Chapter 5

Linear forms in logarithms

Literature:

A. Baker, Transcendental Number Theory, Cambridge University Press, 1975.
T.N. Shorey, R. Tijdeman, Exponential Diophantine equations, Cambridge University Press, 1986; reprinted 2008.

5.1 Lower bounds for linear forms in logarithms

We fix $\varphi_0 \in \mathbb{R}$, and define the complex logarithm by $\log z = \log|z| + i\arg z$ with $\varphi_0 < \arg z \leqslant \varphi_0 + 2\pi$. In the results below, the choice of φ_0 does not matter.

We recall Baker's transcendence result from the previous chapter.

Theorem 5.1 (A. Baker, 1966)**.** *Let* $\alpha_1, \ldots, \alpha_m \in \overline{\mathbb{Q}} \backslash \{0, 1\}$, $\gamma \in \overline{\mathbb{Q}}$ *and* $\beta_1, \ldots, \beta_m \in \overline{\mathbb{Q}} \setminus \{0\}$*. Assume that*

$$\log \alpha_1, \ldots, \log \alpha_m \text{ are linearly independent over } \mathbb{Q}.$$

Then $\gamma + \beta_1 \log \alpha_1 + \cdots + \beta_m \log \alpha_m \neq 0$*.*

One may ask about quantitative versions of this theorem, i.e., can we give a strictly positive lower bound for the absolute value of the left-hand side? In 1967, Baker indeed obtained such a lower bound, which we conveniently refer to as a 'lower bound for a linear form in logarithms'. Baker's lower bound turned out to be an

extremely powerful tool, not only in transcendence theory, but also in applications which do not have anything to do with transcendence, such as Diophantine equations and Gauss' class number 1 problem. For this reason, Baker's lower bound from 1967 was improved by Baker himself and others. We will give some applications to certain Diophantine equations.

We recall a lower bound for linear forms in logarithms by Baker from 1975. Recall that the height $H(\alpha)$ of $\alpha \in \overline{\mathbb{Q}}$ is the maximum of the absolute values of the coefficients of the primitive minimal polynomial F_α of α.

Theorem 5.2 (A. Baker, 1975)**.** *Let $\alpha_1, \ldots, \alpha_m \in \overline{\mathbb{Q}} \setminus \{0, 1\}$ and $\gamma, \beta_1, \ldots, \beta_m \in \overline{\mathbb{Q}}$. Assume that*

$$\Lambda := \gamma + \beta_1 \log \alpha_1 + \cdots + \beta_m \log \alpha_m \neq 0.$$

Then

$$|\Lambda| \geqslant (eB)^{-C} \quad (e = 2.7182\ldots)$$

where $B = \max(H(\gamma), H(\beta_1), \ldots, H(\beta_m))$, and where C is an effectively computable positive number depending on m, the degrees and heights of $\alpha_1, \ldots, \alpha_m$, and φ_0.

The assertion that C is effectively computable means that by going through the proof of Theorem 5.2 one can compute an explicit value of C.

For our applications, we restrict ourselves to the case that $\gamma = 0$ and $\beta_i = b_i \in \mathbb{Z}$ for $i = 1, \ldots, m$. In that case, we can get rid of the logarithms.

Corollary 5.3. *Let $\alpha_1, \ldots, \alpha_m \in \overline{\mathbb{Q}} \setminus \{0, 1\}$ and let $b_1, \ldots, b_m \in \mathbb{Z}$ such that*

$$\alpha_1^{b_1} \cdots \alpha_m^{b_m} \neq 1.$$

Then

$$|\alpha_1^{b_1} \cdots \alpha_m^{b_m} - 1| \geqslant (eB)^{-C'},$$

where $B := \max(|b_1|, \ldots, |b_m|)$ and where C' is an effectively computable number depending only on m and on the degrees and heights of $\alpha_1, \ldots, \alpha_m$.

