ACTA ARITHMETICA
XLI (1982)

On the error term in multidimensional diophantine approximation

by

A. M. Ostrowski (Basel)

§ 1. Introduction

1. Consider the euclidean space $\boldsymbol{R}^m$ and m real irrational numbers a_μ $(\mu = 1, \ldots, m)$ such that $a_1, \ldots, a_m$, 1 are linearly independent with respect to $\boldsymbol{Z}$, that is for integers $g_1, \ldots, g_m, g_0$ the relation

$$g_1 a_1 + \ldots + g_m a_m + g_0 = 0 \tag{1.1}$$

only holds if all g_μ vanish.

Denote generally for any real a by the symbol $\|a\|$ the distance of a from the nearest integer. Further, consider a variable m-vector, z, with integer components, $z = (z_1, \ldots, z_m)$, which is assumed never to vanish and put

$$|z|_\infty := \max_\mu |z_\mu|.$$

Then we can use as the "measure of independence" of the a_μ any continuous and strictly monotonically decreasing $\psi(\sigma) \downarrow 0$ $(\sigma \geqslant 1)$ such that

$$\psi(\sigma) \leqslant \min \Big\| \sum_{\mu=1}^{m} z_\mu a_\mu \Big\| \quad (1 \leqslant |z|_\infty \leqslant \sigma), \ (\sigma \geqslant 1). \tag{1.2}$$

2. By $P(ka_\mu)$ we denote generally for an integer k the point

$$P(ka_\mu) := (ka_1, \ldots, ka_m) \bmod 1.$$

We define a "proper interval" J in $\boldsymbol{R}^m$ as a cartesian product of m linear segments, open in the direction of increasing coordinates,

$$a_\mu \leqslant x_\mu < b_\mu \quad (\mu = 1, \ldots, m). \tag{1.3a}$$

Then the *volume of* J, $|J|$, is the product of the lengths of these segments. If an interval is considered mod 1, we take generally two points of $\boldsymbol{R}^m$ as *identical* if their corresponding coordinates differ by integers. Then usually the points of J have to be taken with a convenient multiplicity. If all points of J, taken mod 1, are simple, J can be considered as a part of the unit cube, $0 \leqslant x_\mu < 1$ $(\mu = 1, \ldots, m)$ and is called *simple*. Then $|J|$ is defined as cartesian product of m segments mod 1, $0 \leqslant a_\mu \leqslant x_\mu < b_\mu \leqslant 1$ or $(0 \leqslant x < b_\mu) \cup (a_\mu \leqslant x < 1)$. Denoting by δ_μ in the first case 0 and in the second case 1, the length of the μ-th edge of J is

$$b_\mu - a_\mu + \delta_\mu, \tag{1.3}$$

while the volume of J becomes

$$|J| = \prod_{\mu=1}^{m} (b_\mu - a_\mu + \delta_\mu).$$

We put generally

$$\tau(J) := \min_{1 \leqslant \mu \leqslant m} (b_\mu - a_\mu + \delta_\mu).$$

3. The essential point of Kronecker's theory of irrationals is the result that for any J and for at least one integer q:

$$P(q\alpha) \in J. \tag{1.4}$$

This result was sharpened by Weyl [7] who proved that for a given $\varepsilon > 0$ for any J with $\tau(J) \geqslant \varepsilon$ the integer $q = q_J$ in (1.4) can be choosen $\leqslant y(\varepsilon)$, with an $y(\varepsilon)$ independent of the special J with $\tau(J) \geqslant \varepsilon$:

$$P(q_J \alpha) \in J, \quad q_J \leqslant y(\varepsilon). \tag{1.5}$$

4. We will denote by $N(x, J)$ the number of all $P(\nu\alpha)$ with $1 \leqslant \nu \leqslant x$ lying mod 1 in J,

$$N(x, J) := N\big(\nu\colon P(\nu\alpha) \in J,\ 1 \leqslant \nu \leqslant x\big). \tag{1.6}$$

Then it follows immediately from Weyl's theorem (1.5), that for any fixed simple interval J mod 1 the relation holds

$$N(x, J) = x|J| + o(x) \quad (x \to \infty). \tag{1.7}$$

5. We introduce, for $x \geqslant 1$, $A(x)$ by

$$A(x) := \sup_{\substack{1 \leqslant y \leqslant x \\ J}} \big|N(y, J) - y|J|\big|, \tag{1.8}$$

where J runs through all simple intervals mod 1 in the unit cube in $\boldsymbol{R}^m$, and denote by $\sigma = \varphi(\tau)$ the inverse function of $\varphi = \psi(\sigma)$ in (1.2). We put further

$$\gamma_0 := 2^{m+1}/\big((m+1)!\big)^2, \quad \varepsilon(x) := \frac{1}{\gamma_0 \varphi\left(\dfrac{1}{\gamma_0 x}\right)}. \tag{1.9}$$

6. It is easily seen that $A(x) = o(x)$ $(x \to \infty)$. Our aim is to improve this estimate using the function $\varepsilon(y)$. Our essential result is the

THEOREM 1. *$A(x)$, as defined by* (1.8), *satisfies for $m > 1$ an inequality*

$$A(x) \leqslant \alpha A(y) + \beta x \varepsilon(y), \quad 1 \leqslant y \leqslant x,\ \alpha > 1,\ \beta > 0 \tag{1.10}$$

with constants $\alpha \geqslant 1$ and $\beta > 0$ depending only on m.

This theorem is proved in the §§ 2–5 while the inequality (1.10) is discussed in § 6 under different assumptions about $\varepsilon(y)$. In the case $m = 1$ an inequality (1.10) holds even with $\alpha = 1$. This case however has been already discussed in [6].

7. Before attacking the problem of $o(x)$ in (1.7) we have to obtain a relationship between ε and $y(\varepsilon)$ in (1.5) and $\psi(\sigma)$ in (1.2). This relationship follows in a particularly simple and fundamental way from a special case of an important theorem due to Khintchine [4](1). We obtain from this theorem for the constant γ_0 from (1.9) the relation (Sec. 13, (2.10)):

$$y = \frac{1}{\gamma_0 \psi\left(\dfrac{1}{\gamma_0 \varepsilon}\right)}. \tag{1.11}$$

8. It is well known since Dirichlet that $\psi(y) = O(1/y^m)$. If there exists a ϱ, $0 < \varrho < 1$, such that

$$\psi(y) \geqslant \frac{c}{y^{1/\varrho}} \quad (y \geqslant y_0) \tag{1.12}$$

then we show that (Sec. 38)

$$A(y) = O(y^{1-\varrho}). \tag{1.13}$$

This is in particular always the case if the $a_1, \ldots, a_m$ are algebraic. However the estimate of $A(x)$ with the exponent $1-\varrho$ could only be obtained using Khintchine's theorem, published 1949. In [5], 1930, we used a weaker result than (1.13), due to Landau.

