Approximation to real numbers by algebraic numbers of bounded degree

(Review of existing results)

Vladislav Frank
University Bordeaux 1
ALGANT Master program

May,2007

Contents

1 Approximation by rational numbers.

It seems, that the problem of approximation of given number by numbers of given class was firstly stated by Dirichlet. So, we may call his theorem as "the beginning of diophantine approximation".

Theorem 1.1. *(Dirichlet, 1842) For every irrational number ζ there are infinetely many rational numbers $\frac{p}{q}$, such that*

$$0 < \left|\zeta - \frac{p}{q}\right| < \frac{1}{q^2}.$$

Proof. Take a natural number N and consider numbers $\{q\zeta\}$ for all q, $1 \leq q \leq N$. They all are in the interval $(0,1)$, hence, there are two of them with distance not exceeding $\frac{1}{q}$. Denote the corresponding q's as q_1 and q_2. So, we know, that there are integers $p_1, p_2 \leq N$ such that $|(q_2\zeta - p_2) - (q_1\zeta - p_1)| < \frac{1}{N}$. Hence, for $q = q_2 - q_1$ and $p = p_2 - p_1$ we have $|q\zeta - p| < \frac{1}{N}$. Division by q gives $\left|\zeta - \frac{p}{q}\right| < \frac{1}{qN} \leq \frac{1}{q^2}$. So, for every N we have an approximation with precision $\frac{1}{qN} < \frac{1}{N}$. Due to irrationality of ζ every rational number may satisfy this bound only for finitely many N. Hence, the inequality has infinitely many solutions.

Remark 1.1.1. It is possible to give slightly better bound: $\left|\zeta - \frac{p}{q}\right| < \frac{c}{q^2}$, here $c = \frac{1}{\sqrt{5}} \approx 0.4472$. This may be proved using countinious fractions. This c is the exact bound and cannot be inproved for ζ equal to the golden ratio $\frac{1+\sqrt{5}}{2} \approx 1.618$. Those statements are due to Hurvitz. See [1, 2] for details.

Remark 1.1.2. The irrationality of ζ is neccecary: if $\zeta = x/y$, and p/q is an approximation, we obtain $\left|\frac{x}{y} - \frac{p}{q}\right| = \left|\frac{xq-yp}{yq}\right| \geq \left|\frac{1}{yq}\right| \geq \frac{1}{q^2}$ for all denominators of approximations except the finite number of them. But surely there are only finitely many good approximations with bounded denominator!

Of course, some numbers may have much better approximations. This was used by Liouville to prove existence of transcendental numbers.

Theorem 1.2. *(Liouville, 1844). Let α be an algebraic number of degree $d \geq 1$. Then there is a constant $c = c(\alpha) > 0$ such that*

$$\left|\alpha - \frac{p}{q}\right| > \frac{c}{q^d} \quad \textit{for every choice of integers } p \textit{ and } q.$$

Proof. Suppose, that $\left|\alpha - \frac{p}{q}\right| \leq 1$, in other case we may change p and decrease the absolute value of the difference for the same q.

Let $F(x) = a_0(x - \alpha_1)\dots(x - \alpha_d)$ be the irreducible (over $\mathbb{Q}$) polynomial with root $\alpha = \alpha_1$. Then $F(\frac{p}{q})$ is non-zero fraction with denominator q^n, whence $\left|F(\frac{p}{q})\right| \geq \frac{1}{q^d}$. On the other hand, we have

$\left|\alpha_i - \frac{p}{q}\right| = \left|\alpha_i - \alpha + \alpha - \frac{p}{q}\right| \leq |\alpha_i - \alpha| + \left|\alpha - \frac{p}{q}\right| \leq \max_{i \neq 1} |\alpha_i - \alpha| + 1 = C.$
This implies

$$\left|\alpha - \frac{p}{q}\right| = \frac{F(p/q)}{|a_0| \left|\alpha_2 - \frac{p}{q}\right| \ldots \left|\alpha_d - \frac{p}{q}\right| \ldots} \geq \frac{1/q^n}{|a_0| \, C^{d-1}}.$$

So, we may take $c = \frac{1}{|a_0| C^{d-1}}$ and we are done.

Corollary 1.2.1. *There exists transcendental (non-algebraic) number.*

Proof. Take the number $\alpha = \frac{1}{2^1} + \frac{1}{2^2} + \frac{1}{2^6} + \ldots + \frac{1}{2^{n!}} + \ldots$. This series converges and defines a real number. Denote by a_n the sum of first n summands of this series. It is rational number with denominator $q = 2^{n!}$. The remainder of the series may be bounded as $\sum\limits_{i=n+1}^{\infty} \frac{1}{2^{i!}} < \sum\limits_{i=(n+1)!}^{\infty} \frac{1}{2^i} = \frac{2}{2^{(n+1)!}} = \frac{2}{q^{n+1}}$. On the other hand, if α is algebraic of degree d, we must have $\left|\alpha - \frac{p}{q}\right| > \frac{c}{q^d}$ for some c. Now choosing $n \geq d$ such that $\frac{2}{q} < c$ leads to contradiction.

Remark 1.2.1. In 1871 Kantor proved (using the set theory just invented) that the set of algebraic numbers is countable, but the set of transcendental numbers is uncountable. Hence in fact almost all real numbers are transcendental. See [3] for details.

The next direction of improving was the exponent d in Liouville's theorem. We state four theorems without proofs, which may be found in [2]. Those statements will be as follows: for every algebraic number α of degree d, every positive ε and any constant c there are only finitely many solutions of the inequality

$$\left|\alpha - \frac{p}{q}\right| < \frac{c}{q^{k+\varepsilon}}$$

and will differ only in the value of k. Note, that Liouville's theorem gives $k = d$.

It is easy to see, that finiteness of the number of solutions does not depend on the value of c

Theorem 1.3. *(Thue Theorem, 1909)* $k = \frac{n}{2} + 1$.

Theorem 1.4. *(Siegel Theorem, 1921)* $k = 2\sqrt{n}$.

Theorem 1.5. *(Dyson Theorem, 1947)* $k = \sqrt{2n}$.

Theorem 1.6. *(Roth Theorem, 1955)* $k = 2$.

Remark 1.6.1. We have seen, that for some algebraic numbers α and some c even the inequality $\left|\zeta - \frac{p}{q}\right| > \frac{c}{q^2}$ may always hold, but now there is no known example of such α other, than quadratic irrationalities.

Remark 1.6.2. If for some c we have finitely many solutions, then we may decrease it and find c_1 (depending on α and ε) such that the inequality $\left|\alpha - \frac{p}{q}\right| < \frac{c_1}{q^{2+\varepsilon}}$ has no solutions. We have seen, that Liouville theorem gives us a possible value of such c_1. On the other hand, even Thue theorem is ineffective: there is no known way to determine such c in general.

Remark 1.6.3. Lets state one of effective results (Bennett 1995): the inequality $\left|\sqrt[3]{2} - \frac{p}{q}\right| < \frac{1}{4\sqrt[6]{32}q^{2.5}}$ has no integer solutions.

2 Wirsing conjecture and Wirsing theorem

From now on and until the end of the memoir we will use the following notations:

$f \ll g$ iff for some $c = c(\zeta, d)$ we have $f \le cg$.

$f \sim g$ iff $f \ll g$ and $g \ll f$.

All implied constants may depend on ζ and d.

Lets reformulate Dirichlet theorem:

Theorem 2.1. *(Dirichlet, revisited) For every real number ζ, which is not an algebraic number of degree 1, there are infinitely many algebraic numbers of degree at most 1 (denoted by α), such that*

$$|\zeta - \alpha| < cH(\alpha)^{-2}$$

for some positive $c = c(\zeta)$. Here H denotes the height of algebraic number, i.e. maximum of absolute values of its minimal integer polynomial coefficients.

It's quite clear, that this theorem is the same, that the original one. Algebraic numbers of degree at most one are rational numbers, their height may be evaluated as $H(p/q) = \max(p, q)$ if p and q are coprime, and if $\zeta - p/q$ has small absolute value, then $q\zeta \approx p$ and $q^{-2} \le H(p/q)^{-2} \le \min(1, 4\zeta^{-2})q^{-2}$ for good approximations, so it doesn't matter what to write - q^{-2} or $H(p/q)^{-2}$.

So, now it seems logical to generalise it from degree one to higher degrees.

Wirsing Conjecture, 1960

For every real number ζ, which is not an algebraic number of degree at most d, there are infinitely many algebraic numbers α of degree at most d, such that

$$|\zeta - \alpha| < cH(\alpha)^{-f(d)}$$

for some positive $c = c(\zeta, n)$ and $f(d) = d + 1$.

This conjecture still unproved now.

One of first result was a theorem by Wirsing himself, for which we will give a sketch of proof here.

Theorem 2.2. *(Wirsing, 1961) Wirsing conjecture is true for all d and $f(d) = (d + 3)/2 - \varepsilon$.*

Remark 2.2.1. In fact, Wirsing obtained a slightly better result. He showed, that $f(d)$ may be taken equal to $\frac{d+6}{4} + \frac{1}{4}\sqrt{d^2+4d-4} - \varepsilon \approx \frac{d}{2} + 2$ for d large enough. See the next chapter or [5] for further details.

Proof. (Sketch) Let ζ, d be fixed. For every polynomial $P(x) = b_0 \prod_{i=1}^{k}(x-\beta_i)$ of degree $k = \deg(P) \le d$ we define

$$p_i = |\zeta - \beta_i|$$

and

$$M_\zeta(P) = b_0 \prod_{i=1}^{k} \max(1, p_i).$$

Obviously, $M_\zeta(P) = M(P(x+\zeta))$, where $M(P)$ is the Mahler measure of P. Then $M_\zeta(P) \sim H(P)$ and $|P(\zeta)| \sim H(P) \prod_{i=1}^{k} \min(1, p_i)$.

Lemma 2.2.1. *Let $P(x) = b_0 \prod_{i=1}^{k}(x-\beta_i)$ and $Q(x) = b_0 \prod_{j=1}^{l}(x-\gamma_j)$ be coprime polynomials, $1 \le k, l \le d$.*

Putting $p_i = |\zeta - \beta_i|$ and $q_j = |\zeta - \gamma_j|$, assume that $p_1 \le p_2 \le \ldots \le p_k$ and $q_1 \le q_2 \le \ldots \le q_l$. Assume also that $p_1 \le q_1 \le 1$.

Then one of the following holds:

1. $p_1 \ll |P(\zeta)|\, H(P)^{-1}$ *and* $\beta_1 \in \mathbb{R}$,
2. $q_1 \ll |Q(\zeta)|\, H(Q)^{-1}$ *and* $\gamma_1 \in \mathbb{R}$,
3. $1 \ll |Q(\zeta)|^2\, H(P)^d H(Q)^{d-2}$,
4. $p_1^2 \ll |P(\zeta)|\, |Q(\zeta)|^2\, H(P)^{d-1} H(Q)^{d-2}$ *and* $\beta_1 \in \mathbb{R}$,
5. $p_1^2 \ll |P(\zeta)|^2\, |Q(\zeta)|\, H(P)^{d-2} H(Q)^{d-1}$ *and* $\beta_1 \in \mathbb{R}$.

The proof is straightforward and uses only lower and upper bounds for the resultant of P and Q.

The proof of the Wirsing theorem may now be finished as follows: due to the theorem of Minkowski there are infinitely many polynomials P, such that $|P(\zeta)| \ll H(P)^{-d}$. Then, of course, this estimate is true for at least one irreducible factor of P. Taking in every case the factor with the maximal height, we may deduce, that we have infinitely many irreducible polynomials with such property.

If some polynomial $R(x)$ is divisible by $P(x)$, then we know, that $H(R) \gg H(P)$. Assume, for example, that $H(R) > c_1 H(P)$ in that case and take the symmetric convex body in $\mathbb{R}^{d+1}$:

$$|r_0|, |r_1|, \ldots, |r_d| \le \frac{c_1}{2} H(P), \left|r_0\zeta^d + \ldots + r_d\right| < c_2 H(P)^{-d}.$$

If $c_2 c_1^d$ is large enough, then there is a non-zero integer point in it, hence, there is a polynomial Q, such that $H(Q) \ll H(P)$, and $Q(\zeta) \ll H(P)^{-d}$. This polynomial is not divisible by P because absolute values of all its coefficients do not exceed $\frac{c_1}{2}H(P)$, hence $H(Q) < c_1 H(P)$. Hence P and Q are coprime (P is irreducible).