Proof. For the logarithm of a complex number z we choose $\log z = \log |z| + i \arg z$ with $-\pi < \arg z \leqslant \pi$. With this choice of log we have

$$\log(1 + z) = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n} \cdot z^n \quad \text{for } z \in \mathbb{C} \text{ with } |z| < 1.$$

Using this power series expansion, one easily shows that

$$\log(1+z)| \leqslant |z|(1+|z|+|z|^2+\cdots) \leqslant 2|z| \quad \text{for } z \in \mathbb{C} \text{ with } |z| \leqslant \tfrac{1}{2}.$$

We apply this with $z := \alpha_1^{b_1} \cdots \alpha_m^{b_m} - 1$. If $|z| > 1/2$ we are done, so we suppose that $|z| \leqslant 1/2$. We have to estimate from below $|\log(1+z)|$.

Recall that the complex logarithm is additive only modulo $2\pi i$. That is,

$$\log(1+z) = b_1 \log \alpha_1 + \cdots + b_m \log \alpha_m + 2k\pi i = b_1 \log \alpha_1 + \cdots + b_m \log \alpha_m + 2k \log(-1)$$

for some $k \in \mathbb{Z}$, since $\log(-1) = \pi i$. Applying Theorem 5.2 we get

$$|\log(1+z)| \geqslant \big(e \max(B, |2k|)\big)^{-C_1}$$

where C_1 is an effectively computable constant depending only on m and $\alpha_1, \ldots, \alpha_m$. Since $|\log(1+z)| \leqslant 2|z| \leqslant 1$ we have

$$|2k\pi i| \leqslant 1 + \sum_{j=1}^{m} |\log \alpha_j| \cdot |b_j| \leqslant \Big(1 + \sum_{j=1}^{m} |\log \alpha_j|\Big) \cdot B = C_2 B,$$

say. Hence $|2k| \leqslant C_2 B$ and so $|\log(1+z)| \geqslant (eC_2B)^{-C_1}$. This implies

$$|z| \geqslant \tfrac{1}{2}(eC_2B)^{-C_1} \geqslant (eB)^{-C'}$$

for a suitable C', as required. □

For completeness, we give a completely explicit version of Corollary 5.3 in the case that $\alpha_1, \ldots, \alpha_m$ are integers. Recall that the height of a rational number $a = x/y$ with $x, y \in \mathbb{Z}$ coprime, is given by $H(a) := \max(|x|, |y|)$.

Theorem 5.4 (Matveev, 2000)**.** *Let $a_1, \ldots, a_m$ be non-zero rational numbers and let $b_1, \ldots, b_m$ be integers such that*

$$a_1^{b_1} \cdots a_m^{b_m} \neq 1.$$

Then $|a_1^{b_1} \cdots a_m^{b_m} - 1| \geqslant (eB)^{-C'}$, where

$$B = \max(|b_1|, \ldots, |b_m|), \qquad C' = \tfrac{1}{2} e \cdot m^{4.5} 30^{m+3} \prod_{j=1}^{m} \max\big(1, \log H(a_j)\big).$$

To illustrate the power of the above results we give a quick application.

Corollary 5.5. *let a, b be integers with $a \geqslant 2, b \geqslant 2$. Then there is an effectively computable number $C_1 > 0$, depending only on a, b, such that for any two positive integers m, n,*

$$|a^m - b^n| \geqslant \frac{\max(a^m, b^n)}{(e \max(m, n))^{C_1}}.$$

Consequently, for any non-zero integer k, there exists an effectively computable number C_2, depending on a, b, k such that if m, n are positive integers with $a^m - b^n = k$, then $m, n \leqslant C_2$.

Proof. Let m, n be positive integers. Put $B := \max(m, n)$. Assume without loss of generality that $a^m \geqslant b^n$. By Corollary 5.3 or Theorem 5.4 we have

$$|1 - b^n a^{-m}| \geqslant (eB)^{-C_1},$$

where C_1 is an effectively computable number depending only on a, b. Multiplying with a^m gives our first assertion.

