On a Problem in Diophantine Approximations

By KURT MAHLER in Manchester

Frequent use is made in the theory of Diophantine approximations of the following result which goes back to DIRICHLET[1]):

Theorem 1: *Let $L_1(x), \ldots, L_m(x)$ be m linear forms in n variables $x_1, x_2, \ldots, x_n$; assume that $m < n$, and that the coefficients of these forms are real numbers of absolute value not greater than the positive integer a. Then, for every positive integer N, there exist integral values not all zero of $x_1, x_2, \ldots, x_n$ such that*

$$|L_\mu(x)| \leqq \frac{1}{N} \text{ for } \mu = 1, 2, \ldots, m, \text{ and } |x_\nu| \leqq 2\,(anN)^{\frac{m}{n-m}} \text{ for } \nu = 1, 2, \ldots, n.$$

In this theorem, the variable factor $(aN)^{\frac{m}{n-m}}$ occurring in the upper bound for $|x_\nu|$ cannot be replaced by any lower power of aN, as can be proved by considering special forms. On the other hand, the factor $2\,n^{\frac{m}{n-m}}$ independent of a and N is not best-possible, and the problem arises of replacing this factor by the smallest possible functions of m and n. Unfortunately, this problem seems to be extremely hard, even when m and n are small.

In the present note I treat only the case $m = 1$ of a single linear form $L(x)$ in n variables and determine an improved constant which, however, is still not the best. Instead of using the DIRICHLET "Schubfach-Prinzip", I apply the much more powerful *theorem of* MINKOWSKI *on convex bodies* and so obtain the following result:

Theorem 2: *Let $L(x)$ be a linear form in n variables $x_1, x_2, \ldots, x_n$; assume that $n \geq 2$, and that the coefficients of this form are real numbers of absolute value not greater than the real number $a \geq 1$. Denote by v_{n-1} the volume of the $(n-1)$-dimensional convex polyhedron*

$$|x_1| \leq 1,\ |x_2| \leq 1, \ldots, |x_{n-1}| \leq 1,\ |x_1 + x_2 + \cdots + x_{n-1}| \leq 1.$$

Then, for every real number $N \geqq 1$, there exist integers $x_1, x_2, \ldots, x_n$ not all zero such that

$$|L(x)| \leq \frac{1}{N},\quad \max(|x_1|, |x_2|, \ldots, |x_n|) \leq 2\left(\frac{aN}{v_{n-1}}\right)^{\frac{1}{n-1}}.$$

[1]) See, e.g., C. L. SIEGEL, Transcendental Numbers (Princeton 1949), 35—36; A. O. GEL'FOND, Transcendental and Algebraic Numbers (In Russian), (Moscow 1952), 18—19.

This theorem is stronger than the case $m = 1$ of Theorem 1 because

$$v_{n-1} \geq 2 \quad \text{for } n \geq 2, \text{ and } v_{n-1} \sim \sqrt{\frac{3}{2\pi n}}\, 2^n \quad \text{as } n \to \infty,$$

and it is nearly best-possible when $n = 2$.

There is evidently no loss of generality if the parameter a in Theorem 2 is put equal to 1; this assumption is therefore made throughout this note.

1. For $n \geqq 2$ let $a_1, a_2, \ldots, a_{n-1}$ be $n-1$ non-negative numbers, and let $K(a) = K(a_1, a_2, \ldots, a_{n-1})$ be the convex polyhedron in $(n-1)$-dimensional space defined by

$$|x_1| \leq 1, \quad |x_2| \leq 1, \ldots, |x_{n-1}| \leq 1, \quad |a_1 x_1 + a_2 x_2 + \ldots + a_{n-1} x_{n-1}| \leq 1.$$

Further denote by $v(a) = v(a_1, a_2, \ldots, a_{n-1})$ the volume of this body. The proof of Theorem 2 depends essentially on the following lemma.