In the case $m = 1$ the fact that (1.13) follows from (1.12) has been already proved by Hecke 1922, however with a method which apparently cannot be generalized to $m > 1$. 1930 we announced the results corresponding to (1.12) and (1.13), however in the form $A(x) = O(x^a)$, $0 < a < 1$ (see [6]).

9. More generally, assume $k(y)$ as a positive constant or a continuous

(1) My attention was drawn kindly to this theorem by J. W. S. Cassels ([1], pp. 97–99).

positive function strictly increasing to ∞, such that

$$\psi(y) \geqslant \frac{c}{y^{1/\varrho} k\big((Dy)^{1/\varrho}\big)} \quad (y \geqslant y_0), \quad c \wedge D > 0, \text{ const}.$$

Then for a conveniently defined $l(x)$:

$$A(x) = O\big(x^{1-\varrho} l(x)^\varrho\big). \tag{1.14}$$

This follows from Lemma 5, Sec. 43.

If in particular $xk'(x) = o\left(\frac{k(x)}{\log k(x)}\right)$, then we obtain even

$$A(x) = O\big(x^{1-\varrho} k(x)^\varrho\big) \quad \text{(Theorem 2, Sec. 50)}. \tag{1.15}$$

10. Our proof of (1.10) was given for $m = 1$, 1930, in Ostrowski [5]. Its essential point was our Lemma 3 of § 3 which we developed l.c. for $m = 1$, but indicated that the whole discussion can be generalized to $m > 1$. In the mean time, 1950, S. Hartman [2] has developed in a very careful way the corresponding generalization of the Lemma 3 to $m > 1$, discussing also the limiting cases. As we need only a part of this argument, we give in § 3 our original proof, which is a straightforward generalization of that given for $m = 1$ in [5].

§ 2. Use of Khintchine's lemma

11. We formulate first one part of Khintchine's theorem in the form in which it was given by Cassels [1], p. 99, but changing conveniently the notation. We will denote generally for an n-vector, $\xi = (x_1, \ldots, x_n)$, by $|\xi|_\infty$ the norm $|\xi|_\infty := \max |x_\nu|$.

LEMMA 1. *Let m and n be natural integers with $m+n = l$. Consider a real $(n\times m)$-matrix, $A = (a_{\nu\mu})$, $\mu = 1, \ldots, m$; $\nu = 1, \ldots, n$, and the linear forms*

$$M_\mu(\xi) = \sum_{\nu=1}^{n} a_{\nu\mu} x_\nu \quad (\mu = 1, \ldots, m), \tag{2.1}$$

$$N_\nu(\zeta) = \sum_{\mu=1}^{m} a_{\nu\mu} z_\mu \quad (\nu = 1, \ldots, n), \tag{2.2}$$

where the x_ν and the z_μ are respectively the components of the n-vector ξ and the m-vectors ζ. Consider two positive constants ε, y and a real m-vector β with components $b_1, \ldots, b_m$.

Then, in order that there exists an integral vector ξ satisfying the relations

$$\|M_\mu(\xi) - b_\mu\| \leqslant \varepsilon \quad (\mu = 1, \ldots, m), \quad |\xi|_\infty \leqslant y \tag{2.3}$$

it is sufficient that for $\gamma := 2^{l-1}/(l!)^2$ the following relation holds for every integer m-vector ζ:

$$\Big\|\sum_{\mu=1}^{m} b_\mu z_\mu\Big\| \leqslant \gamma \max\big(y \max_\nu \|N_\nu(\zeta)\|,\ \varepsilon|\zeta|_\infty\big). \tag{2.4}$$

As a matter of fact the complete formulation of Khintchine's theorem contains also the necessary condition for (2.3), which we however do not need.

12. For our purpose we must now specialize the assumptions of Khintchine's theorem.

Assume $n = 1$, $l = m+1$ and observe that γ_0: $2^{m+1}/\big((m+1)!\big)^2 = 2\gamma$. The n-vector ξ becomes a scalar which we will denote by q, the elements of the matrix A become $a_{\mu\nu} := a_\mu$, so that $M_\mu(\xi)$ becomes qa_μ and the linear forms $N_\nu(\zeta)$ become $N(\zeta) = \sum_{\mu=1}^{m} a_\mu z_\mu$. The requirements (2.3) of Khintchine's theorem become

$$\|qa_\mu - b_\mu\| \leqslant \varepsilon \quad (\mu = 1, \ldots, m), \quad |q| \leqslant y. \tag{2.5}$$

It follows then from the condition (2.4) of Khintchine's theorem that (2.5) can be certainly realized by a rational integer q if for any m-vector ζ we have

$$\Big\|\sum_{\mu} b_\mu z_\mu\Big\| \leqslant \frac{\gamma_0}{2} \max\big(y\|N(\zeta)\|,\ \varepsilon|\zeta|_\infty\big), \quad \gamma_0 = 2^{m+1}/\big((m+1)!\big)^2. \tag{2.6}$$

13. The condition (2.6) is sharpened replacing the left side expression by $1/2$. As it is certainly satisfied if $\gamma_0 \varepsilon |\zeta|_\infty > 1$, it suffices to consider ζ with

$$\varepsilon|\zeta|_\infty \leqslant 1/\gamma_0. \tag{2.7}$$

Thence our condition becomes:

$$1 \leqslant \gamma_0 y \|N(\zeta)\| \text{ follows always from } |\zeta|_\infty \leqslant 1/(\gamma_0\varepsilon). \tag{2.8}$$

If we now assume that (1.1) holds and use the definition (1.2) of $\psi(t)$, (2.8) is satisfied if

$$\psi\left(\frac{1}{\gamma_0\varepsilon}\right) \geqslant \frac{1}{\gamma_0 y},$$

and we can take y in (2.5) as

$$y = y(\varepsilon) := \frac{1}{\gamma_0\big(\psi 1/(\gamma_0\varepsilon)\big)}. \tag{2.9}$$

Using the inverse function to ψ, φ, it follows as in (1.9)

$$\frac{1}{\gamma_0\varepsilon} = \varphi\left(\frac{1}{\gamma_0 y}\right), \quad \varepsilon = \varepsilon(y) := \frac{1}{\gamma_0\varphi(1/(\gamma_0 y))}. \tag{2.10}$$

LEMMA 2. *For any* $y \geqslant 1/(\gamma_0\psi(1))$ *there exists an integer* q *with*

$$|q| \leqslant y, \quad \|q a_\mu - b_\mu\| \leqslant \varepsilon(y) = \frac{1}{\gamma_0\varphi(1/(\gamma_0 y))} \quad (\mu = 1, \ldots, m). \tag{2.11}$$

§ 3. A lemma

14. In what follows we will consider a sequence of points

$$P(\nu a_\mu) \quad (\nu = 1, \ldots, n) \tag{3.1}$$

for a fixed integer $n \geqslant 1$. We define the symbol $[a-0]$ as $[a]$ if a is not integer and $a-1$ if a is integer.