Now let m, n be positive integers with $a^m - b^n = k$. Put again $B := \max(m, n)$. Since $a, b \geqslant 2$ we have $a^m \geqslant 2^m$, $b^n \geqslant 2^n$, hence $a^m = \max(a^m, b^n) \geqslant 2^B$. So,

$$|k| \geqslant 2^B \cdot (eB)^{-C_1}.$$

This proves that B is bounded above by an effectively computable number depending on a, b, k. □

Exercise 5.1. *In 1995, Laurent, Mignotte and Nesterenko proved the following explicit estimate for linear forms in two logarithms. Let a_1, a_2 be two positive rational numbers $\neq 1$. Further, let b_1, b_2 be non-zero integers. Suppose that $\Lambda := b_1 \log a_1 - b_2 \log a_2 \neq 0$. Then*

$$\log |\Lambda| \geqslant$$
$$-24.34 \left(\max \left\{ \log \left(\frac{|b_1|}{\log H(a_2)} + \frac{|b_2|}{\log H(a_1)} \right) + 0.14\,,\ 21 \right\} \right)^2 \log H(a_1) \log H(a_2)\,.$$

Using this estimate, compute an upper bound C, such that for all positive integers m, n with $97^m - 89^n = 8$ we have $m, n \leqslant C$.
Hint. *Use $|\log(1 + z)| \leqslant 2|z|$ if $|z| \leqslant \frac{1}{2}$.*

In 1844, Catalan conjectured that the equation in four unknowns,

$$x^m - y^n = 1 \quad \text{in } x, y, m, n \in \mathbb{Z} \text{ with } x, y, m, n \geqslant 2$$

has only one solution, that is, $3^2 - 2^3 = 1$. In 1976, as one of the striking consequences of the results on linear forms in logarithms mentioned above, Tijdeman proved that there is an effectively computable constant C, such that for every solution (x, y, m, n) of Catalan's equation, one has $x^m, y^n \leqslant C$. The constant C can be computed but it is extremely large. Several people tried to prove Catalan's conjecture, on the one hand by reducing Tijdeman's constant C using sharper linear forms in logarithm estimates, on the other hand by showing with techniques from algebraic number theory that x^m, y^n have to be very large as long as $(x^m, y^n) \neq (3^2, 2^3)$, and finally using heavy computations. This didn't lead to success. In 2000 Mihailescu managed to prove Catalan's conjecture by an algebraic method which is completely independent of linear forms in logarithms.

We give another application. Consider the sequence $\{a_n\}$ with $a_n = 2^n$ for $n = 0, 1, 2, \ldots$. Note that $a_n - a_{n-1} = \frac{1}{2}a_n$. Similarly, we may consider the increasing sequence $\{a_n\}$ of numbers which are all composed of primes from $\{2, 3\}$, i.e., $1, 2, 3, 4, 6, 8, 9, 12, 16, 18, 24, 27, 32, \ldots$ and ask how the gap $a_n - a_{n-1}$ compares with a_n as $n \to \infty$. More generally, we may take a finite set of primes and ask this question about the sequence of consecutive integers composed of these primes.

Theorem 5.6 (Tijdeman, 1974)**.** *Let $S = \{p_1, \ldots, p_t\}$ be a finite set of distinct primes, and let $a_0 < a_1 < a_2 < \cdots$ be the sequence of consecutive positive integers composed of primes from S. Then there are effectively computable positive numbers c_1, c_2, depending on $t, p_1, \ldots, p_t$, such that*

$$a_n - a_{n-1} \geqslant \frac{a_n}{c_1(\log a_n)^{c_2}} \quad \textit{for } n = 1, 2, \ldots.$$

