+1}.$$

Consider the equation (1.5) for $\nu=k$. Then the expression

$$S=\sum_{k<\sigma<n} g_{n\sigma}\varepsilon_{\sigma k}$$

can be considered as already known and our equation is reduced to

$$a_{nk}-S=g_{nk}+\varepsilon_{nk},$$

from which g_{nk} and ε_{nk} can be uniquely determined corresponding to our conditions. We see that all unknown (1.4) can be found in a unique way and the Lemma 1 is proved.

It need hardly be mentioned that a result completely corresponding to Lemma 1 is also correct in which we replace the left triangular matrices of this Lemma by the right triangular ones, and the proof is essentially the same. From this it follows immediately that the statement of the Lemma 1 remains true if we interchange the right-hand factors in the decomposition (1.2). Of course, the matrices G, E are then in the general case different from those in the Lemma 1.

5. The second Lemma will be formulated and proved for general matrices, while the special result about triangular matrices needed in the following will be stated in the Corollary to the Lemma 2.

We remind the reader that for a general $(m\times n)$-matrix $A\,(a_{\mu\nu})$, its *Frobenius norm*, $F(A)$, is defined as

$$F(A)=\sqrt{\sum_{\mu,\nu}|a_{\mu\nu}|^2} \tag{1.6}$$

and has the property that if the product AB of two matrices A, B exists, then

$$F(AB)\leqslant F(A)F(B). \tag{1.7}$$

We will consider a $(k\times(k+m))$-matrix $B\,(b_{\varkappa\nu})$. For such a matrix the square root of the sum of the squares of the moduli of all minors of B of order k will be denoted by $\|B\|$.

LEMMA 2. *Consider the $(k\times(k+m))$-matrix of the form*

$$B=\begin{pmatrix} a_{11} & \dots & a_{1k} & u_{11} & \dots & u_{1m}\\ \dots & \dots & \dots & \dots & \dots & \dots\\ a_{k1} & \dots & a_{kk} & u_{k1} & \dots & u_{km}\end{pmatrix}=(A,\,U) \tag{1.8}$$

where the $(k\times k)$-matrix $A\,(a_{\varkappa\lambda})$ has a determinant $|A|\neq 0$ and $U\,(u_{\varkappa\lambda})$ is a $(k\times m)$-matrix. Then we have

$$F(B)\leqslant\frac{\|B\|}{\|A\|}F(A). \tag{1.9}$$

6. Proof. Denote by S the inverse of A. Then, passing from B to SB, all minors of the order k of B are multiplied by $|S|$, so that

$$\|SB\|=\|S\|\cdot\|B\|. \tag{1.10}$$

On the other hand, put

$$SU=V=(v_{\varkappa\mu})\qquad(\varkappa=1,\dots,k;\ \mu=1,\dots,m). \tag{1.11}$$

Then we have

$$SB=\begin{pmatrix} 1 & 0 & \dots & 0 & v_{11} & \dots & v_{1m}\\ \dots & \dots & \dots & \dots & \dots & \dots & \dots\\ 0 & 0 & \dots & 1 & v_{k1} & \dots & v_{km}\end{pmatrix}. \tag{1.12}$$

We need only to compute some of the minors of the matrix (1.12). For the combination of the column indices $K=(1,2,\dots,k)$ the corresponding minor is $=1$. If we delete in the combination $(1,2,\dots,k)$ an index $\varkappa$ and add at the end the index $k+\mu$, the corresponding minor becomes $(-1)^{k-\varkappa}v_{\varkappa\mu}$. We have therefore

$$1+\sum_{\varkappa,\mu}|v_{\varkappa\mu}|^2=1+F(V)^2\leqslant\|SB\|^2=\frac{\|B\|^2}{\|A\|^2}.$$

On the other hand, we have from (1.11) $U=AV$ and therefore

$$F(B)^2=F(A)^2+F(U)^2\leqslant F(A)^2\left(1+F(V)^2\right)$$

and (1.9) follows immediately.

7. We specialize now the matrix B in (1.12) to the matrix

$$B_0=\begin{pmatrix} a_1 & a_{12} & \dots & a_{1k} & u_{11} & \dots & u_{1m}\\ 0 & a_2 & \dots & a_{2k} & u_{21} & \dots & u_{2m}\\ \dots & \dots & \dots & \dots & \dots & \dots & \dots\\ 0 & 0 & \dots & a_k & u_{k1} & \dots & u_{km}\end{pmatrix} \tag{1.13}$$

where the $a_\varkappa$ and $a_{\varkappa\lambda}$ satisfy the relations

$$|a_\varkappa|\geqslant 1\ \ (\varkappa=1,\dots,k);\ \ |a_{\varkappa\lambda}|\leqslant\tfrac{1}{2}|a_\varkappa|\ \ (\varkappa=1,\dots,k;\ \ \lambda=\varkappa+1,\dots,k). \tag{1.14}$$

In this case $F(A)^2$ is

$$\leqslant|a_1|^2\left(1+\frac{k-1}{4}\right)+|a_2|^2\left(1+\frac{k-2}{4}\right)+\dots+|a_k|^2$$

$$<\frac{k+3}{4}\sum_{\varkappa=1}^{k}|a_\varkappa|^2<k\sum_{\varkappa=1}^{k}|a_\varkappa|^2.$$

(1.9) becomes now

$$F(B_0)^2\leqslant k\frac{\sum_{\varkappa=1}^{k}|a_\varkappa|^2}{|a_1\dots a_k|^2}\|B_0\|^2$$

and using (1.14) we obtain now

$$F(B_0)\leqslant k\|B_0\|. \tag{1.15}$$

It follows the

COROLLARY TO THE LEMMA 2. *If the matrix B of the Lemma 2 is in particular a matrix B_0 in* (1.13) *satisfying* (1.14), *we have* (1.15).

§ 2. Additive vector moduli M. $D_M^{(k)}$.

8. We consider in what follows a set M of points (or vectors) in the n-dimensional real space R^n, for which the difference of two elements of M belongs again to M. Such a set will be called a *k-dimensional additive vector modulus* if it contains k and not more than k linearly independent vectors.

The length of a vector P, that is the distance of the point P from the origin, will be generally denoted by $|P|$. The distance of a point P from

a set S will be denoted by the symbols $|P, S| = |S, P|$ and correspondingly the distance of two sets of points, S_1, S_2, by $|S_1, S_2|$.

If L is the symbol of a linear manifold in R^n we consider L also as the symbol of an *operator* reducing a point or a set to its *orthogonal projection on* L. A linear manifold in R^n spanned by points of M will be called *rational* in M.

9. If $P_1, \ldots, P_k$ are k linearly independent vectors in R^n the set G of vectors

$$g_1P_1+\ldots+g_kP_k, \tag{2.1}$$

where $g_1, \ldots, g_k$ run through all integers, is a special k-dimensional additive vector modulus, which is called a *k-dimensional lattice*. Each k-dimensional additive vector modulus contains k-dimensional lattices.

If the lattice G in (2.1) is formed by the independent vectors $P_1, \ldots, P_k$, they are called a *base* of this lattice; the parallelepipedon C_0 formed by the vectors $P_1, \ldots, P_k$, is a *cell* of G. Adding to C_0 all vectors from G we obtain the set of parallelepipeda which form the *complete net of cells of* G, corresponding to the choice of the base $P_1, \ldots, P_k$. These cells have no interior points in common and cover completely the whole k-dimensional manifold spanned by $P_1, \ldots, P_k$. The length of the greatest diagonal of C_0, that is its diameter, is called the *diameter* of the base $P_1, \ldots, P_k$. The volume of C_0 is an invariant of the lattice; it is called the *determinant* of G and we will denote it by $\Delta(G)$. *From now on, let M be an n-dimensional additive vector modulus.*

10. For $k = 0, 1, \ldots, n-1$ we put [2]

$$D_M^{(k)} = \operatorname*{Sup}_{L^{(k)}} |M-L^{(k)}, L^{(k)}| = \operatorname*{Sup}_{L^{(k)}} \operatorname*{Inf}_{P\in M-L^{(k)}} |P, L^{(k)}|, \tag{2.2}$$

where $L^{(k)}$ runs through all k-dimensional linear manifolds (for $k = 0$: points) in R^n, rational in M. We write in particular

$$D_M = D_M^{(n-1)} \tag{2.3}$$

and call D_M the *free radius* of M.