15. LEMMA 3. *Consider a simple interval* J_0 mod 1 *contained in the unit cube, as characterized in Section 2, and assume the* a_μ *as in Section* 1. *Then there exist two intervals* J_0' *and* J_0'' mod 1 *in* $\boldsymbol{R}^m$, *obtained from* J_0 *by parallel translations, such that*

$$N(n, J_0') \leqslant n|J_0|, \quad N(n, J_0'') \geqslant n|J_0|. \tag{3.2}$$

16. Proof. Without loss of generality we can assume, that all a_μ lie in the open interval $(0, 1)$ and further, that J is not identical with the unit cube, but "begins" at the origin, that is that all a_μ in (1.3a) vanish. Denote the length of the μ-edge of J_0 by d_μ, where

$$0 < d_\mu \leqslant 1 \quad (\mu = 1, \ldots, m), \quad |J_0| = d_1 \ldots d_m < 1.$$

17. If we shift J_0 in the directions of the x_μ by the integers q_μ, we obtain a proper interval which will be denoted by $J_{q_1,\ldots,q_m}$. Then the original J_0 can be written as $J_{0,\ldots,0}$. Obviously $J_{q_1,\ldots,q_m}$ is the cartesian product of the segments $q_1, \ldots, q_m$

$$\langle q_\mu d_\mu, (q_\mu+1) d_\mu) \quad (\mu = 1, \ldots, m).$$

18. We let now, for positive integers $Q_1, \ldots, Q_m$, each q_μ run through $0, 1, \ldots, Q_\mu - 1$. Then all intervals obtained in this way form together an interval J^* with the edges $Q_\mu d_\mu$ $(\mu = 1, \ldots, m)$ and its volume is

$$Q|J_0| = \prod_{\mu=1}^{m} (Q_\mu d_\mu), \quad Q := Q_1 \ldots Q_m.$$

Put further

$$N_{q_1,\ldots,q_m} := N(\nu : P(\nu a) \in J_{q_1,\ldots,q_m};\ \nu \leqslant n) \tag{3.3}$$

and denote by N the sum of all $N_{q_1,\ldots,q_m}$ $(q_1 = 0, 1, \ldots, Q_1; \ldots; q_m = 0, 1, \ldots, Q_m)$,

$$N := \sum_{q_1,\ldots,q_m} N_{q_1,\ldots,q_m}. \tag{3.4}$$

19. Denote by $f(\nu)$ the number of all points in J^* which, considered mod 1, coincide with the fixed $P(\nu a_\mu)$ from (3.1). These points have the coordinates

$$[g_1 + R(\nu a_1), g_2 + R(\nu a_2), \ldots, g_m + R(\nu a_m)]$$

where

$$0 \leqslant g_1 < [Q_1 d_1 - R(\nu a_1) - 0], \quad \ldots, \quad 0 \leqslant g_m \leqslant [Q_m d_m - R(\nu a_m) - 0].$$

Therefore we have

$$f(\nu) = \prod_{\mu=1}^{m} [Q_\mu d_\mu + 1 - R(\nu a_\mu) - 0]. \tag{3.5}$$

By summation over $\nu = 1, \ldots, n$ we obtain the number of all points in J^* equal mod 1 to the points (3.1), that is N. Dividing by Q we obtain finally

$$\frac{1}{Q}\sum_{\nu=1}^{n}\prod_{\mu=1}^{m} [Q_\mu d_\mu + 1 - R(\nu a_\mu) - 0] = \frac{1}{Q}\sum_{q_1,\ldots,q_m} N_{q_1,\ldots,q_m}. \tag{3.6}$$

20. If we let all Q_μ increase to ∞, the left side expression in (3.6) tends to $n|J_0|$. Therefore the same holds in the right-hand expression and we obtain

$$\frac{1}{Q}\sum_{q_1,\ldots,q_m} N_{q_1,\ldots,q_m} \to n|J_0| \quad (\forall Q_\mu \to \infty). \tag{3.7}$$

21. Assume first that $n|J_0|$ is *not integer* and lies in $\langle k, k+1)$. Then obviously, as soon as the left side expression in (3.7) lies in the open interval $(k, k+1)$, it is impossible that all $N_{q_1,\ldots,q_m}$ in (3.7) are $\geqslant k+1$. Neither can all these $N_{q_1,\ldots,q_m}$ be $\leqslant k$. Therefore there exist at least two different $J_{q_1,\ldots,q_m}$, say J_0', J_0'', so that (3.2) is satisfied.

22. Assume now that $n|J_0|$ is an integer. If there exist two $J_{q_1,\ldots,q_m}$, say J_0' and J_0'', so that

$$N(n, J_0') < n|J_0|, \quad N(n, J_0'') > n|J_0|,$$

(3.2) is again satisfied. Otherwise for all $J_{q_1,\ldots,q_m}$ the corresponding $N_{q_1,\ldots,q_m}$ in (3.7) are equal to $n|J_0|$ and then we can take $J_0' = J_0'' = J_0$, and the relations (3.2) are satisfied with the equality sign.

§ 4. An upper limit for $N(n, J)-n|J|$

23. Consider a simple interval $J \bmod 1$ with the edges $d_1, \ldots, d_m$ and a positive $\varepsilon < 1/4$. Assume first that

$$(4.1)\qquad 1-d_\mu > 2\varepsilon \quad (\mu = 1, \ldots, m).$$

Let J_0 be an interval concentric to J with the edges $d_1+2\varepsilon, \ldots, d_m+2\varepsilon$. (See Fig. 1, for $m = 2$.)

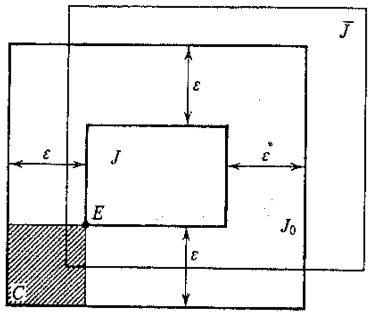

Fig. 1

By (4.1) J_0 is also a simple interval and we have

$$(4.2)\qquad |J_0| = \prod_{\mu=1}^{m} (d_\mu+2\varepsilon), \qquad |J| = \prod_{\mu=1}^{m} d_\mu .$$

24. By the first inequality (3.2) there exists an interval J_0' congruent with J_0 such that

$$(4.3)\qquad N(n, J_0') \leqslant n|J_0| .$$

Consider a cube, C, with the edges parallel to the axes and of the length ε, placed so that it has with J only one point of the boundary in common, the vertex $E = (e_1, \ldots, e_m)$, and lies completely in J_0 (see the hatched square C in Fig. 1).