Now we may use the previous lemma for them. Assume, for example, that (4) holds. Then we obtain:

$|\zeta - \beta_1| \ll H(P)^{-d/2} H(P)^{-d} H(P)^{(d-1)/2} H(P)^{(d-2)/2} = H(P)^{-(d+3)/2} \ll H(\beta_1)^{-(d+3)/2}$.

Case (5) gives the same, cases (1) and (2) give even better bounds (with the exponent $d+1$, as in the original hypothesis), and case 3 leads to a contradiction. See [2] for further details.

3 Mahler and Koksma functions and original Wirsing idea

Lets introduce two new definitions.

Definition. Let ζ be real number. *Mahler function* $\omega_d(\zeta)$ is the supremum of all real numbers ω, for which there are infinitely many polynomials $P(x)$ (of degree at most d, with integer coefficients)satisfying the inequality

$$0 < |P(\zeta)| \ll H(P)^{-\omega}.$$

Remark 3.0.1. The Minkowski theorem gives $\omega_d(\zeta) \geq d$ if ζ is not an algebraic number of degree at most d. On the other hand, it is known, that for almost all (in sence of Lebesgue measure) ζ we have $\omega_d(\zeta) = d$ (Sprindzuk, 1963, see [10]).

Definition. Let ζ be real number. *Koksma function* $\omega_d^*(\zeta)$ is the supremum of all real numbers ω, for which there are infinitely real algebraic numbers α of degree at most d satisfying the inequality

$$0 < |\alpha - \zeta| \ll H(\alpha)^{-\omega-1}.$$

We will write ω and ω^* (without arguments) instead of $\omega_d(\zeta)$ and $\omega_d^*(\zeta)$ in this section. Moreover, we suppose that ζ is transcendental.

Looking through arguments in Wirsing proof we see, that in fact we proved $\omega^* \geq \frac{1}{2}(\omega + 1)$. To see it we need only change d to $\omega_d(\zeta)$ in all exponents.

Theorem 3.1. $\omega^* \geq \frac{\omega}{\omega - d + 1}$.

Proof. Choose pairwise distinct real numbers $\zeta_0 = \zeta, \zeta_1, \ldots, \zeta_d$.

Let $\gamma = \omega(1+\varepsilon)^2 > d$. For every H we may find (due to the Minkowski theorem) an integer polynomial $P(x)$, such that

$$\begin{cases} |P(\zeta)| \ll H^{-\gamma}, \\ |P(\zeta_i)| \ll H \quad for\ i = 1, \ldots, d-1, \\ |P(\zeta_d)| \ll H^{\gamma - d + 1}. \end{cases}$$

Notice, that $P(\zeta) \neq 0$.

From this system we may find coefficients of P. Because the determinant of the corresponding system is a non-zero constant, we have

$$H(P) \ll \max_{i=0}^{d} P(\zeta_i) \ll H^{\gamma-d+1}.$$

On the other hand $H^{-\omega(1+\varepsilon)^2} \gg |P(\zeta)| \gg H(P)^{\omega(1+\varepsilon)}$ (the latter is true due to definition of ω if $H$is large enough). Hence $H(P) \gg H^{1+\varepsilon}$.

Lemma 3.1.1. *Let A be positive real number. Let $P(x) = a_0x^s + \dots$ be integral polynomial with roots $\alpha_1, \dots, \alpha_s$. Then*

$$\prod_{|\zeta-\alpha_i|>A} |\zeta - \alpha_i| \sim \frac{H(P)}{a_0}$$

and

$$\prod_{|\zeta-\alpha_i|<A} |\zeta - \alpha_i| \sim \frac{|P(\zeta)|}{H(P)}.$$

(Constants in $\sim$ may depend on A)

Proof. The first part after suitable linear transformation becomes the statement, that Mahler measure and height of polynomial are equivalent norms, which is well-known. The second part follows from the first obviously.

Now choose such small A, that $|\zeta_i - \zeta_j| > 2A$ for $i \neq j$. Then we have

$$\prod_{|\zeta_j-\alpha_i|<A} |\zeta_j - \alpha_i| \sim \frac{|P(\zeta_j)|}{H(P)} \ll \frac{H}{H^{1+\varepsilon}} = H^{-\varepsilon}.$$

Hence, every small circle with center in ζ_j, $j = 0..d-1$, contains a root of $P(x)$. Hence, every such circle contains exactly one root. This gives for the nearest to ζ root α of $P(x)$

$$|\zeta - \alpha| \sim \frac{|P(\zeta)|}{H(P)} \ll H(P)^{-1}H^{-\gamma} \ll H(P)^{-1-\frac{\gamma}{\gamma-n+1}} \ll H(\alpha)^{-1-\frac{\gamma}{\gamma-n+1}}.$$

So, we are done.

Now we have two estimates for ω^*. The first one becomes bigger while ω increases, the second one – becomes smaller. Hence, the minimal value of $\max(\frac{1}{2}(\omega+1), \frac{\omega}{\omega-d+1})$ occurs then both expressions are equal.

This gives

$$(\omega + 1)(\omega - d + 1) = 2\omega$$

which has a root $\omega = \frac{d+\sqrt{d^2+4d-4}}{2}$. This leads to

$$\omega^* \geq \frac{1}{2}(\omega + 1) = \frac{d + 2 + \sqrt{d^2 + 4d - 4}}{4}.$$

So, we may take $f(d) = \frac{d+6+\sqrt{d^2+4d-4}}{4} - \varepsilon$, as it was announced in the previous section.

Remark 3.1.1. If $\omega = d$, we obtain $\omega^* \geq \frac{d}{d-d+1} = d$, so in this case the original Wirsing conjecture is proven. Due to the Sprindzhuk theorem (see remark 3.0.1) Wirsing conjecture is proven for almost all numbers.

4 Davenport-Schmidt method for the case $d = 2$

In 1967 H.Davenport and W.M.Schmidt wrote a paper about Wirsing conjecture for $d = 2$, i.e. approximations by quadratic irrationals. Here we give a short sketch of their ideas, and we send reader to [4] for complete proof.

Theorem 4.1. *(Davenport-Schmidt,1967)*
For every real number ζ, which is neither rational nor quadratic irrationality, there are infinitely many real quadratic irrationals α, which satisfy

$$|\zeta - \alpha| \leq CH(\alpha)^{-3},$$

where $C > \frac{160}{9}\max(1, |\zeta|^2)$.

Remark 4.1.1. There is no special sence of $\frac{160}{9}$, it may be reduced. Exact value (as $1/\sqrt{5}$ for rationals) is unknown.

Remark 4.1.2. The proof doesn't show an algorithm of constructing such good approximations (like continued fractions for rationals). It is only an existence proof.

Proof. (Sketch) We restrict ourselves to case $0 < \zeta < 1$. Suppose, that for sufficiently large $H(\alpha)$ and some $C_1 > 160/9$ we always have $|\zeta - \alpha| > C_1 H(\alpha)^{-3}$.

Let $\mathbf{x} = (x, y, z)$ be a vector in $\mathbb{Z}^3$ with $||\cdot||_\infty$-norm. We consider two linear forms:

$P(\mathbf{x}) = 2\zeta x + y$ and $L(\mathbf{x}) = \zeta^2 x + \zeta y + z$.

Lemma 4.1.1. *There is a number $C_2 < 9/160$, such that if $gcd(x, y, z) = 1$ and* $\max(|x|, |y|, |z|)$ *is sufficiently large, then* $|P(\mathbf{x})| < C_2 |\mathbf{x}|^3 L(\mathbf{x})$.

Note, that technical proof of this lemma uses the exact expressions for P and L very much.

After this lemma we introduce a *sequence of minimal numbers*, that is the sequence of $\mathbf{y_i}$, such that $L(\mathbf{y_i})$ is the smallest positive value of L on the cube $|\mathbf{x}| \leq i$. The sequence $L(\mathbf{y_i})$ is decreasing. And we select its maximal subsequence, which contains every value only once. That gives us new sequence $\{L\mathbf{x_i}\}$, the corresponding sequence of points $\{\mathbf{x_i}\}$ and a sequence of integers X_i ($\mathbf{x}$ is minimum for $|\mathbf{x}| \leq X$ if $X_i \leq X < X_{i+1}$).

Denote $L_i = L(\mathbf{x})$ and $P_i = P(\mathbf{x})$.

Due to the lemma, for sufficiently large numbers i we obtain $|P_i| < C_2 X_i^3 L_i$. One more inequality, $L_i < \frac{4}{3} X_{i+1}^{-2}$, may be obtained from the Minkowski theorem. It leads, in particular, to $x_i \neq 0$, otherwise $\frac{-z_i}{y_i}$ is very good approximation to ζ.

Lemma 4.1.2. *If three consecutive terms*

$$\mathbf{x_{i-1}}, \mathbf{x_i}, \mathbf{x_{i+1}}$$

are linearly dependent, then for at least one sign we have

$$\mathbf{x_{i-1}} \pm \mathbf{x_{i-1}} = u\mathbf{x_i}$$

with integral u.

Lemma 4.1.3. *There are infinitely many* i *such that* $\mathbf{x_{i-1}}$,$\mathbf{x_i}$,$\mathbf{x_{i+1}}$ *are linearly independent.*

Proof. Assume they are linearly dependent for $i > N$. From previous lemma we obtain $x_{i-1}L(\mathbf{x_i}) - x_iL(\mathbf{x_{i-1}}) = \pm(x_iL(\mathbf{x_{i+1}}) - x_{i+1}L(\mathbf{x_i}))$, which gives us for every $m > n > N$

$$\begin{aligned}|x_nL(\mathbf{x_{n+1}}) - x_{n+1}L(\mathbf{x_n})| &= |x_{m-1}L(\mathbf{x_m}) - x_mL(\mathbf{x_{m-1}})| \\ &< \frac{4}{3}X_{m-1}X_{m+1}^{-2} + \frac{4}{3}X_mX_m^{-2} \to 0\end{aligned}$$

So, $L(x_n\mathbf{x_{n+1}} - x_{n+1}\mathbf{x_n}) = 0$, and all minimums are on the same line, which is impossible.

Lemma 4.1.4. *For linearly independent points* $\mathbf{x_{i-1}}$,$\mathbf{x_i}$,$\mathbf{x_{i+1}}$ *we have* $P_{n+1}L_nX_{n-1} > \frac{1}{2} - \frac{32}{9}C_2$

To prove it, we need to look at the determinant

$$\begin{vmatrix} x_{n-1} & P_{n-1} & L_{n-1} \\ x_n & P_n & L_n \\ x_{n+1} & P_{n+1} & L_{n+1} \end{vmatrix}$$

which is nonzero integer and estimate it.

Now we may prove the theorem. Lets take a large m such that

$$\mathbf{x_{m-1}}, \mathbf{x_m}, \mathbf{x_{m+1}}$$

are linearly independent and $n > m$ - the minimal number, such that

$$\mathbf{x_{n-1}}, \mathbf{x_n}, \mathbf{x_{n+1}}$$

are linearly independent too. Using identities from lemma 4 we obtain

$$\begin{aligned}|P_mL_{m+1} - P_{m+1}L_m| &= |P_{n-1}L_n - L_{n-1}P_n| \\ &< C_2X_{n-1}^3L_{n-1}L_n + C_2X_n^3L_nL_{n-1} < \frac{8}{3}C_2X_nL_n.\end{aligned}$$

But also we have

$$|P_m L_{m+1} - P_{m+1} L_m| > (\frac{1}{2} - \frac{32}{9} C_2) X_m^{-1} L_{m-1}^{-1} L_m - \frac{4}{3} C_2 X_m L_m$$
$$> (\frac{3}{8} - \frac{8}{3} C_2) X_m L_m - \frac{4}{3} C_2 X_m L_m = (\frac{3}{8} - 4C_2) X_m L_m.$$

This gives $X_m L_m < X_n L_n$, so we have an increasing subsequence of $X_i L_i$, which in fact converges to zero — a contradiction.

5 Linear forms and the subspace theorem

One of the most useful ideas for proving the Wirsing conjecture is the following one:

Let α be a good approximation for ζ and let $P(x) = a_0(x - \alpha_1) \dots (x - \alpha_k)$ be a minimal polynomial of $\alpha = \alpha_1$, $k \le d$. Moreover, suppose that α_1 is the best approximation to ζ among α_i. Then we have

$$\left| \frac{P'(\zeta)}{P(\zeta)} \right| = \sum_{i=1}^{k} \frac{1}{|\zeta - \alpha_i|} \sim \frac{1}{\zeta - \alpha}$$

so, to produce good approximations we need to construct polynomials with small ratio $\frac{P(\zeta)}{P'(\zeta)}$.