[2] In the following formula and in what follows the difference $A-B$, where B is a set consisting of more than one point, is to be understood in the *set-theoretic sense*. $A-B$ is obtained from A by suppressing all points which also belong to B. On the contrary, if Q is a point, the difference $A-Q$ is to be understood in the *vectorial sense*. $A-Q$ is obtained from A by subtracting from every point of A the vector Q.

Clearly there exists for each $k = 0, 1, \ldots, n-1$ a sequence of points $P^{(\nu)}$ of M and of linear k-dimensional manifolds L_ν of R^n, rational in M, so that

$$P^{(\nu)} \in M-L_\nu, \qquad D_M^{(k)} = \lim_{\nu\to\infty} |P^{(\nu)}, L_\nu|. \tag{2.4}$$

11. Let G be an k-dimensional lattice contained in M, C a cell of G and D_0 its diameter. Then, without changing the values of $|P^{(\nu)}, L_\nu|$ in (2.4), we can assume that the foot of the perpendicular from $P^{(\nu)}$ on L_ν in (2.4) lies in C and that therefore $P^{(\nu)}$ lies in the sphere around a point of C with the radius $2D_0$. We can therefore assume further that in (2.4),

$$P^{(\nu)} \to P_0, \qquad L_\nu \to L_0, \tag{2.4^0}$$

where L_0 is a k-dimensional linear manifold having points in C.

It follows in particular that all $D_M^{(k)}$ are *finite*.

12. If $D_M^{(0)}$ is positive, M is called *discrete*, otherwise *dense*. If M is dense, then every point of M is an accumulation point of this set. We are going to show, that *if M is discrete, then M is a lattice* [3].

Let $P_1, \ldots, P_n$ be n linearly independent vectors from M. Then every vector P from M can be written in the form $P = x_1P_1+\ldots+x_nP_n$ and the vectors in the cell C formed by $P_1, \ldots, P_n$ have the coordinates $x_1, \ldots, x_n$ between 0 and 1. From $D_M^{(0)} > 0$ follows that the number of vectors from M in C is finite. Since the coordinates x_ν can be arbitrarily varied modulo 1 without P getting out of M, all x_ν must be *rational*, as otherwise, taking the multiples of a point of M with an irrational coordinate and reducing modulo 1, we would obtain an infinite number of different vectors of M in C. The common denominator of the coordinates of the vectors of M in C, N, is then the common denominator of all coordinates of points of M.

13. Let now a_1 be the smallest positive value of the coordinate x_1 corresponding to the points of M and P_1' a point of M with this coordinate. Then all x_1, corresponding to the points of M are multiples of a_1. This proves already, in the case $n = 1$, our assertion. We can therefore assume that this assertion is already proved for all smaller values of n.

Every vector of M can be now written in the form $yP_1'+P'$, with an integer y and a P' from M, the first coordinate of which vanishes and all other coordinates are multiples of $1/N$. These P' form therefore an additive $(n-1)$-dimensional modulus (in the R^{n-1} defined by $x_1 = 0$) which is discrete and for which our assertion can be assumed as proved. All these P' can be therefore represented as linear forms with integer

[3] This result is known. See Cassels, [1], pp. 78-80.

coefficients in $n-1$ of them, and we obtain for the general vector from M the corresponding representation, as asserted.

For a lattice formed by the n independent vectors $P_1, \dots, P_n$ every $D_M^{(k)}$ is positive and at least equal to the maximal height of a cell C, since, if this height corresponds to the $(n-1)$-dimensional face of C, we can take $L^{(k)}$ lying in the $(n-1)$-dimensional linear manifold through this face.

§ 3. The structure of dense moduli.

14. Assume now that M is dense. For a given $\varepsilon > 0$ let q_ε be the greatest number with the property that there exist in M q_ε independent vectors $P_1, \dots, P_{q_\varepsilon}$ with $|P_\varkappa| \leqslant \varepsilon$ $(\varkappa = 1, \dots, q_\varepsilon)$. Denote by L_ε the linear manifold spanned by $P_1, \dots, P_{q_\varepsilon}$. L_ε obviously depends only on ε and not on the choice of the vectors $P_1, \dots, P_{q_\varepsilon}$. If $\varepsilon_1 > \varepsilon > 0$, then obviously $L_\varepsilon \subset L_{\varepsilon_1}$, $q_\varepsilon \leqslant q_{\varepsilon_1}$. With $\varepsilon \downarrow 0$, the q_ε tend to an integer $q \geqslant 1$ and are, from an $\varepsilon = \varepsilon_0$ on, all equal to q, while L_ε for $\varepsilon \leqslant \varepsilon_0$ is a constant q-dimensional manifold L_0. Denote the set of the points of M in L_0 by M^*. Then M^* is a q-dimensional additive vector modulus and, since there exist in M, to any $\varepsilon > 0$, q independent vectors $P_1, \dots, P_q$ with $|P_1| + \dots + |P_q| \leqslant \varepsilon$, the lattice formed by these vectors has cells of diameter $\leqslant \varepsilon$. It follows that M^* *is everywhere dense in* L_0. The number q, the dimension of M^*, will be called the *density dimension* of M.

Two vectors of R^n the difference of which belongs to M^* will be called *congruent modulo* M^*.

The q-dimensional linear manifold spanned by the points of M^* will be called R^*. The $(n-q)$-dimensional linear manifold through the origin orthogonal on R^* will be called $\bar{R}$. The orthogonal projection of a point P or a set S of points from R^n upon $\bar{R}$ will be denoted resp. by $\bar{P}, \bar{S}$. The projection of M upon $\bar{R}$, $\bar{M}$, is again an additive vector modulus (in $\bar{R}$).

15. Introduce an orthogonal system of coordinates, $x_1, \dots, x_n$ in R^n so that the manifolds R^* and $\bar{R}$ are resp.

$$x_1 = \dots = x_{n-q} = 0, \qquad x_{n-q+1} = \dots = x_n = 0.$$

We prove now:

a) $\bar{M}$ *is* $(n-q)$*-dimensional.* Otherwise all points of $\bar{M}$ would satisfy an equation of the form

$$\sum_{\nu=1}^{n-q} a_\nu x_\nu = 0$$

and the same equation would hold for all points of M, while M is n-dimensional.

b) $\bar{M}$ *is discrete.* Otherwise there would exist in $\bar{M}$, for every $\varepsilon > 0$ a vector $\bar{P} \neq 0$ with the coordinates $x_\nu^{(0)}$ $(\nu = 1, \dots, n-q)$ such that

$|x_\nu^{(0)}| < \varepsilon$ $(\nu = 1, \dots, n-q)$. The coordinates of a corresponding vector P from M may be $x_1^{(0)}, \dots, x_{n-q}^{(0)}, x_{n-q+1}, \dots, x_n$. Since M^* is everywhere dense in R^* there exists in M a vector $P' \equiv P \pmod{M^*}$ with coordinates $x_1^{(0)}, \dots, x_{n-q}^{(0)}; x_{n-q+1}^{(0)}, \dots, x_n^{(0)}$ such that all $|x_\nu^{(0)}|$ are $< \varepsilon$. As P' lies outside of R^*, the density dimension of M would be $> q$.

Let now $\bar{P}_1, \dots, \bar{P}_{n-q}$ be a base of $\bar{M}$ and $P_1, \dots, P_{n-q}$ a set of points from M congruent to the $\bar{P}_\nu \pmod{M^*}$. Then the points P_ν $(\nu = 1, \dots, n-q)$ have the property that every point P of M is congruent mod M^* to a linear combination of these points:

$$\sum_{\nu=1}^{n-q} a_\nu P_\nu$$

with integer coefficients a_ν which are uniquely determined by P, as soon as the $\bar{P}_\nu$ $(\nu = 1, \dots, n-q)$ are chosen. We have therefore

$$P = \sum_{\nu=1}^{n-q} a_\nu P_\nu + P^*$$

where the integers a_ν and the element P^* of M^* are uniquely determined by P as soon as the $\bar{P}_\nu$ $(\nu = 1, \dots, n-q)$ are chosen. Such a system of $n-q$ points P_ν from M is called a *base of* M *relative to* M^*.

