ACTA ARITHMETICA
LXXXIX.2 (1999)

On the number of good approximations of algebraic numbers by algebraic numbers of bounded degree

by

HELMUT LOCHER (Marburg)

1. Introduction. Let α be an algebraic number. Roth's celebrated theorem [13] says that for any $\delta > 0$ there are only a finite number of rational approximations x/y of α with

$$|\alpha - x/y| < 1/y^{2+\delta}, \quad y > 0. \tag{1.1}$$

In this paper we consider approximations of α by algebraic numbers of bounded degree. More precisely, let $d \in \mathbb{N}$ and suppose $\mu > 2$. We look for solutions in algebraic numbers β of degree $\leq d$ of the inequality

$$|\alpha - \beta| < H_0(\beta)^{-\mu}, \tag{1.2}$$

where $H_0(\beta)$ denotes the maximum modulus of the coefficients of the minimal defining polynomial of α over $\mathbb{Z}$. For rational β, say $\beta = x/y$, we have $H_0(\beta) = \max\{|x|, |y|\}$ and hence for $d = 1$ the inequality (1.2) is essentially equivalent to (1.1).

Wirsing [19] proved that (1.2) *has for*

$$\mu > 2d \tag{1.3}$$

only a finite number of solutions.

As a consequence of his famous subspace theorem W. M. Schmidt [15] was able to prove the best possible result ([16], p. 278): (1.2) *has for*

$$\mu > d + 1 \tag{1.4}$$

only a finite number of solutions.

Unfortunately, the underlying method of Thue–Siegel–Roth is ineffective in the sense that it does not provide upper bounds for y or $H_0(\beta)$ respectively. However, it allows giving an explicit upper bound for the *number* of $x/y \in \mathbb{Q}$ satisfying (1.1). A first result was proved by Davenport and Roth ([3], 1955). This bound was improved by Bombieri and van der Poorten ([1], 1987) and independently by Luckhardt ([10], 1989) using the modified proof

1991 *Mathematics Subject Classification*: Primary 11J68.

of Roth's Theorem presented by Esnault and Viehweg ([4], 1984). The latest results are due to Evertse ([7], 1996, [8], 1998).

It is the purpose of this paper to prove such a quantitative result of Wirsing's theorem.

To state our theorems we have to define the height of an algebraic number. Let K be a number field and $M(K)$ its set of places. For $v \in M(K)$ denote by $|\cdot|_v$ the associated absolute value, normalized so that on $\mathbb{Q}$ we have $|\cdot|_v = |\cdot|$ (standard absolute value) if v is archimedean, whereas for v non-archimedean $|p|_v = p^{-1}$ if v lies above the rational prime p. We put

$$\|\cdot\|_v = |\cdot|_v^{[K_v:\mathbb{Q}_p]/[K:\mathbb{Q}]},$$

where K_v denotes the completion of $(K, |\cdot|_v)$ and $\mathbb{Q}_p$ denotes the completion of $(\mathbb{Q}, |\cdot|_p)$. We also denote the unique extensions of $|\cdot|_v$ and $\|\cdot\|_v$ to $\overline{K_v}$ by $|\cdot|_v$ and $\|\cdot\|_v$ respectively. For $x \in K$ we define the *height of* x by

$$H(x) = \prod_{v \in M(K)} \max\{1, \|x\|_v\}. \tag{1.5}$$

Let $|\cdot|$ denote the standard absolute value of the complex numbers $\mathbb{C}$, and $\overline{\mathbb{Q}}$ the algebraic closure of $\mathbb{Q}$ in $\mathbb{C}$. For any positive number x we define $\log^+ x = \log x$ if $x \geq e$ and 1 otherwise.

The following is a quantitative version of the result (1.3) of Wirsing ([19], Theorem 1).

THEOREM 1. *Let $0 < \delta \leq 1$, $d \in \mathbb{N}$ and α be an algebraic number of degree f. Consider the inequality*

$$|\alpha - \beta| < H(\beta)^{-2d^2-\delta} \tag{1.6}$$

to be solved in elements $\beta \in \overline{\mathbb{Q}}$ with

$$\deg \beta \leq d. \tag{1.7}$$

(i) *There are at most*

$$e^{26} \cdot \frac{d^{15}\log(6f)}{\delta^5} \log \frac{d\log(6f)}{\delta}$$

solutions $\beta \in \overline{\mathbb{Q}}$ of (1.6) *and* (1.7) *with* $H(\beta) \geq \max\{4^{4d^2/\delta}, H(\alpha)\}$.

(ii) *There are at most*

$$\frac{\log^+ \log H(\alpha)}{\log(1 + \delta/(4d^2))} + \frac{2^{15d^2}}{\delta}$$

solutions $\beta \in \overline{\mathbb{Q}}$ of (1.6) *and* (1.7) *with* $H(\beta) < \max\{4^{4d^2/\delta}, H(\alpha)\}$.

We suppose every number field to be embedded in $\overline{\mathbb{Q}}$ and every valuation of the number field to be extended to $\overline{\mathbb{Q}}$. The following generalizes Theorem 1 to include non-archimedean primes.

THEOREM 2. *Let $0 < \delta \le 1$, $d \in \mathbb{N}$ and F/K be an extension of number fields of degree f. Let S be a finite set of places of K of cardinality s. Suppose that for each $v \in S$ we are given a fixed element $\alpha_v \in F$. Let H be a real number with $H \ge H(\alpha_v)$ for all $v \in S$. Consider the inequality*

$$\prod_{v\in S} \min\{1, \|\alpha_v - \beta\|_v\} < H(\beta)^{-2d^2-\delta} \tag{1.8}$$

to be solved in elements $\beta \in \overline{\mathbb{Q}}$ with

$$[K(\beta) : K] \le d. \tag{1.9}$$

Then there are at most

$$e^{7s+19} \cdot \frac{d^{2s+13}\log(6f)}{\delta^{s+4}} \log \frac{d\log(6f)}{\delta}$$

solutions $\beta \in \overline{\mathbb{Q}}$ of (1.8) *and* (1.9) *with*

$$H(\beta) \ge \max\{H, 4^{4d^2/\delta}\}. \tag{1.10}$$

We have claimed above that Theorems 1 and 2 are quantitative versions of Wirsing's result (1.3). But in our theorems we have the exponent $2d^2$ instead of $2d$. The reason is that our height $H(\cdot)$ as defined in (1.5) is normalized in a different way than the height $H_0(\cdot)$ in (1.2). For algebraic numbers β of degree $\le d$ we have ([17], Chapter I, Lemma 7B)

$$H(\beta)^d \ll_d H_0(\beta) \ll_d H(\beta)^d.$$

Therefore we get an additional factor d in the exponent. For the height $H(\cdot)$ the best possible exponent in (1.6) and (1.8) would be $d(d+1)$.

To prove the best possible result Schmidt uses an induction argument which depends upon his subspace theorem. It is not clear how this argument can be used to obtain a quantitative result.

Independently, J.-H. Evertse [8] also proved a quantitative version of Wirsing's theorem (1.3). Moreover, he gave an explicit upper bound for the number of solutions of a more general problem considered by Wirsing [19]. His upper bounds for (1.3) are similar to ours.

2. The auxiliary polynomial

2.1. *A generalization of the index.* Let P be a non-zero polynomial in m variables $X_1, \ldots, X_m$ with complex coefficients. Roth [13] introduced the index of a polynomial at a certain point to measure to what extent the polynomial vanishes at that point. In this section we will define a different measure for this need. It was introduced by W. M. Schmidt ([18], p. 139).

Let $\boldsymbol{\alpha} \in \mathbb{C}^m$ and $\boldsymbol{r} \in \mathbb{N}^m$. We write P in the form

$$P(X_1, \ldots, X_m) = \sum_{\boldsymbol{i}} a_{\boldsymbol{i}}(\boldsymbol{\alpha})(X_1 - \alpha_1)^{i_1} \ldots (X_m - \alpha_m)^{i_m}$$

with $\boldsymbol{i} = (i_1, \ldots, i_m)$ and unique coefficients $a_{\boldsymbol{i}}(\boldsymbol{\alpha})$. Let M be a subset of $\mathbb{R}^m$ and put

$$k_{\boldsymbol{\alpha},\boldsymbol{r}}(P) = \{(i_1/r_1, \ldots, i_m/r_m) : a_{\boldsymbol{i}}(\boldsymbol{\alpha}) \neq 0\}.$$

We say P *is* M*-centered at* $\boldsymbol{\alpha}$ *with respect to* $\boldsymbol{r}$ if $k_{\boldsymbol{\alpha},\boldsymbol{r}}(P) \subseteq M$.

2.2. *Estimation of volumes.* Suppose $0 \le \gamma \le 1$ and $\varepsilon > 0$. Let $m \in \mathbb{N}$. We put

$$\xi_\gamma(\boldsymbol{x}) = |\{h \in \{1, \ldots, m\} : 0 \le x_h \le \gamma\}| = \sum_{h=1}^{m} \chi_{[0,\gamma]}(x_h),$$

where $\chi_{[0,\gamma]}$ denotes the characteristic function of the closed interval $[0, \gamma]$. The sets

$$M_\varepsilon(m, \gamma) = \{\boldsymbol{x} \in [0,1]^m : \xi_\gamma(\boldsymbol{x}) \le m(\gamma + \varepsilon)\}, \qquad M_\varepsilon(m) = \bigcap_{\gamma \in [0,1]} M_\varepsilon(m, \gamma)$$

are the main objects of this section. We always consider the complement of $M_\varepsilon(m, \gamma)$ and $M_\varepsilon(m)$ with respect to $[0,1]^m$. More precisely, we put

$$M_\varepsilon^{\mathrm{c}}(m, \gamma) = [0,1]^m - M_\varepsilon(m, \gamma) \quad \text{and} \quad M_\varepsilon^{\mathrm{c}}(m) = [0,1]^m - M_\varepsilon(m).$$

LEMMA 2.1. *Suppose* $0 \le \gamma \le 1$, $\varepsilon > 0$ *and let* $m \in \mathbb{N}$. *Then*

$$\int\limits_{M_\varepsilon^{\mathrm{c}}(m,\gamma)} d\boldsymbol{x} \le e^{-\gamma(1-\gamma)\varepsilon^2 m}. \tag{2.1}$$

The line of the proof is the same as the proof of [16], Chapter V, Lemma 4C.

P r o o f. The integral on the left-hand side of (2.1) exists, since the boundary of $M_\varepsilon^{\mathrm{c}}(m, \gamma)$ lies in a finite union of hyperplanes. For all $\boldsymbol{x} \in M_\varepsilon^{\mathrm{c}}(m, \gamma)$ we have $\xi_\gamma(\boldsymbol{x}) - m\gamma > m\varepsilon$. Therefore

$$\begin{aligned} (2.2)\quad & e^{\gamma\varepsilon^2 m} \int\limits_{M_\varepsilon^{\mathrm{c}}(m,\gamma)} d\boldsymbol{x} \\ &\le \int\limits_{M_\varepsilon^{\mathrm{c}}(m,\gamma)} e^{\gamma\varepsilon(\xi_\gamma(\boldsymbol{x}) - m\gamma)}\, d\boldsymbol{x} = \int\limits_{M_\varepsilon^{\mathrm{c}}(m,\gamma)} e^{\gamma\varepsilon((\sum_{h=1}^m \chi_{[0,\gamma]}(x_h)) - m\gamma)}\, d\boldsymbol{x} \\ &= \int\limits_{M_\varepsilon^{\mathrm{c}}(m,\gamma)} e^{\gamma\varepsilon \sum_{h=1}^m (\chi_{[0,\gamma]}(x_h) - \gamma)}\, d\boldsymbol{x} \le \int\limits_{[0,1]^m} \prod_{h=1}^{m} e^{\gamma\varepsilon(\chi_{[0,\gamma]}(x_h) - \gamma)}\, d\boldsymbol{x} \\ &= \Big(\int_0^1 e^{\gamma\varepsilon(\chi_{[0,\gamma]}(x) - \gamma)}\, dx\Big)^m. \end{aligned}$$

Note that $e^y \le 1 + y + y^2$ for $|y| \le 1$. Hence we get

$$(2.3)\qquad \int_0^1 e^{\gamma\varepsilon(\chi_{[0,\gamma]}(x)-\gamma)}\,dx$$
$$\le \int_0^1 (1+\gamma\varepsilon(\chi_{[0,\gamma]}(x)-\gamma)+\gamma^2\varepsilon^2(\chi_{[0,\gamma]}(x)-\gamma)^2)\,dx$$
$$\le 1+\gamma\varepsilon\int_0^1 (\chi_{[0,\gamma]}(x)-\gamma)\,dx+\gamma^2\varepsilon^2 = 1+\gamma^2\varepsilon^2.$$

(2.2) and (2.3) together give

$$e^{\gamma\varepsilon^2 m}\int\limits_{M_\varepsilon^c(m,\gamma)} d\boldsymbol{x} \le (1+\gamma^2\varepsilon^2)^m \le e^{\gamma^2\varepsilon^2 m}$$

and the lemma follows.