Lemma 1: *The volume $v(a)$ is a non-increasing function in each of the $n-1$ parameters $a_1, a_2, \ldots, a_{n-1}$.*

Proof: The assertion is trivially true if $n = 2$ because then $v(a_1) = 2 \min\left(1, \frac{1}{a_1}\right)$. Let therefore now $n \geq 3$. Evidently $v(a)$ is a symmetric function of the a's. It thus suffices to prove the assertion with respect to the parameter a_{n-1}.

The transformation

$$y_1 = a_1 x_1, \quad y_2 = a_2 x_2, \ldots, \quad y_{n-1} = a_{n-1} x_{n-1}$$

changes $K(a)$ into the new polyhedron $K'(a) = K'(a_1, a_2, \ldots, a_{n-1})$ given by

$$|y_1| \leq a_1, \ |y_2| \leq a_2, \ldots, |y_{n-1}| \leq a_{n-1}, \ |y_1 + y_2 + \ldots + y_{n-1}| \leq 1$$

and, say, of volume $v'(a) = v'(a_1, a_2, \ldots, a_{n-1})$, where evidently

$$v'(a) = a_1 a_2 \ldots a_{n-1} v(a).$$

The assertion is thus equivalent to the inequality,

$$\frac{v'(a_1, \ldots, a_{n-2}, a_{n-1})}{a_{n-1}} \geq \frac{v'(a_1, \ldots, a_{n-2}, a'_{n-1})}{a'_{n-1}} \quad \text{if} \quad 0 \leq a_{n-1} \leq a'_{n-1}.$$

To prove this assertion, denote by t a real parameter, by $\varkappa(t)$ the $(n-2)$-dimensional polyhedral section

$$|y_1| \leq a_1, \ldots, |y_{n-2}| \leq a_{n-2}, \ y_{n-1} = t, \ |y_1 + y_2 + \ldots + y_{n-1}| \leq 1$$

of $K'(a)$, and by $\omega(t)$ the $(n-2)$-dimensional volume of $\varkappa(t)$. Since $K'(a)$ is symmetrical in the origin, it is obvious that $\omega(t)$ is an even function of t, i.e. that

$$\omega(-t) = \omega(t),$$

and that therefore

$$v'(a_1, \ldots, a_{n-2}, a_{n-1}) = 2 \int_0^{a_{n-1}} \omega(t)\, dt.$$

Now, if t and τ are arbitrary real numbers and if $0 \leq \Theta \leq 1$, then by the BRUNN-MINKOWSKI theorem on the sections of a convex body[2]),

$$\omega(\Theta t+(1-\Theta)\tau)^{\frac{1}{n-2}} \geq \Theta\omega(t)^{\frac{1}{n-2}} + (1-\Theta)\,\omega(\tau)^{\frac{1}{n-2}}.$$

On substituting $\tau = -t$ and $\Theta = 1/2$, we find thus that

$$\omega(t) \leq \omega(0) \qquad \text{for all } t.$$

Also, for $\tau = 0$, the BRUNN-MINKOWSKI inequality gives

$$\omega(\Theta t)^{\frac{1}{n-2}} \geq \Theta\omega(t)^{\frac{1}{n-2}} + (1-\Theta)\,\omega(0)^{\frac{1}{n-2}} \geq \omega(t)^{\frac{1}{n-2}},$$

and so, on replacing Θt by τ, we obtain the inequality

$$\omega(\tau) \geq \omega(t) \qquad \text{if} \quad 0 \leq \tau \leq t.$$

Assume now that $0 \leq a_{n-1} \leq a'_{n-1}$. Then, in particular,

$$\omega(t) \geq \omega(a_{n-1}) \qquad \text{for} \quad 0 \leq t \leq a_{n-1},$$

$$\omega(t) \leq \omega(a_{n-1}) \qquad \text{for} \quad a_{n-1} \leq t \leq a'_{n-1}.$$