16. We are now going to prove the

LEMMA 3. *We have in the notations of the section* 10, *if* q *is the density dimension of* M,

$$D_M^{(k)} = 0 \quad (k = 0, 1, \dots, q-1), \qquad D_M^{(q)} = D_{\bar{M}}^{(0)}, \tag{3.1}$$
$$D_M^{(k)} \geqslant D_{\bar{M}}^{(k-q)} \quad (k = q+1, \dots, n-1).$$

Proof. Let L be a k-dimensional linear manifold from R^n, rational in M, and Q_0 a point of M from L.

Subtracting Q_0 from all points of L we obtain a manifold L' containing the origin. Since this operation does not change M, we will only, in the discussion of $D_M^{(k)}$, consider the manifolds L containing the origin.

Now, if L does not contain the whole M^*, any neighborhood of the origin contains points of M^* not on L, and we see that $|L, M-L| \leqslant |L, M^*-L| = 0$. This proves in particular that $D_M^{(k)} = 0$ $(k = 0, 1, \dots, q-1)$.

17. Assume now that L contains M^*, and therefore R^*. If $k \neq q$, obviously $L = R^*$. We have therefore

$$D_M^{(q)} = \inf_{P \in M - M^*} |P, R^*|.$$

Projecting here P and R^* upon $\bar{R}$, we have $|P, R^*| = |\bar{P}, 0| = |\bar{P}|$, and therefore

$$D_M^{(q)} = \operatorname*{Inf}_{P \in M - M^*} |\bar{P}|.$$

On the other hand, if P runs through $M - M^*$, $\bar{P}$ runs through $\bar{M}$, excluding the origin. But then $\operatorname{Inf} |\bar{P}| = D_{\bar{M}}^{(0)}$, and we see that indeed $D_M^{(q)} = D_{\bar{M}}^{(0)}$.

18. From now on we assume that $k \geqslant q+1$. A general point $\bar{P}$ of $\bar{M} - \bar{L}$ may be the projection of a point P of M. If P were $\in L$, then $\bar{P}$ would be $\in \bar{L}$. Therefore we would have $P \in M - L$. The foot of the perpendicular from P upon L may be denoted by F. Then the projection of the segment FP upon $\bar{L}$ is the segment $\bar{F}\bar{P}$, so that

$$0 < |\bar{F}\bar{P}| \leqslant |FP| = |P, L| \leqslant D_M^{(k)}.$$

On the other hand $|\bar{F}\bar{P}| \geqslant |\bar{P}, \bar{L}|$, so that

$$|\bar{P}, \bar{L}| \leqslant D_M^{(k)}, \qquad |\bar{M} - \bar{L}, \bar{L}| \leqslant D_M^{(k)}$$

and therefore

$$D_{\bar{M}}^{(k-q)} = \operatorname*{Sup}_{\bar{L}} |\bar{M} - \bar{L}, \bar{L}) \leqslant D_M^{(k)};$$

(3.1) is completely proved.

§ 4. Bases of sublattices.

19. LEMMA 4. *Let M be an n-dimensional lattice and Λ a sublattice of M of dimension k. For a certain constant $K(M)$ depending only on M there exists a base of Λ, $P_1^*, \ldots, P_k^*$, satisfying the inequality*

$$\sqrt{|P_1^*|^2 + \ldots + |P_k^*|^2} \leqslant K(M)\, \Delta(\Lambda). \tag{4.1}$$

Proof. Let P_μ $(\mu = 1, \ldots, n)$ be a base of M. Introducing an orthogonal system of coordinates in R^n, let generally the coordinates of the point P_μ be $x_{\mu 1}, \ldots, x_{\mu n}$. The $(n \times n)$-matrix $(x_{\mu\nu})$ may be denoted by X.

For each fixed $k = 1, \ldots, n$ we consider all minors of the order k from the inverse matrix X^{-1} of X and form the square root of the sum of their squares. This is T_k.

In particular, T_n^{-1} is the modulus of the determinant of X. T_n^{-1} is also, as is well known, the volume of the parallelepipedon formed by $P_1, \ldots, P_n$, so that $T_n = \Delta(M)^{-1}$.

20. A base $P_1^*, \ldots, P_k^*$ of Λ can be expressed linearly with integer coefficients in terms of $P_1, \ldots, P_n$. This base can be transformed by a well known procedure in such a way that we have, interchanging if necessary the order of the P_ν,

$$P_\varkappa^* = a_\varkappa P_\varkappa + \sum_{\nu=\varkappa+1}^{k} a_{\varkappa\nu} P_\nu + \sum_{\nu=k+1}^{n} a_{\varkappa\nu} P_\nu \qquad (\varkappa = 1, \ldots, k) \tag{4.2}$$

with the matrix B_0:

$$B_0 = \begin{pmatrix} a_1 & a_{12} & \ldots & a_{1k} & a_{1,k+1} & \ldots & a_{1n} \\ 0 & a_2 & \ldots & a_{2k} & a_{2,k+1} & \ldots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \ldots & a_k & a_{k,k+1} & \ldots & a_{kn} \end{pmatrix}. \tag{4.3}$$

The $a_\varkappa$ are here positive integers, the $a_{\varkappa\nu}$ are integers satisfying in particular the relations

$$|a_{\varkappa\nu}| \leqslant \tfrac{1}{2} a_k \qquad (\varkappa < \nu \leqslant k). \tag{4.4}$$

To obtain a convenient inequality for $\|B_0\|$ we use the (k-dimensional) volume $\Delta(\Lambda)$ of a cell of Λ.

21. Indeed, we obtain the volume of the parallelepipedon formed by the $P_\varkappa^*$ in the following way:

Denote by K the general combination

$$K = (\mu_1, \ldots, \mu_k), \qquad \mu_1 < \mu_2 < \ldots < \mu_k, \tag{4.5}$$

of k among the n elements $(1, 2, \ldots, n)$ and denote by B_K the determinant formed by the corresponding columns of the matrix B_0.

The minor of the order k of the matrix X corresponding to the rows with the indices of the combination K (4.5) and the columns with the indices of the combination K' $(\nu_1 < \nu_2 < \ldots < \nu_k)$ may be denoted by $X_{KK'}$:

$$X_{KK'} = \begin{vmatrix} x_{\mu_1\nu_1} & \ldots & x_{\mu_1\nu_k} \\ \cdot & \cdot & \cdot \\ x_{\mu_k\nu_1} & \ldots & x_{\mu_k\nu_k} \end{vmatrix}.$$

Denote the coordinates of the vector $P_\varkappa^*$ in (4.2) by $y_{\varkappa 1}, \ldots, y_{\varkappa n}$ and by Y_K the minor of the matrix $Y = (y_{\varkappa\nu})$, corresponding to (4.5):

$$Y_K = \begin{vmatrix} y_{1\mu_1} & \ldots & y_{1\mu_k} \\ \cdot & \cdot & \cdot \\ y_{k\mu_1} & \ldots & y_{k\mu_k} \end{vmatrix}.$$

Then $\Delta(\Lambda)$ is $\sqrt{\sum_K Y_K^2}$.

22. The set of combinations (4.5) may be denoted in an arbitrary but fixed order by $K_1, \dots, K_{\binom{n}{k}}$. Then we have

$$Y_{K_\mu} = \sum_{\nu=1}^{\binom{n}{k}} X_{K_\mu K_\nu} B_{K_\nu} \quad (\mu = 1, \dots, \tbinom{n}{k}), \tag{4.6}$$

where the $X_{K_\mu K_\nu}$ form the *k-th compound* $X^{(k)}$ of the matrix X. This compound matrix is non-singular, since the determinant of X is $\neq 0$.