A very powerful tool to study linear forms is

Theorem 5.1. *(Schmidt subspace theorem, ????) Let $L_1, L_2, \dots, L_n$ be linearly independent linear forms in $\mathbb{R}^n$ with algebraic coefficients. Then the set of integral solutions of the inequality*

$$|L_1(v) L_2(v) \dots L_n(v)| < c \, ||v||^{-\varepsilon}$$

for every c and every $\varepsilon > 0$ is contained in finite number of hyperplanes.

Remark 5.1.1. The proof is unconstructive, but Voita in 1989 proved, that all solutions are contained in $S_1 \cup S_2 \cup \dots \cup S_l \cup F$, where hyperplanes S_i may be expressed constructively, l and S_i does not depend on ε and F depends on ε but is finite (and now cannot be expressed constructively).

Exapmple. Let

$$L_1 = x_1 + \sqrt{2} x_2 + \sqrt{3} x_3,$$
$$L_2 = x_1 - \sqrt{2} x_2 + \sqrt{3} x_3,$$
$$L_3 = x_1 - \sqrt{2} x_2 - \sqrt{3} x_3.$$

Take, for example, $x_3 = 0$ and $x_1^2 - 2x_2^2 = 1$ (there are infinitely many such positive x_1 and x_2 - solutions of Pell equation). Then

$$|(L_1 \cdot L_2 \cdot L_3)(x_1, x_2, x_3)| = \left| x_1 - \sqrt{2} x_2 \right| = \frac{1}{|x_1 + \sqrt{x_2}|} < \max(x_1, x_2, x_3)^{-1}.$$

So, we have infinitely many solutions in the hyperplane $x_3 = 0$ for $c = 1$, $\varepsilon = 1$. The same may be done for $x_1 = 0$ or $x_2 = 0$. On the other hand, it is possible to show, that only finitely many solutions satisfy $x_1 x_2 x_3 \neq 0$.

Using of Schmidt subspace theorem allows us to prove Wirsing conjecture for algebraic numbers ζ and $f(d) = d + 1 - \varepsilon$.

Theorem 5.2. *Let ζ be algebraic number of degree at least $d+1$, $c > 0$ and $\varepsilon > 0$ be given numbers. Then there are infinitely many algebraic numbers α of degree at most d, such that $|\zeta - \alpha| < cH(\alpha)^{-d-1+\varepsilon}$*

Proof. Denote $|x| = \max(x_0, \ldots, x_d)$ - the norm on $\mathbb{R}^{d+1}$. Consider linear form $L(x) = \zeta^d x_d + \ldots + x_0$. We know from Minkowski theorem, that there are infinitely many integer solutions of $|L(x)| < c_1 |x|^{-d}$ for some c_1. Hence, we need only to prove, that there are only finitely many solutions of $|L'(x)| = \left|d\zeta^{d-1}x_d + \ldots + d_1\right| < \frac{nc_1}{c} |x|^{1-\varepsilon}$ among them (due to remark in the beginning of this section, in all other cases we will have

$$|\zeta - \alpha| < \frac{nP(\zeta)}{P'(\zeta)} < c_1 |x|^{-d} \frac{c}{nc_1} |x|^{-1+\varepsilon} = c |x|^{-d-1+\varepsilon} \leq cH(\alpha)^{-d-1+\varepsilon}$$

for α the nearest to ζ root of $x_d t^d + \ldots x_0 = 0$).

Take $d+1$ linear forms: $L(x), L'(x)$ and first $d-1$ coordinates (denote as $L_d, \ldots L_2$ corresponding linear forms). They are linearly independent linear forms with algebraic coefficients (here we use, that ζ is algebraic), so Schmidt subspace theorem may be used for them. Because all solutions of the system $L(x) < c_1 |x|^{-d}, L'(x) < \frac{nc_1}{c} |x|^{1-\varepsilon}, L_d(x) \leq |x|, \ldots L_2(x) \leq |x|$ satisfy $|(L \cdot L' \cdot L_d \cdot \ldots \cdot L_2)| < C |x|^{-\varepsilon}$ it gives, that they are contained in finite number of hyperplanes. Now we will prove, that every hyperplane may contain only finite number of solutions even of the first inequality - $|L(x)| < c_1 |x|^{-d}$.

Lemma 5.2.1. *For every linear form $L(x) = \beta_0 x_0 + \ldots + \beta_d x_d$, every constant c and every hyperplane in $\mathbb{R}^{d+1}$ there are only finitely many solutions of $|L(x)| < c|x|^{-d}$ in that hyperplane, if L vanishes on integer points only at zero (in other words, β_i are linearly independent over $\mathbb{Q}$).*

Proof. On our hyperplane we have a discrete subgroup of $\mathbb{Z}^{d+1}$. We may suppose, that its rank is exactly d (in other case we take some new elements to this subgroup and then span hyperplane on them - it will only increase the number of solutions). So, there is a basis, consisting of d linearly independent vectors. It defines a norm on the hyperplane, which is equivalent to the induced (from $\mathbb{R}^{d+1}$) norm, because every two norms are equivalent. Hence we may suppose, that our hyperplane is a horizontal one (we change coordinates that way), and we have $L_1(x) < c_1 |x|_1^{-d} < c_1 |x|_1^{-d-1}$ for new norm and new linear form.

Using of Schmidt theorem gives, that all solutions are contained in finitely many hyperplanes (now they will have dimension $d-1$), and we need only to prove, that every such hyperplane contains only finitely many solutions. So, we have the same problem for dimension d and hyperplane of dimension $d-1$. Induction finishes the proof. The statement for $d=1$ is "$|\beta_0 x_0| < c_1$ has finitely many solutions", which is obvious.

So, the lemma is proved and the theorem too.

6 Hopeless approach and Schmidt counterexample

The next logical question must be "is it possible to improve Schmidt theorem to make it true for forms with arbitrary given coefficients"? If it would be possible, we may use exactly the same way to prove Wirsing conjecture for $f(d) = d+1-\varepsilon$ and every (not only algebraic) ζ. Unfortunately, the answer is "NO".

First obvious reason to believe in it is that Schmidt theorem for dimension 2 is equivalent to Roth theorem. But we know, that Roth theorem cannot be proved for transcendental numbers - it is wrong for Liouville counterexample.

Moreover, Schmidt in [6] gave a counterexample, and showed, that quotient of linear forms (in any number of variables) MAY be always large. So, every proof of Wirsing conjecture must use specific forms (coming from polynomial and its derivative).

Theorem 6.1. *(Davenport, Schmidt, 1968)*

Let $m \geq 1$, $n \geq m+2$ and let $L, P_1, \dots P_m$ be linearly independent forms in $\mathbf{x} = (\mathbf{x_1}, \dots \mathbf{x_n})$. Then there are infinitely many integer solutions of the inequality

$$|L(\mathbf{x})| \leq c(L, P_1, \dots, P_m) \max(|P_1(\mathbf{x})|, \dots, |P_m(\mathbf{x})|) \, |\mathbf{x}|^{-m-2}.$$

On the other hand, there are linearly independent forms $L, P_1, \dots, P_m$ such that for every $\varepsilon > 0$ and every integer point $(\mathbf{x})$ we have

$$|L(\mathbf{x})| \geq c(\varepsilon) \max(|P_1(\mathbf{x})|, \dots, |P_m(\mathbf{x})|) \, |\mathbf{x}|^{-m-2-\varepsilon}.$$

This theorem shows, that we may take $f(d) = 3-\varepsilon$ for all d in Wirsing conjecture, which is not very interesting for $d > 2$.

The proof may be found in [6].

7 Modern approach to Wirsing conjecture

Since 1993 some new results in Wirsing conjecture were obtained by byelorussian mathematician K.Tishenko. In his papers [11, 12, 13, 14] he proved many improvements in Wirsing conjecture, and the main part of his papers is about small values of d. Lets give at first the short review of his results.

Theorem 7.1. *For $d \geq 11$ function $f(d)$ in Wirsing conjecture may be taken equal to $\frac{d}{2} + \gamma_d$ where $\gamma_d \to 4$ if $d \to \infty$. More precisely, $f(d)$ may be taken equal to maximal real root of polynomial $2X^4 - (d+10)X^3 + (d+20)X^2 + (3d-20)X + 6 = 0$*

All other theorems improve results for small values of d. We suppose also, that ζ is transcendental number.

Theorem 7.2. *For $d \leq 10$ we may take $f(d)$ equal to the maximal real root of $(3d-5)X^2 - (2d^2+d-9)X - d - 3 = 0$.*

Theorem 7.3. *For $d \leq 5$ we may take $f(d)$ equal to the maximal real root of $(5d-9)X^3 - (3d^2+5d-22)X^2 + (d^2-7d+20)X + d^2 - 2d + 5 = 0$.*

Theorem 7.4. *For $d = 3$ we may take $f(d)$ equal to the maximal real root of $4x^5 - 30X^4 + 72X^3 - 70X^2 + 43X - 4 = 0$.*

Remark 7.4.1. For $d = 3$ every next theorem gives better bound: approximately $3.43, 3.60, 3.73$. Surely, those results still far enough from 4 which is Wirsing conjecture.

Here we give only the main guideline of work [11] and show where it may be improved to obtain the result of 7.3. Moreover, we restrict ourselves to the case $d = 3$. For proof details see original works [11, 12, 13] if you have nothing else to do. All polynomials in this chapter have integer coefficients and degree at most 3, if other is not stated directly.

Suppose,that there exists number ζ (not algebraic of degree 3 or less) such that

$\forall c > 0 \exists H_0 > 0 \forall \alpha \in \mathbb{A}, \deg(\alpha) \leq 3, H(\alpha) \geq H_0 \, |\zeta - \alpha| > cH(\alpha)^{-A}$.

(The exact value of A will be determined during the proof).

This gives, that for every $Q(x), H(Q) \geq H_0$ the inequality

$$\frac{|Q(\zeta)|}{|Q'(\zeta)|} > C_T H(Q)^{-A}$$

holds. Here $C_T = 4^9 \cdot 6^{108} \cdot \zeta^{-972}$ but we will not think about constants very much until the end of proof. Every numerated constant will be defined when introduced.

We may suppose that $0 < \zeta < 1/4$.

We start with some preparations.

Lemma 7.2.1. *Let $L(x) = c_3x^3 + \ldots + c_0, |L(\zeta)| < 1/2$. Then $H(L) = c_i$ for some $i \neq 0$.*

Lemma 7.2.2. *Let $L(x) = c_3x^3 + \ldots + c_0, |L(\zeta)| < 1/2$. Suppose, that for two different indexes i, j we have $|c_i|, |c_j| < \zeta^2 H(L)$. Then $H(L) < \zeta^{-2} |L'(\zeta)|$.*

Lemma 7.2.3. *For integer polynomials $P(x)$ and $Q(x)$ without common roots, $\deg(P) = m, \deg(Q) = n, 2 \leq m, n \leq 3$ one of the following holds:*

1) $1 < C_R \max(|P(\zeta)|, |Q(\zeta)|)^2 \max(H(P), H(Q))^{m+n-2})$,
2) $1 < C_R \max(|P(\zeta)P'(\zeta)Q'(dz)|, \left|Q(\zeta)P'(\zeta)^2\right|)H(P)^{n-2}H(Q)^{m-1})$,
3) $1 < C_R \max(|Q(\zeta)P'(\zeta)Q'(dz)|, \left|P(\zeta)Q'(\zeta)^2\right|)H(P)^{n-1}H(Q)^{m-2})$.
Constant C_R depends only on ζ.

For proof we need only consider the resultant of P and Q. C_R may be taken equal to $6! \cdot 4^6$.

Lemma 7.2.4. *Let $G(x) = G_1(x) \dots G_v(x)$ and denote* $\deg(G) = l$. *Then $e^{-l}H(G_1) \dots H(G_v) \leq H(G) \leq (l+1)^{v-1}H(G_1) \dots H(G_v)$.*

Lemma 7.2.5. *Let $G_1(x), G_2(x)$ be integer polynomials of degree $\leq l$ and $H(G_2) < e^{-l}H(G_1)$. Then G_1, G_2 have no common roots.*

Lemma 7.2.4 is well-known, and lemma 7.2.5 is easy consequence of it.