Hence, by the first mean value theorem of integral calculus,

$$\int_0^{a_{n-1}} \omega(t)\,dt \geq a_{n-1}\,\omega(a_{n-1}), \qquad \int_{a_{n-1}}^{a'_{n-1}} \omega(t)\,dt \leq (a'_{n-1}-a_{n-1})\,\omega(a_{n-1}),$$

and therefore

$$\int_0^{a'_{n-1}} \omega(t)\,dt \leq \int_0^{a_{n-1}} \omega(t)\,dt \left(1+\frac{a'_{n-1}-a_{n-1}}{a_{n-1}}\right) = \frac{a'_{n-1}}{a_{n-1}} \int_0^{a_{n-1}} \omega(t)\,dt,$$

whence the assertion.

Corollary: *If* $0 \leq a_1 \leq 1, \ldots, 0 \leq a_{n-1} \leq 1$, *then*

$$v(a_1, a_2, \ldots, a_{n-1}) \geq v(1, 1, \ldots, 1) = v_{n-1}.$$

2. It is obvious that $v_1 = 2$. More generally, the following result holds.

Lemma 2: $v_n \geq v_{n-1}$, *and therefore* $v_{n-1} \geq 2$ *for* $n = 2, 3, 4, \ldots$.

Proof: Let $n \geq 2$. By definition, v_n is the volume of the n-dimensional polyhedron K_n defined by

$$|x_1| \leq 1,\ |x_2| \leq 1, \ldots, |x_{n-1}| \leq 1,\ |x_n| \leq 1,\ |x_1+x_2+\cdots+x_n| \leq 1.$$

This polyhedron K_n contains as a subset the polyhedron $K_n{}^*$ given by the additional condition

$$|x_1+x_2+\cdots+x_{n-1}| \leq 1.$$

Denote by $v_n{}^*$ the volume of $K_n{}^*$; the assertion will be proved if it can be shown that $v_n{}^* \geq v_{n-1}$. This may be done as follows.

To every point $X_n = (x_1, x_2, \ldots, x_{n-1}, x_n)$ of $K_n{}^*$ corresponds its projection $X_{n-1} = (x_1, x_2, \ldots, x_{n-1})$ in K_{n-1}. Assume now, say, that

$$0 \leq x_1+x_2+\cdots+x_{n-1} \leq 1.$$

[2]) See BONNESEN-FENCHEL, Konvexe Körper (Berlin 1934), 71—73, 87f.

The n-th coordinate x_n of X_n is then restricted solely by the two inequalities

$$-1 \leq x_n \leq 1\,, \quad -1 \leq (x_1 + x_2 + \cdots + x_{n-1}) + x_n \leq 1$$

and so may certainly assume all values for which

$$-1 \leq x_n \leq 0\,.$$

Similarly x_n may run over the whole interval

$$0 \leq x_n \leq 1 \qquad \text{if} \qquad -1 \leq x_1 + x_2 + \cdots + x_{n-1} \leq 0\,.$$

It follows then that the volume of K_n^* is not less than that of a prism of height 1 over K_{n-1} as its basis, and so $v_n^* \geq 1 \cdot v_{n-1}$, whence the assertion.

3. It is not difficult to find explicit formulae for v_{n-1}, if use is made of DIRICHLET's discontinuous integral

$$\int\limits_0^\infty \sin x \cos \lambda x \frac{dx}{x}\,.$$

In this way, one first obtains the identity

$$v_{n-1} = \frac{2^n}{\pi} \int\limits_0^\infty \left(\frac{\sin x}{x}\right)^n dx\,,$$

and from this derives, secondly, that

$$v_{n-1} = 2^{n-1} - \frac{2}{(n-1)!} \sum_{k=0}^{[n/2]-1} (-1)^k \binom{n-1}{k} (n-2k-2)^{n-1}\,.$$