Solving (4.6) with respect to the B_{K_ν} we obtain

$$B_{K_\nu} = \sum_{\mu=1}^{\binom{n}{k}} U_{K_\nu K_\mu} Y_{K_\mu} \quad (\nu = 1, \dots, \tbinom{n}{k}) \tag{4.7}$$

where the matrix $(U_{K_\nu K_\mu})$ of the order $\binom{n}{k}$ is the kth compound of the matrix X^{-1}. But then we have from (4.7)

$$\sum_{\nu=1}^{\binom{n}{k}} B_{K_\nu}^2 \leqslant \Big(\sum_{\nu,\mu=1}^{\binom{n}{k}} U_{K_\nu K_\mu}^2 \Big) \sum_{\mu=1}^{\binom{n}{k}} Y_{K_\mu}^2 .$$

Here the last factor is $\Delta(\Lambda)^2$ while the first right-hand factor is the square of the expression T_k introduced in the section 19.

Thus we obtain

$$\sum_{\nu=1}^{\binom{n}{k}} B_{K_\nu}^2 \leqslant T_k^2 \Delta(\Lambda)^2 ,$$

or, observing that the left-hand sum is, in the notation of the section 5, $\|B_0\|^2$,

$$\|B_0\| \leqslant T_k \Delta(\Lambda). \tag{4.8}$$

But in our case the matrix B_0 in (4.3) has the form of B_0 in (1.13) if we identify the $u_{\varkappa\mu}$ with the $a_{\varkappa, k+\mu}$ and use (4.4). Applying (1.15) we have

$$F(B_0) \leqslant k T_k \Delta(\Lambda). \tag{4.9}$$

Further, the left-hand expression in (4.1) is $F(Y)$ and since $Y = B_0 X$, we have from (4.9)

$$F(Y) \leqslant F(X) F(B_0) \leqslant k T_k F(X) \Delta(\Lambda) \tag{4.10}$$

and this is (4.1) with $K(M) = k T_k F(X)$. The Lemma 4 is proved.

23. LEMMA 5. *For an n-dimensional lattice M and a k-dimensional linear manifold L, rational in M, assume that we have for all points P of $M-L$*

$$|P, L| > D \quad (P \in M-L). \tag{4.11}$$

Then if Q_0 is a point of M in L, the intersection of M and L can be written in the form

$$ML = Q_0 + \Lambda \tag{4.12}$$

where Λ is a k-dimensional lattice and has a base $P_1^*, \dots, P_k^*$ such that, for a constant $c(M)$ depending only on M, we have

$$\sqrt{|P_1^*|^2 + \dots + |P_k^*|^2} \leqslant \frac{c(M)}{D^{n-k}}. \tag{4.13}$$

24. Proof. Without loss of generality we can assume that $Q_0 = 0$, as we can replace L by $L - Q_0$. We have to derive an estimate of $\Delta(\Lambda)$ in order to be able to use the Lemma 4. Consider a base $V_1, \dots, V_k$ of Λ and the cell C formed by these k vectors. Then C has one vertex in the origin and the opposite vertex is $W = V_1 + \dots + V_k$. All 2^k vertices of C lie on the boundary of C (the boundary with respect to L), while there are no points of M in the interior of C. Consider now the "doubled cell" C^* obtained from C by (2:1)-dilatation from the origin. C^* is a k-dimensional parallelepipedon formed by the vectors $2V_1, \dots, 2V_k$. It has in its interior from the lattice M only the point W, while all other points of M belonging to C^* lie on the boundary of C^*.

25. Translating now C^* by $-W$ we obtain the parallelepipedon $C^* - W$ which has the origin in its interior and is symmetric with respect to the origin, but does not contain any other points of M in its interior. Applying therefore a dilatation from the origin in the ration $(1-\varepsilon):1$ we obtain a k-dimensional parallelepipedon

$$F \equiv (1-\varepsilon)(C^* - W),$$

which is symmetric with respect to the origin and does not contain any points of M save the origin. For its k-dimensional volume $|F|$ we have

$$|F| = (1-\varepsilon)^k 2^k \Delta(\Lambda). \tag{4.14}$$

26. Through each point P of F we consider the $(n-k)$-dimensional linear manifold *orthogonal* on L, $E(P)$, and consider the points of $E(P)$ in the distance $\leqslant D$ from P.

The set of all these points forms an n-dimensional convex body F^* which lies in the Riemannian product of F and the $(n-k)$-dimensional sphere around the origin with the radius D, $\Phi_D^{(n-k)}$:

$$F^* = F \times \Phi_D^{(n-k)}.$$

For the n-dimensional volume $|F^*|$ of F^*, we have

$$|F^*| = |F|\,|\Phi_D^{(n-k)}| = (1-\varepsilon)^k 2^k \Delta(\Lambda)\,\varrho_{n-k} D^{n-k}, \tag{4.15}$$

where ϱ_{n-k} is a positive constant (4).

27. On the other hand, it follows from (4.11) that F^* does not contain any points of M save the origin. By Minkowski's Theorem we have therefore

$$|F^*| \leqslant 2^n \Delta(M),$$

and, using (4.15), obtain

$$\Delta(\Lambda) \leqslant \frac{2^{n-k}}{(1-\varepsilon)^k} \cdot \frac{1}{\varrho_{n-k}} \cdot \frac{\Delta(M)}{D^{n-k}},$$

and finally for $\varepsilon \downarrow 0$

$$\Delta(\Lambda) \leqslant \frac{c_k(M)}{D^{n-k}}, \qquad c_k(M) = \frac{2^{n-k}}{\varrho_{n-k}} \Delta(M). \tag{4.16}$$

Since $c_k(M)$ has, for a given lattice M, only a finite number of different values, we obtain, introducing (4.16) into (4.1), the estimate (4.13) and the Lemma 5 is proved.

§ 5. Relations between the $D_M^{(k)}$.

28. LEMMA 6. *Let, for a fixed $k = 0, 1, \ldots, n-1$, L be a k-dimensional linear manifold of R^n, rational in M. Then, if M is a lattice, we have for a convenient point P_0 of $M-L$:*

$$D \equiv |M-L, L| = \operatorname*{Inf}_{P \in M-L} |P, L| = |P_0, L|, \qquad P_0 \in M-L. \tag{5.1}$$

Proof. Since the points of M lying in L form a discrete k-dimensional additive modulus, this modulus is a lattice and L is covered by the net of congruent parallelepipedic cells of diameter Δ. By definition, we have for a sequence of points P_ν from $M-L$:

$$D+1 > D_\nu \equiv |P_\nu, L| \to D, \qquad P_\nu \in M-L. \tag{5.2}$$

29. Denote by F_ν the foot of the perpendicular from P_ν upon L and assume that F_ν lies in the cell C_ν of L:

$$P_\nu F_\nu \perp L, \qquad F_\nu \in C_\nu.$$

Denote by C^* a fixed cell of the lattice M. Then we have for each ν a vector V_ν from M so that $F_\nu - V_\nu$ lies in C^*. Denote by U^* the set of all points of R^n which have from the points of C^* a distance not greater than $\operatorname{Max}(\Delta, D+1)$.

30. U^* contains obviously only a finite number, say N, of points of the lattice M. Then the points $P_\nu - V_\nu$ lie in U^* and the same holds for the set of points $C_\nu - V_\nu$.

On the other hand, we have obviously, as $C_\nu \subset L$,

$$D_\nu = |P_\nu, F_\nu| = |P_\nu, C_\nu| = |P_\nu - V_\nu, C_\nu - V_\nu|. \tag{5.3}$$

But here $P_\nu - V_\nu$ is a point of M lying in U^* and belongs therefore to a finite set of N points.

On the other hand, $C_\nu - V_\nu$ is a k-dimensional parallelepipedon with vertices belonging to M and to U^*, so that there are only a finite number of different among these $C_\nu - V_\nu$. We see that the D_ν in (5.3) can only have a finite number of different values, and it follows now that in (5.2) we have from a certain ν on the equality $|P_\nu, L| = D$. The Lemma 6 is proved.

