A note on two linear forms

by Klaus Moshchevitin

1. Diophantine exponents.
Let \( \theta_1, \theta_2 \) be real numbers such that

\[ 1, \theta_1, \theta_2 \text{ are linearly independent over } \mathbb{Z}. \quad (1) \]

We consider linear form

\[ L(x) = x_0 + x_1 \theta_1 + x_2 \theta_2, \quad x = (x_0, x_1, x_2) \in \mathbb{Z}^3. \]

By \( |z| \) we denote the Euclidean length of a vector \( z = (z_0, z_1, z_2) \in \mathbb{R}^3 \). Let

\[ \hat{\omega} = \hat{\omega}(\theta_1, \theta_2) = \sup \left\{ \gamma : \limsup_{t \to \infty} \left( t^\gamma \min_{0 < |x| \leq t} |L(x)| \right) < \infty \right\} \quad (2) \]

be the uniform Diophantine exponent for the linear form \( L \).

We consider another linear form \( P(x) \). The main result of the present paper is as follows.

**Theorem 1.** Suppose that linear forms \( L(x) \) and \( P(x) \) are independent and the exponent \( \hat{\omega} \) for the form \( L \) are defined in (2). Then for the Diophantine exponent

\[ \omega_{LP} = \sup \left\{ \gamma : \text{there exist infinitely many } x \in \mathbb{Z}^3 \text{ such that } |L(x)| \leq |P(x)| \cdot |x|^{-\gamma} \right\} \]

we have a lower bound

\[ \omega_{LP} \geq \hat{\omega}^2 - \hat{\omega} + 1. \]

**Remark.** Of course in the definition (2) and in Theorem 1 instead of the Euclidean norm \( |x| \) we may consider the value \( \max_{i=1,2} |x_i| \) as it was done by the most of authors.

Consider a real \( \theta \) which is not a rational number and not a quadratic irrationality. Define

\[ \omega_* = \omega_*(\theta) = \sup \{ \gamma : \text{there exist infinitely many algebraic numbers } \xi \text{ of degree } \leq 2 \]

such that \( |\theta - \xi| \leq H(\xi)^{-\gamma} \}

(here \( H(\xi) \) is the maximal value of the absolute values of the coefficients for canonical polynomial to \( \xi \)). Then for linear forms

\[ L(x) = x_0 + x_1 \theta + x_2 \theta^2, \quad P(x) = x_1 + 2x_2 \theta \]

one has

\[ \omega_* \geq \omega_{LP}. \quad (3) \]

So Theorem 1 immediately leads to the following corollary.

**Theorem 2.** For a real \( \theta \) which is not a rational number and not a quadratic irrationality one has

\[ \omega_* \geq \hat{\omega}^2 - \hat{\omega} + 1 \quad (4) \]

\[ ^1 \text{Research is supported by RFBR grant No.12-01-00681-a and by the grant of Russian Government, project 11. G34.31.0053.} \]
with $\hat{\omega} = \hat{\omega}(\theta, \theta^2)$.

2. Some history.

In 1967 H. Davenport and W. Schmidt [2] (see also Ch. 8 from Schmidt’s book [11]) proved that for any two independent linear forms $L, P$ there exist infinitely many integer points $x$ such that

$$|L(x)| \leq C|P(x)||x|^{-3},$$

with a positive constant $C$ depending on the coefficients of forms $L, P$. From this result they deduced that for any real $\theta$ which is not a rational number and not a quadratic irrationality the inequality

$$|\theta - \xi| \leq C_1 H(\xi)^{-3}$$

has infinitely many solutions in algebraic $\xi$ of degree $\leq 2$.

We see that for any two pairs of forms one has $\omega_{LP} \geq 3$. But from the Minkowski convex body theorem it follows that under the condition (1) one has $\hat{\omega} \geq 2$. Moreover

$$\min_{\hat{\omega} \geq 2} (\hat{\omega}^2 - \hat{\omega} + 1) = 3.$$

So our Theorems 1, 2 may be considered as generalizations of Davenport-Schmidt’s results.