25. Consider the vertex of J_0 corresponding to E and the corresponding vertex of J_0', $E' = (e_1', \ldots, e_m')$. Then by what has been proved in Section 13 about the relation (1.5) it follows that for a convenient positive integer $q \leqslant y(\varepsilon) = 1/\left[\gamma_0 \psi\left(\frac{1}{\gamma_0\varepsilon}\right)\right]$ and convenient integers $q_1, \ldots, q_m$ the relations hold

$$(4.4)\qquad e_\mu - e_\mu' = R(q\alpha_\mu)+q_\mu+\theta_\mu\varepsilon, \quad 0<|\theta_\mu|<1 \ (\mu = 1, \ldots, m).$$

It follows that if we apply the parallel translation first by the vector $q\alpha$ and then by the integer vector (q_μ) to the interval J_0', this interval goes over into a congruent intervall $\bar{J}$, which has the property that the vertex of $\bar{J}$ corresponding to E' lies in the hatched domain C. Obviously, *J is contained* in $\bar{J}$.

26. Consider the translation from J_0' to $\bar{J}$. To the points $P(\nu\alpha)$, $1 \leqslant \nu \leqslant n$, lying in J_0' correspond the points congruent mod 1 to $P\big((\nu+q)\alpha\big)$ that is to the points $P(\nu\alpha)$, $q+1 \leqslant \nu \leqslant n+q$. Their number is

$$N(n, J_0) = N(n+q, \bar{J}) - N(q, \bar{J}).$$

But the minuend here can be written as

$$N(n+q, \bar{J}) = N(n, \bar{J}) + N\big(R(\nu\alpha), R(\nu\alpha) \in \bar{J}\big) \qquad (n+1 \leqslant \nu \leqslant n+q).$$

Here the last summand can be again written as $N(q, J^*)$ if we denote by J^* the interval obtained from $\bar{J}$ by the parallel translation with the vector $R(n\alpha)$. And obviously $|J^*| = |J_0| = |\bar{J}|$.

27. We obtain therefore from (4.3)

$$(4.5)\qquad N(n, J_0') = N(n, \bar{J}) - N(q, \bar{J}) + N(q, J^*) \leqslant n|J_0|,$$
$$N(n, \bar{J}) \leqslant n|J_0| + N(q, \bar{J}) - N(q, J^*).$$

As $J \in \bar{J}$, it follows further

$$N(n, J) \leqslant N(n, \bar{J}) \leqslant n|J_0| - N(q, \bar{J}) + N(q, J^*).$$

Since $|J| = |J^*|$, we can write this in the form

$$N(n, J) \leqslant n|J_0| - [N(q, \bar{J}) - n|\bar{J}|] + [N(q, J^*) - n|\bar{J}|].$$

The two last bracket terms on the right have moduli $\leqslant A(q)$ and we obtain further

$$N(n, J) - n|J_0| \leqslant 2A(q).$$

On the other hand it follows from (4.2):

$$|J_0| - |J| \leqslant \prod_{\mu=1}^{m} (d_\mu+2\varepsilon) - \prod_{\mu=1}^{m} d_\mu \leqslant (1+2\varepsilon)^m - 1 < 2^{m+1}\varepsilon,$$

as $2\varepsilon < 1$ and the development of $|J_0|-|J|$ in products of the d_μ has *positive* coefficients.

Since $q \leqslant y(\varepsilon)$ and $A(q)$ is not decreasing we can finally write

$$(4.6)\qquad N(n, J) - n|J| \leqslant 2A\big(y(\varepsilon)\big) + 2^{m+1} n\varepsilon .$$

28. (4.6) has been derived assuming the condition (4.1). If this condition is not satisfied, we can by halfing each edge of the unit cube decompose the unit cube into the sum of 2^m cubes of the edge length $1/2$. Correspondingly J is decomposed into at the most 2^m intervals $J^{(\nu)}$ $(\nu = 1, 2, \ldots)$ with the edge length $\leqslant 1/2$.

For each of the intervals $J^{(\nu)}$ the condition (4.1) is satisfied so that we can write

$$N(n, J^{(\nu)}) - n|J^{(\nu)}| \leqslant 2A\big(y(\varepsilon)\big) + 2^{m+1} n\varepsilon.$$

Summing over ν it follows

$$N(n, J) - n|J| \leqslant 2^{m+1} A\big(y(\varepsilon)\big) + 2^{2m+1} n\varepsilon. \tag{4.7}$$

§ 5. A lower limit for $N(n, J) - n|J|$

29. We consider again the simple interval J mod 1 of the Section 23 with the edges $d_1, \ldots, d_m$, but assume first that for a positive $\varepsilon < 1/4$ the relations hold:

$$d_\mu > 2\varepsilon \qquad (\mu = 1, \ldots, m). \tag{5.1}$$

Let now J_0 be an interval concentric to J with the edges $d_1 - 2\varepsilon, \ldots$ $\ldots, d_m - 2\varepsilon$. J_0 is again a simple interval with

$$|J_0| = \prod_{\mu=1}^{m} (d_\mu - 2\varepsilon), \qquad |J| = \prod_{\mu=1}^{m} d_\mu. \tag{5.2}$$

By the second inequality in (3.2) there exists an interval J_0' congruent with J_0 and such that

$$N(n, J_0') \geqslant n|J_0|. \tag{5.3}$$

Consider a cube, C, with the edges parallel to the axes and of the length ε which has with J_0 only an edge $E = (e_1, \ldots, e_m)$ in common (see the hatched square in Fig. 2).

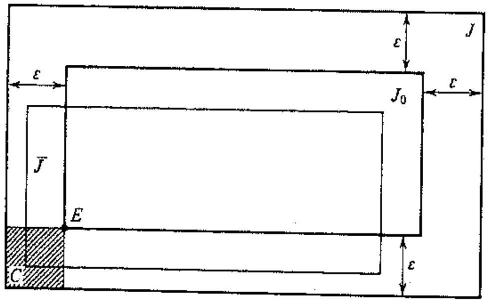

Fig. 2

30. Consider the vertex of J_0', $E' = (e_1', \ldots, e_m')$, which corresponds to E. Then, by what has been proved in Section 13 about the relation (1.3), it follows that for a convenient positive integer $q \leqslant y(\varepsilon)$ and convenient integers $q_1, \ldots, q_m$ the relations hold:

$$e_\mu - e_\mu' = R(q\alpha) + q_\mu + \theta_\mu \varepsilon, \qquad 0 \leqslant \theta_\mu < 1 \ (\mu = 1, \ldots, m).$$

We see that if we apply to the interval J_0' the parallel translations first by the vector $q\alpha$ and then by the integer vector $q^* = (q_\mu)$, this interval goes over into a congruent interval $\bar{J}$ which has the property that the vertex of $\bar{J}$ corresponding to E_0' lies in the hatched domain C. Obviously *J contains $\bar{J}$* (see Fig. 2 for $m = 2$).