Lemma 2.2. *Suppose* $0 < \varepsilon \le 2/3$ *and let* $m \in \mathbb{N}$. *Then*

$$\int\limits_{M_\varepsilon^c(m)} d\boldsymbol{x} < 2e^{-(m\varepsilon^3/16+\log\varepsilon)}.$$

Proof. In analogy to Lemma 2.1 the integral on the left-hand side exists since the boundary lies in a finite union of hyperplanes. We put

$$n = \begin{cases} \frac{2}{\varepsilon}(1-\varepsilon) & \text{if } \frac{2}{\varepsilon}(1-\varepsilon)\in\mathbb{N},\\ \left[\frac{2}{\varepsilon}(1-\varepsilon)\right]+1 & \text{otherwise;}\end{cases}$$

$$\gamma_i = i\cdot\frac{\varepsilon}{2}\quad \left(1\le i\le\left[\frac{2}{\varepsilon}(1-\varepsilon)\right]\right),\qquad \gamma_n = 1-\varepsilon.$$

For every $\gamma\in[0,1-\varepsilon]$ there exists an $i\in\{1,\ldots,n\}$ with

$$(2.4)\qquad \gamma_i-\varepsilon/2\le\gamma\le\gamma_i.$$

Next we show that

$$(2.5)\qquad M_\varepsilon(m)\supseteq\bigcap_{i=1}^{n} M_{\varepsilon/2}(m,\gamma_i).$$

Trivially, $\xi_\gamma(\boldsymbol{x})\le m$, and so we have

$$M_\varepsilon(m)=\bigcap_{\gamma\in[0,1]} M_\varepsilon(m,\gamma)=\bigcap_{\gamma\in[0,1-\varepsilon]} M_\varepsilon(m,\gamma).$$

Now let $\gamma\in[0,1-\varepsilon]$. Take $i\in\{1,\ldots,n\}$ satisfying (2.4). Since $\xi_\gamma(\boldsymbol{x})$ is non-decreasing in γ, for all $\boldsymbol{x}\in\bigcap_{j=1}^n M_{\varepsilon/2}(m,\gamma_j)$ we get

$$\xi_\gamma(\boldsymbol{x})\le\xi_{\gamma_i}(\boldsymbol{x})\le m(\gamma_i+\varepsilon/2)\le m(\gamma+\varepsilon/2+\varepsilon/2)=m(\gamma+\varepsilon)$$

and hence $\boldsymbol{x}\in M_\varepsilon(m,\gamma)$. Thus we have verified (2.5).

From (2.5) we get by De Morgan's formulae $M_\varepsilon^{\mathrm{c}}(m) \subseteq \bigcup_{i=1}^n M_{\varepsilon/2}^{\mathrm{c}}(m,\gamma_i)$. We now apply Lemma 2.1 to get

$$\int\limits_{M_\varepsilon^{\mathrm{c}}(m)} d\boldsymbol{x} \le \int\limits_{\bigcup_{i=1}^n M_{\varepsilon/2}^{\mathrm{c}}(m,\gamma_i)} d\boldsymbol{x} \le \sum_{i=1}^n e^{-\gamma_i(1-\gamma_i)\varepsilon^2 m/4}.$$

Since

$$\min_{1\le i\le n} \gamma_i(1-\gamma_i) = \frac{\varepsilon}{2}\left(1-\frac{\varepsilon}{2}\right) \ge \frac{\varepsilon}{4} \quad \text{and} \quad n \le \left[\frac{2}{\varepsilon}(1-\varepsilon)\right] + 1 \le \left[\frac{2}{\varepsilon}\right] - 1 \le \frac{2}{\varepsilon}$$

we conclude

$$\int\limits_{M_\varepsilon^{\mathrm{c}}(m)} d\boldsymbol{x} \le n e^{-\varepsilon^3 m/16} \le \frac{2}{\varepsilon} e^{-\varepsilon^3 m/16} \le 2e^{-(\varepsilon^3 m/16 + \log\varepsilon)}.$$

The following lemma is one of the main reasons for the exponent $2d^2$.

Lemma 2.3 ([18], Lemma 7.2.1). *Let $I_1,\ldots,I_D$ be subsets of $\{1,\ldots,m\}$ and let $\widetilde{d} \in \mathbb{N}$ with $\sum_{k=1}^D |I_k| \ge Dm/\widetilde{d}$. Then*

$$\sum_{k=1}^D \inf\Big\{\sum_{h\in I_k} x_h : \boldsymbol{x} \in M_\varepsilon(m)\Big\} \ge \frac{Dm}{2\widetilde{d}^2}(1 - 2\varepsilon\widetilde{d}^2).$$

Lemma 2.4. *Let $r_1,\ldots,r_m \in \mathbb{N}$. The number of tuples $\boldsymbol{i} \in \mathbb{Z}^m$ with $0 \le i_h \le r_h$ $(1 \le h \le m)$ and $(i_1/r_1,\ldots,i_m/r_m) \notin M_\varepsilon(m)$ is*

$$r_1\ldots r_m \int\limits_{M_\varepsilon^{\mathrm{c}}(m)} d\boldsymbol{x} + O_m\left(\frac{r_1\ldots r_m}{\min_{1\le h\le m} r_h}\right).$$

Proof. We put $\xi_{\gamma,\boldsymbol{r}}(\boldsymbol{x}) = |\{h \in \{1,\ldots,m\} : x_h/r_h \le \gamma\}|$ and

$$\mathcal{M} = \{\boldsymbol{x} \in [0,r_1]\times\ldots\times[0,r_m] : \xi_{\gamma,\boldsymbol{r}}(\boldsymbol{x}) \le m(\gamma+\varepsilon), \forall\gamma\in[0,1]\},$$
$$\mathcal{M}^{\mathrm{c}} = \{\boldsymbol{x} \in [0,r_1]\times\ldots\times[0,r_m] : \exists\gamma\in[0,1] : \xi_{\gamma,\boldsymbol{r}}(\boldsymbol{x}) > m(\gamma+\varepsilon)\}.$$

Observe that $\int_{\mathcal{M}^{\mathrm{c}}} d\boldsymbol{x} = r_1\ldots r_m \int_{M_\varepsilon^{\mathrm{c}}(m)} d\boldsymbol{x}$. We denote by $\mathcal{G}^{\mathrm{c}}$ the set of integer points of $\mathcal{M}^{\mathrm{c}}$, thus

$$\mathcal{G}^{\mathrm{c}} = \{\boldsymbol{i} \in \mathbb{Z}^m : (i_1/r_1,\ldots,i_m/r_m) \notin M_\varepsilon(m),\ 0 \le i_h \le r_h,\ 1 \le h \le m\}.$$

For $\boldsymbol{i} \in \mathbb{Z}^m$ we put

$$\mathcal{Q}_{\boldsymbol{i}} = [i_1, i_1+1]\times\ldots\times[i_m, i_m+1].$$

Now we can write the assertion as

$$\int\limits_{\bigcup_{\boldsymbol{i}\in\mathcal{G}^{\mathrm{c}}} \mathcal{Q}_{\boldsymbol{i}}} d\boldsymbol{x} = \int\limits_{\mathcal{M}^{\mathrm{c}}} d\boldsymbol{x} + O_m\left(\frac{r_1\ldots r_m}{\min_{1\le h\le m} r_h}\right).$$

For $\boldsymbol{x} \in \mathcal{M}^{\mathrm{c}}$ it follows that $([x_1], \ldots, [x_m]) \in \mathcal{M}^{\mathrm{c}}$ and hence $\boldsymbol{x} \in \bigcup_{\boldsymbol{i} \in \mathcal{G}^{\mathrm{c}}} \mathcal{Q}_{\boldsymbol{i}}$. In other words, $\mathcal{M}^{\mathrm{c}} \subseteq \bigcup_{\boldsymbol{i} \in \mathcal{G}^{\mathrm{c}}} \mathcal{Q}_{\boldsymbol{i}}$. Therefore it suffices to show

$$\int\limits_{\bigcup_{\boldsymbol{i}\in\mathcal{G}^{\mathrm{c}}} \mathcal{Q}_{\boldsymbol{i}} - \mathcal{M}^{\mathrm{c}}} d\boldsymbol{x} = O_m\biggl(\frac{r_1 \ldots r_m}{\min_{1 \le h \le m} r_h}\biggr). \tag{2.6}$$

We have

$$\bigcup_{\boldsymbol{i}\in\mathcal{G}^{\mathrm{c}}} \mathcal{Q}_{\boldsymbol{i}} - \mathcal{M}^{\mathrm{c}} = \{\boldsymbol{x} \in [0, r_1] \times \ldots \times [0, r_m] :$$
$$\exists \gamma \in [0,1] : \xi_{\gamma, \boldsymbol{r}}(([x_1], \ldots, [x_m])) > m(\gamma + \varepsilon)\} \cap \mathcal{M}.$$

Let $\boldsymbol{x} \in \bigcup_{\boldsymbol{i}\in\mathcal{G}^{\mathrm{c}}} \mathcal{Q}_{\boldsymbol{i}} - \mathcal{M}^{\mathrm{c}}$. There exists some $\widetilde{\gamma} \in [0,1]$ with

$$|\{h \in \{1, \ldots, m\} : [x_h]/r_h \le \widetilde{\gamma}\}| > m(\widetilde{\gamma} + \varepsilon). \tag{2.7}$$

On the other hand, for all $\gamma \in [0,1]$ we have

$$|\{h \in \{1, \ldots, m\} : x_h/r_h \le \gamma\}| \le m(\gamma + \varepsilon). \tag{2.8}$$

Observe that for all permutations π of $\{1, \ldots, m\}$,

$$(x_{\pi(1)}, \ldots, x_{\pi(m)}) \in \bigcup_{\boldsymbol{i}\in\mathcal{G}^{\mathrm{c}}} \mathcal{Q}_{\boldsymbol{i}} - \mathcal{M}^{\mathrm{c}}. \tag{2.9}$$

Thus additionally we can assume

$$x_1/r_1 \le \ldots \le x_m/r_m \tag{2.10}$$

and therefore we also have

$$[x_1]/r_1 \le \ldots \le [x_m]/r_m. \tag{2.11}$$

We put $\widetilde{h} = |\{h \in \{1, \ldots, m\} : [x_h]/r_h \le \widetilde{\gamma}\}|$. Then (2.7) and (2.11) together give

$$\widetilde{h} > m(\widetilde{\gamma} + \varepsilon) \ge m([x_{\widetilde{h}}]/r_{\widetilde{h}} + \varepsilon). \tag{2.12}$$

If we choose $\gamma = x_{\widetilde{h}}/r_{\widetilde{h}}$, then (2.10) and (2.8) imply

$$\widetilde{h} \le |\{h \in \{1, \ldots, m\} : x_h/r_h \le x_{\widetilde{h}}/r_{\widetilde{h}}\}| \le m(x_{\widetilde{h}}/r_{\widetilde{h}} + \varepsilon). \tag{2.13}$$

The combination of (2.13) and (2.12) gives

$$r_{\widetilde{h}}(\widetilde{h}/m - \varepsilon) \le x_{\widetilde{h}} < r_{\widetilde{h}}(\widetilde{h}/m - \varepsilon) + 1. \tag{2.14}$$

The value $\widetilde{h}$ depends on $\boldsymbol{x}$, but the possible values of $\widetilde{h}$ range between 1 and m, since $\widetilde{h}$ is positive. As $\bigcup_{\boldsymbol{i}\in\mathcal{G}^{\mathrm{c}}} \mathcal{Q}_{\boldsymbol{i}} - \mathcal{M}^{\mathrm{c}} \subseteq [0, r_1] \times \ldots \times [0, r_m]$ we finally conclude from (2.14) that

$$\int\limits_{\{x \in \bigcup_{\boldsymbol{i}\in\mathcal{G}^{\mathrm{c}}} \mathcal{Q}_{\boldsymbol{i}} - \mathcal{M}^{\mathrm{c}} : x_1/r_1 \le \ldots \le x_m/r_m\}} d\boldsymbol{x} \le \sum_{\widetilde{h}=1}^{m} \frac{r_1 \ldots r_m}{r_{\widetilde{h}}} = O_m\biggl(\frac{r_1 \ldots r_m}{\min_{1 \le h \le m} r_h}\biggr).$$

Now (2.6) follows immediately using (2.9).