Proof. let $n \geqslant 1$. We have $a_n = p_1^{k_1} \cdots p_t^{k_t}$, and $a_{n-1} = p_1^{l_1} \cdots p_t^{l_t}$ with non-negative integers k_i, l_i. By Corollary 5.3,

$$\left|1 - \frac{a_{n-1}}{a_n}\right| = |1 - p_1^{l_1 - k_1} \cdots p_t^{l_t - k_t}| \geqslant (eB)^{-C},$$

where $B := \max(|l_1 - k_1|, \ldots, |l_t - k_t|)$ and C is effectively computable and depends only on $t, p_1, \ldots, p_t$. Note that

$$k_i \leqslant \frac{\log a_n}{\log p_i} \leqslant \frac{\log a_n}{\log 2}, \quad l_i \leqslant \frac{\log a_{n-1}}{\log p_i} \leqslant \frac{\log a_n}{\log 2} \quad \text{for } i = 1, \ldots, t,$$

hence $B \leqslant \log a_n / \log 2$. It follows that $a_n - a_{n-1} \geqslant a_n (e \log a_n / \log 2)^{-C}$. $\square$

5.2 Dirichlet's Unit Theorem

We want to apply the results from the previous section to certain Diophantine equations, and for this, we need some facts on units in algebraic number fields.

Let K be an algebraic number field of degree d. Recall that K has precisely d embeddings in $\mathbb{C}$. An embedding σ of K in $\mathbb{C}$ is called *real* if $\sigma(K) \subset \mathbb{R}$, and *complex* otherwise. If σ is a complex embedding of K, then so is $\overline{\sigma} : x \mapsto \overline{\sigma(x)}$, i.e., the composition of σ and complex conjugation. Hence the complex embeddings of K occur in complex conjugate pairs $\{\sigma, \overline{\sigma}\}$, and so, the number of complex embeddings of K is even. Let us denote by r_1 the number of real embeddings of K, and by $2r_2$ the number of complex embeddings of K. Thus,

$$r_1 + 2r_2 = d.$$

Further, we order the embeddings $\sigma_1, \ldots, \sigma_d$ of K in such a way that

$$\begin{aligned} &\sigma_1, \ldots, \sigma_{r_1} \text{ are the real embeddings,} \\ &\sigma_{r_1+r_2+1} = \overline{\sigma_{r_1+1}}, \ldots, \sigma_{r_1+2r_2} = \overline{\sigma_{r_1+r_2}}. \end{aligned}$$

We denote as usual by O_K the ring of integers of K, and by O_K^* the group of units of O_K. Further, we define the *norm* and *house* of $\alpha \in O_K$ by respectively,

$$N_{K/\mathbb{Q}}(\alpha) := \sigma_1(\alpha) \cdots \sigma_d(\alpha), \qquad |\alpha| := \max_{1 \leqslant i \leqslant d} |\sigma_i(\alpha)|.$$

Recall that the norm is multiplicative, and that $N_{K/\mathbb{Q}}(\alpha) \in \mathbb{Z}$ for $\alpha \in O_K$.

Lemma 5.7. *Let $\alpha \in O_K$. Then $\alpha \in O_K^* \iff N_{K/\mathbb{Q}}(\alpha) = \pm 1$.*

Proof. $\Longrightarrow$. Let $\alpha \in O_K^*$. Then $\alpha, \alpha^{-1} \in O_K$. Hence $N_{K/\mathbb{Q}}(\alpha) \in \mathbb{Z}$, $N_{K/\mathbb{Q}}(\alpha^{-1}) \in \mathbb{Z}$. But the product of these two integers is $N_{K/\mathbb{Q}}(1) = 1$, hence both integers are ± 1.