Later Davenport and Schmidt generalized their theorems to the case of several linear forms [3]. In the next paper [4] they showed that the value of the uniform exponent for simultaneous approximations to any point $(\theta, \theta^2)$ is not greater than $\sqrt{5} - 1$. This together with Jarník’s transference equality (see [5]) leads to the bound $\hat{\omega} \leq 3 + \sqrt{5}$ which holds for all linear forms with coefficients of the form $\theta, \theta^2$. So for a linear form with coefficients $\theta, \theta^2$ one has

$$2 \leq \hat{\omega} \leq \frac{3 + \sqrt{5}}{2}. \quad (5)$$

D. Roy [9, 10] showed that the set of values $\hat{\omega}$ for linear forms under our consideration form a dense set in the segment [5]. Moreover he constructed a countable set of numbers $\theta$ such that

$$\hat{\omega}(\theta, \theta^2) = \frac{3 + \sqrt{5}}{2} \quad \text{and} \quad \omega_*(\theta) = 3 + \sqrt{5}.$$

This shows that our bound (1) from Theorem 2 is optimal in the right endpoint of the segment (5), namely for $\hat{\omega} = \frac{3 + \sqrt{5}}{2}$.

Other results on approximation by algebraic numbers are discussed in W. Schmidt’s book [11], in wonderful book by Y. Bugeaud [1] and in M. Waldschmidt’s survey [12].

Our proof of Theorem 1 generalizes ideas from [2, 3, 4] and uses Jarník’s inequalities [6, 7].

3. Minimal points.

In the sequel we may suppose that $\hat{\omega} > 2$ as the case $\hat{\omega} = 2$ follows from Davenport-Schmidt’s theorem (in this case our Theorem 1 claims that $\omega_{LP} \geq 3$). We take $\alpha < \hat{\omega}$ close to $\hat{\omega}$ so that $\alpha > 2$.

A vector $x = (x_0, x_1, x_2) \in \mathbb{Z}^3 \setminus \{0\}$ is defined to be a minimal point (or best approximation) if

$$\min_{x': 0 < |x'| \leq |x|} |L(x')| = L(x).$$

As $1, \theta_1, \theta_2$ are linearly independent, all the minimal points form a sequence $x_\nu = (x_{0, \nu}, x_{1, \nu}, x_{2, \nu})$, $\nu = 1, 2, 3, \ldots$ such that for $X_\nu = |x_{\nu}|$, $L_\nu = L(x_\nu)$ where one has

$$X_1 < X_2 < \ldots < X_\nu < X_{\nu+1} < \ldots, \quad L_1 > L_2 > \ldots > L_\nu > L_{\nu+1} > \ldots.$$
Here we should note that

\[ L_j \leq X_j^{-\alpha} \]

for all \( j \) large enough. Of course each vector \( x_j \) is primitive and each couple \( x_j, x_{j+1} \) form a basis of the two-dimensional lattice \( \mathbb{Z}^2 \cap \text{span} \{ x_j, x_{j+1} \} \).

Let \( F(x) \) be a linear form linearly independent with \( L \) and \( P \). Then

\[ \max \{ |L(x)|, |P(x)|, |F(x)| \} \asymp |x|. \]

We also use the notation \( P_\nu = P(x_\nu), F_\nu = F(x_\nu) \). In the sequel we need to consider determinants

\[ \Delta_j = \begin{vmatrix} L_{j-1} & P_{j-1} & F_{j-1} \\ L_j & P_j & F_j \\ L_{j+1} & P_{j+1} & F_{j+1} \end{vmatrix} = A \begin{vmatrix} x_{0,j-1} & x_{1,j-1} & x_{2,j-1} \\ x_{0,j} & x_{1,j} & x_{2,j} \\ x_{0,j+1} & x_{1,j+1} & x_{2,j+1} \end{vmatrix}, \]

where \( A \) is a non-zero constant depending on the coefficients of linear forms \( L, P, F \). We take into account \( (7), (6) \) and the inequality \( \alpha > 2 \) to see that

\[ \Delta_j = L_{j-1}P_jF_{j+1} - L_{j-1}P_{j+1}F_j + O(L_jX_{j+1}^2) = L_{j-1}P_jF_{j+1} - L_{j-1}P_{j+1}F_j + o(1), \quad j \to \infty. \]  