Therefore, in particular,

$$v_1 = 2, \quad v_2 = 3, \quad v_3 = \frac{16}{3}, \quad v_4 = \frac{115}{12}, \quad v_5 = \frac{88}{5}, \quad v_6 = \frac{5887}{180}, \text{ etc.}$$

Thirdly, if in the integral for v_{n-1} a new variable $t = x\sqrt{n}$ is introduced, then the integral changes into

$$v_{n-1} = \frac{2^n}{\pi\sqrt{n}} \int\limits_0^\infty \left(\frac{\sin t/\sqrt{n}}{t/\sqrt{n}}\right)^n dt\,,$$

and here, uniformly in every finite interval for t,

$$\lim_{n\to\infty} \left(\frac{\sin\ /\sqrt{n}}{t/\sqrt{n}}\right)^n = e^{-t^2/6}\,.$$

It is therefore not difficult to prove that

$$v_{n-1} \sim \frac{2^n}{\pi\sqrt{n}} \int\limits_0^\infty e^{-t^2/6}\, dt = \sqrt{\frac{3}{2\pi n}}\, 2^n \qquad \text{as } n \to \infty\,.$$

There is no need for giving detailed proofs of these statements which serve only as an interpretation of Theorem 2.

4. We procede now to the proof of Theorem 2.

By the hypothesis, the linear form $L(x)$ is of the form

$$L(x) = a_1 x_1 + a_2 x_2 + \dots + a_n x_n$$

where the coefficients $a_1, a_2, \dots, a_n$ are real numbers such that

$$\max(|a_1|, |a_2|, \dots, |a_n|) \leq 1 .$$

The aim is to show that there exist integers $x_1, x_2, \dots, x_n$ such that

$$|L(x)| \leq \frac{1}{N}, \qquad 0 < \max(|x_1|, |x_2|, \dots, |x_n|) \leq 2\left(\frac{N}{v_{n-1}}\right)^{\frac{1}{n-1}} .$$

There is clearly no loss of generality in assuming, (i) that all coefficients a_ν are non-negative, and (ii)

$$\max(a_1, a_2, \dots, a_n) = 1 .$$

Moreover, we may demand that this maximum is attained by a_n, so that

$$0 \leq a_1 \leq 1, \qquad 0 \leq a_2 \leq 1, \dots, \qquad 0 \leq a_{n-1} \leq 1, \qquad a_n = 1 .$$

We distinguish now two cases. As a first case, assume that at least one of the coefficients $a_1, a_2, \dots, a_{n-1}$ be greater than $1 - \frac{1}{N}$; let, say,

$$1 - \frac{1}{N} < a_1 \leq 1 .$$

Then the assertion is satisfied by the choice

$$x_1 = 1,\ x_2 = x_3 = \dots = x_{n-1} = 0,\ x_n = -1 ,$$

because

$$|L(x)| = |a_1 - 1| < \frac{1}{N}, \quad 0 < \max(|x_1|, |x_2|, \dots, |x_n|) = 1 \leq 2\left(\frac{N}{v_{n-1}}\right)^{\frac{1}{n-1}};$$

for $N \geq 1$, and it is obvious that $v_{n-1} \leq 2^{n-1}$.

As a second case, assume that no coefficient $a_1, a_2, \dots, a_{n-1}$ is greater than $1 - \frac{1}{N}$. Put

$$b_\nu = \frac{N}{N-1} a_\nu \qquad (\nu = 1, 2, \dots, n-1),$$

so that the new numbers $b_1, b_2, \dots, b_{n-1}$ are non-negative and

$$\max(b_1, b_2, \dots, b_{n-1}) \leq 1 .$$

We must show that the region R_n in n-dimensional space defined by the inequalities

$$\left|\left(1 - \frac{1}{N}\right)(b_1 x_1 + b_2 x_2 + \dots + b_{n-1} x_{n-1}) + x_n\right| \leq \frac{1}{N},$$

$$\max(|x_1|, |x_2|, \dots, |x_n|) \leq 2\left(\frac{N}{v_{n-1}}\right)^{\frac{1}{n-1}}$$

contains at least one point $(x_1, x_2, \ldots, x_n)$ with integral coordinates not all zero. This region is a convex polyhedron symmetrical in the origin; hence, by MINKOWSKI's theorem on convex bodies, it certainly contains such a lattice point if its volume is not less than 2^n. This condition is in fact satisfied, as shall now be proved.