$\Longleftarrow$ Suppose $\sigma_1 = \mathrm{id}$. Then $\alpha^{-1} = \pm \prod_{i=2}^{d} \sigma_i(\alpha)$ is an algebraic integer in K, hence in O_K. So $\alpha \in O_K^*$. $\square$

To study the units of $\mathcal{O}_K$, it is useful to consider the absolute values of their conjugates. Clearly, for $\varepsilon \in \mathcal{O}_K^*$ we have

$$|\sigma_{r_1+r_2+i}(\varepsilon)| = |\sigma_{r_1+i}(\varepsilon)| \text{ for } i = 1, \ldots, r_2,$$
$$\prod_{i=1}^{r_1} |\sigma_i(\varepsilon)| \prod_{i=r_1+1}^{r_1+r_2} |\sigma_i(\varepsilon)|^2 = 1,$$

so $|\sigma_i(\varepsilon)|$ $(i = 1, \ldots, r_1 + r_2 - 1)$ determine $|\sigma_i(\varepsilon)|$ $(i = r_1 + r_2, \ldots, d)$.

The following result is known as Dirichlet's Unit Theorem. For a proof we refer to any textbook on algebraic number theory.

Theorem 5.8 (Dirichlet)**.** *Let $r := r_1 + r_2 - 1$ and define the map*

$$\overrightarrow{\log}\colon \mathcal{O}_K^* \to \mathbb{R}^r : \varepsilon \mapsto (\log|\sigma_1(\varepsilon)|, \ldots, \log|\sigma_r(\varepsilon)|)^T.$$

Then $\overrightarrow{\log}$ is a group homomorphism. The kernel of $\overrightarrow{\log}$ is the group U_K of roots of unity of K, and this group is finite. The image of $\overrightarrow{\log}$ is a lattice in $\mathbb{R}^r$.

Choose units $\varepsilon_1, \ldots, \varepsilon_r$ such that $\overrightarrow{\log}(\varepsilon_1), \ldots, \overrightarrow{\log}(\varepsilon_r)$ form a basis of the lattice $\overrightarrow{\log}(O_K^*)$ (we call such $\varepsilon_1, \ldots, \varepsilon_r$ a *system of fundamental units for K*). Then for every $\varepsilon \in O_K^*$, there are unique integers $b_1, \ldots, b_r$ such that

$$\overrightarrow{\log}(\varepsilon) = b_1 \overrightarrow{\log}(\varepsilon_1) + \cdots + b_r \overrightarrow{\log}(\varepsilon_r)$$

Hence $\varepsilon \in O_K^*$ can be expressed uniquely as

$$\zeta \varepsilon_1^{b_1} \cdots \varepsilon_r^{b_r} \text{ with } \zeta \in U_K,\ b_1, \ldots, b_r \in \mathbb{Z}. \tag{5.1}$$

Further, the matrix

$$M := \begin{pmatrix} \log|\sigma_1(\varepsilon_1)| & \cdots & \log|\sigma_1(\varepsilon_r)| \\ \vdots & & \vdots \\ \log|\sigma_r(\varepsilon_1)| & \cdots & \log|\sigma_r(\varepsilon_r)| \end{pmatrix} \tag{5.2}$$

is invertible.

We deduce some consequences.

Lemma 5.9. *There is an effectively computable number $C > 0$ depending on $K, \varepsilon_1, \ldots, \varepsilon_r$, such that for every $\varepsilon \in O_K^*$ we have*

$$\max(|b_1|, \ldots, |b_r|) \leqslant C \cdot \log \lceil \varepsilon \rceil,$$

where $b_1, \ldots, b_r$ are the integers defined by (5.1).

Proof. Let $\mathbf{b} := (b_1, \ldots, b_r)^T$ (column vector). Then $\overrightarrow{\log}\ (\varepsilon) = M\mathbf{b}$, hence $\mathbf{b} = M^{-1} \overrightarrow{\log}\ (\varepsilon)$. Writing $M^{-1} = (a_{ij})$, we obtain

$$b_i = \sum_{j=1}^{r} a_{ij} \log |\sigma_j(\varepsilon)| \quad (i = 1, \ldots, r).$$

Notice that $\lceil \varepsilon \rceil \geqslant 1$ and $|\sigma_j(\varepsilon)| \geqslant \lceil \varepsilon \rceil^{1-d}$ for $j = 1, \ldots, d$. Hence

$$\big| \log |\sigma_j(\varepsilon)| \big| \leqslant d \log \lceil \varepsilon \rceil \text{ for } j = 1, \ldots, d.$$