The following statement is a variant of Davenport-Schmidt’s lemma. We give it without a proof. It deals with three consecutive minimal points \( x_{j-1}, x_j, x_{j+1} \) lying in a two-dimensional linear subspace, say \( \pi \). We should note that our definition of minimal points differs from those in \([2, 3, 11]\). However the main argument is the same. It is discussed in our survey \([8]\). One may look for the approximation of the one dimensional subspace \( \ell = \pi \cap \{ \mathbf{z} : L(\mathbf{z}) = 0 \} \) by the points of two-dimensional lattice \( \Lambda_j = \langle x_{j-1}, x_j \rangle \) Then the points \( x_{j-1}, x_j, x_{j+1} \in \Lambda_j \) are the consecutive best approximations to \( \ell \) with respect to the induced norm on \( \pi \) (see \([8]\), Section 5.5).

**Lemma 1.** If for some \( j \) the points \( x_{j-1}, x_j, x_{j+1} \) are linearly dependent then

\[ x_{j+1} = tx_j + x_{j-1} \]

for some integer \( t \).

The next statement is known for long time. It comes from Jarník’s papers \([6, 7]\). It was rediscovered by Davenport and Schmidt in \([11]\) and discussed in our survey \([8]\).

**Lemma 2.** there exist infinitely many indices \( j \) such that the vectors \( x_{j-1}, x_j, x_{j+1} \) are linearly independent.

The following lemma is due to Jarník \([6, 7]\) (see also Section 5.3 from our paper \([8]\)).

**Lemma 3.** Suppose that \( j \) is large enough and the points \( x_{j-1}, x_j, x_{j+1} \) are linearly independent. Then

\[ X_{j+1} \gg X_j^{\alpha - 1} \]

and

\[ L_j \ll X_j^{-\alpha(\alpha - 1)} \]

Now we take large \( \nu \) and \( k \geq \nu + 1 \) such that

- vectors \( x_{\nu-1}, x_{\nu}, x_{\nu+1} \) are linearly independent;
- vectors \( x_{k-1}, x_k, x_{k+1} \) are linearly independent;
- vectors \( x_j, \nu \leq j \leq k \) belong to the two-dimensional lattice \( \Lambda_\nu = \mathbb{Z}^2 \cap \text{span} (x_\nu, x_{\nu+1}) \).

From Lemma 1 it follows that for \( j \) from the range \( \nu \leq j \leq k - 1 \) one has

\[ L_{j+1} = t_{j+1}L_j + L_{j-1}, \quad P_{j+1} = t_{j+1}P_j + P_{j-1}, \]

and

\[ X_{j+1} = t_{j+1}X_j + X_{j-1}, \quad X_{j+1} = t_{j+1}X_j + X_{j-1}, \]
with some integers $t_{j+1}$, and hence
\[ L_{\nu}P_{\nu+1} - L_{\nu+1}P_{\nu} = \pm (L_{k-1}P_{k} + L_{k}P_{k-1}). \] (11)

**Lemma 4.** Consider positive $r$ under the condition
\[ r < \alpha^2 - \alpha + 1 < \hat{\omega}^2 - \hat{\omega} + 1. \] (12)

Suppose that
\[ |P_{\nu}| \leq L_{\nu}X_{\nu}^r \] (13)
and $\nu$ is large. Then
\[ |P_{\nu+1}| \gg X_{\nu}^{\alpha-1}. \] (14)

Proof. For $j = \nu$ consider the second term in the r.h.s of (8). From (6,7,12) and the inequality (9) of Lemma 3 we have
\[ |L_{\nu-1}P_{\nu}F_{\nu+1}| \ll |L_{\nu-1}L_{\nu}X_{\nu}^r| X_{\nu+1} \ll X_{\nu}^{r-\alpha}X_{\nu+1}^{1-\alpha} \ll X_{\nu}^{r-\alpha+1} = o(1). \]

As $\Delta_{\nu} \neq 0$ we see that
\[ 1 \ll |L_{\nu-1}P_{\nu+1}F_{\nu}| \ll L_{\nu-1}|P_{\nu+1}| X_{\nu} \ll X_{\nu}^{1-\alpha}|P_{\nu+1}| \]
(in the last inequalities we use (7) and (6). Everything is proved. \(\square\).