For shortness, denote by $\boldsymbol{K}_{n-1}$ the $(n-1)$-dimensional region

$$\max\left(|x_1|, |x_2|, \ldots, |x_{n-1}|, |b_1 x_1 + b_2 x_2 + \cdots + b_{n-1} x_{n-1}|\right) \leq 2\left(\frac{N}{v_{n-1}}\right)^{\frac{1}{n-1}}$$

which, in the notation of section 1, is of volume

$$v(b_1, b_2, \ldots, b_{n-1}) \frac{2^{n-1} N}{v_{n-1}}.$$

Here, by Lemma 1,

$$v(b_1, b_2, \ldots, b_{n-1}) \geq v_{n-1},$$

and so the volume of $\boldsymbol{K}_{n-1}$ is not less than $2^{n-1} N$.

Consider now a point $(x_1, x_2, \ldots, x_n)$ such that its projection $(x_1, x_2, \ldots, x_{n-1})$ lies in $\boldsymbol{K}_{n-1}$ while the n-th coordinate x_n satisfies the inequality

$$\left|\left(1 - \frac{1}{N}\right)(b_1 x_1 + b_2 x_2 + \cdots + b_{n-1} x_{n-1}) + x_n\right| \leq \frac{1}{N};$$

hence x_n may describe an interval of length $2/N$. Then

$$|x_n| \leq \left(1 - \frac{1}{N}\right)|b_1 x_1 + b_2 x_2 + \cdots + b_{x-1} x_{n-1}| + \frac{1}{N}$$

$$\leq \left(1 - \frac{1}{N}\right) \cdot 2\left(\frac{N}{v_{n-1}}\right)^{\frac{1}{n-1}} + \frac{1}{N} \leq 2\left(\frac{N}{v_{n-1}}\right)^{\frac{1}{n-1}}$$

because

$$2\left(\frac{N}{v_{n-1}}\right)^{\frac{1}{n-1}} \geq 1.$$

It follows that $(x_1, x_2, \ldots, x_n)$ lies in R_n, hence that the volume of R_n is not less than

$$\frac{2}{N} \cdot 2^{n-1} N = 2^n,$$

whence the assertion.

5. In the lowest case $n = 2$, Theorem 2 gives a result which may be expressed as follows:

Let $-1 \leq a \leq 1$ *and* $N \geq 1$. *Then there exist integers* x, y *not both zero such that*

$$|ax + y| \leq \frac{1}{N}, \quad \max(|x|, |y|) \leq N.$$

Here, in the second inequality, the coefficient 1 of N cannot be replaced by any smaller number. For take for N any sufficiently large positive integer, and choose a such that

$$\frac{1}{N} < a \leq \frac{1}{N-1}.$$

Then

$$\frac{1}{N} < a\,x \leq \frac{N-2}{N-1} < 1 - \frac{1}{N} \quad \text{if } x = 1, 2, \ldots, N-2,$$

and so, whatever the integer y,

$$|\,a\,x + y\,| > \frac{1}{N} \quad \text{if } x = 1, 2, \ldots, N-2.$$

It follows that if $0 < \Theta < 1$, then there are not always integral solutions x, y of the inequalities

$$|\,a\,x + y\,| \leq \frac{1}{N}, \quad 0 < \max(|\,x\,|, |\,y\,|) \leq \Theta N$$

provided N is so large that

$$\Theta N \leq N - 2, \quad \text{i. e.} \quad N \geq \frac{2}{1-\Theta}.$$

Manchester University.

Eingegangen am 30. 9. 1954