Now an application of the triangle inequality gives

$$\max_{1 \leqslant i \leqslant r} |b_i| \leqslant \left(\max_{1 \leqslant i \leqslant r} \sum_{j=1}^{r} |a_{ij}| \right) \cdot d \log \lceil \varepsilon \rceil = C \cdot \log \lceil \varepsilon \rceil.$$

□

The next lemma states that given $\alpha \in O_K \setminus \{0\}$, we can find $\varepsilon \in O_K^*$ such that all conjugates of $\varepsilon\alpha$ have about the same absolute value. Then the maximum of these absolute values, which is $\lceil \varepsilon\alpha \rceil$, is about the d-th root of the product of these absolute values, which is $|N_{K/\mathbb{Q}}(\varepsilon\alpha)| = |N_{K/\mathbb{Q}}(\alpha)|$ since $N_{K/\mathbb{Q}}(\varepsilon) = \pm 1$.

Lemma 5.10. *There is an effectively computable number $c > 0$ with the following property: for every non-zero $\alpha \in O_K$ there is $\varepsilon \in O_K^*$ such that*

$$c^{-1} |N_{K/\mathbb{Q}}(\alpha)|^{1/d} \leqslant \lceil \varepsilon\alpha \rceil \leqslant c |N_{K/\mathbb{Q}}(\alpha)|^{1/d}.$$

Proof. In general, if L is a lattice in $\mathbb{R}^r$, then for every point $\mathbf{x} \in \mathbb{R}^r$ there is a point $\mathbf{u} \in L$ such that $\|\mathbf{u} - \mathbf{x}\|_2 \leqslant c(L)$ for some number $c(L)$ depending only on L. This $c(L)$ can be computed in terms of a basis of L. By applying this with $L = \overrightarrow{\log}\ (O_K^*)$, we see that there is an effectively computable number $c_1 > 0$,

depending on the lattice $\overrightarrow{\log}\,(O_K^*)$, such that for every $\mathbf{x} \in \mathbb{R}^r$ there is $\varepsilon \in O_K^*$ with $\|\mathbf{x} - \overrightarrow{\log}\,(\varepsilon)\|_2 \leqslant c_1$. This implies

$$\big|x_i - \log|\sigma_i(\varepsilon)|\big| \leqslant c_1 \quad \text{for } i = 1, \dots, r,$$

where $\mathbf{x} = (x_1, \dots, x_r)^T$. We apply this with

$$x_i := -\log|\sigma_i(\alpha)| + \tfrac{1}{d}\log|N_{K/\mathbb{Q}}(\alpha)| \quad (i = 1, \dots, r).$$

It follows that there is $\varepsilon \in O_K^*$ such that

$$\Big|\log|\sigma_i(\varepsilon)| + \log|\sigma_i(\alpha)| - \tfrac{1}{d}\log|N_{K/\mathbb{Q}}(\alpha)|\Big| \leqslant c_1 \quad \text{for } i = 1, \dots, r,$$

i.e.,

$$\Big|\log|\sigma_i(\varepsilon\alpha)| - \tfrac{1}{d}\log|N_{K/\mathbb{Q}}(\alpha)|\Big| \leqslant c_1 \quad \text{for } i = 1, \dots, r.$$

Put

$$\xi_i := \log|\sigma_i(\varepsilon\alpha)| - \tfrac{1}{d}\log|N_{K/\mathbb{Q}}(\alpha)| \quad (i = 1, \dots, d),$$
$$e_i := 1 \text{ for } i = 1, \dots, r_1, \quad e_i := 2 \text{ for } i = r_1 + 1, \dots, r_1 + r_2 = r + 1.$$