LEMMA 2.5. *Suppose $\varepsilon > 0$. Let $P \in \mathbb{C}[X_1,\dots,X_m]$, $\boldsymbol{\alpha} \in \mathbb{C}^m$ and $\boldsymbol{r} \in \mathbb{N}^m$. Let $\boldsymbol{j} \in \mathbb{Z}^m$ with $0 \le j_h \le r_h$ $(1 \le h \le m)$ and $j_1/r_1 + \dots + j_m/r_m \le \varepsilon$. Suppose P is $M_\varepsilon(m)$-centered at $\boldsymbol{\alpha}$ with respect to $\boldsymbol{r}$. Then*

$$\frac{\partial^{j_1+\dots+j_m}}{\partial X_1^{j_1}\dots\partial X_m^{j_m}} P$$

is $M_{2\varepsilon}(m)$-centered at $\boldsymbol{\alpha}$ with respect to $\boldsymbol{r}$.

Proof. Since $j_h/r_h \le \varepsilon$ for all $h \in \{1,\dots,m\}$ the lemma is an easy consequence of the definition of the set $M_\varepsilon(m)$.

2.3. *Heights and Siegel's Lemma.* Let K be a number field and $M(K)$ its set of places. Let $n \in \mathbb{N}$. For $\boldsymbol{x} \in K^n$ and $v \in M(K)$ we put

$$|\boldsymbol{x}|_v = \max\{|x_1|_v,\dots,|x_n|_v\} \quad \text{and} \quad \|\boldsymbol{x}\|_v = |\boldsymbol{x}|_v^{[K_v:\mathbb{Q}_p]/[K:\mathbb{Q}]}.$$

If v is archimedean we put

$$|\boldsymbol{x}|_{v,\mathrm{E}} = (|x_1|_v^2 + \dots + |x_n|_v^2)^{1/2} \quad \text{and} \quad \|\boldsymbol{x}\|_{v,\mathrm{E}} = |\boldsymbol{x}|_{v,\mathrm{E}}^{[K_v:\mathbb{Q}_p]/[K:\mathbb{Q}]}.$$

The *height* and the *euclidean height* of $\boldsymbol{x} \in K^n$ are defined by

$$H(\boldsymbol{x}) = \prod_{v\in M(K)} \|\boldsymbol{x}\|_v, \qquad H_{\mathrm{E}}(\boldsymbol{x}) = \Big(\prod_{\substack{v\in M(K)\\ v\mid\infty}} \|\boldsymbol{x}\|_{v,\mathrm{E}}\Big) \prod_{\substack{v\in M(K)\\ v\nmid\infty}} \|\boldsymbol{x}\|_v.$$

We have $H(\boldsymbol{x}) \le H_{\mathrm{E}}(\boldsymbol{x}) \le \sqrt{n}\, H(\boldsymbol{x})$. The *height of a polynomial* is defined as the height of its coefficient vector. We use the notation

$$\Delta^{\boldsymbol{i}} = \frac{1}{i_1!\dots i_m!}\frac{\partial^{i_1+\dots+i_m}}{\partial X_1^{i_1}\dots\partial X_m^{i_m}}.$$

Let $P \in \overline{\mathbb{Q}}[X_1,\dots,X_m]$ have degree $\le r_h$ in X_h $(1 \le h \le m)$. Let $\boldsymbol{j} \in \mathbb{Z}^m$ with $j_h \ge 0$ $(1 \le h \le m)$. We have

$$H(\Delta^{\boldsymbol{j}} P) \le 2^{r_1+\dots+r_m} H(P). \tag{2.15}$$

Finally, we are able to construct the auxiliary polynomial.

LEMMA 2.6. *Suppose $0 < \varepsilon < 1$. Let F/K be an extension of number fields of degree f, let $\alpha_1,\dots,\alpha_s \in F$ and $m \in \mathbb{N}$. Suppose*

$$m \ge \frac{16}{\varepsilon^3}(\log(6sf) + \log\varepsilon^{-1}).$$

There is a constant $R = R(m)$ such that for all $\boldsymbol{r} = (r_1,\dots,r_m) \in \mathbb{N}^m$ with $r_h \ge R$ $(1 \le h \le m)$ there exists a non-zero polynomial $P \in K[X_1,\dots,X_m]$ such that

(i) $\deg_{X_h} P \le r_h$ $(1 \le h \le m)$;

(ii) *P is $M_\varepsilon(m)$-centered at the points $\boldsymbol{\alpha}_k = (\alpha_k,\dots,\alpha_k)$ $(1 \le k \le s)$ with respect to $\boldsymbol{r}$*;

(iii) $H(P) \le C(F)(4H)^{r_1+\ldots+r_m}$, *where* $H = \max\{H(\alpha_1),\ldots,H(\alpha_s)\}$ *and* $C(F)$ *denotes a constant depending only on* F.

Proof. We put $N = (r_1+1)\ldots(r_m+1)$ and

$$M = |\{\boldsymbol{i} \in \mathbb{Z}^m : (i_1/r_1,\ldots,i_m/r_m) \notin M_\varepsilon(m),\ 0 \le i_h \le r_h,\ 1 \le h \le m\}|.$$

Let $P \in K[X_1,\ldots,X_m]$ with (i). We need to determine the coefficients of P such that (ii) and (iii) hold. (ii) says that

$$\Delta^{\boldsymbol{i}} P(\boldsymbol{\alpha}_k) = 0 \tag{2.16}$$

for all $\boldsymbol{i} \in \mathbb{Z}^m$ with $(i_1/r_1,\ldots,i_m/r_m) \notin M_\varepsilon(m)$, $0 \le i_h \le r_h$, $1 \le h \le m$ and $1 \le k \le s$. (2.16) is a system of linear equations, where the unknowns are the coefficients of P. To solve (2.16) we will apply Siegel's Lemma in the form given by Bombieri and Vaaler [2].

Lemma 2.4 says

$$M = r_1\ldots r_m \int\limits_{M_\varepsilon^c(m)} d\boldsymbol{x} + O_m\left(\frac{r_1\ldots r_m}{\min_{1\le h\le m} r_h}\right)$$

and therefore it follows from Lemma 2.2 that

$$\begin{aligned}\frac{M}{N} &= \int\limits_{M_\varepsilon^c(m)} d\boldsymbol{x} + \frac{1}{(r_1+1)\ldots(r_m+1)}\cdot O_m\left(\frac{r_1\ldots r_m}{\min_{1\le h\le m} r_h}\right)\\ &< 2e^{-(m\varepsilon^3/16+\log\varepsilon)} + \frac{1}{(r_1+1)\ldots(r_m+1)}\cdot O_m\left(\frac{r_1\ldots r_m}{\min_{1\le h\le m} r_h}\right).\end{aligned}$$

Hence for large $r_1,\ldots,r_m$ we get

$$M/N < 3e^{-(m\varepsilon^3/16+\log\varepsilon)}. \tag{2.17}$$

By assumption,

$$m \ge \frac{16}{\varepsilon^3}(\log(6sf) + \log\varepsilon^{-1}) = \frac{16}{\varepsilon^3}\log(6sf/\varepsilon).$$

This is equivalent to

$$3sfe^{-(m\varepsilon^3/16+\log\varepsilon)} \le 1/2. \tag{2.18}$$

The inequalities (2.17) and (2.18) together give

$$sfM < Nsf3e^{-(m\varepsilon^3/16+\log\varepsilon)} \le N/2. \tag{2.19}$$

If we denote by A the matrix corresponding to (2.16), then by [2], Theorem 12 and (2.19) we get a non-zero polynomial $P \in K[X_1,\ldots,X_m]$ satisfying (i), (2.16) and

$$H(P) \le C(F)(\max_{\boldsymbol{a}\text{ row of }A} H_{\mathrm{E}}(\boldsymbol{a}))^{sfM/(N-sfM)} \le C(F)\max_{\boldsymbol{a}\text{ row of }A} H_{\mathrm{E}}(\boldsymbol{a}),$$

where $C(F)$ denotes a constant only depending on F. By standard estimates we know that $H_{\mathrm{E}}(\boldsymbol{a}) \le (4H)^{r_1+\ldots+r_m}$ and the lemma follows.

3. Roth's Lemma. The essential ingredient to Roth's Theorem in [13] is the so-called Roth's Lemma. We quote its version proved by J. H. Evertse [6], which is a quantitative improvement on the original. J. H. Evertse proved this result by using Faltings' Product Theorem [9].

Let P be a non-zero polynomial in unknowns $X_1, \ldots, X_m$ with complex coefficients. Let $\boldsymbol{\alpha} \in \mathbb{C}^m$ and $\boldsymbol{r} \in \mathbb{N}^m$. We define, as in [13],

$$\begin{aligned}&\mathrm{Ind}_{\boldsymbol{\alpha},\boldsymbol{r}} P\\ &\quad = \min\{i_1/r_1 + \ldots + i_m/r_m : \Delta^{\boldsymbol{i}} P(\boldsymbol{\alpha}) \neq 0,\ \boldsymbol{i} \in \mathbb{Z}^m,\ i_h \geq 0,\ 1 \leq h \leq m\}\end{aligned}$$

and say that P *has index* $\mathrm{Ind}_{\boldsymbol{\alpha},\boldsymbol{r}} P$ *at* $\boldsymbol{\alpha}$ *with respect to* $\boldsymbol{r}$.

PROPOSITION 3.1 ([6], Theorem 3). *Let m be an integer ≥ 2, let $\boldsymbol{r} = (r_1, \ldots, r_m)$ be a tuple of positive integers, let $P \in \overline{\mathbb{Q}}[X_1, \ldots, X_m]$ be a non-zero polynomial of degree $\leq r_h$ in X_h for $h = 1, \ldots, m$ and let $0 < \varepsilon \leq m+1$ be such that*

$$r_h/r_{h+1} \geq 2m^3/\varepsilon \quad \text{for } h = 1, \ldots, m-1. \tag{3.1}$$

Further, let $\beta_1, \ldots, \beta_m$ be algebraic numbers with

$$H_{\mathrm{E}}((1, \beta_h))^{r_h} > \{e^{r_1 + \ldots + r_m} H_{\mathrm{E}}(P)\}^{(3m^3/\varepsilon)^m} \quad (1 \leq h \leq m). \tag{3.2}$$

Then $\mathrm{Ind}_{\boldsymbol{\beta},\boldsymbol{r}} P < \varepsilon$.

4. A quantitative result. Suppose $0 < \varepsilon < 1$. Let F/K be an extension of number fields of degree f. Let S be a finite subset of $M(K)$ of cardinality s. Suppose that for each $v \in S$ we are given fixed elements $\alpha_v \in F$. Suppose $H(\alpha_v) \leq H$ $(v \in S)$. Let $m \in \mathbb{N}$ with $m \geq (16/\varepsilon^3)(\log(6sf) + \log \varepsilon^{-1})$.