31. We can now prove the

THEOREM 1. *Let M be an n-dimensional lattice. Then for every $k = 1, \ldots, n-1$, there exists a point $P^{(k)}$ and a linear k-dimensional manifold $L^{(k)}$ from R^n, rational in M, so that*

$$D_M^{(k)} = |P^{(k)}, L^{(k)}|, \qquad P^{(k)} \in M - L^{(k)}. \tag{5.4}$$

Proof. We have, for any fixed k, by definition (2.2)

$$D_M^{(k)} = \lim_{\nu \to \infty} \operatorname*{Inf}_{P \in M - L_\nu} |L_\nu, P|, \tag{5.5}$$

where L_ν runs through certain k-dimensional linear manifolds from R^n, rational in M.

Observe that for any Q_0 from M we have

$$|L_\nu, P| = |L_\nu - Q_0, P - Q_0|, \qquad \operatorname*{Inf}_{P \in M - L_\nu} |L_\nu, P| = \operatorname*{Inf}_{P \in M - (L_\nu - Q_0)} |L_\nu - Q_0, P|.$$

We can therefore assume, without loss of generality, that in (5.5) all L_ν contain the origin. Further, since $D_M^{(k)} > 0$, it can be assumed that for all L_ν, $\nu = 1, 2, \ldots$,

$$\operatorname*{Inf}_{P \in M - L_\nu} |L_\nu, P| > D \equiv \tfrac{1}{2} D_M^{(k)}.$$

32. But then, by the Lemma 5, the sublattices ML_ν have each a base $P_1^{*(\nu)}, \ldots, P_k^{*(\nu)}$ satisfying (4.13). The vectors $P_\varkappa^{*(\nu)}$ belong to M and as M is assumed to be a lattice and therefore discrete, there is only a finite number of possibilities for the vectors $P_\varkappa^*$.

(4) As a matter of fact we have $\varrho_{n-k} = 2\pi^{(n-k)/2}/[(n-k)\Gamma((n-k)/2)]$.

We can therefore assume, replacing the sequence L_ν by a convenient subsequence, that all bases $P_1^{*(\nu)}, \dots, P_k^{*(\nu)}$ are the same. But then the L_ν must coincide with a fixed linear manifold L_0 and we have

$$D_M^{(k)} = \operatorname*{Inf}_{P \in M-L_0} |L_0, P|.$$

And now the assertion of the Theorem follows immediately from the Lemma 6.

33. We are now going to prove

THEOREM 2. *We have for any additive n-dimensional vector modulus M and for any $k < n-1$:*

$$D_M^{(k)} \leqslant \sqrt{\tfrac{1}{3}} D_M^{(k+1)} \quad (0 \leqslant k \leqslant n-2). \tag{5.6}$$

Proof. We can assume that $D_M^{(k)}$ is positive. Put $D_M^{(k)} = D$, $D_M^{(k+1)} = D^*$ and choose an arbitrary positive $\delta < D$. It is then sufficient to prove

$$(D^* + \delta)^2 \geqslant (D-\delta)^2 - \tfrac{1}{4}(D+\delta)^2 \tag{5.7}$$

since for $\delta \downarrow 0$ (5.6) follows immediately.

Choose the k-dimensional linear manifold L through the origin so that

$$D \geqslant |M-L, L| > D-\delta$$

and a point P_1 of $M-L$ so that

$$|P_1, L| < |M-L, L| + \delta \leqslant D+\delta.$$

We can then write for a Q from L and a vector π orthogonal on L:

$$P_1 = \pi + Q, \quad \pi \perp L, \quad Q \in L.$$

Then obviously $|P_1, L| = |\pi|$ and therefore

$$|\pi| < D+\delta. \tag{5.8}$$

34. Let L_1 be the $(k+1)$-dimensional linear manifold spanned by L and P_1. Then there exists in $M-L_1$ a point P_2 such that

$$|P_2, L_1| < D^* + \delta.$$

P_2 can be written, for a point Q_1 of L, a vector π_1 orthogonal on L_1 and a scalar a in the form

$$P_2 = \pi_1 + a\pi + Q_1, \quad \pi_1 \perp L_1, \quad Q_1 \in L.$$

For an integer g we have

$$P_2 - gP_1 = \pi_1 + (a-g)\pi + Q_1 - gQ.$$

Integer g can be chosen so that $|a-g| \leqslant \frac{1}{2}$. Then, putting $\varepsilon = a-g$, $Q_2 = Q_1 - gQ$:

$$P^* \equiv P_2 - gP_1 = \pi_1 + \varepsilon\pi + Q_2, \quad Q_2 \in L, \quad |\varepsilon| \leqslant \tfrac{1}{2}. \tag{5.9}$$

Further, since P_1 lies in L_1,

$$|P^*, L_1| = |P_2, L_1| = |\pi_1| < D^* + \delta. \tag{5.10}$$

35. Let now E be the 2-dimensional manifold spanned by π and π_1, which is orthogonal on L. Applying to (5.9) the operator E, that is projecting on E, we obtain

$$EP^* = E\pi_1 + \varepsilon E\pi + EQ_2$$

and therefore, as π and π_1 lie in E, while Q_2 is orthogonal on E:

$$EP^* = \pi_1 + \varepsilon\pi.$$

It follows further, since π and π_1 are orthogonal:

$$|EP^*|^2 = |\pi|^2 + \varepsilon^2 |\pi_1|^2. \tag{5.11}$$

But here the right-hand expression is, by (5.8), (5.9) and (5.10) $< (D+\delta)^2 + \frac{1}{4}(D^*+\delta)^2$, while $|EP^*|$, by (5.11), is > 0 and therefore

$$|EP^*| = |P^*, L| \geqslant |M-L, L| > D-\delta.$$

(5.7) and our Theorem is proved.

§ 6. Reduced base of a lattice.

36. If $P_1, \dots, P_n$ is a base of an n-dimensional lattice M we can obtain from this base by the *shearing reduction* (Erhard Schmidt procedure) a system of n *orthogonal vectors* π_μ in the form

$$\pi_\mu = P_\mu + \sum_{\nu=1}^{\mu-1} a'_{\mu\nu} P_\nu \quad (\mu = 1, \dots, n) \tag{6.1}$$

or, introducing the triangular matrix $A_0(a'_{\mu\nu})$, $a'_{\mu\nu} = 0$ $(\nu \geqslant \mu)$,

$$(\pi_1, \dots, \pi_n)' = (I + A_0)(P_1, \dots, P_n)', \tag{6.2}$$

where the $a'_{\mu\nu}$ are uniquely defined, as soon as the order of the P_ν in the base of M is fixed.

The π_ν are a base of R^n but not necessarily one of M since the $a'_{\mu\nu}$ need not be all integers. However, a base of M which is a kind of "approximation" to the base (π_ν), can be constructed in the following way.

37. Taking

$$(I + A_0)^{-1} = I + A, \quad A = (a_{\mu\nu}), \tag{6.3}$$

where A is again a left triangular matrix with zeros along the main diagonal, we use the decomposition of the Lemma 1,

$$I+A = (I+G)(I+E). \tag{6.4}$$

The inverse of $I+G$,

$$I+G_0 = I-G+G^2-G^3+\ldots+(-1)^{n-1}G^{n-1}$$

is again a left triangular matrix $G_0 = (g_{\mu\nu}^{(0)})$ *with integer elements* and zeros along the main diagonal. From (6.2), (6.3) and (6.4) we have now

$$(I+G_0)(P_1, \ldots, P_n)' = (I+E)(\pi_1, \ldots, \pi_n)'$$

or, putting $(P_\mu^*)' = (I+G_0)(P_\nu)'$:

$$P_\mu^* = P_\mu + \sum_{\nu=1}^{\mu-1} g_{\mu\nu}^{(0)} P_\nu \qquad (\mu = 1, \ldots, n): \tag{6.5}$$

$$P_\mu^* = \pi_\mu + \sum_{\nu=1}^{\mu-1} \varepsilon_{\mu\nu}\pi_\nu, \qquad |\varepsilon_{\mu\nu}| \leqslant \tfrac{1}{2} \qquad (\nu < \mu = 1, \ldots, n). \tag{6.6}$$

By (6.5) the (P_μ^*) form a base of M. We call this base the *reduced base of* M and the base (π_μ) the *reduced base of* E^n, *corresponding to the base* (P_μ).