Lemma 7.2.6. *Consider the following system of inequalities:*

$$\begin{cases} |a_{11}x_1 + a_{12}x_2 + a_{13}x_3| \leq A_1 \\ |a_{21}x_1 + a_{22}x_2 + a_{23}x_3| \leq A_2 \\ |a_{31}x_1 + a_{32}x_2 + a_{33}x_3| \leq A_3 \end{cases}$$

with the following properties of coefficients:
1) $\forall j \quad |a_{2j}|, |a_{3j}| \leq B_j$ *and* $B_1, B_2 > B_3 > 0$.
2) $|a_{11}|, |a_{21}| \leq |a_{31}|$ *and* $a_{31} \neq 0$.
3) $A_2 \leq A_3$.
4) $|\Delta| \geq cB_1B_2|a_{1n}| > 0$ *(here Δ is the determinant of matrix $A = (a_{ij})$).*
Then for every solution $(x_1, x_2, x_3)$of this system we have
$|x_i| < \frac{6}{c}B_i^{-1} \max(\frac{A_1B_3}{|a_{1n}|}, A_3)$.

Proof follows from the Cramer formula after some computations.

Now we start construction of polynomials, that will be used in proof. We will need two sequences of polynomials.

Fix $h > H > (6!)^{15} \cdot 4^{90} \cdot e^{540}H_0$. Take as $\tilde{P}_0(x)$ the polynomial of degree at most 3 and of height at most h with minimal absolute value at ζ. Then take polynomial $\tilde{P}_1(x)$ of minimal height such that $\tilde{P}_1(\zeta) < c_0^{-1}\left|\tilde{P}_0(\zeta)\right|$ and so on. We obtain sequence of polynomials, whose heights are increasing and values at ζ decrease at least as geometric progression. Finally, we renumerate $\tilde{P}_i$, starting from i such that $H(\tilde{P}_i)$ is large enough: if $H(\tilde{P}_k) \leq H_0 < H(\tilde{P}_{k+1})$, then we denote $P_i(x) = \tilde{P}_{k+i}(x)$. Here $c_0 = 6\zeta^{-9}$.

Lemma 7.2.7. $\left|\tilde{P}_i(\zeta) < H(\tilde{P}_{i+1})\right|^{-2\frac{A-1}{A-2}}$.

This follows from Minkowski theorem and "minimality" of P_i in the sence of value at ζ.

Lemma 7.2.8. $|P_i(\zeta)| < H(P_{i+1})^{-3}$.

Proof. Simply take $A = 4$. We may take A bigger, than real value. If real value is bigger than 4, we are already done.

Remark 7.2.2. In the proof we will use lemma 7.2.8. But lemma 7.2.7 may give better bound. Using it, we may obtain result from [12]. This is almost unique change in proof between two articles.

Lemma 7.2.9. *For every number i $|a_{i,j}| \geq \zeta^2 H(P_i)$ for at least two indexes $j \in \{1,2,3\}$.*

Note, that first lemma of this chapter shows, that there is at least one such index.

Lemma 7.2.10. *$P_i(x)$ is irreducible over $\mathbb{Z}$ for every i.*

Lemma 7.2.11. *$|P_i(\zeta)|^{-1} < H(P_{i+1})^{(2A+1)/3} H(P_i)^{2/3}$ for every i.*

Lemma 7.2.9 will be used to construct one more polynomial sequence, Q_i. Take index j such that $|a_{i,j}| \geq \zeta^2 H(P_i)$. Let $Q_i(x) = b_3x^3 + b_2x^2 + b_1x + b_0$ be polynomial of minimal height, such that $|Q_i(\zeta)| < c_0^{-1} |P_{i-1}(\zeta)|$ and $b_j \leq c_0^{-1} H(P_i)$. It is easy to show, that height of Q_i is realized by b_k, where $k \neq 0, j$.

Lemma 7.2.12. *$H(Q_i) < c_0^{-2} H(P_i)^{-1/2} |P_{i-1}(\zeta)|^{-1/2}$.*

For proof we use Minkowski theorem to construct some polynomial Q with needed properties and use the minimality condition in definition of Q_i.

Lemma 7.2.13. *There are at least two indexes (j,k) such that $|b_j|, |b_k| \geq \zeta^2 H(Q_i)$.*

Lemma 7.2.14. *Polynomials P_i, P_{i-1}, Q_i are linearly independent and the determinant*

$$\begin{vmatrix} b_j & b_k & Q_i(\zeta) \\ a_{i,l} & a_{i,m} & P_i(\zeta) \\ a_{i-1,n} & a_{i-1,p} & P_{i-1}(\zeta) \end{vmatrix}$$

is greater than $\zeta^9 |P_{i-1}(\zeta)| H(P_i) H(Q_i)$. Here indexes are chosen to satisfy conditions of lemmas 7.2.9 and 7.2.13.

Proof follows from writing determinant as sum of six summands and noting, that one of them may be bounded from below and all others from above, that gives the needed bound. Note, that in lemma 7.2.6 we will need determinant bounded from below.

Now we may state main inequalities, that we will need.

Lemma 7.2.15. *Let $L(x)$ be polynomial such that*

$|L(\zeta)| < |P_{i-1}(\zeta)|^{1/2} H(P_i)^{1/2} H(P_{i-m})^{-2}$,
$|L'(\zeta)| < |P_{i-1}(\zeta)|^{1-A/2} H(P_i)^{1-A/2} H(P_{i-m})^{-1}$,
$H(L) < \zeta^{-2} |L'(\zeta)|$,
where m is chosen strictly between 1 and $i-1$ and satisfy
$H(P_{i-1}) \leq \max(c_3 H(P_{i-m-1}), H(P_{i-m}))$, $c_3 = 144 c_0^4 \zeta^{-4}$.
Then $\frac{|L(\zeta)|}{|L'(\zeta)|} < (c_3\zeta^{-1})^{2A} H(L)^{-A}$.

To prove it we write sequence of inequalities, which gives
$|L(\zeta)| < c_3^2 |P_{i-1}(\zeta)|^{1/2+\alpha_1-\alpha_2} H(P_i)^{(2A+1)\cdot\alpha_1/3-3\alpha_2+1/2} H(P_{i-1})^{2\alpha_1/3-2}$
for any choice of positive constants α_1, α_2. We take $\alpha 1 = 3(A+1)/2, \alpha_2 = 3A+1-A^2$. Then first and last powers transform respectively to $(A/2-1)(A-1)$ and $A-1$ (in fact we choose them to be solutions of corresponding linear system), and the second – into $(A/2-1)(A-1)$ because A is a root of suitable polynomial (this is how the polynomial was constructed). Now everything may be easily obtained from lemma conditions.

Remark 7.2.3. To obtain the first inequality we use in particular the bound from 7.2.8. Using 7.2.7 instead of it gives another linear system and another polynomial. This is the main change in [12] comparing with [11].

Remark 7.2.4. In general case we need to take

$$\alpha_2 = \frac{7}{2} + \frac{3(d-2)(A-1)}{d-1} - \frac{(A-1)(A-2)}{2} < \frac{-A^2+9A-1}{2} < 0$$

for $A \geq 9$.

Wirsing original result gives $A > d/2$, hence this method is unsuitable for large d (if we want to have for them results greater than 9). In fact, several other lemma proofs work only for A satisfying some inequalities. That's the main reason, why almost all results in Tishenko papers stated only for small values of d.

Lemma 7.2.16. *Let A_1, A_2, A_3 be positive numbers with following properties:*
$A_1A_2A_3 = 6|P_{i-1}(\zeta)| H(P_i)H(Q_i)$,
$A_1 \leq 6\zeta^{-4} |P_{i-1}(\zeta)| H(P_i)H(Q_i)H(P_{i-m})^{-2}$,
$H(\tilde{P}_1) \leq A_2 \leq A_3 \leq |P_{i-1}(\zeta)|^{-1/2} H(P_i)H(P_{i-m})^{-1}$.
And m is chosen to satisfy $H(P_{i-1}) \leq \max(c_3H(P_{i-m-1}), H(P_{i-m}))$, $1 < m < i-1$.

Then we may find non-trivial polynomial $L(x) = u_0x^3 + u_1x^2 + u_2x + u_3$ such that

$$|L(\zeta)| < A_1, \left|u_{i_j}\right| \leq A_{j+1} for \{i1, i2\} = \{1, 2\}, H(L) < \zeta^{-2}A_3.$$

The proof is quite simple – we choose suitable system of linear inequalities and use lemma 7.2.6.

This lemma gives immediately two new inequalities:
$|P_{i-m-1}(\zeta)| < c_0 \cdot 6\zeta^{-4} |P_{i-1}(\zeta)| H(P_i)H(Q_i)H(P_{i-m})^{-2}$,
$|P_{i-m-1}(\zeta)| < |P_{i-1}(\zeta)|^{-1/2} H(P_i)^{1/2}H(P_{i-m})^{-2}$,
where m is chosen to satisfy $H(P_{i-1}) \leq \max(c_3H(P_{i-m-1}), H(P_{i-m}))$, $1 < m < i-1$.

To prove them we take $A_1 = 6\zeta^{-4} |P_{i-1}(\zeta)| H(P_i)H(Q_i)H(P_{i-m})^{-2}$ and $A_2 = A_3 = \zeta^2 H(P_{i-m})$. It is easy to check, that all conditions of previous lemma hold, hence there exists corresponding polynomial L with needed property. Same property for P_{i-m-1} follows from minimality condition of sequence

P_i. To prove second inequality we use the first one and the bound obtained in 7.2.12.

Finally, we state one more inequality:

Lemma 7.2.17.

$$|P_{i-m-1}(\zeta)|\, H(P_{i-m})H(Q_{i-m}) < 6c_0^2\zeta^{-2}\, |P_{i-1}(\zeta)|\, H(P_i)H(Q_i),$$

Where, as usual, m is chosen to satisfy

$$H(P_{i-1}) \leq \max(c_3 H(P_{i-m-1}), H(P_{i-m})), \quad 1 < m < i-1.$$

After long preparations we may start proof of the main theorem. It will be very short.

Lets choose sequence $1 = m_1 < m_2 < m_3 < \ldots$ such that

$$H(P_{m_{k+1}}) \leq \max(c_3 H(P_{m_k}), H(P_{m_k+1})) < H(P_{m_{k+1}+1}).$$

We find easily that $H(P_{m_{k+1}+1})^{-1} < c_3^{-1} H(P_{m_{k-1}+1})^{-1}$, hence for every even k

$$H(P_{m_{k+1}})^{-1} < c_3^{-k/2} H(P_2)^{-1}.$$

Taking $i = m_k$ in statement of lemma 7.2.17 and multplying all inequalities for $m_1, m_2, \ldots, m_k$ we obtain

$|P_1(\zeta)|\, H(P_2)H(Q_2) < 6^k c_0^{2k}\zeta^{-2k} \left|P_{m_{k+1}}(\zeta)\right| H(P_{m_{k+1}+1})H(Q_{m_{k+1}+1})$.

Hence $|P_1(\zeta)| < 6^k c_0^{2k}\zeta^{-2k} \left|P_{m_{k+1}}(\zeta)\right| H(P_{m_{k+1}+1})H(Q_{m_{k+1}+1})$.

Now, using lemma 7.2.12, bound from lemma 7.2.8 and the bound

$$H(P_{m_{k+1}})^{-1} < c_3^{-k/2} H(P_2)^{-1}$$

we have

$$\begin{aligned} |P_1(\zeta)| &< 6^k c_0^{2k}\zeta^{-2k} \left|P_{m_{k+1}}(\zeta)\right|^{1/2} H(P_{m_{k+1}+1})^{1/2} \\ &< 6^k c_0^{2k}\zeta^{-2k} H(P_{m_{k+1}+1})^{-1} < 6^k c_0^{2k}\zeta^{-2k} c_3^{-k/2} H(P_2)^{-1} < (1/2)^k. \end{aligned}$$

This is a contradiction with transcendence of ζ.

Remark 7.2.5. This technical proof becomes much more complicated for $d > 3$ because we need to construct not one but $d-2$ sequences of auxiliary polynomials $Q_i^{(v)}$. See details in [11].

Those results (including last Tishenko works) are best known in Wirsing problem. Here we stop our investigation in homogeneous problem and start with integral approximations.

8 Integral approximation

Consider now approximations by algebraic integers. The same argumentation as for algebraic numbers leads to the quotient $\frac{P(x)}{P'(x)}$, where $P(x)$ is monic. On the first sight it seems logical to expect that we may build approximation with precision H^{-d} – in the numerator we have d parameters (coefficients) which run from $-H$ to H, so, we expect numerator sometimes be of order H^{-d+1}, and denominator to be independent of it (in some sense), so denominator is expected to be or order H and this gives the bound that we need.