4. **The main estimate.**

The following Lemma presents our main argument.

**Lemma 5.** Suppose that $r$ satisfies (12). Suppose that (6) holds for all indices $j$ and suppose that for a certain $\beta_0$ one has
\[ L_{\nu} \gg X_{\nu}^{-\beta_0}. \] (15)

Suppose that simultaneously we have
\[ |P_{\nu}| \leq L_{\nu}X_{\nu}^r, \] (16)
\[ |P_{k-1}| \leq L_{k-1}X_{k-1}^r, \] (17)
\[ |P_{k}| \leq L_{k}X_{k}^r. \] (18)

Then
\[ r \geq \alpha^2 + 1 - \frac{\beta_0}{\alpha - 1}. \] (19)
and
\[ L_{k} \gg X_{k}^{-\beta'}, \text{ with } \beta' = r - \alpha - 1 + \frac{\beta_0}{\alpha - 1} < \beta_0. \] (20)

First of all we note that
\[ L_{\nu+1}|P_{\nu}| \leq L_{\nu}L_{\nu+1}X_{\nu}^r \ll L_{\nu}X_{\nu+2}^{-\alpha}X_{\nu}^r \ll L_{\nu}X_{\nu+1}^{-\alpha}X_{\nu}^r \ll L_{\nu}X_{\nu}^{r-\alpha+1} = o(L_{\nu}X_{\nu}^{\alpha-1}). \]

Here the first inequality comes from (10). The second inequality is (6) with $j = \nu + 1$. The third one is simply $X_{\nu+2} \gg X_{\nu+1}$. The fourth one is (9) of Lemma 3 for $j = \nu$. The last inequality here follows from (12) as $r < \alpha^2 - \alpha + 1 < \alpha^2 - 1$ (because $\alpha > 2$). We see that the conditions of Lemma 4 are satisfies and by Lemma 4 we see that
\[ L_{\nu}|P_{\nu+1}| \gg L_{\nu}X_{\nu}^{\alpha-1}. \]
So in the l.h.s. of (11) the first summand is larger than the second. Now from (11) we have
\[ L_\nu X_\nu^{\alpha-1} \ll L_{k-1}[P_k] + L_k|P_{k-1}|. \] (21)
We apply (17,18) to see that
\[ \max(L_{k-1}|P_k|, L_k|P_{k-1}|) \leq L_{k-1}L_kX_k^r \ll X_k^{r-\alpha}X_{k+1}^{-\alpha} \leq X_k^{r-\alpha^2} \ll X_{\nu+1}^{r-\alpha^2} \ll X_\nu^{(r-\alpha^2)(\alpha-1)}. \] (22)
Here the second inequality comes from (11) for \( j = k - 1 \) and \( j = k \). The third inequality is Lemma 3 with \( j = k \). The fourth one is just \( X_k \geq X_{\nu+1} \). The fifth one is Lemma 3 for \( j = \nu \).
Now from estimates (21,22) and (15) we have
\[ X_\nu^{-\beta_0 + \alpha - 1} \ll X_\nu^{(r-\alpha^2)(\alpha-1)}. \]
This gives
\[ r \geq \alpha^2 + 1 - \frac{\beta_0}{\alpha - 1}. \]
So (19) is proved.

To get (20) we combine the estimate (21) with the left inequality of (22), the bound (15) for \( j = \nu \) and the bound (1) for \( j = k - 1 \). This gives
\[ X_\nu^{\alpha - 1 - \beta_0} \ll L_\nu X_\nu^{\alpha-1} \ll L_{k-1}L_kX_k^r \ll L_kX_k^{r-\alpha}, \]
or
\[ L_k \gg X_k^{\alpha-r}X_k^{\alpha-1-\beta_0}. \]
But \( \beta_0 > \alpha(\alpha - 1) \geq (\alpha - 1) \) by inequality (10) of Lemma 3 and \( X_k \geq X_{\nu+1} \) \( \gg X_\nu^{-\alpha} \) by inequality (9) of Lemma 3. So
\[ L_k \gg X_k^{\alpha-r + \frac{\alpha-1-\beta_0}{\alpha-1}}, \]
and this is the first inequality form (20).
Moreover as \( \beta_0 > \alpha(\alpha - 1) \), from (12) we deduce \( \beta' < \beta \). Lemma is proved. \( \square \)