For every k, $k = 1, \ldots, n-1$, denote by

$$L_k = (0, P_1, \ldots, P_k) = (0, P_1^*, \ldots, P_k^*) = (0, \pi_1, \ldots, \pi_k) \tag{6.7}$$

the linear k-dimensional manifold spanned by the $k+1$ points $0, P_1, \ldots, P_k$. As follows from (6.5) and (6.6), L_k is also spanned by the points $0, P_1^*, \ldots, P_k^*$ and by the points $0, \pi_1, \ldots, \pi_k$.

As π_{k+1} is orthogonal on L_k we have from (6.5) and (6.6):

$$|P_{k+1}^*, L_k| = |P_{k+1}, L_k| = |\pi_{k+1}, L_k| = |\pi_{k+1}| \qquad (k = 1, \ldots, n-1). \tag{6.8}$$

38. We construct now a base of M in the following way. As P_1 we take a point of M for which $|P_1| = D_M^{(0)}$, that is which has of all points of M, distinct from the origin, the minimal distance from the origin. L_1 we define as the line through 0 and P_1. As P_2 we take a point of $M-L_1$ for which $|P_2, L_1| = |M-L_1, L_1|$. That such a point exists in M, follows from the Lemma 6.

As L_2 we define then the two-dimensional linear manifold through $0, P_1, P_2$. Then P_3 is defined as a point of $M-L_2$ for which $|P_3, L_2|$ is $|M-L_2, L_2|$. And in the same way we go on, so that L_k is defined as the k-dimensional linear manifold through $0, P_1, \ldots, P_k$ and the P_{k+1} as a point of $M-L_k$ for which $|P_{k+1}, L_k| = |M-L_k, L_k|$. The last point P_n is then a point of $M-L_{n-1}$ which has the minimal distance from L_{n-1}.

39. We show first that the points P_μ form a *base* of M and more generally that the points $P_1, \ldots, P_k$ form a base for the k-dimensional lattice ML_k. This is clear for $k = 1$, since otherwise there would be on the line through 0 and P_1 a point of M which is tP_1 with a non-integer t. But then the point $tP_1 - [t]P_1 = (t-[t])P_1$ would belong to M and would have the distance $(t-[t])|P_1| < |P_1|$ from the origin.

Assume now that our assertion has been already proved for a $k < n$. Then every point P of ML_{k+1} can be written in the form

$$P = tP_{k+1} + t_1P_1 + \ldots + t_kP_k = tP_{k+1} + Q, \tag{6.9}$$

where $Q \in L_k$. If t were not integer, the point $P' = P - [t]P_{k+1}$ would also belong to ML_{k+1}. On the other hand, we would have

$$P' = (t-[t])P_{k+1} + Q, \qquad |P', L_k| = (t-[t])|P_{k+1}, L_k| < |P_{k+1}, L_k|,$$

which contradicts the definition of P_{k+1}.

We see that t is an integer so that

$$P - tP_1 = t_1P_1 + \ldots + t_kP_k$$

is a point of M and therefore also of ML_k. But then the $t_\varkappa$ must be integers and in (6.9) all coefficients are integers so that $(P_1, \ldots, P_{k+1})$ is a base of ML_{k+1}.

Our assertion is now proved by induction.

40. Form now by the procedure of section 37 the reduced base (P_μ^*) of M and the reduced base (π_μ) of E^n. By definition of the P_μ we have $|P_{k+1}, L_k| \leqslant D_M^{(k)}$ and from (6.8)

$$|\pi_k| \leqslant D_M^{(k-1)} \qquad (k = 1, \ldots, n),$$

since for $k = 1$ this follows directly from the definition of $P_1 = P_1^* = \pi_1$.

On the other hand, using D_M as defined in (2.3), we obtain from (5.6), $D_M^{(k)} \leqslant \sqrt{\frac{4}{3}}^{n-k-1} D_M$ and therefore

$$|\pi_\mu| \leqslant \sqrt{\tfrac{4}{3}}^{n-\mu} D_M \qquad (\mu = 1, \ldots, n). \tag{6.10}$$

41. From (6.6) and (6.10) we have further, putting

$$w = \sqrt{\tfrac{4}{3}}, \qquad \delta = \sqrt{\tfrac{3}{4}}, \tag{6.11}$$

for $\mu = 1, \ldots, n$:

$$\frac{|P_\mu^*|}{D_M} \leqslant w^{n-\mu} + \frac{1}{2}\sum_{\nu=1}^{\mu-1} w^{n-\nu} = w^n\left(\delta^\mu + \frac{1}{2}\sum_{\nu=1}^{\mu-1}\delta^\nu\right) = w^n\left(\delta^\mu + \frac{1}{2}\cdot\frac{\delta-\delta^\mu}{1-\delta}\right)$$

and, dividing on both sides by w^{n-1} and using $w\delta = 1$:

$$2\frac{1-\delta}{w^{n-1}}\cdot\frac{|P_\mu^*|}{D_M} \leqslant 1+\delta^{\mu-1}-2\delta^\mu \quad (\mu \geqslant 1). \tag{6.12}$$

We compute now

$$u = \sum_{\mu=1}^{n}[1+(1-2\delta)\,\delta^{\mu-1}]^2 \tag{6.13}$$

and obtain

$$U = \sum_{\mu=1}^{n}[1+2(1-2\delta)\,\delta^{\mu-1}+(1-2\delta)^2\delta^{2\mu-2}]$$

$$= n+2(1-2\delta)\frac{1-\delta^n}{1-\delta}+(1-2\delta)^2\frac{1-\delta^{2n}}{1-\delta^2}.$$

The last right-hand term becomes, using $1-\delta^2 = \frac{1}{4}$, $(1-2\delta)^2 = 4(1-\delta)$ and (6.11), $16(1-\delta)(1-\delta^{2n})$. The second right-hand term is

$$2(1-2\delta)(1+\delta)\frac{1-\delta^n}{1-\delta^2} = -4(1+2\delta)(1-\delta^n)$$

and we obtain

$$U = n-4(1+2\delta)(1-\delta^n)+16(1-\delta)(1-\delta^{2n}),$$

$$U = n+f(\delta^n), \quad f(x) \equiv -4(1+2\delta)(1-x)+16(1-\delta)(1-x^2). \tag{6.14}$$

42. The polynomial $f(x)$ has its derivative

$$f'(x) = -32(1-\delta)x+4(1+2\delta) = -32(1-\delta)\left(x-\frac{1+2\delta}{8(1-\delta)}\right),$$

where the root of $f'(x)$,

$$\frac{1+2\delta}{8(1-\delta)} = \frac{1+3\delta+2\delta^2}{8(1-\delta^2)} = \frac{5+3\sqrt{3}}{4} > 1,$$

so that $f(x)$ monotonically increases for $x \leqslant 1$ and we have $f(x) < f(1) = 0$ for $x < 1$. It follows now from (6.14) that

$$U < n \quad (n = 1, 2, \ldots). \tag{6.15}$$

We have finally from (6.12), using $4(1-\delta)^2 = 7-8\delta = 1/(7+8\delta)$,

$$\frac{1}{D_M^2}\sum_{\mu=1}^{n}|P_\mu^*|^2 \leqslant \frac{w^{2n-2}}{4(1-\delta)^2}\,U \leqslant (7+4\sqrt{3})\,w^{2n-2}n$$

$$= \frac{21+12\sqrt{3}}{4}\left(\frac{4}{3}\right)^n n < 11\left(\frac{4}{3}\right)^n n,$$

$$\sum_{\mu=1}^{n}|P_\mu^*|^2 \leqslant (7+4\sqrt{3})\,n\left(\tfrac{4}{3}\right)^{n-1}D_M^2 < 11\,n\left(\tfrac{4}{3}\right)^n D_M^2. \tag{6.16}$$

We obtain now

THEOREM 3. *If D_M is the free radius of the n-dimensional lattice M there exists a base (P_μ^*) of M for which* (6.16) *holds.*

§ 7. A quantitative theorem on simultaneous Diophantine approximations.