But surely our old methods needs to be changed. Even the first idea (Dirichlet box principle) cannot be used in non-homogeneous problem.

Davenport and Schmidt showed in their article [9], that the following statement holds:

Theorem 8.1. *For $d = 3$ it's possible to take the exponent equal to $\varphi^2 = \varphi+1 \approx 2.618$ (here $\varphi = \frac{1+\sqrt{5}}{2}$ – golden ratio) for ζ not algebraic of degree at most 2.*

Remark 8.1.1. For general case they proved, also, that it is possible to take the exponent equal to $-[\frac{d+1}{2}]$, which gives 2 for $d = 3$ and even for $d = 4$. This result looks like original Wirsing result – it's approximately one half of that we may expect.

Proof. At first we will prove the following "dual" statement: for every ζ neither rational nor quadratic irrational there are arbitrary large X such that the system

$$\begin{cases} |x_0| < X \\ |x_0\zeta - x_1| < cX^{-\varphi+1} \\ \left|x_0\zeta^2 - x_2\right| < cX^{-\varphi+1} \end{cases}$$

has no non-zero integral solutions. Here c denotes constant, depending only on ζ.

Suppose, that for every c and every sufficiently large X this system has solution.

Consider for every real $X > 1$ the set of integer points (x_0, x_1, x_2) with the following properties:

$1 \le x_0 \le X, |x_0\zeta - x_1| < 1, \left|x_0\zeta^2 - x_2\right| < 1.$

And choose among such points the unique point, which minimize the value of

$$\max(|x_0\zeta - x_1|, \left|x_0\zeta^2 - x_2\right|).$$

(Note, that uniqueness holds, because ζ is neither rational or quadratic irrational.) Obviously we may choose a sequence of real numbers X_i, such that every minimal point is minimal exactly for X in $[X_i, X_{i+1})$. We denote this point as $\mathbf{x_i} = (x_{i0}, x_{i1}, x_{i2})$. Plainly

$$x_{i0} = X_i \text{ and } |\mathbf{x_i}| \ll X_i.$$

Denote $L_i = \max(|x_{i0}\zeta - x_{i1}|, \left|x_{i0}\zeta^2 - x_{i2}\right|)$. We have $L_1 > L_2 > \ldots$.

The minimal point for $X = X_{i+1} - \varepsilon$ is $\mathbf{x_i}$, due to our hypothesis it gives $L_i \le c(X_{i+1} - \varepsilon)^{-\varphi+1}$. Taking the limit for $\varepsilon \to 0$ we obtain $L_i \le cX_{i+1}^{-\varphi+1}$.

Lemma 8.1.1. *For sufficiently large i we have $x_{i0}x_{i2} - x_{i1}^2 \neq 0$.*

Proof. x_{i0}, x_{i1}, x_{i2} have no common factors. If $x_{i0}x_{i2} - x_{i1}^2 = 0$ then $x_{i0} = m^2, x_{i1} = mn, x_{i2} = n^2$ for some integers m, n. Suppose, that $m \ge 0$ (in other case change all signs). Then $X_i^{1/2} = m$ and

$$x_i^{-\varphi+1} \gg |x_{i0}\zeta - x_{i1}| = |m(m\zeta - n)| = X_i^{1/2} |m\zeta - n| .$$

So, we have $|m\zeta - n| < X_i^{-1/2} X_{i+1}^{-\varphi+1}$

Vectors $\mathbf{x_i}$ and $\mathbf{x_{i-1}}$ are linearly independent, hence, their matrix has rank 2. For sufficiently large i x_{i1} cannot be zero (in other case $L_i \ge |x_{i0}\zeta| \ge |\zeta|$) and $x_{i-1,1}$ too (due to same reasons). Hence, there is a column in the matrix, which is non-proportional to the middle one. So, at least one of determinants, involving the middle column, is non-zero.

Let it be for example $\begin{pmatrix} x_{i-1,0} & x_{i-1,1} \\ m & n \end{pmatrix}$ (the second case is very similar).

Then we have the following estimate:

$$\begin{aligned} -\begin{vmatrix} x_{i-1,0} & x_{i-1,1} \\ m & n \end{vmatrix} &= \begin{vmatrix} x_{i-1,0} & x_{i-1}\zeta - x_{i-1,1} \\ m & m\zeta - n \end{vmatrix} \\ &\ll X_i^{-1/2} X_{i+1}^{-\varphi+1} X_{i-1} + +X_i^{-\varphi+1} X_i^{1/2} \\ &\ll X_i^{-\varphi+3/2} \to 0 \end{aligned}$$

which contradicts with integrity of initial determinant. The lemma is proven.

Lemma 8.1.2. *For all sufficiently large i we have $X_{i+1}^{\varphi-1} \le 2c(1 + |\zeta|)X_i$.*

Proof. We have the following estimates:

$|x_{i0}\zeta - x_{i1}| \le cX_{i+1}^{-\varphi+1}$ and

$$\begin{aligned} |x_{i1}\zeta - x_{i2}| = \left|xi1\zeta - x_{i0}\zeta^2 + x_{i0}\zeta^2 - x_{i2}\right| &\le |\zeta| \, |xi1 - x_{i0}\zeta| + \left|x_{i0}\zeta^2 - x_{i2}\right| \\ &\le c(1 + |\zeta|)X_{i+1}^{-\varphi+1} \end{aligned}$$

Hence

$$\begin{aligned} -\begin{vmatrix} x_{i0} & x_{i1} \\ x_{i1} & x_{i2} \end{vmatrix} = -\begin{vmatrix} x_{i0} & x_{i1 - \zeta x_{i0}} \\ x_{i1} & x_{i2} - \zeta x_{i1} \end{vmatrix} &\le cX_{i+1}^{-\varphi+1}(|x_{i0}| \, (1 + |\zeta|) + |x_1|) \\ &\le 2c(1 + |\zeta|)X_i X_{i+1}^{-\varphi+1}. \end{aligned}$$

Now the lemma follows, because determinant is integer and non-zero (due to previous lemma).

Lemma 8.1.3. *Suppose, that* $\mathbf{x_{i-1}}, \mathbf{x_i}, \mathbf{x_{i+1}}$ *are linearly independent. Then* $6c^2 X_{i+1}^{2-\varphi} \geq X_i^{\varphi-1}$.

Proof. Consider the determinant

$$\begin{vmatrix} x_{i-1,0} & x_{i-1,1}, x_{i-1,2} \\ x_{i0}, x_{i1} & x_{i2} \\ x_{i+1,0} & x_{i+1,1}, x_{i+1,2} \end{vmatrix} = \begin{vmatrix} x_{i-1,0} & x_{i-1,1} - \zeta x_{i-1,0}, x_{i-1,2} - \zeta^2 x_{i-1,0} \\ x_{i0}, x_{i1} - \zeta x_{i0} & x_{i2} - \zeta^2 x_{i0} \\ x_{i+1,0} & x_{i+1,1} - \zeta x_{i+1,0}, x_{i+1,2} - \zeta^2 x_{i+1,0} \end{vmatrix}$$

$$\leq 6 \cdot X_{i+1} \cdot c X_i^{-\varphi+1} X_{i+1}^{-\varphi+1} = 6c^2 X_i^{-\varphi+1} X_{i+1}^{-\varphi+2}.$$

Initial determinant has absolute value at least 1, hence lemma is proven.

Lemma 8.1.4. *For infinitely many* i *points* $\mathbf{x_{i-1}}, \mathbf{x_i}, \mathbf{x_{i+1}}$ *are linearly independent.*

Proof. If the converse is true, then every three consecutive points lie in same plane, but every two consecutive points define plane (because we know, that two points cannot be dependent, even if they are non-consecutive). Hence, all points lie in the same plane. Let it be $ax_{i0} + bx_{i1} + cx_{i2} = 0$. We know, that $x_{i1} = \zeta x_{i0} + O(X_{i+1}^{-\varphi+1})$ and $x_{i2} = \zeta^2 x_{i0} + O(X_{i+1}^{-\varphi+1})$. Hence $0 = ax_{i0} + bx_{i1} + cx_{i2} = X_i(a\zeta^2 + b\zeta + c) + O(X_{i+1}^{-\varphi+1})$. Second term tends to zero, first is a constant multiple of increasing sequence, hence, constant is zero, that is possible only for $a = b = c = 0$.

Now we may prove the statement. We know, that suggestions of first lemmas hold simultaneously for infinitely many i. So, for those i we have

$$X_i^{(\varphi-1)^2} \leq (6c^2)^{\varphi-1} X_{i+1}^{(\varphi-1)(2-\varphi)} \leq (6c^2)^{\varphi-1} (2c(1+|\zeta|))^{2-\varphi} X_i^{2-\varphi}$$

which is impossible for sufficiently small c because $(\varphi-1)^2 = 2-\varphi$.

Remark 8.1.2. Note, that this is the first place, where we really need, that φ is a root of $x^2 - x - 1 = 0$.

So, the dual statement is proven and we may now start with main theorem.

Suppose that for some $c > 0$ there are X arbitrary large such that the system of inequalities

$$\begin{cases} |x_0| < X \\ |x_0\zeta - x_1| < cX^{-\lambda} \\ \left|x_0\zeta^2 - x_2\right| < cX^{-\lambda} \end{cases}$$

has no integer solutions except of zero-solution. We may suppose, that $c < 1$ and let $Y = X^{\frac{\lambda+1}{3}}$. We will prove, that there are infinitely many algebraic integers α of degree at most 3, satisfying $0 < |\zeta - \alpha| \ll H(\alpha)^{-1-1/\lambda}$. This will be enough - our statement is proven for $\lambda = \varphi - 1$ and $-1 - 1/\lambda = -\frac{\varphi}{\varphi - 1} = -(\varphi + 1)$.

Let $K(Y)$ be the parallelepiped, defined by

$$\begin{cases} |x_0| < Y^2 \\ |x_0\zeta - x_1| < Y^{-1} \\ \left|x_0\zeta^2 - x_2\right| < Y^{-1} \end{cases}$$

Then the first minimum of $K(Y)$ (denoted by $\tau_1(Y)$) satisfies $\tau_1(Y) \geq cY^{-\delta}$, where $\delta = \frac{3\lambda}{\lambda+1} - 1$.

(Because otherwise we will have non-trivial solution of
$|x_0| \leq \tau_1 Y^2 < Y^{2-\delta} = X$ and
$\left|x_0\zeta^i - x_i\right| \leq \tau_1(Y)Y^{-1} < cY^{-1-\delta} = cX^{-\lambda}$
which is a contradiction with choice of X).

Mahler theorem for polar reciprocal body $K^*(Y) = \{\left|x_2\zeta^2 + x_1\zeta + x_0\right| \leq Y^{-2}, |x_i| \leq Y\}$ gives us $\tau_n^*(Y) \leq c_1 Y^{\delta}$. So, there are 3 linearly independent points $\mathbf{x_i} = (x_0^{(i)}, x_1^{(i)}, x_2^{(i)})$ such that

$$\begin{cases} \left|x_2^{(i)}\zeta^2 + x_1^{(i)}\zeta + x_0^{(i)}\right| \leq c_1 Y^{-2+\delta} \\ \left|x_m^{(i)}\right| \leq Y^{1+\delta} \end{cases}$$

Denote $L^{(i)} = x_2^{(i)}\zeta^2 + x_1^{(i)}\zeta + x_0^{(i)}, P^{(i)} = 2x_2^{(i)}\zeta + x_1^{(i)}$.

There are real numbers $\theta_1, \theta_2, \theta_3$ satisfying the following system:

$$\begin{cases} \theta_1 x_2^{(1)} + \theta_2 x_2^{(2)} + \theta_3 x_2^{(3)} = 0 \\ 3\zeta^2 + \theta_1 P^{(1)} + \theta_2 P^{(2)} + \theta_3 P^{(3)} = Y^{1+\delta} + \left|P^{(1)}\right| + \left|P^{(2)}\right| + \left|P^{(3)}\right| \\ \zeta^3 + \theta_1 L^{(1)} + \theta_2 L^{(2)} + \theta_3 L^{(3)} = 4c_1 Y^{\delta - n + 1} \end{cases}$$

(Determinant of this system is in fact determinant of $(\mathbf{x^{(3)}}, \mathbf{x^{(2)}}, \mathbf{x^{(1)}})$, hence non-zero, because points were linearly independent).