5. Proof of Theorem 1.
Suppose that \( r \) satisfies (12). We take infinite sequence indices \( \nu_1 < \nu_2 < ... < \nu_i < ... \) such that
- for every \( i = 1, 2, ... \) vectors \( x_{\nu_i-1}, x_{\nu_i}, x_{\nu_i+1} \) are linearly independent;
- for \( i = 1, 2, ... \) vectors \( x_j, \nu_i \leq j \leq \nu_{i+1} \) belong to the two-dimensional lattice \( \Lambda_{\nu_i} = \mathbb{Z}^2 \cap \text{span}(x_{\nu_i}, x_{\nu_i+1}). \)

Now we suppose that three inequalities (16,17,18) hold for all triples \( (\nu, k-1, k) = (\nu_i, \nu_{i+1} - 1, \nu_{i+1}) \) for all \( i \geq 1 \).
Define recursively
\[ \beta_{i+1} = r - \alpha - 1 + \frac{\beta_i}{\alpha - 1}. \]
Then
\[ \beta_i = \alpha(\alpha - 1) + \frac{\beta_0}{(\alpha - 1)^i} \to \alpha(\alpha - 1), \quad i \to \infty. \]
We apply of Lemma 5 to the first \( w \) triple of indices. Then we get (20) for \( k = \nu_{i+1} \), and in particular for \( k = \nu_w \) with \( \beta_w \) close to \( \alpha(\alpha - 1) \). Now we apply Lemma 5 to \( \nu = \nu_w \). In (15) we have \( \beta_w \) instead of \( \beta \). So (19) gives
\[ r \geq \alpha^2 - \alpha + 1 - \frac{\beta_w}{\alpha - 1} \]
We take limit $w \to \infty$ to see that 

$$r \geq \alpha^2 - \alpha + 1.$$ 

This contradicts to \footnote{12}. So there exists $j \in \bigcup_{i=1}^{\infty} \{\nu_i, \nu_{i+1} - 1, \nu_{i+1}\}$ such that $L_j \leq |P_j|X_j^{-r}$. Theorem is proved. □

**Acknowledgement.** The author is grateful to Igor Rochev for important comments on the proof and for pointing out certain inaccuracies in the manuscript.

**References**

[1] Y. Bugeaud, Approximation by algebraic numbers, Cambridge University Press, 2004.

[2] H. Davenport, W.M. Schmidt, Approximation to real numbers by quadratic irrationalities, Acta Arithmetica, 13 (1967), 169 - 176.

[3] H. Davenport, W.M. Schmidt, A theorem on linear forms, Acta Arithmetica, 14 (1968), 209 - 223.

[4] H. Davenport, W.M. Schmidt, Approximation to real numbers by algebraic integers, Acta Arithmetica, 15 (1969), 393 - 416.

[5] V. Jarník Zum Khintcineschen "Ubertragungssats", Travaux de l’Institut Mathematique de Tbilissi, 3, 193 - 216 (1938).

[6] V. Jarník, Une remarque sur les approximation diophantiiennes lineaires, Acta Scientarium Mathem. Szeged, 12 (pars B), 82 - 86 (1949).

[7] B. Ярник, К теории однородных линейных диофантовых приближений, Чехословацкий математический журнал, т. 4 (79), 330 - 353 (1954).

[8] N.G. Moshchevitin, Khintchine’s singular Diophantine systems and their applications, Russian Mathematical Surveys, 2010, 65:3, 433–511.

[9] D. Roy, Approximation simultanée d’un nombre et de son carré. C. R. Acad. Sci., Paris 336 (2003) no 1, 1–6.

[10] D. Roy, On Two Exponents of Approximation Related to a Real Number and Its Square, Canad. J. Math. Vol. 59 (1), 2007, 211 - 224.

[11] W.M. Schmidt, Diophantine Approximations, Lect. Not. Math., 785 (1980).

[12] M. Waldschmidt, Recent advances in Diophantine approximation, Number theory, Analysis and Geometry: In memory of Serge Lang, Springer (2012), 659-704; preprint available at arXiv:0908.3973 (2009).