43. THEOREM 4. *Assume that the n real numbers $\alpha_1, \ldots, \alpha_n$ and two positive numbers ε, η have the property that we have always*

$$|m+m_1\alpha_1+\ldots+m_n\alpha_n| \geqslant \eta \tag{7.1}$$

whenever the integers $m, m_1, \ldots, m_n$ satisfy the conditions

$$-|m| < \sqrt{\sum_{\nu=1}^{n} m_\nu^2} < \frac{Q}{\varepsilon}, \quad Q = 2\sqrt{n}\left(\frac{4}{3}\right)^{(n-1)/2}; \tag{7.2}$$

further assume that

$$\varepsilon < \sqrt{\frac{3}{4n}}, \quad \eta < Q\sqrt{\frac{3}{4n}}. \tag{7.3}$$

Then for any set of n real numbers $x_1, \ldots, x_n$ there exists a positive integer g satisfying the inequality

$$g < \frac{3}{\eta\varepsilon^n} \tag{7.4}$$

and n integers $g_1, \ldots, g_n$, so that

$$\sum_{\nu=1}^{n}(x_\nu - g\alpha_\nu - g_\nu)^2 < \tfrac{16}{3}\,n\varepsilon^2. \tag{7.5}$$

44. LEMMA 7. *Assume ε, η and γ as positive numbers and n real numbers $\alpha_1, \ldots, \alpha_n$ satisfying the condition that always*

$$\Big|m+\sum_{\nu=1}^{n} m_\nu\alpha_\nu\Big| \geqslant \eta \tag{7.6}$$

whenever the integers $m, m_1, \ldots, m_n$ *are such that*

$$-m^2 < \sum_{\nu=1}^{n} m_\nu^2 < \frac{1}{\gamma^2\varepsilon^2}. \tag{7.7}$$

Assume further that for a positive integer q *and* n *integers* $q_1, \ldots, q_n$ *we have*

$$\sum_{\nu=1}^{n}\left(\alpha_\nu - \frac{q_\nu}{q}\right)^2 < \gamma^2\eta^2\varepsilon^2. \tag{7.8}$$

Denote by M *the additive vector modulus in the real* n*-dimensional space* R*, formed by the set of points*

$$\left(g\frac{q_1}{q}+g_1,\ g\frac{q_2}{q}+g_2, \ldots, g\frac{q_n}{q}+g_n\right), \tag{7.9}$$

if $g, g_1, \ldots, g_n$ *run through all integers. Then the free radius of* M *is* $\leqslant \gamma\varepsilon$:

$$D_M \leqslant \gamma\varepsilon. \tag{7.10}$$

45. Proof. Since all points of M have rational coordinates the directional cosines of the normals of all $(n-1)$-dimensional hyperplanes, rational in M, have rational quotients. As the distance of a point from a hyperplane L through the origin can be obtained projecting the corresponding vector upon the normal to L, (7.10) will be proved if we show that to any set of n integers $k_1, \ldots, k_n$ with

$$k \equiv \sqrt{k_1^2+\ldots+k_n^2} > 0 \tag{7.11}$$

there exists a point (7.9) of M such that we have

$$0 < W \equiv \sum_{\nu=1}^{n}\left(g\frac{q_\nu}{q}+g_\nu\right)\frac{k_\nu}{k} \leqslant \gamma\varepsilon. \tag{7.12}$$

46. Put

$$k_0 = \sum_{\nu=1}^{n} q_\nu k_\nu,$$

and denote by h the greatest common divisor of $k_0, qk_1, \ldots, qk_n$,

$$h = (k_0, qk_1, \ldots, qk_n).$$

Then we have identically

$$\sum_{\nu=1}^{n} q\frac{k_\nu}{h}\alpha_\nu - \frac{k_0}{h} = \sum_{\nu=1}^{n} q\frac{k_\nu}{h}\left(\alpha_\nu - \frac{q_\nu}{q}\right)$$

and therefore, by the Cauchy-Schwarz inequality, in virtue of (7.8) and (7.11)

$$\left|\sum_{\nu=1}^{n}\frac{qk_\nu}{h}\alpha_\nu - \frac{k_0}{h}\right|^2 \leqslant \sum_{\nu=1}^{n}\left(\alpha_\nu-\frac{q_\nu}{q}\right)^2\sum_{\nu=1}^{n}\left(\frac{qk_\nu}{h}\right)^2 \leqslant \gamma^2\varepsilon^2\eta^2\frac{q^2k^2}{h^2}. \tag{7.13}$$

If now $\frac{qk}{h}$ were $< \frac{1}{\gamma\varepsilon}$, it would follow from (7.13) that

$$\left|\sum_{\nu=1}^{n}\frac{qk_\nu}{h}\alpha_\nu - \frac{k_0}{h}\right| < \eta,$$

in contradiction to (7.6). We see therefore that

$$\frac{qk}{h} \geqslant \frac{1}{\gamma\varepsilon}. \tag{7.14}$$

47. Choose now the integers $g, g_1, \ldots, g_n$ such that

$$gk_0 + \sum_{\nu=1}^{n} g_\nu qk_\nu = h.$$

For these integers we have from the definition of W in (7.12)

$$W = \frac{1}{k}\left(g\frac{k_0}{q} + \sum_{\nu=1}^{n} g_\nu k_\nu\right) = \frac{h}{kq},$$

and (7.12) follows now from (7.14). Our Lemma is proved.

48. Lemma 8. *In the hypotheses of the Lemma 7 to any point* $P = (x_1, \ldots, x_n)$ *of the real* n*-dimensional space there exists a point* Γ *of the additive modulus* M *given by* (7.9) *such that*

$$|P, \Gamma| < Q_0\gamma\varepsilon, \quad Q_0 = 2n\left(\tfrac{4}{3}\right)^{n/2}. \tag{7.15}$$

49. Proof. Let $(P_1^*, \ldots, P_n^*)$ be the base of the modulus M, the existence of which is asserted in the Theorem 3, and for which we have the relation (6.16).

Since in the Lemma 7 for our modulus M we have (7.10) the relation (6.16) becomes, using (6.15)

$$\sum_{\nu=1}^{n} |P_\nu^*|^2 < 11n\left(\tfrac{4}{3}\right)^n\gamma^2\varepsilon^2. \tag{7.16}$$

Express the point P linearly in terms of the points P_ν^*:

$$P = \sum_{\nu=1}^{n} \gamma_\nu P_\nu^*.$$

Write each γ_ν in the form $\beta_\nu+g'_\nu$, where g'_ν is an integer and β_ν satisfies the relation $|\beta_\nu|\leqslant\frac{1}{2}$. Then our representation of P becomes

$$P=\sum_{\nu=1}^{n}g'_\nu P^*_\nu+\sum_{\nu=1}^{n}\beta_\nu P^*_\nu.$$

50. The first sum on the right is a point Γ of M. Consider the point K, given by the second sum on the right, $K=\sum_{\nu=1}^{n}\beta_\nu P^*_\nu$, and denote its $\varkappa$th coordinate by $k_\varkappa$ and generally, the $\varkappa$th coordinate of P^*_ν by $p_{\nu\varkappa}$. Then we have by the Cauchy-Schwarz inequality

$$k_\varkappa^2=\Big(\sum_{\nu=1}^{n}\beta_\nu p_{\nu\varkappa}\Big)^2\leqslant\Big(\sum_{\nu=1}^{n}\beta_\nu^2\Big)\Big(\sum_{\nu=1}^{n}p_{\nu\varkappa}^2\Big),$$

where the first sum on the right is $\leqslant n/4$. Summing this relation over $\varkappa$ from 1 to n, we have

$$|K|^2\leqslant\frac{n}{4}\sum_{\nu=1}^{n}\sum_{\varkappa=1}^{n}p_{\nu\varkappa}^2=\frac{n}{4}\sum|P^*_\nu|^2,$$

and therefore, using (7.16)

$$|K|^2<3n^2\left(\tfrac{4}{3}\right)^n\gamma^2\varepsilon^2$$

and (7.15) follows immediately. The Lemma 8 is proved.