Under these assumptions the hypotheses of Lemma 2.6 are satisfied. Let $R = R(m)$ be the constant given by Lemma 2.6. Suppose $r_h \geq R$ $(1 \leq h \leq m)$. Then there is a polynomial P with

$$P \in K[X_1, \ldots, X_m], \qquad P \neq 0; \tag{4.1}$$

$$\deg_{X_h} P \leq r_h \quad (1 \leq h \leq m); \tag{4.2}$$

(4.3) P is $M_\varepsilon(m)$-centered with respect to $\boldsymbol{r}$ at the points $\boldsymbol{\alpha}_v = (\alpha_v, \ldots, \alpha_v)$ $(v \in S)$;

$$H(P) \leq C(4H)^{r_1 + \ldots + r_m}, \tag{4.4}$$

where $C = C(F)$ is a constant just depending on F.

LEMMA 4.1. *Suppose $0 < \delta \leq 1$, $d \in \mathbb{N}$ and $0 < \varepsilon \leq \delta/(20d^4)$. Let $\boldsymbol{\Gamma}$ be a tuple of non-negative integers with*

$$\sum_{v \in S} \Gamma_v = 1 - \delta/(24d^2).$$

Suppose there are elements $\beta_1, \ldots, \beta_m \in \overline{\mathbb{Q}}$ *satisfying*

$$[K(\beta_h) : K] \le d,$$

$$H(\beta_1)^{r_1} \le H(\beta_h)^{r_h} \le H(\beta_1)^{(1+\varepsilon)r_1} \quad (1 \le h \le m), \tag{4.5}$$

$$\|\alpha_v - \beta_h\|_v < H(\beta_h)^{-\Gamma_v(2d^2+\delta)} \quad (1 \le h \le m,\ v \in S) \tag{4.6}$$

and

$$H(\beta_h)^{\varepsilon/2} \ge \max\{C^{1/r_1}, 2^7 H^{3fs}\} \quad (1 \le h \le m). \tag{4.7}$$

Then $\mathrm{Ind}_{\boldsymbol{\beta},\boldsymbol{r}} P > \varepsilon$.

Proof. Let $\boldsymbol{j} \in \mathbb{Z}^m$ with $0 \le j_h \le r_h$, $1 \le h \le m$ and

$$j_1/r_1 + \ldots + j_m/r_m \le \varepsilon. \tag{4.8}$$

Put

$$T(\boldsymbol{X}) = \sum_{\boldsymbol{i}} a_{\boldsymbol{i}} X_1^{i_1} \ldots X_m^{i_m} = \Delta^{\boldsymbol{j}} P(\boldsymbol{X}). \tag{4.9}$$

We have to show

$$T(\boldsymbol{\beta}) = 0. \tag{4.10}$$

First we establish an inequality for the height of T. From (2.15), (4.2), (4.4) and (4.7) we get

$$(16H^{2fs})^{r_1+\ldots+r_m} H(T) \le (2^5 H^{2fs})^{r_1+\ldots+r_m} H(P) \le C(2^7 H^{3fs})^{r_1+\ldots+r_m}$$
$$\le C\Big(\prod_{h=1}^m H(\beta_h)^{r_h}\Big)^{\varepsilon/2}.$$

By (4.7) we have $C \le H(\beta_1)^{r_1\varepsilon/2} \le \prod_{h=1}^m H(\beta_h)^{r_h\varepsilon/2}$ and therefore

$$(16H^{2fs})^{r_1+\ldots+r_m} H(T) \le \Big(\prod_{h=1}^m H(\beta_h)^{r_h}\Big)^{\varepsilon}. \tag{4.11}$$

We will need (4.11) later on.

Put $E = K(\beta_1, \ldots, \beta_m)$. We denote by $E \overset{K}{\hookrightarrow} \overline{K}_v$ the set of K-embeddings of E into $\overline{K}_v$, i.e. the homomorphisms of E in $\overline{K}_v$ which are the identity on K. For each place w of E which lies over v of K, there exists a $\lambda \in E \overset{K}{\hookrightarrow} \overline{K}_v$ with $|a|_w = |\lambda(a)|_v$ for all $a \in E$. There are in fact $[E_w : K_v]$ such embeddings. With these notations the product formula reads

$$\prod_{p\in M(\mathbb{Q})} \prod_{\substack{w\in M(E)\\ w|p}} |x|_w^{[E_w:\mathbb{Q}_p]} = \prod_{p\in M(\mathbb{Q})} \prod_{\substack{v\in M(K)\\ v|p}} \prod_{\lambda\in E\overset{K}{\hookrightarrow}\overline{K}_v} |\lambda(x)|_v^{[K_v:\mathbb{Q}_p]} = 1$$

for all $x \in E^*$. From (4.1) and (4.9) we know that T has coefficients in K and hence $T(\boldsymbol{\beta}) \in E$. Therefore to prove (4.10) it suffices to show

$$\prod_{p\in M(\mathbb{Q})} \prod_{\substack{v\in M(K)\\ v|p}} \prod_{\lambda\in E\overset{K}{\hookrightarrow}\overline{K}_v} |T(\lambda(\boldsymbol{\beta}))|_v^{[K_v:\mathbb{Q}_p]} < 1. \tag{4.12}$$

Let $v \in M(K)$. In the sequel we estimate $\prod_{\lambda\in E\overset{K}{\hookrightarrow}\overline{K}_v} |T(\lambda(\boldsymbol{\beta}))|_v$. Put

$$\kappa_v = \begin{cases} 1 & \text{if } v \mid \infty, \\ 0 & \text{if } v \nmid \infty \end{cases} \tag{4.13}$$

and $r = r_1 + \ldots + r_m$. For $v \notin S$, by trivial estimates of (4.9) we get

$$\begin{aligned} \prod_{\lambda\in E\overset{K}{\hookrightarrow}\overline{K}_v} |T(\lambda(\boldsymbol{\beta}))|_v &\le \prod_{\lambda\in E\overset{K}{\hookrightarrow}\overline{K}_v} \Big(2^{\kappa_v r} \max_{\boldsymbol{i}} |a_{\boldsymbol{i}}|_v \prod_{h=1}^m |(1,\lambda(\beta_h))|_v^{r_h}\Big) \\ &= 2^{\kappa_v [E:K] r} \max_{\boldsymbol{i}} |a_{\boldsymbol{i}}|_v^{[E:K]} \prod_{\substack{\lambda\in E\overset{K}{\hookrightarrow}\overline{K}_v\\ 1\le h\le m}} |(1,\lambda(\beta_h))|_v^{r_h}. \end{aligned} \tag{4.14}$$

Now, let $v \in S$. We can write (4.6) as

$$\begin{aligned} &|\mu_v(\alpha_v) - \mu_v(\beta_h)|_v \\ &\qquad < H(\beta_h)^{-\Gamma_v(2d^2+\delta)[K:\mathbb{Q}]/[K_v:\mathbb{Q}_p]} \qquad (1 \le h \le m,\ v \in S), \end{aligned} \tag{4.15}$$

where μ_v denotes a fixed K-embedding of $\overline{\mathbb{Q}}$ in $\overline{K}_v$. Let $\lambda \in E \overset{K}{\hookrightarrow} \overline{K}_v$. We expand $T(\boldsymbol{X})$ around the point $\mu_v(\boldsymbol{\alpha}_v) = (\mu_v(\alpha_v), \ldots, \mu_v(\alpha_v))$ in a Taylor series to get

$$|T(\lambda(\boldsymbol{\beta}))|_v \le 2^{\kappa_v r} \max_{\boldsymbol{i}} \Big| \Delta^{\boldsymbol{i}} T(\mu_v(\boldsymbol{\alpha}_v)) \prod_{h=1}^m (\lambda(\beta_h) - \mu_v(\alpha_v))^{i_h} \Big|_v. \tag{4.16}$$

By trivial estimates we get

$$|\Delta^{\boldsymbol{i}} T(\mu_v(\boldsymbol{\alpha}_v))|_v \le 4^{\kappa_v r} \max_{\widetilde{\boldsymbol{i}}} |a_{\widetilde{\boldsymbol{i}}}|_v |(1, \mu_v(\alpha_v))|_v^r. \tag{4.17}$$

The main term we have to look at is $\max^*_{\boldsymbol{i}} \prod_{h=1}^m |\lambda(\beta_h) - \mu_v(\alpha_v)|_v^{i_h}$, where the maximum is taken over all $\boldsymbol{i}$ with $\Delta^{\boldsymbol{i}} T(\mu_v(\boldsymbol{\alpha}_v))_v \ne 0$. By (4.3), (4.8), Lemma 2.5 and $\mu_v(\Delta^{\boldsymbol{i}} T(\boldsymbol{\alpha}_v)) = \Delta^{\boldsymbol{i}} T(\mu_v(\boldsymbol{\alpha}_v))$, for tuples $\boldsymbol{i}$ satisfying $\Delta^{\boldsymbol{i}} T(\mu_v(\boldsymbol{\alpha}_v))_v \ne 0$ we have $(i_1/r_1, \ldots, i_m/r_m) \in M_{2\varepsilon}(m)$. Therefore it suffices to consider the term $\sup_{\boldsymbol{x}\in M_{2\varepsilon}(m)} \prod_{h=1}^m |\lambda(\beta_h) - \mu_v(\alpha_v)|_v^{x_h r_h}$.

If $\lambda(\beta_h) = \mu_v(\beta_h)$, we can estimate the factor satisfying (4.15) nontrivially. Hence we treat the cases $\lambda(\beta_h) = \mu_v(\beta_h)$ and $\lambda(\beta_h) \ne \mu_v(\beta_h)$ separately. Put

$$I_\lambda = \{h \in \{1, \ldots, m\} : \lambda(\beta_h) = \mu_v(\beta_h)\}.$$

We have

$$\sup_{\boldsymbol{x}\in M_{2\varepsilon}(m)}\prod_{h=1}^{m}|\lambda(\beta_h)-\mu_v(\alpha_v)|_v^{x_h r_h} \tag{4.18}$$
$$\leq\Big(\sup_{\boldsymbol{x}\in M_{2\varepsilon}(m)}\prod_{h\notin I_\lambda}|\lambda(\beta_h)-\mu_v(\alpha_v)|_v^{x_h r_h}\Big)$$
$$\times\Big(\sup_{\boldsymbol{x}\in M_{2\varepsilon}(m)}\prod_{h\in I_\lambda}|\lambda(\beta_h)-\mu_v(\alpha_v)|_v^{x_h r_h}\Big).$$

We estimate the first factor of (4.18) trivially and get

$$\sup_{\boldsymbol{x}\in M_{2\varepsilon}(m)}\prod_{h\notin I_\lambda}|\lambda(\beta_h)-\mu_v(\alpha_v)|_v^{x_h r_h} \tag{4.19}$$
$$\leq\prod_{h=1}^{m}2^{\kappa_v r_h}|(1,\lambda(\beta_h))|_v^{r_h}|(1,\mu_v(\alpha_v))|_v^{r_h}.$$

For the second factor of (4.18) we use (4.15) and (4.5) to get

$$\sup_{\boldsymbol{x}\in M_{2\varepsilon}(m)}\prod_{h\in I_\lambda}|\lambda(\beta_h)-\mu_v(\alpha_v)|_v^{x_h r_h}$$
$$=\sup_{\boldsymbol{x}\in M_{2\varepsilon}(m)}\prod_{h\in I_\lambda}|\mu_v(\alpha_v)-\mu_v(\beta_h)|_v^{x_h r_h}$$
$$<\sup_{\boldsymbol{x}\in M_{2\varepsilon}(m)}\prod_{h\in I_\lambda}H(\beta_h)^{-x_h r_h\Gamma_v(2d^2+\delta)[K:\mathbb{Q}]/[K_v:\mathbb{Q}_p]}$$
$$\leq\sup_{\boldsymbol{x}\in M_{2\varepsilon}(m)}H(\beta_1)^{-r_1\Gamma_v(2d^2+\delta)\frac{[K:\mathbb{Q}]}{[K_v:\mathbb{Q}_p]}\sum_{h\in I_\lambda}x_h}$$
$$=H(\beta_1)^{-r_1\Gamma_v(2d^2+\delta)\frac{[K:\mathbb{Q}]}{[K_v:\mathbb{Q}_p]}\inf_{\boldsymbol{x}\in M_{2\varepsilon}(m)}\sum_{h\in I_\lambda}x_h}.$$