Let $t_i = [\theta_i]$. They cannot all be zero (due to second equation, all θ_i cannot be in $[0, 1]$), hence $\mathbf{x} = t_1 \cdot \mathbf{x^{(1)}} + t_2 \cdot \mathbf{x^{(2)}} + t_3 \cdot \mathbf{x^{(3)}} \neq \mathbf{0}$.

Lets obtain bound for its coordinates.

Due to first equation $|x_2| \leq 3c_1 Y^{1+\delta}$.

Due to second equation

$$Y^{1+\delta} \leq \left|3\zeta^2 + 2x_2\zeta + x_1\right| \leq Y^{1+\delta} + 2\left|P^{(1)}\right| + 2\left|P^{(2)}\right| + 2\left|P^{(3)}\right| \ll Y^{1+\delta}.$$

Due to last one, finally, $0 < \left|\zeta^3 + x_2\zeta^2 + x_1\zeta + x_0\right| \leq 7c_1 Y^{\delta-2}$.

Hence we may see, that polynomial $Q(t) = t^3 + x_2 t^2 + x_1 t + x_0$ has height $H(Q(t)) \ll Y^{1+\delta}$, value at point ζ at most $Y^{\delta-2}$ and derivative at that point $\gg Y^{1+\delta}$. Last inequality means, that derivative is $\gg Y^{1+\delta}$ on segment of length comparable with 1 (because height of polynomial is small enough). Hence, Q has a root α, satisfying $0 < |\zeta - \alpha| \ll Y^{\delta-2-(\delta+1)} = Y^{-3}$.

Since $H(\alpha) \ll H(Q) \ll Y^{1+\delta}$, we have $0 < |\zeta - \alpha| \ll H(\alpha)^{-3/(\delta+1)} = H(\alpha)^{1+1/\lambda}$.

We are done.

9 Extremal numbers due to Damien Roy

In 2002 Damien Roy in his articles [7] and [8] proved, that Schmidt estimate is exact in general case:

Theorem 9.1. *(Damien Roy) There exists a real transcendental number ζ and a constant $c_1 > 0$, such that for any algebraic integer α of degree at most 3 we have $|\zeta - \alpha| > c_1 H(\alpha)^{-\varphi-1}$*

Here we will give only a sketch of proof, because many statements in proofs are pure technical.

Proof. At first we will state and prove another theorem. We had seen in previous section, that due to duality arguments its useful to study solutions of the system

$$\begin{cases} |x_0| < X \\ |x_0\zeta - x_1| < cX^{1-\varphi} \\ \left|x_0\zeta^2 - x_2\right| < cX^{1-\varphi} \end{cases}$$

Lets call *extremal* those real numbers, for which this system has non-zero solutions for some $c > 0$ and all X. We had seen, that for every real number there is some $c > 0$ such that system has no solutions for some arbitrary large X. Extremal numbers (if they exist) show us, that estimates cannot be improved. Our first goal will be

Theorem 9.2. *Extremal numbers exist. Moreover, they are transcendental and set of extremal numbers is at most countable.*

We will use the following notation: point $\mathbf{x} = (x_0, x_1, x_2)$ will be identified with matrix $\begin{pmatrix} x_0 & x_1 \\ x_1 & x_2 \end{pmatrix}$, and we write $\det(\mathbf{x}) = \mathbf{x_0 x_2} - \mathbf{x_1^2}$. Let $J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$.

Define $\|\mathbf{x}\| = \max\{|x_0|, |x_1|, |x_2|\}$, $L(\mathbf{x}) = \max\{|x_0\zeta - x_1|, \left|x_0\zeta^2 - x_2\right|\}$ and $\|A\| = |\det(A)|$ for 2×2-matrix.

Lemma 9.2.1. *Let $\mathbf{x}, \mathbf{y}, \mathbf{z}$ be three integer points. Then*

1. *$tr(J\mathbf{x}J\mathbf{z}J\mathbf{y}) = \det(\mathbf{x}, \mathbf{y}, \mathbf{z})$.*
2. *$\mathbf{x}J\mathbf{z}J\mathbf{y}$ is symmetric if and only if $\det(\mathbf{x}, \mathbf{y}, \mathbf{z})$ is zero. In this case we denote point, corresponding to $\mathbf{x}J\mathbf{z}J\mathbf{y}$ as $[\mathbf{x}, \mathbf{y}, \mathbf{z}]$.*
3. *for $\mathbf{x}, \mathbf{y}, \mathbf{z}$ with $\det(\mathbf{x}, \mathbf{y}, \mathbf{z}) = 0$ and any $\mathbf{w}$ we have*
   - *3.1.* $\det[\mathbf{x}, \mathbf{y}, \mathbf{z}] = \det(\mathbf{x}) \det(\mathbf{y}) \det(\mathbf{z})$
   - *3.2.* $\det(\mathbf{w}, \mathbf{y}, [\mathbf{x}, \mathbf{y}, \mathbf{z}]) = \det(y) \det(\mathbf{w}, \mathbf{z}, \mathbf{x})$
   - *3.3.* $\det(\mathbf{x}, \mathbf{y}, [\mathbf{x}, \mathbf{y}, \mathbf{z}]) = 0$

3.4. $[\mathbf{x},\mathbf{y},[\mathbf{x},\mathbf{y},\mathbf{z}]] = \det(\mathbf{x})\det(\mathbf{y})\mathbf{z}$

4. for any $\mathbf{x},\mathbf{y},\mathbf{z}$

4.1. for any choice of $r,s,t,u \in \{0,1,2\}$*, such that* $s-r = u-t$*, we have* $\|\begin{matrix} x_r & x_s \\ y_t & y_u \end{matrix}\| \ll \|\mathbf{x}\|L(\mathbf{y}) + \|\mathbf{y}\|L(\mathbf{x})$.

4.2. $|\det(\mathbf{x},\mathbf{y},\mathbf{z})| \ll \|\mathbf{x}\|L(\mathbf{y})L(\mathbf{z}) + \|\mathbf{y}\|L(\mathbf{x})L(\mathbf{z}) + \|\mathbf{z}\|L(\mathbf{x})L(\mathbf{y})$

4.3 Let $\mathbf{w} = [\mathbf{x},\mathbf{x},\mathbf{y}]$*. then* $\|\mathbf{w}\| \ll (\|\mathbf{x}\|L(\mathbf{y}) + \|\mathbf{y}\|L(\mathbf{x}))L(\mathbf{x})$

First statements are pure computation. Statement 4 uses multilinearity of determinants.

We may define the sequence of minimal points, as it was done in previous chapter. Our next goal will be to study some properties of minimal points.

We define *height* of 2×3-matrix A as maximum of absolute values of its minors 2×2 and denote it $H(A)$. For any 2-dimensional subspace V of Q^3 we define $H(V)$ as height of any matrix, whose rows form a basis of $V \cap Z^3$.

Lemma 9.2.2. *Let* $\mathbf{x}$ *be minimal point,* $\mathbf{y}$ *be next minimal point and* $V = \langle \mathbf{x},\mathbf{y}\rangle$*. Then* $\{\mathbf{x},\mathbf{y}\}$ *form basis for* $V\cap Z^3$ *and* $H(V) \sim \|\mathbf{y}\|L(\mathbf{x})$

Proof. Let $\mathbf{x} = (x_0,x_1,x_2), \mathbf{y} = (y_0,y_1,y_2)$. If they are not a basis for lattice, we may find an integral point of form $\mathbf{z} = q_1\mathbf{x} + q_2\mathbf{y}$. Obviously, we may take $|q_1|, |q_2| \le 1/2$. Hence $z_0 \le |q_1| x_0 + |q_2| y_0 < y_0$ and $L(\mathbf{z}) \le |q_1| L(\mathbf{x}) + |q_2| L(\mathbf{y}) < L(\mathbf{x})$. This is a contradiction, because $\mathbf{y}$ was the next minimal point after $\mathbf{x}$.

Due to 9.2.1 part 4.1, we find that $H(V) = H(\begin{pmatrix} x_0 & x_1 & x_2 \\ y_0 & y_1 & y_2 \end{pmatrix}) \ll \|\mathbf{x}\|L(\mathbf{y}) + \|\mathbf{y}\|L(\mathbf{x}) \ll \|\mathbf{y}\|L(\mathbf{x})$.

To prove lower bound we denote $u_i = x_i - x_0\zeta^i, v_i = y_i - y_0^i$ and choose j such that $|u_i| = L(\mathbf{x})$. Then we obtain

$$H(V) \ge \|\begin{matrix} x_0 & x_j \\ y_0 & y_j \end{matrix}\| = |x_0 v_j - y_0 u_j| \ge y_0 |u_j| - x_0 |v_j| \ge (y_0 - x_0)L(\mathbf{x}).$$

The point $\mathbf{z} = \mathbf{y} - \mathbf{x}$ has first coordinate less, than y_0, that means that there exists index i such that $\left|z_i - z_0\zeta^i\right| = |v_i - u_i| > L(\mathbf{x})$. Hence

$$H(V) \ge |x_0 v_i - y_0 u_i| = |x_0(v_i - u_i) - (y_0 - x_0)u_i| \ge x_0 L(\mathbf{x}) - (y_0 - x_0)L(\mathbf{x}).$$

Those inequalities give $3H(V) \ge y_0 L(\mathbf{x})$ and lower bound is proven.

Lets give now another characterization of extremal numbers.

Lemma 9.2.3. *Real number* ζ *is extremal if and only if there exists an increasing sequence of integers* Y_k *and sequence of points* $\mathbf{y_k}$*, such that:*

$Y_{k+1} \sim Y_k^{\varphi}, \|y_k\| = Y_k, L(Y_k) \sim Y_k^{-1}, 1 \le |\det(\mathbf{y_k})| \ll 1,$
$1 \le |\det(\mathbf{y_k},\mathbf{y_{k+1}},\mathbf{y_{k+2}})| \ll 1.$

Proof. Suppose at first that such sequences exist. Then ζ cannot be rational or quadratic irrational, because if we have $p + q\zeta + r\zeta^2 = 0$, then $|py_{k,0} + qy_{k,1} + ry_{k,2}| = \left|q(y_{k,1} - y_{k,0}\zeta^2) + r(y_{k,2} - y_{k,0}\zeta^2)\right| \ll Y_k^{-1} \to 0$. But this is an integer number, so equals 0 for k large enough, and almost all points lie in plane $p\mathbf{y_{k,0}} + q\mathbf{y_{k,1}} + r\mathbf{y_{k,2}} = 0$, which is impossible, because every three consecutive points are linearly independent.

For every X sufficiently large we may find k such that $Y_k \le X \le Y_{k+1}$. Hence $\|\mathbf{y_k}\| \ll Y_k \ll X$ and $L(\mathbf{y_k}) \ll Y_k^{-1} \ll Y_{k+1}^{-1/\varphi} \ll X^{-1/\varphi}$. So, ζ is extremal.

Conversely, let ζ be extremal. Consider the corresponding sequence of minimal points, $\mathbf{x_i}$, denote X_i to be first coordinate of $\mathbf{x_i}$ and $L_i = L(\mathbf{x_i})$. Choose also corresponding constant c. So, $L_i \le cX_{i+1}^{1/\varphi}$.

Due to lemma 8.1.1 $\det(\mathbf{x_i}) \neq 0$ for i sufficiently large. Due to lemma 8.1.4 there are infinitely many indexes i, such that $\mathbf{x_{i-1}}, \mathbf{x_i}, \mathbf{x_{i+1}}$ are linearly independent over Q. Lets take all such i and construct sequence $\mathbf{y_j} = \mathbf{x_{i_j}}$. We claim, that $\mathbf{y_j}$ and $Y_j = X_{i_j}$ satisfy all properties.

We have $L_{i-1} \ll X_i^{-1/\varphi}$ and $L_i \ll X_{i+1}^{-1/\varphi}$.

Hence we obtain $1 \le |\det(\mathbf{x_i})| \ll L_i X_i \ll X_i X_{i+1}^{-1/\varphi}$ and $X_{i+1} \ll X_i^{\varphi}$. But also we have

$$1 \le |\det(\mathbf{x_{i-1}}, \mathbf{x_i}), \mathbf{x_{i+1}}| \ll X_{i+1} L_i L_{i-1} \ll X_{i+1}^{1/\varphi^2} L_{i-1} \le X_{i+1}^{1/\varphi^2} X_i^{1/\varphi}.$$

Both estimates together give $X_{i+1} \sim X_i^{\varphi}$ and $L_i \sim X_i^{-1}$.