31. By parallel translation from J_0' to J, to the points $P(\nu\alpha)$, $1 \leqslant \nu \leqslant n$, lying in J_0' correspond points congruent mod 1 to $P\big((\nu+q)\alpha\big)$ that is to the $P(\nu\alpha)$, $q+1 \leqslant \nu \leqslant n+q$.

Their number is

$$N(n, J_0') = N(n+q, \bar{J}) - N(q, \bar{J}).$$

But the minuend can be written as

$$N(n+q, \bar{J}) = N(n, \bar{J}) + N\big(R(\nu\alpha) \in \bar{J},\ n+1 \leqslant \nu \leqslant n+q\big).$$

Here the last summand can be again written as $N(q, J^*)$ if we denote by J^* the interval obtained from $\bar{J}$ by the parallel translation with the vector $R(n\alpha)$. And obviously $|J^*| = |J_0| = |\bar{J}|$.

We obtain therefore

$$N(n, J_0') = N(n, \bar{J}) - N(q, \bar{J}) + N(q, J^*),$$

$$N(n, \bar{J}) = N(n, J_0') + [N(q, \bar{J}) - n|\bar{J}|] - [N(q, J^*) - n|\bar{J}|].$$

But here $N(n, \bar{J})$ is $\leqslant N(n, J)$ while $N(n, J_0')$ is, by (5.3), $\geqslant n|J_0|$. As both bracket expressions are $\leqslant A(q)$ we obtain

$$N(n, J) \geqslant n|J_0'| - 2A(q). \tag{5.4}$$

32. On the other hand, similarly as in Section 27,

$$|J_0'| = |J_0| = \prod_{\mu=1}^{m} (d_\mu - 2\varepsilon) \geqslant \prod_{\mu=1}^{m} d_\mu - 2\varepsilon 2^m = |J| - 2^{m+1}\varepsilon.$$

Introducing this into (5.4) we obtain finally

$$N(n, J) - n|J| \geqslant -\big(2A(q) + 2^{m+1} n\varepsilon\big). \tag{5.5}$$

33. If we now drop the restriction (5.1) and assume that at least one of the d_μ is $\leqslant 2\varepsilon$, obviously $|J| \leqslant 2\varepsilon$. But then the relation (5.5) holds again and is therefore now proved independently of the restriction (5.1).

Combining (4.7) and (5.5) it follows

$$\big|N(n, J) - n|J|\big| \leqslant 2^{m+1} A\big(y(\varepsilon)\big) + 2^{2m+1} n\varepsilon.$$

Refering now to the definition (1.8) of $A(x)$ it follows now since obviously $A(x) = A([x])$, while $y(\varepsilon)$ in (2.10) is continuous,

$$A(x) \leqslant 2^{m+1}A(y)+2^{2m+1}x\varepsilon(y) \quad (x \geqslant y \geqslant 1). \tag{5.6^0}$$

This functional inequality the derivation of which is the essential point of our method, is a special case of the following inequality

$$A(x) \leqslant \alpha A(y)+\beta x\varepsilon(y) \quad (\alpha \geqslant 1,\ \beta > 0,\ x \geqslant y \geqslant 1), \tag{5.6}$$

where α and β are given constants.

§ 6. Discussion of the fundamental inequality

34. We are going first to treat the general inequality (5.6). We assume generally about $\varepsilon(y)$ that it is positive and monotonically decreasing to 0 with $y \to \infty$ while $A(x)$ is assumed to be positive and monotonically increasing for $x \geqslant 1$.

LEMMA 4. *Assume that for four constants g, g', L and x_0 with*

$$g > 1, \quad 0 < g' < g/\alpha, \quad L > 0, \quad x_0 \geqslant 1 \tag{6.1}$$

the following relations are satisfied

$$\varepsilon(x) \leqslant g'\varepsilon(gx) \quad (x_0 \leqslant x), \tag{6.2}$$

$$\varepsilon(x) \geqslant L \quad (1 \leqslant x \leqslant x_0). \tag{6.3}$$

Under these conditions

$$A(x) = O\big(x\varepsilon(x)\big) \quad (x \uparrow \infty) \tag{6.4}$$

and more precisely

$$A(x) \leqslant Dx\varepsilon(x) \quad (x \geqslant 1), \tag{6.5}$$

where D is defined by

$$D := \max\left(\frac{A(x_0)}{L}, \frac{gg'\beta}{g-g'\alpha}\right). \tag{6.6}$$

35. Proof. If we first assume that $1 \leqslant x \leqslant x_0$, it follows by (6.3) and (6.6) as $A(x)$ is increasing,

$$A(x) \leqslant A(x_0) \leqslant DL \leqslant Dx\varepsilon(x)$$

and we see that (6.5) holds for $1 \leqslant x \leqslant x_0$.

It is therefore sufficient to prove that, if (6.5) holds for an $x \geqslant 1$ it also holds for gx. But replacing in (5.6) x with gx and y with x it follows

$$A(gx) \leqslant \alpha A(x)+\beta gx\varepsilon(x) = (\alpha D+\beta g)x\varepsilon(x) \leqslant \frac{g'}{g}(\alpha D+\beta g)(gx)\varepsilon(gx)$$

and here the factor $\frac{g'}{g}(\alpha D+\beta g)$ is $\leqslant D$ as follows immediately from

$$D \geqslant \frac{gg'\beta}{g-g'\alpha}.$$

The inequality (6.5) is completely proved.