Taking the product over all K-embeddings of E into $\overline{K}_v$ gives

$$\prod_{\lambda\in E\overset{K}{\hookrightarrow}\overline{K}_v}\sup_{\boldsymbol{x}\in M_{2\varepsilon}(m)}\prod_{h\in I_\lambda}|\lambda(\beta_h)-\mu_v(\alpha_v)|_v^{x_h r_h} \tag{4.20}$$
$$<H(\beta_1)^{-r_1\Gamma_v(2d^2+\delta)\frac{[K:\mathbb{Q}]}{[K_v:\mathbb{Q}_p]}\sum_{\lambda\in E\overset{K}{\hookrightarrow}\overline{K}_v}\inf_{\boldsymbol{x}\in M_{2\varepsilon}(m)}\sum_{h\in I_\lambda}x_h}.$$

To apply Lemma 2.3 we need a lower bound for $\sum_{\lambda\in E\overset{K}{\hookrightarrow}\overline{K}_v}|I_\lambda|$. Let $\delta_{x,y}$ denote the Kronecker symbol. We have

$$\sum_{\lambda\in E\overset{K}{\hookrightarrow}\overline{K}_v}|I_\lambda|=\sum_{\lambda\in E\overset{K}{\hookrightarrow}\overline{K}_v}|\{h\in\{1,\ldots,m\}:\lambda(\beta_h)=\mu_v(\beta_h)\}|$$
$$=\sum_{\lambda\in E\overset{K}{\hookrightarrow}\overline{K}_v}\sum_{1\leq h\leq m}\delta_{\lambda(\beta_h),\mu_v(\beta_h)}=\sum_{1\leq h\leq m}\sum_{\lambda\in E\overset{K}{\hookrightarrow}\overline{K}_v}\delta_{\lambda(\beta_h),\mu_v(\beta_h)}$$
$$=\sum_{1\leq h\leq m}[E:K(\beta_h)]\geq\sum_{1\leq h\leq m}\frac{[E:K]}{d}=\frac{m[E:K]}{d}.$$

Now we apply Lemma 2.3 to (4.20) with $[E:K]$ in place of D and 2ε in place of ε to get

$$\prod_{\lambda\in E\overset{K}{\hookrightarrow}\overline{K}_v} \sup_{\boldsymbol{x}\in M_{2\varepsilon}(m)} \prod_{h\in I_\lambda} |\lambda(\beta_h)-\mu_v(\alpha_v)|_v^{x_h r_h} < H(\beta_1)^{-\frac{[E:\mathbb{Q}]}{[K_v:\mathbb{Q}_p]}mr_1(1-4\varepsilon d^2)(1+\frac{\delta}{2d^2})\Gamma_v}. \tag{4.21}$$

The combination of (4.18), (4.19) and (4.21) gives

$$\begin{aligned}\prod_{\lambda\in E\overset{K}{\hookrightarrow}\overline{K}_v} \sup_{\boldsymbol{x}\in M_{2\varepsilon}(m)} \prod_{h=1}^{m} |\lambda(\beta_h)-\mu_v(\alpha_v)|_v^{x_h r_h} &< \Big(\prod_{\substack{\lambda\in E\overset{K}{\hookrightarrow}\overline{K}_v\\ 1\le h\le m}} 2^{\kappa_v r_h}|(1,\mu_v(\alpha_v))|_v^{r_h}|(1,\lambda(\beta_h))|_v^{r_h}\Big)\\ &\quad\times H(\beta_1)^{-\frac{[E:\mathbb{Q}]}{[K_v:\mathbb{Q}_p]}mr_1(1-4\varepsilon d^2)(1+\frac{\delta}{2d^2})\Gamma_v}\\ &< (2^{\kappa_v}|(1,\mu_v(\alpha_v))|_v)^{r[E:K]}\Big(\prod_{\substack{\lambda\in E\overset{K}{\hookrightarrow}\overline{K}_v\\ 1\le h\le m}} |(1,\lambda(\beta_h))|_v^{r_h}\Big)\\ &\quad\times H(\beta_1)^{-\frac{[E:\mathbb{Q}]}{[K_v:\mathbb{Q}_p]}mr_1(1-4\varepsilon d^2)(1+\frac{\delta}{2d^2})\Gamma_v}\end{aligned} \tag{4.22}$$

and the combination of (4.16), (4.17) and (4.22) gives

$$\begin{aligned}&\prod_{\lambda\in E\overset{K}{\hookrightarrow}\overline{K}_v} |T(\lambda(\boldsymbol{\beta}))|_v\\ &\quad< \prod_{\lambda\in E\overset{K}{\hookrightarrow}\overline{K}_v} 2^{\kappa_v r}\max_{\boldsymbol{i}}|\Delta^{\boldsymbol{i}}T(\mu_v(\boldsymbol{\alpha}_v))|_v \prod_{h=1}^{m}|\lambda(\beta_h)-\mu_v(\alpha_v)|_v^{i_h}\\ &\quad\le 8^{\kappa_v[E:K]r}\max_{\widetilde{\boldsymbol{i}}}|a_{\widetilde{\boldsymbol{i}}}|_v^{[E:K]}|(1,\mu_v(\alpha_v))|_v^{[E:K]r}\\ &\quad\quad\times \prod_{\lambda\in E\overset{K}{\hookrightarrow}\overline{K}_v} {\max_{\boldsymbol{i}}}^{*}\prod_{h=1}^{m}|\lambda(\beta_h)-\mu_v(\alpha_v)|_v^{i_h}\\ &\quad< 16^{\kappa_v[E:K]r}\max_{\boldsymbol{i}}|a_{\boldsymbol{i}}|_v^{[E:K]}|(1,\mu_v(\alpha_v))|_v^{2[E:K]r}\\ &\quad\quad\times\Big(\prod_{\substack{\lambda\in E\overset{K}{\hookrightarrow}\overline{K}_v\\ 1\le h\le m}}|(1,\lambda(\beta_h))|_v^{r_h}\Big)H(\beta_1)^{-\frac{[E:\mathbb{Q}]}{[K_v:\mathbb{Q}_p]}mr_1(1-4\varepsilon d^2)(1+\frac{\delta}{2d^2})\Gamma_v}.\end{aligned} \tag{4.23}$$

Finally, if we take the product over all valuations of E, then (4.14) and (4.23) together lead to

$$(4.24)\quad \prod_{p\in M(\mathbb{Q})}\prod_{\substack{v\in M(K)\\ v|p}}\Big(\prod_{\lambda\in E\overset{K}{\hookrightarrow}\overline{K}_v}|T(\lambda(\boldsymbol{\beta}))|_v\Big)^{[K_v:\mathbb{Q}_p]}$$
$$< \Big(\prod_{p\in M(\mathbb{Q})}\Big(16^{[E:K]r\sum_{v\in M(K),v|p}\kappa_v[K_v:\mathbb{Q}_p]}\prod_{v\in S,\, v|p}|(1,\mu_v(\alpha_v))|_v^{2[E:K]r[K_v:\mathbb{Q}_p]}\Big)\Big)$$
$$\times\Big(\prod_{p\in M(\mathbb{Q})}\Big(\prod_{\substack{v\in M(K)\\ v|p}}\max_{\boldsymbol{i}}|a_{\boldsymbol{i}}|_v^{[K_v:\mathbb{Q}_p]}\Big)^{[E:K]}\Big)$$
$$\times\Big(\prod_{p\in M(\mathbb{Q})}\prod_{\substack{v\in M(K)\\ v|p}}\Big(\prod_{\substack{\lambda\in E\overset{K}{\hookrightarrow}\overline{K}_v\\ 1\le h\le m}}|(1,\lambda(\beta_h))|_v^{r_h}\Big)^{[K_v:\mathbb{Q}_p]}\Big)$$
$$\times H(\beta_1)^{-[E:\mathbb{Q}]mr_1(1-4\varepsilon d^2)(1+\frac{\delta}{2d^2})\sum_{v\in S}\Gamma_v}.$$

For the middle term of the right-hand side of (4.24) we have

$$(4.25)\quad \Big(\prod_{\substack{p\in M(\mathbb{Q})\\ v\in M(K),v|p}}\max_{\boldsymbol{i}}|a_{\boldsymbol{i}}|_v^{[K_v:\mathbb{Q}_p]}\Big)^{[E:K]}$$
$$\times\prod_{\substack{p\in M(\mathbb{Q})\\ v\in M(K),v|p}}\Big(\prod_{\substack{\lambda\in E\overset{K}{\hookrightarrow}\overline{K}_v\\ 1\le h\le m}}|(1,\lambda(\beta_h))|_v^{r_h}\Big)^{[K_v:\mathbb{Q}_p]}$$
$$=\Big(\prod_{\substack{p\in M(\mathbb{Q})\\ v\in M(K),v|p}}\max_{\boldsymbol{i}}|a_{\boldsymbol{i}}|_v^{[K_v:\mathbb{Q}_p]/[K:\mathbb{Q}]}\Big)^{[E:\mathbb{Q}]}$$
$$\times\prod_{\substack{p\in M(\mathbb{Q})\\ v\in M(K),v|p}}\Big(\prod_{\substack{w\in M(E),w|v\\ 1\le h\le m}}|(1,\beta_h)|_w^{r_h[E_w:K_v]}\Big)^{[K_v:\mathbb{Q}_p]}$$
$$=\Big(\prod_{v\in M(K)}\max_{\boldsymbol{i}}\|a_{\boldsymbol{i}}\|_v\Big)^{[E:\mathbb{Q}]}\Big(\prod_{h=1}^{m}\prod_{w\in M(E)}\|(1,\beta_h)\|_w^{r_h}\Big)^{[E:\mathbb{Q}]}$$
$$=H(T)^{[E:\mathbb{Q}]}\Big(\prod_{h=1}^{m}H(\beta_h)^{r_h}\Big)^{[E:\mathbb{Q}]}.$$

Before we estimate the first term of the right-hand side of (4.24) we make some remarks: For each $v\in S$ there exists some $w_v\in M(F)$ such that $|\mu_v(x)|_v=|x|_{w_v}$ for all $x\in F$, hence $|(1,\mu_v(\alpha_v))|_v\le H(\alpha_v)^{[F:\mathbb{Q}]/[F_{w_v}:\mathbb{Q}_p]}$. Further from (4.13) we have

$$\prod_{p\in M(\mathbb{Q})}16^{[E:K]r\sum_{v\in M(K),v|p}\kappa_v[K_v:\mathbb{Q}_p]}=16^{[E:\mathbb{Q}]r}.$$

Therefore for the first term of (4.24) we get

$$\begin{aligned}
(4.26)\quad & \prod_{p\in M(\mathbb{Q})}\Big(16^{[E:K]r\sum_{v\in M(K)}\kappa_v[K_v:\mathbb{Q}_p]}\prod_{v\in S,\,v\mid p}|(1,\mu_v(\alpha_v))|_v^{2[E:K]r[K_v:\mathbb{Q}_p]}\Big)\\
&\le \Big(16^{[E:\mathbb{Q}]}\prod_{p\in M(\mathbb{Q})}\prod_{v\in S,\,v\mid p}H(\alpha_v)^{2[E:K][K_v:\mathbb{Q}_p][F:\mathbb{Q}]/[F_{w_v}:\mathbb{Q}_p]}\Big)^r\\
&= \Big(16^{[E:\mathbb{Q}]}\prod_{v\in S}H(\alpha_v)^{2[E:K][F:\mathbb{Q}]/[F_{w_v}:K_v]}\Big)^r\\
&= \Big(16^{[E:\mathbb{Q}]}\prod_{v\in S}H(\alpha_v)^{2[E:\mathbb{Q}][F:K]/[F_{w_v}:K_v]}\Big)^r\\
&\le \Big(16\prod_{v\in S}H(\alpha_v)^{2f}\Big)^{[E:\mathbb{Q}]r}\le (16H^{2fs})^{[E:\mathbb{Q}]r}.
\end{aligned}$$