51. Denote now the points with the coordinates q_ν/q and α_ν resp. by P_q, π:

$$P_q=\left(\frac{q_1}{q},\dots,\frac{q_n}{q}\right),\qquad \pi=(\alpha_1,\dots,\alpha_n). \tag{7.17}$$

Then our point Γ in (7.15) and (7.9) can be written as $\Gamma=gP_q+\Gamma_1$, where g can be reduced mod q and therefore can be assumed positive and $\leqslant q$, while all coordinates of Γ_1 are integers. By the triangle inequality we have

$$|P,g\pi+\Gamma_1|\leqslant|P,gP_q+\Gamma_1|+|gP_q,g\pi|\leqslant|P,gP_q+\Gamma_1|+q|\pi,P_q|,$$

and therefore, by (7.15):

LEMMA 9. *In the hypotheses of the Lemma* 8 *and in the above notations we have*

$$|P,g\pi+\Gamma_1|\leqslant Q_0\gamma\varepsilon+q|\pi,P_q|, \tag{7.18}$$

for a conveniently chosen positive integer $g\leqslant q$ *and a point* Γ_1 *with integer coordinates.*

In order to make the expression on the left in (7.18) of the order of ε, we must now choose P_q in such a way as to make $q|\pi,P_q|$ of this order.

52. LEMMA 10. *Take positive* ε_1,η_1 *both* <1. *To any system of* n *real numbers* α_ν $(\nu=1,\dots,n)$ *there exists a positive integer* q *and integers* q_ν $(\nu=1,\dots,n)$ *such that, in notation* (7.17),

$$|q\pi,qP_q|<\varepsilon_1,\qquad |\pi,P_q|<\varepsilon_1\eta_1, \tag{7.19}$$

$$0<q<\frac{2}{\eta_1}\left(\frac{\sqrt{n}}{\varepsilon_1}\right)^n. \tag{7.20}$$

Proof. By a Theorem of Dirichlet (see [2], p. 68, Satz 7) there exists a positive $p<\varepsilon_1^{-n}\sqrt{n}^n$ and integers p_ν $(\nu=1,\dots,n)$ for which $|p\alpha_\nu-p_\nu|<\varepsilon_1/\sqrt{n}$ $(\nu=1,\dots,n)$ and therefore, if we denote by P_p the point $(p_1/p,\dots,p_n/p)$,

$$|p\pi,pP_p|<\varepsilon_1,\qquad p\leqslant\sqrt{n}^n/\varepsilon_1^n. \tag{7.21}$$

If such a p is $\geqslant1/\eta_1$, $1/p\leqslant\eta_1$, we obtain from (7.21), putting $q=p$, $q_\nu=p_\nu$ $(\nu=1,\dots,n)$ and dividing the first relation (7.21) by q, the relations (7.19), (7.20).

53. If however we have for each set of solutions p, p_ν of the inequalities (7.21) the relation $p<1/\eta_1$, the inequalities

$$|p\pi,pP_p|<\varepsilon_1,\qquad p\leqslant1/\eta_1 \tag{7.22}$$

have a solution with a positive integer p and integers p_ν.

Assume this solution chosen in such a way that p has its greatest possible value. Putting $\beta_\nu=(p+1)\alpha_\nu$ $(\nu=1,\dots,n)$, apply Dirichlet's Theorem to the β_ν. We obtain a positive integer p' and integers p'_ν so that

$$|p'(p+1)\pi,p'P_{p'}|<\varepsilon_1,\qquad p'\leqslant\sqrt{n}^n/\varepsilon_1^n. \tag{7.23}$$

Here we have certainly $p'(p+1)\geqslant1/\eta_1$ since p is the Maximum for a solution of (7.22). Putting now $q=p'(p+1)$, $q_\nu=p'_\nu$ $(\nu=1,\dots,n)$, we have from (7.23), dividing by q, the relations (7.19), while

$$0<q=p'(p+1)<\frac{1+1/\eta_1}{\varepsilon_1^n}\sqrt{n}^n$$

so that (7.20) is satisfied and the Lemma 10 is proved.

54. Proof of the Theorem 4. We put in the Lemmata 9 and 10:

$$\gamma=\frac{1}{Q},\qquad Q=\sqrt{\frac{3}{4n}}Q_0=\sqrt{3n}\left(\frac{4}{3}\right)^{\frac{n}{2}},\qquad \varepsilon_1=\varepsilon\sqrt{\frac{4n}{3}},\qquad \eta_1=\frac{\eta}{Q}\sqrt{\frac{4n}{3}}.$$

From (7.3) we have then $\varepsilon_1 < 1$, $\eta_1 = \eta\sqrt{\frac{4n}{3}}/Q < 1$ and by (7.19) the condition (7.8) is satisfied, so that the Lemmata 7, 8 and 9 can be applied, while

$$q < \frac{2Q}{\eta\varepsilon^n\sqrt{n}}\sqrt{\frac{3}{4}}^{n+1} = \frac{2}{\eta\varepsilon^n}\left(\sqrt{3n}\left(\frac{4}{3}\right)^{\frac{n}{2}}\frac{1}{\sqrt{n}}\right)\sqrt{\frac{3}{4}}^{n+1} = \frac{2}{\eta\varepsilon^n}\sqrt{3}\sqrt{\frac{3}{4}} = \frac{3}{\eta\varepsilon^n},$$

$$q < \frac{3}{\eta\varepsilon^n}, \quad 0 < g \leqslant q < \frac{3}{\eta\varepsilon^n}. \tag{7.24}$$

We have therefore in the formula (7.18) of the Lemma 9, by the first inequality in (7.19),

$$\frac{1}{\varepsilon}|P, g\pi + \Gamma_1| \leqslant \frac{Q_0}{Q} + \sqrt{\frac{4n}{3}} = \frac{4}{\sqrt{3}}\sqrt{n},$$

and (7.5) follows. The Theorem 4 is proved.

References

[1] J. W. S. Cassels, *An introduction to the Geometry of Numbers*, Berlin 1959.

[2] J. F. Koksma, *Diophantische Approximationen*, Ergebnisse der Mathematik, Berlin 1936.

[3] E. Landau, *Ueber Diophantische Approximationen*, Scripta Universitatis atque Bibliothecae Hierosolymitanarum, 1923.

[4] L. H. Thomas, *An extended form of Kronecker's Theorem with an application which shows that Burger's Theorem on adiabatic invariants is statistically true for an assembly*, Proc. Cambr. Phil. Soc. 22 (1925), pp. 886-903.

Reçu par la Rédaction le 5. 2. 1964

ACTA ARITHMETICA
IX (1964)

Remarks on two theorems of Siegel

by

S. Chowla (University Park, Pa.)

Dedicated to L. J. Mordell

1. Siegel found that for $n \geqslant 3$ the number of solutions of

$$ax^n - by^n = c$$

(i) does not exceed 1 if $|ab|$ exceeds a certain limit depending on c and n alone [1],

(ii) does not exceed 75 if $c = 1$ [2].

Assuming a certain conjecture (Conjecture C below) we shall reduce 75 to 2 in (ii), provided $n \geqslant 1$ (in fact, even when $c \neq 1$). As for (i) we obtain the result (i) without the restriction that "$|ab|$ should exceed a certain limit depending on c and n" but at the cost of replacing 1 by 2, and again demand $n \geqslant 7$.

The hypothesis is as follows:

Conjecture C. *The equation* (*where n is a fixed position integer* $\geqslant 3$)

$$\sum_{m=1}^{n-1} \pm X_m^n = 0$$

is impossible in positive integers X_m, unless it is trivially possible (*i.e. by "cancellation"*).

Both our results concerning (i) and (ii) are covered by the

Theorem. *Suppose Conjecture* C *is true. Then*

$$ax^n - by^n = C \tag{1}$$

has at most 2 solutions for $n \geqslant 7$.

(1) Abh. d. Preuss. Akad. d. Wiss. 1929, 1-70, p. 70.

(2) According to Erdös, Proc. Boulder No. Theory Conference, 1958, p. 238, Siegel's proof was never published.