In particular, $\|\mathbf{y_k}\| \sim Y_k, L(\mathbf{y_k}) \sim Y_k^{-1}$ and $1 \le |\det(\mathbf{y_k})| \ll 1$.

Now, let $i = i_k, V = \langle \mathbf{x_i}, \mathbf{x_{i+1}} \rangle_{\mathbb{Q}}$. Let $\mathbf{x_i}, \mathbf{x_{i+1}}, \ldots, \mathbf{x_j} \in V, \mathbf{x_{j+1}} \notin V$. Then $\langle \mathbf{x_i}, \mathbf{x_{i+1}} \rangle_{\mathbb{Q}} = \langle \mathbf{x_{j-1}}, \mathbf{x_j} \rangle_{\mathbb{Q}}$ and hence points $\mathbf{x_{j-1}}, \mathbf{x_j}, \mathbf{x_{j+1}}$ are linearly independent. Moreover, j is the smallest index with such property, hence $j = i_{k+1}$. Then we obtain due to 9.2.2 two estimates for one subspace:

$H(\langle \mathbf{x_i}, \mathbf{x_{i+1}} \rangle_{\mathbb{Q}}) \sim X_{i+1} L_i \sim X_i^{1/\varphi}$,

$H(\langle \mathbf{x_{j-1}}, \mathbf{x_j} \rangle_{\mathbb{Q}}) \sim X_j L_{j-1} \sim X_j^{1/\varphi^2}$.

This gives $Y_{k+1} \sim Y_k^{\varphi}$.

We know already, that $\langle \mathbf{y_k}, \mathbf{y_{k+1}} \rangle_{\mathbb{Q}}$ contains $\mathbf{x_i}$ and $\mathbf{x_{j-1}}$, and $\langle \mathbf{y_{k+1}}, \mathbf{y_{k+2}} \rangle_{\mathbb{Q}}$ contains $\mathbf{x_{j+1}}$. So, $\langle \mathbf{y_k}, \mathbf{y_{k+1}}, \mathbf{y_{k+2}} \rangle_{\mathbb{Q}}$ contains three linearly independent points, hence $\mathbf{y_k}, \mathbf{y_{k+1}}, \mathbf{y_{k+2}}$ are linearly independent. Now we may prove the last estimate, that we need:

$1 \le |\det(\mathbf{y_k}, \mathbf{y_{k+1}}), \mathbf{y_{k+2}}| \ll \|\mathbf{y_{k+2}}\| L(\mathbf{y_k}) L(\mathbf{y_{k+1}}) \ll Y_{k+2} Y_k^{-1} Y_{k+1}^{-1} \ll 1$.

We are done.

Corollary 9.2.1. *Let ζ be extremal number and $\mathbf{y_k}, Y_k$ – corresponding sequences. Then for sufficiently large k the vector $\mathbf{y_{k+1}}$ is a rational multiple of* $[\mathbf{y_k}, \mathbf{y_k}, \mathbf{y_{k-2}}]$.

Proof. Let $\mathbf{w} = [\mathbf{y_k}, \mathbf{y_k}, \mathbf{y_{k+1}}]$. We know that $\det(\mathbf{w}) = \det(\mathbf{y_k})^2 \det(\mathbf{y_{k+1}})$.

So, this determinant is non-zero and

$$\|\mathbf{w}\| \ll Y_k^2 Y_{k+1}^{-1} \sim Y_{k-2}, L(\mathbf{w}) \ll Y_{k+1} Y_k^{-2} \sim Y_{k-2}^{-1}.$$

But then we may see immediately, that

$|\det(\mathbf{w}, \mathbf{y_{k-3}}, \mathbf{y_{k-2}})| \ll Y_{k-2} Y_{k-2}^{-1} Y_{k-3}^{-1} \sim Y_{k-3}^{-1}$ and

$|\det(\mathbf{w}, \mathbf{y_{k-2}}, \mathbf{y_{k-1}})| \ll Y_{k-1} Y_{k-2}^{-2} \sim Y_{k-3}^{-1/\varphi}$

Right parts converges to zero, left part are integers. So, they are zero integers. We know, that $\mathbf{y_{k-3}}, \mathbf{y_{k-2}}, \mathbf{y_{k-1}}$ are linearly independent, so $\mathbf{w}$ is rational multiple of $\mathbf{y_{k-2}}$. then due to 9.2.1 $[\mathbf{y_k}, \mathbf{y_k}, \mathbf{w}] = \det(\mathbf{y_k})^2 \mathbf{y_{k+1}}$ and hence $\mathbf{y_{k+1}}$ is rational multiple of $[\mathbf{y_k}, \mathbf{y_k}, \mathbf{y_{k-2}}]$.

Corollary 9.2.2. *We have*

$$\det(\mathbf{y_{k-2}}, \mathbf{y_{k-1}}, \mathbf{y_k})\mathbf{y_{k+1}} = \det(\mathbf{y_{k-2}}, \mathbf{y_{k-1}}, \mathbf{y_{k+1}})\mathbf{y_k} + \det(\mathbf{y_{k-1}}, \mathbf{y_k}, \mathbf{y_{k+1}})\mathbf{y_{k-2}}$$

.

Proof. due to 9.2.1 we see, that

$$\det(\mathbf{y_{k-2}}, \mathbf{y_k}, [\mathbf{y_k}, \mathbf{y_k}, \mathbf{y_{k-2}}]) = 0.$$

Hence points $\mathbf{y_{k-2}}, \mathbf{y_k}$ and $\mathbf{y_{k+1}}$ which is rational multiple of $[\mathbf{y_k}, \mathbf{y_k}, \mathbf{y_{k-2}}]$ due to previous corollary, are linearly dependent. No two of them may be proportional, hence $\mathbf{y_{k+1}} = a\mathbf{y_k} + b\mathbf{y_{k-2}}$. This gives

$\det(\mathbf{y_{k-2}}, \mathbf{y_{k-1}}, \mathbf{y_{k+1}}) = a \cdot \det(\mathbf{y_{k-2}}, \mathbf{y_{k-1}}, \mathbf{y_k})$ and

$\det(\mathbf{y_{k-1}}, \mathbf{y_k}, \mathbf{y_{k+1}}) = b \cdot \det(\mathbf{y_{k-1}}, \mathbf{y_k}, \mathbf{y_{k-2}})$ due to multilinearity of determinant. This gives the formula that we need.

Corollary 9.2.3. *The set of extremal numbers is at most countable.*

Proof is very simple - we will construct injective map from this set to $(\mathbb{Z}^3)^3$. For each extremal number we take corresponding sequence $\mathbf{y_k}$ and start it from such index, that $\mathbf{y_{k+1}}$ is a rational multiple of $[\mathbf{y_k}, \mathbf{y_k}, \mathbf{y_{k-2}}]$ for all k. So, if we know first three points, we may determine all other points up to rational multiples. Now we make correspondence between this sequence and extremal number $\lim\limits_{k\to\infty} \frac{y_{k,1}}{y_{k,0}}$.

Now we shall prove

Theorem 9.3. *Extremal numbers exist, and set of extremal number is infinite.*

To prove the theorem, we need some preparations:

Lemma 9.3.1. *Let A and B be non-commuting symmetric matrixes in $Gl_2(Z)$. Consider sequence of points (note, that point in Z^3 may be identified with symmetric matrix)* $\mathbf{y_k}$*:*

$\mathbf{y_{-1}} = B^{-1}, \mathbf{y_0} = E, \mathbf{y_1} = A, \mathbf{y_k} = [\mathbf{y_{k-1}}, \mathbf{y_{k-1}}, \mathbf{y_{k-3}}]$.

Then $|\det(\mathbf{y_k})| = 1$ *and* $|\det(\mathbf{y_k}, \mathbf{y_{k+1}}, \mathbf{y_{k+2}})| = |tr(JAB)| \neq 0$.

Moreover, $\mathbf{y_k} = \pm \mathbf{y_{k-1}} S \mathbf{y_{k-2}}$*, where* $S = AB$ *for odd* k *and* $S = BA$ *for even* k*.*

Proof. Simply by induction using recurrence relation. The only interesting place in proof is that $|\det(\mathbf{y_{-1}}, \mathbf{y_0}, \mathbf{y_1})| = |\det(B)tr(JAB)| = |tr(JAB)| \neq 0$, because A and B do not commute.

Lemma 9.3.2. *Let A, B and $\mathbf{y_k}$ be as in previous lemma. Suppose, that A has non-negative and AB - positive coefficients. Write $\mathbf{y_k} = (y_{k,0}, y_{k,1}, y_{k,2})$ and $\zeta = \lim\limits_{k\to+\infty} \frac{y_{k,1}}{y_{k,0}}$. Then ζ is defined and extremal. Moreover, $\mathbf{y_k}$ and $Y_k = \|\mathbf{y_k}\|$ form sequences with all properties of extremality criterion (norm here is a norm in 3-dimentional space, not determinant of matrix).*

Proof. Due to previous lemma, last two conditions of criterion are satisfied. The recurrence $\mathbf{y_k} = \pm\mathbf{y_{k-1}}S\mathbf{y_{k-2}}$ shows, that all coefficients of $\mathbf{y_k}$ have the same sign (it is true for $\mathbf{y_0}$ and $\mathbf{y_1}$, also for AB and $BA = B^T A^T = (AB)^T$, induction completes the proof). So, $Y_k > 0$ and $Y_{k-2}Y_{k-1} \leq Y_k \leq c_1 Y_{k-2}Y_{k-1}$ (here c_1 may be taken equal to $\det(AB)$). In particular, Y_k is monotone and unbounded. Similar arguments shows the same for $y_{k,0}$.

Denote by I_k the interval with ends $\frac{y_{k,1}}{y_{k,0}}$ and $\frac{y_{k,2}}{y_{k,1}}$. Due to recurrence equation, for some rational numbers r, s, t, u we have $\frac{y_{k,1}}{y_{k,0}} = \frac{ry_{k-1,1}+ty_{k-1,2}}{ry_{k-1,0}+ty_{k-1,1}} \in I_{k-1}$ and $\frac{y_{k,2}}{y_{k,1}} = \frac{sy_{k-1,1}+uy_{k-1,2}}{sy_{k-1,0}+uy_{k-1,1}} \in I_{k-1}$.

Hence $I_k \subset I_{k-1}$ and we have a decreasing sequence of segments, whose lengths are $\frac{1}{y_{k,0}y_{k,1}}$ and tend to zero. So, they have exactly one point of intersection. It is $\lim\limits_{k\to+\infty} \frac{y_{k,1}}{y_{k,0}} = \zeta$. So, $y_{k,0} \sim y_{k,1} \sim y_{k,2} \sim Y_k$, length of I_k is $\sim Y_k^{-2}$ and $L(\mathbf{y_k}) \sim Y_k\,|I_k| \sim Y_k^{-1}$.

We need now only to check $Y_k \sim Y_{k-1}^{\varphi}$. This follows from $Y_{k-2}Y_{k-1} \leq Y_k \leq c_1 Y_{k-2}Y_{k-1}$. Let $q_k = Y_k Y_{k-1}^{-\varphi}$. Then $q_{k-1}^{-1/\varphi} \leq q_k \leq c_1 q_{k-1}^{-1/\varphi}$. Hence q_k is bounded and we are done.

Finally, we may give an example of infinite family of extremal numbers. Define sequence of words in two letters a and b by $w_0 = b, w_1 = a, w_k = w_{k-1}w_{k-2}$. Because every next word begins with previous one, we may define their limit $w = abaabab\ldots$.

Now take $\zeta = [0, a, b, a, a, b, a, b, \ldots] = 1/(a + 1/(b + \ldots))$. This number is extremal and may be produced from matrixes $A = \begin{pmatrix} a & 1 \\ 1 & 0 \end{pmatrix}$ and $B = \begin{pmatrix} b & 1 \\ 1 & 0 \end{pmatrix}$. Moreover, it is easy to see, that if ζ is extremal, then $\frac{x\zeta+y}{z\zeta+t}$ is also extremal if $xt \neq yz$, that gives infinitely many examples.