Now we simplify (4.24) using (4.25), (4.26), $\sum_{v\in S}\Gamma_v = 1-\delta/(24d^2)$, (4.11) and (4.5) to

$$\begin{aligned}
(4.27)\quad & \prod_{p\in M(\mathbb{Q})}\prod_{\substack{v\in M(K)\\ v\mid p}}\Big(\prod_{\lambda\in E\overset{K}{\hookrightarrow}\overline{K}_v}|T(\lambda(\boldsymbol{\beta}))|_v\Big)^{[K_v:\mathbb{Q}_p]}\\
&< (16H^{2fs})^{[E:\mathbb{Q}]r}H(T)^{[E:\mathbb{Q}]}\Big(\prod_{h=1}^m H(\beta_h)^{r_h}\Big)^{[E:\mathbb{Q}]}\\
&\quad\times H(\beta_1)^{-[E:\mathbb{Q}]mr_1(1-4\varepsilon d^2)(1+\frac{\delta}{2d^2})\sum_{v\in S}\Gamma_v}\\
&\le \Big(\prod_{h=1}^m H(\beta_h)^{r_h}\Big)^{\varepsilon[E:\mathbb{Q}]}\Big(\prod_{h=1}^m H(\beta_h)^{r_h}\Big)^{[E:\mathbb{Q}]}\\
&\quad\times H(\beta_1)^{-[E:\mathbb{Q}]mr_1(1-4\varepsilon d^2)(1+\frac{\delta}{2d^2})(1-\frac{\delta}{24d^2})}\\
&\le \Big(\prod_{h=1}^m H(\beta_h)^{r_h}\Big)^{[E:\mathbb{Q}](1+\varepsilon)}\Big(\prod_{h=1}^m H(\beta_h)^{r_h}\Big)^{-[E:\mathbb{Q}](1+\varepsilon)^{-1}(1-4\varepsilon d^2)(1+\frac{\delta}{2d^2})(1-\frac{\delta}{24d^2})}\\
&\le \Big(\prod_{h=1}^m H(\beta_h)^{r_h}\Big)^{[E:\mathbb{Q}](1+\varepsilon)^{-1}((1+\varepsilon)^2-(1-4\varepsilon d^2)(1+\frac{\delta}{2d^2})(1-\frac{\delta}{24d^2}))}.
\end{aligned}$$

Since $\varepsilon\le\delta/(20d^4)$ and $\delta\le 1$, by elementary estimates we get

$$(1+\varepsilon)^2-(1-4\varepsilon d^2)\Big(1+\frac{\delta}{2d^2}\Big)\Big(1-\frac{\delta}{24d^2}\Big)<0.$$

Therefore the exponent in (4.27) is negative, hence (4.10) holds true and the lemma follows.

Proposition 4.1. *Suppose $0<\delta\le 1$. Let $d\in\mathbb{N}$. Let F/K be an extension of number fields of degree f and let S be a finite subset of $M(K)$*

of cardinality s*. Suppose that for each* $v \in S$ *we are given fixed elements* $\alpha_v \in F$*. Suppose* $H(\alpha_v) \le H$ *and*

$$m \ge e^{14} d^{13} \log(6sf)/\delta^4.$$

Let $E = 42d^4m^3/\delta$ *and let* $\boldsymbol{\Gamma}$ *be a tuple of non-negative reals with*

$$\sum_{v \in S} \Gamma_v = 1 - \frac{\delta}{24d^2}.$$

Then the heights of algebraic numbers $\beta \in \overline{\mathbb{Q}}$ *with*

$$[K(\beta) : K] \le d, \tag{4.28}$$

$$H(\beta) > (16H)^{2m(60d^4m^3/\delta)^m} \tag{4.29}$$

and

$$\|\alpha_v - \beta\|_v < H(\beta)^{-\Gamma_v(2d^2+\delta)} \qquad (v \in S) \tag{4.30}$$

lie in at most $m-1$ *intervals of the type*

$$H_h \le H(\beta) \le H_{h+1}^E \qquad (1 \le h \le m-1).$$

Proof. Suppose the proposition were false. Let H_1 be the infimum of the heights of $\beta \in \overline{\mathbb{Q}}$ satisfying (4.28)–(4.30). If all the heights of the numbers $\beta \in \overline{\mathbb{Q}}$ satisfying (4.28)–(4.30) were in the interval $H_1 \le H(\beta) \le H_1^E$, the assertion would be correct. Hence there are $\beta \in \overline{\mathbb{Q}}$ satisfying (4.28), (4.30) and $H(\beta) > H_1^E$. Let H_2 be their infimum. Hence $H_1^E \le H_2$.

Continuing in this way we find $H_1, \ldots, H_m$ which are defined as follows:

$$H_1 = \inf\{H(\beta) : \beta \in \overline{\mathbb{Q}} \text{ satisfying } (4.28)\text{–}(4.30)\},$$
$$H_{h+1} = \inf\{H(\beta) : \beta \in \overline{\mathbb{Q}} \text{ satisfying } H_h^E < H(\beta), (4.28), (4.30)\}$$

for $1 \le h \le m-1$. Let $\beta_m \in \overline{\mathbb{Q}}$ satisfy (4.28), (4.30) and $H_{m-1}^E < H(\beta_m)$. By the definition of H_{m-1} there exists a $\beta_{m-1} \in \overline{\mathbb{Q}}$ satisfying (4.28), (4.30) and

$$H(\beta_{m-1})^E < H(\beta_m), \qquad H_{m-2}^E < H(\beta_{m-1}).$$

After $m-2$ analogous steps we have $\beta_1, \ldots, \beta_m \in \overline{\mathbb{Q}}$ satisfying (4.28), (4.30),

$$H(\beta_h)^E < H(\beta_{h+1}) \qquad (1 \le h \le m-1) \tag{4.31}$$

and

$$H(\beta_1) > (16H)^{2m(60d^4m^3/\delta)^m}. \tag{4.32}$$

Put $\varepsilon = \delta/(20d^4)$. By trivial estimates we get

$$\frac{16}{\varepsilon^3}(\log(6sf) + \log \varepsilon^{-1}) \le m. \tag{4.33}$$

Hence the hypotheses of Lemma 2.6 are satisfied. Let $R = R(m)$ be the constant given by Lemma 2.6. Now let $r_1 \in \mathbb{N}$ be so large that

$$H(\beta_1)^{\varepsilon r_1} \ge H(\beta_m), \tag{4.34}$$

$$H(\beta_1)^{r_1} \ge H(\beta_m)^R, \tag{4.35}$$

$$H(\beta_1)^{r_1/2} \ge C^{(60d^4m^3/\delta)^m}, \tag{4.36}$$

where C is the constant of (4.4). For $h = 2, \ldots, m$ put

$$r_h = \left[\frac{r_1 \log H(\beta_1)}{\log H(\beta_h)}\right] + 1. \tag{4.37}$$

From (4.34) it follows that

$$H(\beta_1)^{r_1} \le H(\beta_h)^{r_h} \le H(\beta_1)^{r_1(1+\varepsilon)} \quad (1 \le h \le m). \tag{4.38}$$

Moreover, from (4.38) and (4.35) we get $H(\beta_m)^{r_m} \ge H(\beta_1)^{r_1} \ge H(\beta_m)^R$, hence $r_m \ge R$. By (4.31) and (4.37) the sequence $r_1, \ldots, r_m$ is decreasing and therefore $r_h \ge R$ $(1 \le h \le m)$. Lemma 2.6 gives us a polynomial P satisfying (4.1)–(4.4). The inequalities (4.38) are identical with (4.5) of Lemma 4.1. Hence the hypotheses of Lemma 4.1 are satisfied apart from (4.7). But by (4.32), (4.36) and (4.31) also (4.7) holds true. It follows that

$$\mathrm{Ind}_{\boldsymbol{\beta},\boldsymbol{r}} P > \varepsilon. \tag{4.39}$$

Now we verify the hypotheses of Proposition 3.1. From (4.38) and (4.31) we get

$$\begin{aligned}\frac{r_h}{r_{h+1}} &\ge \frac{1}{r_{h+1}}\left(\frac{r_{h+1}\log H(\beta_{h+1})}{(1+\varepsilon)\log H(\beta_h)}\right) = \frac{\log H(\beta_{h+1})}{(1+\varepsilon)\log H(\beta_h)}\\ &> \frac{E\log H(\beta_h)}{(1+\varepsilon)\log H(\beta_h)} = \frac{E}{1+\varepsilon}.\end{aligned}$$

Using $\varepsilon = \delta/(20d^4)$, $d \ge 1$, $\delta \le 1$ and the definition of E yields

$$\frac{r_h}{r_{h+1}} \ge \frac{E}{1+\frac{\delta}{20d^4}} \ge \frac{20E}{21} = \frac{42m^3}{21\cdot\frac{\delta}{20d^4}} = \frac{2m^3}{\varepsilon}. \tag{4.40}$$

Therefore (3.1) holds true. Put $\widetilde{E} = (3m^3/\varepsilon)$, thus $\widetilde{E} = (60d^4m^3/\delta)^m$. Since $r_1 \ge \ldots \ge r_m$ we have $mr_1 \ge r_1 + \ldots + r_m$. Additionally, from (4.32), (4.36) and (4.31) we get

$$\begin{aligned}(e^{r_1+\ldots+r_m}H_{\mathrm{E}}(P))^{(3m^3/\varepsilon)^m} &= (e^{r_1+\ldots+r_m}H_{\mathrm{E}}(P))^{\widetilde{E}} \le (e^{mr_1}2^{mr_1/2}H(P))^{\widetilde{E}}\\ &\le (4^{mr_1}C(4H)^{mr_1})^{\widetilde{E}} \le C^{\widetilde{E}}(16H)^{mr_1\widetilde{E}}\\ &< H(\beta_1)^{r_1/2}H(\beta_1)^{r_1/2} = H(\beta_1)^{r_1}\\ &\le H(\beta_h)^{r_h} \le H_{\mathrm{E}}((1,\beta_h))^{r_h} \quad (1 \le h \le m).\end{aligned}$$

Hence we have also verified (3.2). Therefore the hypotheses of Proposition 3.1 are satisfied and we get $\mathrm{Ind}_{\beta,r}P < \varepsilon$. This contradicts (4.39) and the proposition follows.

5. Gap principles

5.1. *A gap principle for big solutions*

LEMMA 5.1. *Let δ, F, K, S, s, d and α_v be as in Theorem 2. Suppose $4^{4d^2/\delta} \le A < B$ and $\gamma \ge 1 - \delta/(6d^2)$. Let $\boldsymbol{\Gamma}$ be a tuple of non-negative reals with $\sum_{v\in S} \Gamma_v = \gamma$. There are at most*

$$1 + \frac{\log(\log(B)/\log A)}{\log(1+\delta/(4d^2))}$$

elements $\beta \in \overline{\mathbb{Q}}$ such that

(i) $[K(\beta) : K] \le d$;
(ii) $A \le H(\beta) \le B$;
(iii) $\|\alpha_v - \beta\|_v < H(\beta)^{-\Gamma_v(2d^2+\delta)}$ $(v \in S)$.

Proof. Suppose $H \ge A$. First we show that in an interval of the type

$$H \le H(\beta) \le H^{1+\delta/(4d^2)} \tag{5.1}$$

lies at most one $\beta \in \overline{\mathbb{Q}}$ satisfying (i) and (iii).