36. Consider now instead of the assumptions made in Section 34 the assumptions

$$\alpha > 1, \quad \varepsilon(x) = \delta(x)/x, \quad 0 < \delta(x) < 1, \quad A(x) \leqslant x \quad (x \geqslant 1). \tag{6.7}$$

Then choosing

$$\log z := \sqrt{\log\alpha \log x} \tag{6.8}$$

put

$$n := \left[\frac{\log x}{\log z}\right] = \left[\sqrt{\frac{\log x}{\log \alpha}}\right]. \tag{6.9}$$

It follows

$$1 \leqslant xz^{-n} < z,$$

$$\alpha^n \leqslant e^{\log\alpha \frac{\log x}{\log z}} = e^{\sqrt{\log\alpha\log x}} = z. \tag{6.10}$$

37. Put in (5.6) $y = x/z$. We obtain, by (6.7),

$$A(x) \leqslant \alpha A(x/z)+\beta z.$$

Writing this inequality for x/z^ν instead of x and multiplying it with α^ν,

$$\alpha^\nu A\left(\frac{x}{z^\nu}\right) \leqslant \alpha^{\nu+1}A\left(\frac{x}{z^{\nu+1}}\right)+\beta z\alpha^\nu.$$

Summing over $\nu = 0, 1, \ldots, n-1$ it follows

$$A(x) \leqslant \alpha^n A\left(\frac{x}{z^n}\right)+\beta z\sum_{\nu=0}^{n-1}\alpha^\nu.$$

We use now (6.10) and obtain

$$A(x) \leqslant z^2\left(1+\frac{\beta}{\alpha-1}\right) = o\left(e^{2\sqrt{\log\alpha\log x}}\right). \tag{6.11}$$

The reader may be reminded that in the case $\alpha = 1$, from $\varepsilon(x) = O(1/x)$ follows

$$A(x) = O(\log x)$$

as is shown in [6].

38. Replace now the conditions of Section 34 by the conditions

$$(6.12)\qquad \alpha>1,\quad \varepsilon(x)\leqslant\frac{\varkappa^\delta}{x^\delta}\quad (x\geqslant x_0,\ x_0\geqslant e),\qquad \delta>1,\ \varkappa>0.$$

Denote $1/\delta$ by ε and put in (5.6) $y=x^\varepsilon$. As by (6.12) $x\varepsilon(x^\varepsilon)\leqslant\varkappa$ it follows

$$(6.13)\qquad A(x)\leqslant\alpha A(x^\varepsilon)+\beta\varkappa\quad (x\geqslant x_0).$$

Replacing here x by x^{ε^ν} and multiplying by α^ν we obtain

$$\alpha^\nu A(x^{\varepsilon^\nu})\leqslant\alpha^{\nu+1}A(x^{\varepsilon^{\nu+1}})+\beta\varkappa\alpha^\nu\quad (0\leqslant\nu\leqslant n-1).$$

Adding over $\nu=0,1,\dots,n-1$ it follows

$$A(x)\leqslant\alpha^nA(x_0)+\beta\varkappa\sum_{\nu=0}^{n-1}\alpha^\nu\leqslant\alpha^nx_0+\beta\varkappa\frac{\alpha^n-1}{\alpha-1}<\alpha^n\left(x_0+\frac{\beta\varkappa}{\alpha-1}\right),$$

as soon as $x^{\varepsilon^n}\leqslant x_0$, that is, as soon as

$$\log x\leqslant\delta^n\log x_0,\qquad n\log\delta\geqslant\log\frac{\log x}{\log x_0}.$$

39. The last condition is satisfied as soon as $n\log\delta\geqslant\log\log x$,

$$n\geqslant\frac{\log\log x}{\log\delta},\qquad n=\left[\frac{\log\log x}{\log\delta}\right]+1.$$

For this value of n it follows

$$\log\alpha^n=n\log\alpha<\log\alpha\left(1+\frac{\log\log x}{\log\delta}\right)=\log\left[(\log x^\delta)^{\frac{\log\alpha}{\log\delta}}\right],$$

or putting

$$(6.14)\qquad \mu_0:=\frac{\log\alpha}{\log\delta},\qquad \delta^{\mu_0}=\alpha,$$

$$\alpha^n<\delta^{\mu_0}(\log x)^{\mu_0}=\alpha(\log x)^{\mu_0}.$$

Therefore, finally,

$$(6.15)\qquad A(x)\leqslant\left(x_0+\frac{\beta\varkappa}{\alpha-1}\right)\alpha(\log x)^{\mu_0},\qquad \mu_0:=\frac{\log\alpha}{\log\delta}.$$

§ 7. $A(x)$ in dependence on $\psi(x)$

40. Returning now to the functional inequality (5.6) derived under the conditions specified in Section 1 we have to use the value (1.9) of $\varepsilon(x)$,

$$(7.1)\qquad \varepsilon(x)=\frac{1}{\gamma_0\varphi\left(\frac{1}{\gamma_0x}\right)}.$$

Thence, solving with respect to φ and using the inverse function ψ,

$$(7.2)\qquad \varphi\left(\frac{1}{\gamma_0x}\right)=\frac{1}{\gamma_0\varepsilon(x)},\qquad \frac{1}{\gamma_0x}=\psi\left(\frac{1}{\gamma_0\varepsilon(x)}\right).$$

However the cases (6.7) and (6.12) can be discarded. Indeed under the assumption (6.7) it follows from (7.2)

$$\varphi\left(\frac{1}{\gamma_0x}\right)=\frac{1}{\gamma_0\varepsilon(x)}>\frac{x}{\gamma_0},\qquad \psi\left(\frac{x}{\gamma_0}\right)>\frac{1}{\gamma_0x}$$

so that finally $\psi(x)>1/\gamma_0x$. But this is only possible for $m=1$, $\gamma_0=1$ and in this case $\psi(x)$ is always $<1/x$.

In the case of the condition (6.12) we obtain from (7.2)

$$\varphi\left(\frac{1}{\gamma_0x}\right)>\frac{x^\delta}{\gamma_0\varkappa},\qquad \psi\left(\frac{x^\delta}{\gamma_0\varkappa^\delta}\right)>\frac{1}{\gamma_0x},$$

and putting $y:=x^\delta/(\gamma_0\varkappa^\delta)$, $x=\varkappa(\gamma_0\varkappa^\delta y)^{1/\delta}$ it follows

$$\psi(y)>\frac{1}{\gamma_0\varkappa(\gamma_0y)^{1/\delta}}$$

which is impossible for a sufficiently large y, since $1/\delta<1$.

We have therefore only to consider the case of Section 34.

41. The assumptions (6.1), (6.2) and (6.3) in Section 34 can be considerably simplified. Putting $g'>1$,

$$(7.3)\qquad 0<\varrho:=\frac{\log g'}{\log g}<1,\qquad \varepsilon(x)=:\left(\frac{l(x)}{x}\right)^\varrho$$

the relation (6.2) becomes

$$\varepsilon(x)\leqslant g^\varrho\varepsilon(gx),\qquad \left(\frac{l(x)}{x}\right)^\varrho\leqslant g^\varrho\left(\frac{l(gx)}{gx}\right)^\varrho=\left(\frac{l(gx)}{x}\right)^\varrho,$$

$$(7.4)\qquad l(x)\leqslant l(gx)\qquad (x\geqslant x_0).$$

The relations (6.1) and (6.3) become now

$$(7.5)\qquad g>1,\ 0<\varrho<1,\ g^{1-\varrho}>\alpha,\ L>0,\ \varepsilon(x)>L\ (1\leqslant x\leqslant x_0).$$

The inequality (7.4) is in any case satisfied if $l(x)$ is assumed as *non decreasing*. In this particular case (6.2) holds for any sufficiently large $g>1$ and, for a fixed ϱ, (6.2) holds for all $g'=g^\varrho$ from a $g>1$ on. From now on we restrict ourselves to the case (7.3) with ϱ constant, $0<\varrho<1$.