Proof for ζ is straightforward: write

$$\begin{pmatrix} q_i & q_{i-1} \\ p_i & p_{i-1} \end{pmatrix} = \begin{pmatrix} a_1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a_2 & 1 \\ 1 & 0 \end{pmatrix} \cdots \begin{pmatrix} a_i & 1 \\ 1 & 0 \end{pmatrix}$$

(standart formula for another representation in continued fractions form). Now define a map from monoid of words in a, b to $Gl_2(Z)$ by $f(a) = A, f(b) = B$. Let m_k be w_{k+2} without last two letters. Then $m_0 = 1, m_1 = a$ and $m_k = m_{k-1}sm_{k-2}$, where s is equal to ab or ba depending on the pairity of k. So, for sequence of matrixes defined this way we obtain $\mathbf{y_k} = \pm\mathbf{x_k}$, and $\frac{y_{k,1}}{y_{k,0}} = \frac{x_{k,1}}{x_{k,0}}$.

Right part tends to extremal number, left part tends to ζ and that finishes the proof.

10 Exactness of Schmidt result

In previous section we had shown, that result of "dual" theorem cannot be improved. This causes doubts in upgrading of general results, but surely does not prove, that such upgrading does not hold.

But that is true and we need some more technical work to prove it. Proves for estimates may be found in [7]

Theorem 10.1. *Let ζ be extremal number.*

1. There exist $c > 0$ and $t > 0$, such that $|\zeta - \alpha| \geq cH(\alpha)^{-2}(1 + \ln H(\alpha))^{-t}$ has no solutions in rational α.

2. There exist $c_1, c_2 > 0$, such that $|\zeta - \alpha| \leq c_1 H(\alpha)^{-2\varphi^2}$ has infinitely many solutions in quadratic irrationalities, but $|\zeta - \alpha| \leq c_2 H(\alpha)^{-2\varphi^2}$ has no solutions in quadratic irrationalities.

3. There exist $c > 0$ such that for every algebraic integer $\alpha, \deg(\alpha) \leq 3$ we have $|\zeta - \alpha| \leq c_1 H(\alpha)^{-\varphi^2 - 1}$

Remark 10.1.1. Note, that in the first and second parts we speak about algebraic numbers, but in last part - about algebraic integers.

Remark 10.1.2. This result is not strong enough, because $\varphi^2 + 1 \approx 3.618 > 3$. We will give more precise estimate later.

Proof may be found in [7] (theorems 1.3—1.5).

Denote by $\{x\}$ the distance from x to nearest integer.

Lemma 10.1.1. *Assume, that for real numbers $c > 0$ and $\delta \in [0, 1)$ we have $\left\{y_{k,0}\zeta^3\right\} \geq cY_k^{-\delta}$. Then $|\zeta - \alpha| \geq c_1 H(\alpha)^{-\theta}$ for some positive constant c_1 and $\theta = \frac{\varphi^2 + \delta/\varphi}{1-\delta}$*

Proof. Let $P(T) = T^3 + pT^2 + qT + r$ be minimal polynomial of α, multiplued by some power of T if needed. Note that $H(P) = H(\alpha)$. Then

$$\left\{y_{k,0}\zeta^3\right\} \leq |y_{k,0}P(\zeta)| + |p| \left\{y_{k,0}\zeta^2\right\} + |q| \left\{y_{k,0}\zeta\right\} \leq CY_k H(\alpha) |\zeta - \alpha| + Y_k^{-1} H(\alpha)$$

for some $C > 0$. Choose the smallest k such that $H(\alpha) \leq c(2C)^{-1} Y_k^{1-\delta}$. When we obtain $|\zeta - \alpha| \geq c(2C)^{-1} Y_k^{-1-\delta} H(\alpha)^{-1}$. This is that we need, because $Y_k \ll H(\alpha)^{\frac{\varphi}{1-\delta}}$.

Note, that lemma gives $\theta = \varphi^2$ for $\delta = 0$.

Lemma 10.1.2. *For k large enough we have $\left\{y_{k,0}\zeta^3\right\} \geq Y_k^{-\varphi^3}$.*

Proof see in [7].

Lemma 10.1.3. *Let ζ be extremal number and $\mathbf{y_k}$ – the corresponding sequence of points. Then there exists $k_0 \geq 1$ and 2×2–matrix M with integer entries, such that $\mathbf{y_{k+2}}$ is rational multiple of $\mathbf{y_{k+1}} M \mathbf{y_k}$ for odd $k \geq k_0$ and rational multiple of $\mathbf{y_{k+1}} M^T \mathbf{y_k}$ for even $k \geq k_0$.*

Proof. We know, that for some k_0 and all $k \geq k_0$ point $\mathbf{y_{k+2}}$ is rational multiple of $\mathbf{y_{k+1}} \mathbf{y_{k-1}}^{-1} \mathbf{y_{k+1}}$ If S is 2×2–matrix such that $\mathbf{y_{k+1}}$ is rational multiple of $\mathbf{y_k} S \mathbf{y_{k-1}}$ this implies, that $\mathbf{y_{k+2}}$ is rational multiple of $\mathbf{y_k} S \mathbf{y_{k+1}}$ and after taking transpose we obtain that we need. So, we must only construct matrix to satisfy condition for $k = k_0$ and everything will be satisfied by induction.

Remark 10.1.3. Note, that if $\det(\mathbf{y_k}) = 1$ for all k, we may choose S (and M) with integer entries.

Let M be nonsymmetric matrix in $GL_2(Z)$. We denote $E(M)$ the set of extremal numbers, such that M satisfies conditions of the last proposition.

For example, extremal number $\zeta = \zeta_{a,b} = [0, a, b, a, a, b, \ldots]$, constructed in previous section, belongs to $E(M)$ for $M = \begin{pmatrix} a & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} b & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} ab+1 & a \\ b & 1 \end{pmatrix}$

Lemma 10.1.4. *Let $\zeta \in E(M), M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ and $\mathbf{y_k}$ – corresponding sequence of points. Then:*

1) $\mathbf{y_{k+2}} = (ay_{k,0} + (b+c)y_{k_1} + dy_{k,2})\mathbf{y_{k+1}} \pm \mathbf{y_{k-1}}$.

2) $y_{k,0}y_{k+1,2} - y_{k,2}y_{k+1,0} = \pm(ay_{k-1,0} - dy_{k-1,2}) \pm (b-c)x_{k-1,1}$.

Proof. Simple calculations. For some optimization of them see [8].

Now we start to prove the main result of Roy's paper:

Theorem 10.2. *1) Let a be positive integer. Then every $\zeta \in E(\begin{pmatrix} a & 1 \\ -1 & 0 \end{pmatrix})$ satisfies $|\zeta - \alpha| \leq cH(\alpha)^{-\varphi^2}$ for some $c > 0$ and all algebraic integers α of degree at most 3.*

2) For every positive integer m the number $\eta = (m + 1 + \zeta_{m,m+2})^{-1} = [0, m+1, m, m+2, m, m, m+2, \ldots]$ belongs to $E(\begin{pmatrix} 1 & 1 \\ -1 & 0 \end{pmatrix})$

Remark 10.2.1. The main idea of second part is not to construct the number exactly (we don't need its exact value, in fact), but to show, that set in the first part is non-empty for some positive a, so, extremal numbers MAY be built that way.

Proof. We start with second part. It is easy to check, that if $\zeta' \in E(M)$ for some $M, C \in GL_2(Z)$ and η is real number with the following property: $(\eta, -1)$ is proportional to $(\zeta', -1)C$, then $\eta \in E(C^T MC)$ with corresponding sequence $\mathbf{x_k} = C^{-1} \mathbf{y_k}^T C^{-1}$.

Now we simply note, that $\zeta \in E(\begin{pmatrix} m(m+2)+1 & m \\ m+2 & 1 \end{pmatrix})$, and that

$$\begin{pmatrix} 0 & -1 \\ -1 & m+1 \end{pmatrix}^T \begin{pmatrix} m(m+2)+1 & m \\ m+2 & 1 \end{pmatrix} \begin{pmatrix} 0 & -1 \\ -1 & m+1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ -1 & 0 \end{pmatrix}$$

In fact, the set $E(\begin{pmatrix} a & 1 \\ -1 & 0 \end{pmatrix})$ is nonempty for every positive integer a, but we don't need it.

Now the first part. So, we fix a, $\zeta \in E(a)$ and corresponding sequence of points, $\mathbf{x_k}$. Let $X_k = \|\mathbf{x_k}\|$ and $\delta_k = \{x_{k,2}\zeta\}$.

We have $\{x_{k,0}\zeta\} \le |x_{k,0}\zeta - x_{k,1}| \ll X_k^{-1}$ and similarly $\{x_{k,1}\zeta\} \ll X_k^{-1}$.

Due to 10.1.4 we obtain

$x_{k+2,2} = ax_{k,0}x_{k+1,2} \pm x_{k-1,2}$ and

$x_{k,0}x_{k+1,2} = x_{k,2}x_{k+1,0} \pm ax_{k-1,0} \pm 2x_{k-1,1}$.

Second gives the estimate $\{x_{k,0}x_{k+1,2}\zeta\} \le X_k\{x_{k+1,0}\zeta\} + a\{x_{k-1,0}\zeta\} + 2\{x_{k-1,1}\zeta\} \ll X_kX_{k+1}^{-1} + X_{k-1}^{-1} \ll X_{k-1}^{-1}$. Combining with the first we obtain $|\delta_{k+2} - \delta_{k-1}| \ll X_{k-1}^{-1}$.

We see, that sequence δ_{3i+j} has limit (θ_j), because $\sum X_{3i+j}^{-1}$ converges. Moreover, X^k grows faster, than geometric progression, hence this sum is equivalent to the first term.

Note, that $\left|\left\{x_{k,0}\zeta^3\right\} - \delta_k\right| \le \left|x_{k,0}\zeta^3 - x_{k,2}\zeta\right| \ll X_k^{-1}$. Hence, the sequence $\left\{x_{3i+j,0}\zeta^3\right\}$ has the same limit, θ_j and $\left|\left\{x_{3i+j,0}\zeta^3\right\} - \theta_j\right| \ll X_{3i+j}^{-1}$.

This gives, that $\theta_j \ne 0$, because due to 10.1.2 $\left\{x_{3i+j,0}\zeta^3\right\} \gg X_{3i+j}^{-1/\varphi^3}$.

So, the sequence $\left\{x_{k,0}\zeta^3\right\}$ (excluding finitely many indexes for which $x_{k,0} = 0$) consists of three sequences, convergent to non-zero points. Hence, it is separated from zero: $\left\{x_{k,0}\zeta^3\right\} \ge c > 0$. Now we are in assumptions of 10.1.1. This completes the proof.

References

[1] A. Ya. Khintchine, *Continuous fractions*, The university of Chicago Press, 1964.

[2] W.M.Schmidt, *Diophantine approximation*, Lecture Notes in Mathematics 785, Springer, Berlin, 1980.

[3] Kantor, Set theory.

[4] H. Davenport and W.M.Schmidt, *approximation to real numbers by quadratic irrationals*, Acta Arith. 13 (1967), pp.169-176.

[5] E.Wirsing, *Approximation mit algebraischen Zahlen beschränkten grades*, Journ. Math. 206 (1960), pp.67-77.

[6] H. Davenport and W.M.Schmidt, *A theorem on linear forms*, Acta Arith. 14 (1968), pp.208-223.

[7] D. Roy, *Approximation to real numbers by cubic algebraic integers. I*, Proc. London Math. Soc. (3) 88 (2004), pp.42-62.

[8] D. Roy, *Approximation to real numbers by cubic algebraic integers. II*, Ann.Math. 158 (2003), pp.1081-1087.

[9] H.Davenport and W.M.Schmidt, *Approximation to real numbers by algebraic integers*, Acta Arith, 15 (1969), pp.393-416.

[10] V.G.Sprindžuk, *Mahler's Problem in Metric Number Theory*, Izdat. "Nauka i Tehnika", Minsk,1967, in russian.

[11] K.I.Tishenko *On approximation of real numbers by algebraic numbers of bounded degree*, Acta Arith. 94 (2000) pp.1-24.

[12] K.I.Tishenko *On some special cases of Wirsing conjecture*, Vestsī Nats. Akad. Navuk Belarusī Ser. Fīz.Mat.Navuk. №3 (2000) pp.47-52 (in russian).

[13] K.I.Tishenko *On new methods in problem of approximations of real numbers by algebraic numbers of degree at most three*, Vestsī Nats. Akad. Navuk Belarusī Ser. Fīz.Mat.Navuk. №4 (2000) pp.26-31 (in russian).

[14] K.I.Tishenko *On using of linearly independent polynomials in Wirsing problem*, Dokl. Nats. Akad. Navuk Belarusī. 44-5 (2000) pp 34-36 (in russian).