Let $\beta_0, \beta_1 \in \overline{\mathbb{Q}}$ satisfy $\beta_0 \ne \beta_1$, (i), (iii) and $H(\beta_i) \ge A$ $(i = 0, 1)$. Without loss of generality we assume $H(\beta_0) \le H(\beta_1)$. Put $E = K(\beta_0, \beta_1)$. As in the proof of Proposition 4.1 we have

$$|\mu_v(\alpha_v) - \mu_v(\beta_i)|_v^{[K_v:\mathbb{Q}_p]/[K:\mathbb{Q}]} < H(\beta_i)^{-\Gamma_v(2d^2+\delta)} \qquad (v \in S,\ i = 0, 1) \tag{5.2}$$

for some fixed K-embedding μ_v of $\overline{\mathbb{Q}}$ in $\overline{K}_v$. Let $v \in S$, say $v \mid p$. Hence $|\mu_v(\cdot)|_v$ is a valuation of E, which is identical to $|\cdot|_p$ on $\mathbb{Q}$. Therefore, there exists some $w_v \in M(E)$ such that $|\cdot|_{w_v} = |\mu_v(\cdot)|_v$. Put $S_E = \{w_v : v \in S\}$. Thus it follows from (5.2) that

$$\begin{aligned}\prod_{w\in S_E} \|\beta_0 - \beta_1\|_w &= \prod_{v\in S} \|\beta_0 - \beta_1\|_{w_v} \\ &= \prod_{v\in S} |\beta_0 - \beta_1|_{w_v}^{[E_{w_v}:\mathbb{Q}_p]/[E:\mathbb{Q}]} \\ &= \prod_{v\in S} |\mu_v(\beta_0) - \mu_v(\beta_1)|_v^{[E_{w_v}:\mathbb{Q}_p]/[E:\mathbb{Q}]} \\ &\le \prod_{v\in S} \big(|(1,2)|_v^{[E_{w_v}:\mathbb{Q}_p]/[E:\mathbb{Q}]} \\ &\quad \times \max_{i=0,1}\{|\mu_v(\beta_i) - \mu_v(\alpha_v)|_v^{[E_{w_v}:\mathbb{Q}_p]/[E:\mathbb{Q}]}\}\big)\end{aligned}$$

$$
\begin{aligned}
&= \prod_{v\in S}\big(|(1,2)|_{w_v}^{[E_{w_v}:\mathbb{Q}_p]/[E:\mathbb{Q}]}\\
&\quad\times \max_{i=0,1}\{|\mu_v(\beta_i)-\mu_v(\alpha_v)|_v^{[K_v:\mathbb{Q}_p]/[K:\mathbb{Q}]}\}^{[E_{w_v}:K_v]/[E:K]}\big)\\
&< \prod_{v\in S}\big(\|(1,2)\|_{w_v}\max_{i=0,1}\{H(\beta_i)^{-\Gamma_v(2d^2+\delta)}\}^{[E_{w_v}:K_v]/[E:K]}\big).
\end{aligned}
$$

Simplifying this further using $H(\beta_0)\le H(\beta_1)$, the trivial estimate $[E_{w_v}:K_v]\ge 1$ and $[E:K]\le d^2$ gives

$$
\begin{aligned}
(5.3)\quad \prod_{w\in S_E}\|\beta_0-\beta_1\|_w &< \prod_{v\in S}\|(1,2)\|_{w_v}H(\beta_0)^{-\Gamma_v(2d^2+\delta)/[E:K]}\\
&\le 2\prod_{v\in S}H(\beta_0)^{-\Gamma_v(2+\delta/d^2)} = 2H(\beta_0)^{-\gamma(2+\delta/d^2)}.
\end{aligned}
$$

On the other hand, from the product formula we get

$$
\begin{aligned}
(5.4)\quad \prod_{w\in S_E}\|\beta_0-\beta_1\|_w &= \Big(\prod_{w\in M(E)-S_E}\|\beta_0-\beta_1\|_w\Big)^{-1}\\
&\ge \Big(\prod_{w\in M(E)-S_E}\|(1,2)\|_w\|(1,\beta_0)\|_w\|(1,\beta_1)\|_w\Big)^{-1}\\
&\ge (2H(\beta_0)H(\beta_1))^{-1}.
\end{aligned}
$$

The inequalities (5.3) and (5.4) give $H(\beta_1) > \frac{1}{4}H(\beta_0)^{\gamma(2+\delta/d^2)-1}$. By elementary estimates using $\gamma\ge 1-\delta/(6d^2)$, $d\ge 1$ and $\delta\le 1$ we see that

$$
\gamma\Big(2+\frac{\delta}{d^2}\Big)-1\ge 1+\frac{\delta}{2d^2}.
$$

Therefore

$$
H(\beta_1) > \tfrac{1}{4}H(\beta_0)^{1+\delta/(2d^2)}.
$$

Since $4^{4d^2/\delta}\le A\le H(\beta_0)$ we have $1/4\ge H(\beta_0)^{-\delta/(4d^2)}$ and finally

$$
H(\beta_1) > H(\beta_0)^{1+\delta/(4d^2)}.
$$

This proves (5.1). The interval $[A,B]$ can be covered by

$$
1+\frac{\log(\log(B)/\log A)}{\log(1+\delta/(4d^2))}
$$

intervals of the type (5.1) and hence the assertion follows.

5.2. *A gap principle for small solutions.* J. Mueller and W. M. Schmidt [12] used the well-ordering of the rational numbers to prove a gap principle for small solutions. In this section we follow this idea using a packing lemma instead.

LEMMA 5.2. *Suppose $F > 1$ and $r > 0$. Let $x_1, \ldots, x_\mu \in \mathbb{C}$ with*

$$|x_i - x_j| \ge r \quad (1 \le i < j \le \mu) \tag{5.5}$$

and

$$|x_i - x_j| \le Fr \quad (1 \le i, j \le \mu). \tag{5.6}$$

Then $\mu \le (2F + 1)^2$.

Proof. Without loss of generality we assume $x_1 = 0$. The open discs with center x_i and radius $r/2$ are pairwise disjoint because of (5.5). By (5.6) they also lie in the open disc with center 0 and radius $Fr + r/2$. Therefore $\mu\pi(r/2)^2 \le \pi(Fr + r/2)^2$ and the lemma follows.

The contraposition of Lemma 5.2 is:

LEMMA 5.3. *Suppose $F > 1$ and $r > 0$. Let $x_1, \ldots, x_\mu \in \mathbb{C}$ with $|x_i - x_j| \ge r$ $(1 \le i < j \le \mu)$. Let $\mu \in \mathbb{N}$ with $F < \frac{1}{2}(\sqrt{\mu} - 1)$. Then there exist x_i, x_j with $|x_i - x_j| > Fr$.*

LEMMA 5.4. *Suppose $0 < \delta \le 1$. Let $d \in \mathbb{N}$ and $\alpha \in \overline{\mathbb{Q}}$. There are at most $2^{15d^2}/\delta$ elements $\beta \in \overline{\mathbb{Q}}$ with*

(i) $\deg \beta \le d$;

(ii) $H(\beta) \le 2^{(2^{d^2}+6)/\delta}$;

(iii) $|\alpha - \beta| < H(\beta)^{-2d^2-\delta}$.

Proof. Let $u \in \mathbb{Z}$ with $u \ge 0$. We denote by $S(u)$ the set of all $\beta \in \overline{\mathbb{Q}}$ satisfying (i), (iii) and $2^u \le H(\beta) < 2^{u+1}$.

First we estimate $|S(u)|$. Without loss of generality $S(u) \ne \emptyset$. The set of algebraic numbers of bounded height is finite, hence $S(u)$ is finite, say $S(u) = \{\beta_1, \ldots, \beta_{\mu(u)}\}$. To make the notations less clumsy we write μ instead of $\mu(u)$. We have

$$\begin{aligned} |\beta_i - \beta_j| &\le 2\max\{|\beta_i - \alpha|, |\alpha - \beta_j|\} \\ &< 2\max\{H(\beta_i)^{-2d^2-\delta}, H(\beta_j)^{-2d^2-\delta}\} \\ &\le 2 \cdot 2^{-2d^2u-\delta u} = 2^{-2d^2u-\delta u+1} \end{aligned} \tag{5.7}$$

for all $i, j \in \{1, \ldots, \mu\}$. Let now $i \ne j$ and put $E = \mathbb{Q}(\beta_i, \beta_j)$. Denote by $|\cdot|_w$ the valuation of E which is the restriction of the standard absolute value of $\mathbb{C}$ on E. Using the product formula we get, in analogy to (5.4),

$$\begin{aligned} |\beta_i - \beta_j| = \|\beta_i - \beta_j\|_w^{[E:\mathbb{Q}]/[E_w:\mathbb{Q}_\infty]} &\ge (2H(\beta_i)H(\beta_j))^{-[E:\mathbb{Q}]/[E_w:\mathbb{Q}_\infty]} \\ &\ge (2 \cdot 2^{2(u+1)})^{-[E:\mathbb{Q}]/[E_w:\mathbb{Q}_\infty]} = 2^{-[E:\mathbb{Q}](2u+3)/[E_w:\mathbb{Q}_\infty]}. \end{aligned}$$

By (i) we have $[E : \mathbb{Q}] \le d^2$ and therefore

$$|\beta_i - \beta_j| \ge 2^{-2d^2 u - 3d^2}$$

for all distinct $i, j \in \{1, \ldots, \mu\}$.

Suppose $\mu > 16$. Put $F = \frac{1}{3}(\sqrt{\mu} - 1)$ and $r = 2^{-2d^2 u - 3d^2}$. Then $F > 1$ and so we can apply Lemma 5.3 to the set $S(u)$. Hence there exist $i, j \in \{1, \ldots, \mu\}$ with

$$(5.8) \qquad |\beta_i - \beta_j| > \tfrac{1}{3}(\sqrt{\mu} - 1)2^{-2d^2 u - 3d^2} > (\sqrt{\mu} - 1)2^{-2d^2 u - 3d^2 - 2}.$$

The inequalities (5.7) and (5.8) together give $\sqrt{\mu} < 2^{-\delta u + 3d^2 + 3} + 1$. Considering our assumption $\mu > 16$ we have in general

$$(5.9) \quad \begin{aligned} |S(u)| = \mu(u) = \mu &\le \max\{16, 2^{-2\delta u + 6d^2 + 6} + 2^{-\delta u + 3d^2 + 4} + 1\} \\ &\le 2^{-\delta u + 13d^2} + 16. \end{aligned}$$

Note

$$\sum_{u=0}^{[(2^{d^2}+6)/\delta]} 2^{-\delta u} < \sum_{u=0}^{\infty} 2^{-\delta u} = (1 - 2^{-\delta})^{-1} < 1 + 2/\delta.$$

Therefore from (5.9) we get

$$\begin{aligned} \sum_{u=0}^{[(2^{d^2}+6)/\delta]} |S(u)| &\le \sum_{u=0}^{[(2^{d^2}+6)/\delta]} (2^{-\delta u + 13d^2} + 16) \\ &\le 2^{13d^2} \sum_{u=0}^{[(2^{d^2}+6)/\delta]} 2^{-\delta u} + \left(\frac{2^{d^2} + 6}{\delta} + 1\right) \cdot 16 \\ &\le 2^{13d^2}\left(1 + \frac{2}{\delta}\right) + \frac{16 \cdot 2^{d^2} + 112}{\delta} < \frac{2^{15d^2}}{\delta} \end{aligned}$$

and this is the assertion.

6. Conclusion. The following lemma goes back to Mahler [11]. We state it in the form of [14], Lemma 5.1, but we have used the estimate [5], (46) instead of [14], (5.9).