The simplest case is of course $l(x)=c=$ constant,

$$(7.6)\qquad \varepsilon(x)\leqslant(c/x)^\varrho,\qquad 0<\varrho<1,\qquad r:=1/\varrho.$$

By (7.2) it follows

$$\varphi\left(\frac{1}{\gamma_0 x}\right) \geqslant \frac{x^\varrho}{\gamma_0 c^\varrho}, \quad \varphi(x) \geqslant \frac{1}{\gamma_0 (c\gamma_0 x)^\varrho}. \tag{7.7}$$

For the inverse function of $y = \varphi(x)$ it follows now

$$\psi(y) \geqslant \left(\frac{1}{\gamma_0 y}\right)^r \frac{1}{\gamma_0 c}. \tag{7.8}$$

Inversely from (7.8) follows (7.7). From (6.5) we obtain now

$$A(x) = O(x^{1-\varrho}). \tag{7.9}$$

42. *The formula* (7.9) *holds in particular if the* a_μ *in* (1.1) *are all algebraic.* To prove this denote by a some primitive element of the field $R(a_1, \ldots, a_m)$ so that

$$a_\mu = h_\mu(a) \quad (\mu = 1, \ldots, m)$$

where the h_μ are polynomials with rational coefficients. Then, denoting by $u_0, u_1, \ldots, u_m$ independent indeterminates, put

$$P(x) = \sum_{\mu=1}^{m} u_\mu h_\mu(x) + u_0.$$

Let $n+1$ be the degree of a with respect to R. Denoting by $a^{(\nu)}$ $(\nu = 0, 1, \ldots, n)$, $a^{(0)} = a$, the complete set of the conjugates of a, form the expression

$$T(u_0, \ldots, u_m) = \prod_{\nu=0}^{n} P(a^{(\nu)})$$

which is a polynomial with rational coefficients with common denominator N. If we put for the u_μ rational integers g_μ with $\gamma := \max_\mu |g_\mu|$ we have for a fixed natural N:

$$NT(g_0, \ldots, g_m) = G \neq 0,$$

with a rational integer G, so that $|NT(g_0, \ldots, g_m)| \geqslant 1$.

On the other hand

$$T^*(g_0, \ldots, g_m) := T(g_0, \ldots, g_m)/\left[\sum_{\mu=1}^{m} g_\mu h_\mu(a) + g_0\right]$$

is of dimension n and therefore $T^*(g_0, \ldots, g_m) = O(\gamma^n)$. It follows

$$\left|\sum_{\mu=1}^{m} g_\mu h_\mu(a) + g_0\right| > \frac{C}{\gamma^n}, \quad C > 0$$

with a constant C. We obtain from (1.2)

$$\psi(\gamma) > C/\gamma^n$$

which is the relation (7.8) with $r = n$ and thence (7.9) with $\varrho = 1/n$.

43. We can assume now $l(x)$ as strictly monotonically increasing. The essential difficulty in applying (6.5) consists in the necessity to obtain sufficiently good approximation of the inverse functions $\varphi(x)$ and $\psi(x)$. To do this we use the

Lemma 5. *Assume $l(x)$ for an $x \geqslant x_0 > 1$ a positive strictly monotonically increasing function of x such that $x/l(x)$ also strictly monotonically increases. Let $0 < \varrho < 1$ and put $r := 1/\varrho$. Then necessary and sufficient for the inequality*

$$\varepsilon(x) = \frac{1}{\gamma_0 \varphi(1/(\gamma_0 x))} \leqslant C\left(\frac{l(x)}{x}\right)^\varrho \quad (x \geqslant x_0) \tag{7.10}$$

is that $\psi(x)$ satisfies the inequality

$$\psi(y) \geqslant \frac{1}{\gamma_0 (Dy)^{1/\varrho} k\left((Dy)^{1/\varrho}\right)}, \quad D := \gamma_0 C \quad (y \geqslant y_0) \tag{7.11}$$

for a convenient constant $y_0 > 0$, where with

$$z := x/l(x), \quad z \geqslant x_0/l(x_0), \tag{7.12}$$

$k(z)$ is defined by

$$k(z) := l(x), \quad x = zk(z). \tag{7.13}$$

44. Proof. Using $\varepsilon(x)$ from (7.10) it follows

$$\varphi\left(\frac{1}{\gamma_0 x}\right) \geqslant \left(\frac{x}{l(x)}\right)^\varrho \Big/ D =: y. \tag{7.14}$$

Since $z = x/l(x)$ is strictly monotonically increasing, the same holds for $k(z)$ defined by (7.13) and it follows from (7.13) and (7.14) that

$$z = (Dy)^{1/\varrho}, \quad z \geqslant \frac{x_0}{l(x_0)}, \quad y \geqslant \left(\frac{x_0}{l(x_0)}\right)^\varrho \Big/ D =: y_0. \tag{7.15}$$

Applying to both sides of (7.14) the function ψ we obtain

$$\psi(y) \geqslant 1/(\gamma_0 x)$$

and since by (7.13) and (7.15) $x = (Dy)^{1/\varrho} k\left((Dy)^{1/\varrho}\right)$, (7.11) follows.

45. On the other hand, assuming (7.11) for $y \geqslant y_0$ with $(Dy_0)^{1/\varrho} \times \times k\left((Dy_0)^{1/\varrho}\right) > 1$ and defining z by (7.15) we can rewrite (7.11) as

$$\psi(y) \geqslant \frac{1}{\gamma_0 z k(z)}, \quad z \geqslant (Dy_0)^{1/\varrho} =: z_0. \tag{7.16}$$

Put then in (7.16)

(7.17) $$x := zk(z),\qquad x \geqslant z_0 k(z_0) =: x_0 > 1$$

and apply on both sides of (7.16) the function φ. We obtain

(7.18) $$y \leqslant \varphi\big(1/(\gamma_0 x)\big).$$

Defining now $l(x)$ by (7.13) we obtain from (7.12) and (7.15)

$$y = z^{\varrho}/D = \left(\frac{x}{l(x)}\right)^{\varrho}\Big/D.$$

(7.18) becomes now

$$\varphi\left(\frac{1}{\gamma_0 x}\right) \geqslant \left(\frac{x}{l(x)}\right)^{\varrho}\Big/D$$

and the formula (7.10) follows.

46. Applying Lemma 5 and starting from an inequality of the type of (7.11), it is important to find convenient functions $k(y)$. The following lemma allows this in a greater number of cases.

LEMMA 6. *Assume for $x \geqslant x_0 \geqslant e^e$, with $x \to \infty$, $k(x)$ strictly increasing,*

(7.19) $$e < k(x) \uparrow \infty \qquad (x \uparrow \infty),$$

(7.20) $$xk'(x) = o\big(k(x)/\log k(x)\big),$$

and define $Z(x)$ by:

(7.21) $$zk(z) = x,\qquad z = Z(x) \uparrow \infty.$$

Then for an arbitrary small $\varepsilon > 0$ with $x \to \infty$:

(7.22) $$\frac{x}{k\left(\dfrac{x}{k(x)}\right)} > Z(x) > \frac{x}{k(x)},$$

(7.23) $$k(x) = O(e^{\varepsilon\sqrt{\log x}}),$$

(7.24) $$Z(x) = \frac{x}{k(x)}\big(1+o(1)\big).$$

47. Proof. From $k(x_0) > e$ it follows by (7.21) $Z(x_0) < x_0$, $Z(x) < x$:

$$k(z) < k(x) \qquad (x \geqslant x_0)$$

and from

(7.25) $$z = x/k(z),$$

we obtain

(7.26) $$Z(x) > \frac{x}{k(x)},\qquad k(z) > k\left(\frac{x}{k(x)}\right).$$

From (7.25) and (7.26) we obtain further

(7.27) $$Z(x) < \frac{x}{k\big(x/k(x)\big)}$$

and (7.22) is proved.

48. By (7.20) we obtain, for an $\varepsilon > 0$,

$$x\big((\log k(x))^2\big)' < \varepsilon^2 \qquad (x > x_1),$$
$$\big((\log k(x))^2\big)' < \varepsilon^2/x \qquad (x > x_1),$$
$$\big(\log k(x)\big)^2 < \varepsilon^2\log x + \big(\log k(x_1)\big)^2 - \varepsilon^2\log x_1,$$
$$\log k(x) < \sqrt{\varepsilon^2\log x + c} < \varepsilon\sqrt{\log x} + \sqrt{c}$$

for a constant c, and (7.23) follows.

49. Finally using (7.20) we obtain

$$k(x) - k\left(\frac{x}{k(x)}\right) = \int_{x/k(x)}^{x} k'(y)\,dy = o\left(\int_{x/k(x)}^{x} \frac{k(y)}{\log k(y)}\,\frac{dy}{y}\right).$$

But obviously, in virtue of

$$\big(\log k(x)\big)^2\left(\frac{k(x)}{\log k(x)}\right)' = k'(x)\big(\log k(x) - 1\big) \geqslant 0,$$

we can take in the last integral the factor $\dfrac{k(x)}{\log k(x)}$ out of the integral and obtain

$$k(x) - k\left(\frac{x}{k(x)}\right) = o\left(\frac{k(x)}{\log k(x)}\int_{x/k(x)}^{x}\frac{dy}{y}\right) = o\left(\frac{k(x)}{\log k(x)}\log k(x)\right) = o\big(k(x)\big).$$

It follows

$$k(x)\Big/k\left(\frac{x}{k(x)}\right) \to 1 + o(1)$$

and (7.24) follows from (7.22). Lemma 6 is proved.

50. We can now formulate in a particularly simple and important case

THEOREM 2. *Assume $k(x)$ a constant or strictly increasing function satisfying the conditions* (7.19) *and* (7.20). *Assume* (7.11) *for a convenient*

$C > 1$ and a ϱ with $0 < \varrho < 1$. Then

(7.28) $$\varepsilon(x) = O\left(\left(\frac{k(x)}{x}\right)^{\varrho}\right),$$

(7.29) $$A(x) = O\big(x^{1-\varrho}k^{\varrho}(x)\big).$$

Proof. Defining z by (7.21) it follows from $l(x) := k(z)$ and (7.24):

$$\frac{l(x)}{x} = \frac{k(z)}{zk(z)} = \frac{1}{z} = \frac{k(x)}{x}\big(1+o(1)\big),$$

$$l(x) = k(x)\big(1+o(1)\big)$$

and therefore, by (6.4) and (7.10), (7.28) and (7.29).

51. Consider, for instance, the monotonically increasing expressions of the type

$$k(x) := c\log_1^{a_1}x\log_2^{a_2}x \ldots \log_n^{a_n}x \quad (x > x_0),$$

where generally the ν-times iterated logarithm of x is denoted by $\log_\nu x$ and the first non vanishing term in the sequence $a_1, a_2, \ldots, a_n$ is *positive.*

Then we have for the logarithmic derivative of $k(x)$:

$$\frac{k'(x)}{k(x)} = \sum_{\nu=1}^{n} \frac{a_\nu}{x\log_1 x \ldots \log_{\nu-1}x} = O\left(\frac{1}{x\log x}\right).$$

Since $\log k(x) = O(\log_2 x)$ it follows

$$xk'(x)\log k(x)/k(x) = O\left(\frac{\log_2 x}{\log x}\right) = o(1)$$

and the conditions of Lemma 6 are satisfied. It follows

(7.30) $$A(x) = O\left(\frac{x^{1-\varrho}}{k(x)^{\varrho}}\right).$$

References

[1] J. W. S. Cassels, *An introduction to diophantine approximation,* Cambridge Tracts, Cambridge University Press, 1951.

[2] S. Hartman, *Une généralisation d'un théorème de M. Ostrowski sur la répartition des nombres* mod 1, Ann. Soc. Polon. Math. 22 (1949), pp. 169–172.

[3] E. Hecke, *Über analytische Funktionen und die Verteilung von Zahlen mod Eins,* Hamb. Math. Abh. 1 (1921), pp. 54–76; in particular pp. 72–76. Reprinted in Hecke's Math. Werke, Göttingen 1959, pp. 313–335, in particular § 6, pp. 331–335.

[4] A. Khintchine, *The quantitative conception of Kronecker's approximation theory,* Izv. Acad. Nauk USSR 12 (1943), pp. 113–122.

[5] A. Ostrowski, *Zur Theorie der linearen Diophantischen Approximationen,* Jber. Deutsch. Math.-Verein. 39 (1930), pp. 34–46.

[6] — *On rational approximations to an irrational number,* Rendiconti Sem. mat. Milano 47 (1977), pp. 241–256.

[7] H. Weyl, *Über die Gibbssche Erscheinung und verwandte Konvergenzphänomene,* Rend. Circ. Mat. Palermo 30 (1910), pp. 377–407; in particular p. 406. Reprinted in Weyl's collected papers, vol. 1, Springer 1968, pp. 352–353.

Received on 12.2.1980
and in revised form on 20.5.1980 (1200)