LEMMA 6.1. *Let* $1/2 \le \gamma < 1$ *and* $s \in \mathbb{N}$. *Then there exists a subset* $\mathcal{S}$ *of cardinality* $< (e/(1-\gamma))^{s-1}$ *of* $\{(\Gamma_1, \ldots, \Gamma_s) \in \mathbb{R}^s_{\ge 0} : \Gamma_1 + \ldots + \Gamma_s = \gamma\}$ *with the following property*: *For every* $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_s) \in \mathbb{R}^s$ *having* $\xi_i \ge 0$ *for each* i $(1 \le i \le s)$, *there exists* $\boldsymbol{\Gamma} \in \mathcal{S}$ *such that for each* i $(1 \le i \le s)$,

$$(6.1) \qquad \xi_i \ge \Gamma_i(\xi_1 + \ldots + \xi_s).$$

6.1. *Proof of Theorem 2.* Let $\beta \in \overline{\mathbb{Q}}$ satisfy (1.8)–(1.10). For each $v \in S$ we define $\xi_v(\beta) \ge 0$ through

$$(6.2) \qquad \min\{1, \|\alpha_v - \beta\|_v\} = H(\beta)^{-\xi_v(\beta)(2d^2+\delta)}.$$

Since

$$H(\beta)^{-(\sum_{v\in S}\xi_v(\beta))(2d^2+\delta)} = \prod_{v\in S}\min\{1, \|\alpha_v-\beta\|_v\} < H(\beta)^{-(2d^2+\delta)}$$

it follows immediately that

$$\sum_{v\in S}\xi_v(\beta) > 1. \tag{6.3}$$

Put $\gamma = 1-\delta/(24d^2)$. Lemma 6.1 says that there exists a subset $\mathcal{S}$ of $\{(\mathit{\Gamma}_v)_{v\in S}\subseteq\mathbb{R}^s_{\geq 0} : \sum_{v\in S}\mathit{\Gamma}_v = \gamma\}$ with

$$|\mathcal{S}| \leq \left(\frac{e}{1-\gamma}\right)^{s-1} = \left(\frac{24ed^2}{\delta}\right)^{s-1} < \left(\frac{66d^2}{\delta}\right)^{s-1} \tag{6.4}$$

and the following property: For each $(\xi_v)_{v\in S}\subseteq\mathbb{R}^s_{\geq 0}$ there exists a tuple $\boldsymbol{\Gamma}\in\mathcal{S}$ with

$$\xi_v \geq \mathit{\Gamma}_v\Big(\sum_{w\in S}\xi_w\Big) \qquad (v\in S). \tag{6.5}$$

Let $(\mathit{\Gamma}_v(\beta))_{v\in S}$ be such a tuple for $(\xi_v(\beta))_{v\in S}$. We divide the elements $\beta\in\overline{\mathbb{Q}}$ satisfying (1.8)–(1.10) into classes as follows: β and $\widetilde{\beta}$ are in the same class if $\boldsymbol{\Gamma}(\beta) = \boldsymbol{\Gamma}(\widetilde{\beta})$. By (6.4), there are at most

$$(66d^2/\delta)^{s-1} \tag{6.6}$$

such classes. We now fix one class, i.e. let $\boldsymbol{\Gamma}\in\mathcal{S}$ be fixed and let $\mathcal{B}$ be the set of all $\beta\in\overline{\mathbb{Q}}$ satisfying (1.8)–(1.10) and $\boldsymbol{\Gamma}(\beta) = \boldsymbol{\Gamma}$. Put $\widetilde{S} = \{v\in S : \mathit{\Gamma}_v > 0\}$ and $\widetilde{s} = |\widetilde{\mathcal{S}}|$. Observe that $1\leq\widetilde{s}\leq s$ and $\sum_{v\in\widetilde{S}}\mathit{\Gamma}_v = \gamma$. Let $\beta\in\mathcal{B}$. If $\xi_v(\beta) = 0$ for some $v\in S$, we conclude from (6.5) and (6.3) that $\mathit{\Gamma}_v(\beta) = \mathit{\Gamma}_v = 0$, and hence $v\notin\widetilde{S}$. Therefore $\xi_v(\beta) > 0$ for all $v\in\widetilde{S}$. Again from (6.3) and (6.5) we get $\xi_v(\beta) > \mathit{\Gamma}_v(\beta) = \mathit{\Gamma}_v$ $(v\in\widetilde{S})$. By (6.2) this implies

$$\|\alpha_v-\beta\|_v < H(\beta)^{-\mathit{\Gamma}_v(2d^2+\delta)} \qquad (v\in\widetilde{S}). \tag{6.7}$$

Put

$$m = [e^{14}d^{13}\log(6sf)/\delta^4]+1. \tag{6.8}$$

Then $m\geq e^{14}d^{13}\log(6\widetilde{s}f)/\delta^4$ and we can apply Proposition 4.1: either

$$\log H(\beta)\leq 2m(60d^4m^3/\delta)^m\log(16H)$$

or $H(\beta)$ lies in a union of $m-1$ intervals of the type

$$H_h\leq H(\beta)\leq H_{h+1}^{42d^4m^3/\delta} \qquad (1\leq h\leq m-1).$$

In the latter case we count the using Lemma 5.1. In each one of the intervals, the number of $\beta\in\overline{\mathbb{Q}}$ satisfying (1.8)–(1.10) is bounded by $1+$

$\log(42d^4m^3/\delta)/\log(1+\delta/(4d^2))$. By (6.8) and

$$\log\left(1+\frac{\delta}{4d^2}\right) > \frac{\delta}{5d^2} \tag{6.9}$$

this is $< \frac{707d^2}{\delta}\log\frac{d\log(6sf)}{\delta}$. Therefore, the number of elements of $\beta \in \mathcal{B}$ with

$$\log H(\beta) > 2m(60d^4m^3/\delta)^m \log(16H)$$

is bounded by

$$(m-1)\frac{707d^2}{\delta}\log\frac{d\log(6sf)}{\delta} < \frac{e^{21}d^{15}\log(6sf)}{\delta^5}\log\frac{d\log(6sf)}{\delta}. \tag{6.10}$$

By Lemma 5.1 the number of elements $\beta \in \overline{\mathbb{Q}}$ satisfying (1.9), (6.7) and

$$\max\{H, 4^{4d^2/\delta}\} \le H(\beta) \le (16H)^{2m(60d^4m^3/\delta)^m} \tag{6.11}$$

is bounded by

$$1 + \frac{\log(2m(60d^4m^3/\delta)^m \log(16H)\log^{-1}(\max\{H, 4^{4d^2/\delta}\}))}{\log(1+\delta/(4d^2))}. \tag{6.12}$$

Note that

$$\log(16H)\log^{-1}\max\{H, 4^{4d^2/\delta}\} \le \frac{\log(16H)}{\log\max\{16^2, H\}} \le \frac{1}{2}+1 = \frac{3}{2} < 2.$$

This together with (6.9) and (6.8) implies that (6.12) is less than

$$\begin{aligned}(6.13)\qquad 1 + \frac{5d^2}{\delta}\log\left(4m\left(\frac{60d^4m^3}{\delta}\right)^m\right) \\ &\le \frac{5d^2}{\delta}\log\left(\left(\frac{dm}{\delta}\right)^{4m}\right) \\ &\le \frac{20e^{15}d^{15}\log(6sf)}{\delta^5}\log\frac{e^{15}d^{14}\log(6sf)}{\delta^5} \\ &< \frac{e^{22}d^{15}\log(6sf)}{\delta^5}\log\frac{d\log(6sf)}{\delta}.\end{aligned}$$

Therefore the number of elements $\beta \in \mathcal{B}$ satisfying (6.11) is bounded by (6.13). The cardinality of $\mathcal{B}$ is bounded by the sum of (6.10) and (6.13). It is less than

$$\frac{e^{23}d^{15}\log(6sf)}{\delta^5}\log\frac{d\log(6sf)}{\delta}.$$

Finally, we conclude from (6.6) that the number of $\beta \in \overline{\mathbb{Q}}$ satisfying (1.8)–(1.10) is less than

$$\left(\frac{66d^2}{\delta}\right)^{s-1}\frac{e^{23}d^{15}\log(6sf)}{\delta^5}\log\frac{d\log(6sf)}{\delta}$$
$$\le \frac{e^{23}66^{s-1}s^2d^{2s+13}\log(6f)}{\delta^{s+4}}\log\frac{d\log(6f)}{\delta}$$
$$< e^{7s+19}\cdot\frac{d^{2s+13}\log(6f)}{\delta^{s+4}}\log\frac{d\log(6f)}{\delta}.$$

This is our assertion.

6.2. *Proof of Theorem 1.* We apply Theorem 2 with $K = \mathbb{Q}$, $F = \mathbb{Q}(\alpha)$, $H = H(\alpha)$ and $S = \{\infty\}$. Hence $s = 1$. In this situation the inequality (1.6) is identical with (1.8). Therefore (1.6) has at most

$$e^{26}\cdot\frac{d^{15}\log(6f)}{\delta^5}\log\frac{d\log(6f)}{\delta} \tag{6.14}$$

solutions in algebraic numbers β of degree $\le d$ with

$$H(\beta)\ge\max\{4^{4d^2/\delta}, H(\alpha)\}.$$

This proves (i). Since $4^{4d^2/\delta}\le 2^{(2^{d^2}+6)/\delta}$, we can estimate the solutions with $H(\beta)\le 4^{4d^2/\delta}$ by Lemma 5.4: there are at most

$$2^{15d^2}/\delta \tag{6.15}$$

such solutions. If $\max\{4^{4d^2/\delta}, H(\alpha)\} = H(\alpha)$ we have to count the solutions $\beta\in\overline{\mathbb{Q}}$ of (1.6), (1.7) and

$$2^{(2^{d^2}+6)/\delta}\le H(\beta)\le H(\alpha).$$

By Lemma 5.1 the number of those solutions is bounded by

$$1+\frac{\log(\log H(\alpha)/\log 2^{(2^{d^2}+6)/\delta})}{\log(1+\delta/(4d^2))}\le\frac{\log^+\log H(\alpha)}{\log(1+\delta/(4d^2))}. \tag{6.16}$$

The estimates (6.15) and (6.16) show the claimed bound of (ii), and the theorem is proved.

References

[1] E. Bombieri and A. J. van der Poorten, *Some quantitative results related to Roth's Theorem*, Macquarie Math. Reports, Report No. 87-0005, February 1987.

[2] E. Bombieri and J. Vaaler, *On Siegel's Lemma*, Invent. Math. 73 (1983), 11–32.

[3] H. Davenport and K. F. Roth, *Rational approximations to algebraic numbers*, Mathematika 2 (1955), 160–167.

[4] H. Esnault and E. Viehweg, *Dyson's Lemma for polynomials in several variables* (*and the Theorem of Roth*), Invent. Math. 78 (1984), 445–490.

[5] J.-H. Evertse, *On equations in S-units and the Thue–Mahler equation*, ibid. 75 (1984), 561–584.

[6] J.-H. Evertse, *An explicit version of Faltings' Product Theorem and an improvement of Roth's lemma*, Acta Arith. 73 (1995), 215–248.
[7] —, *An improvement of the quantitative Subspace theorem*, Compositio Math. 101 (1996), 225–311.
[8] —, *The number of algebraic numbers of given degree approximating a given algebraic number*, in: Analytic Number Theory, Y. Motohashi (ed.), London Math. Soc. Lecture Notes Ser. 247, Cambridge Univ. Press, 1998, 53–83.
[9] G. Faltings, *Diophantine approximation on abelian varieties*, Ann. of Math. 133 (1991), 549–576.
[10] H. Luckhardt, *Herbrand-Analysen zweier Beweise des Satzes von Roth: polynomiale Anzahlschranken*, J. Symbolic Logic 54 (1989), 234–263.
[11] K. Mahler, *Zur Approximation algebraischer Zahlen I.* (*Über den größten Primteiler binärer Formen*), Math. Ann. 107 (1933), 691–730.
[12] J. Mueller and W. M. Schmidt, *On the number of good rational approximations to algebraic numbers*, Proc. Amer. Math. Soc. 106 (1987), 859–866.
[13] K. F. Roth, *Rational approximations to algebraic numbers*, Mathematika 2 (1955), 1–20.
[14] H. P. Schlickewei, *The quantitative Subspace Theorem for number fields*, Compositio Math. 82 (1992), 245–273.
[15] W. M. Schmidt, *Simultaneous approximation to algebraic numbers by rationals*, Acta Math. 125 (1970), 189–201.
[16] —, *Diophantine Approximations*, Lecture Notes in Math. 785, Springer, 1980.
[17] —, *Diophantine Approximations and Diophantine Equations*, Lecture Notes in Math. 1467, Springer, 1991.
[18] K. B. Stolarsky, *Algebraic Numbers and Diophantine Approximation*, Dekker, 1974.
[19] E. Wirsing, *On approximations of algebraic numbers by algebraic numbers of bounded degree*, in: Number Theory Institute 1969, Proc. Sympos. Pure Math. 20, Amer. Math. Soc., 1971, 213–247.

Fachbereich Mathematik und Informatik
Philipps-Universität Marburg
Hans-Meerwein-Straße
35032 Marburg, Germany
E-mail: locher@mathematik.uni-marburg.de

Received on 10.3.1998
and in revised form on 10.11.1998 (3349)