Since $\log|\sigma_{r_1+r_2+i}(\varepsilon\alpha)| = \log|\sigma_{r_1+i}(\varepsilon\alpha)|$ for $i = 1, \dots, r_2$ and $\sum_{i=1}^{d} \log|\sigma_i(\varepsilon\alpha)| = \log|N_{K/\mathbb{Q}}(\varepsilon\alpha)| = \log|N_{K/\mathbb{Q}}(\alpha)|$, we have

$$\xi_{r_1+r_2+i} = \xi_{r_1+i} \quad \text{for } i = 1, \dots, r_2, \quad \sum_{i=1}^{r_1+r_2} e_i\xi_i = 0.$$

Hence

$$|\xi_{r_1+r_2}| \leqslant e_{r_1+r_2}^{-1} \sum_{i=1}^{r} e_i|\xi_i| \leqslant c_1 \cdot e_{r_1+r_2}^{-1} \sum_{i=1}^{r} e_i =: c_2,$$

and so,

$$\Big|\log|\sigma_i(\varepsilon\alpha)| - \tfrac{1}{d}\log|N_{K/\mathbb{Q}}(\alpha)|\Big| = |\xi_i| \leqslant \max(c_1, c_2) \quad \text{for } i = 1, \dots, d.$$

Picking i with $\overline{|\varepsilon\alpha|} = |\sigma_i(\varepsilon\alpha)|$ we obtain that Lemma 5.10 holds with $c := e^{\max(c_1,c_2)}$. □

Corollary 5.11. *Given $\alpha \in O_K \setminus \{0\}$, one can effectively determine a finite set of divisors $\gamma_1, \dots, \gamma_m$ of α in O_K such that for each divisor β of α in O_K there are $\varepsilon \in O_K^*$ and $i \in \{1, \dots, m\}$ such that $\beta = \varepsilon\gamma_i$.*

Proof. Let β be a divisor of α. Then $N_{K/\mathbb{Q}}(\beta)$ divides $N_{K/\mathbb{Q}}(\alpha)$. By the previous lemma, there is $\varepsilon \in O_K^*$, such that

$$\overline{|\varepsilon\beta|} \leqslant c|N_{K/\mathbb{Q}}(\beta)|^{1/d} \leqslant c|N_{K/\mathbb{Q}}(\alpha)|^{1/d}.$$

By one of the homework exercises, there are only finitely many algebraic integers γ of degree at most d and house at most $c|N_{K/\mathbb{Q}}(\alpha)|^{1/d}$. This implies that there are at most finitely many $\gamma \in O_K$ with

$$\overline{|\gamma|} \leqslant c|N_{K/\mathbb{Q}}(\alpha)|^{1/d}.$$

In fact, these can be determined effectively, and for each of these γ it can be checked whether they divide α in O_K. This leaves us with a finite set $\{\gamma_1, \ldots, \gamma_m\}$ as in the statement of our corollary. □

5.3 Unit equations and Thue equations

Let K be an algebraic number field. We consider the so-called *unit equation*

$$\alpha x + \beta y = 1 \quad \text{in } x, y \in O_K^*, \tag{5.3}$$

where $\alpha, \beta \in K^*$.

Theorem 5.12. *Eq.* (5.3) *has at most finitely many solutions, and these can be determined effectively.*

In 1921, Siegel proved that (5.3) has only finitely many solutions, but his proof is ineffective, in the sense that it shows only that there are only finitely many solutions, but it does not give a method how to determine them. Our proof, based on lower bounds for linear forms in logarithms, does give a method to determine the solutions. This effective proof is already implicit in work of Baker from the 1960's. Győry (1978) made this explicit.

Proof. Let (x, y) be a solution of (5.3). By (5.1), there are $\zeta_1, \zeta_2 \in U_K$, as well as $a_1, \ldots, a_r$, $b_1, \ldots, b_r \in \mathbb{Z}$, such that

$$x = \zeta_1 \varepsilon_1^{a_1} \cdots \varepsilon_r^{a_r}, \quad y = \zeta_2 \varepsilon_1^{b_1} \cdots \varepsilon_r^{b_r}.$$

Thus,

$$\alpha\zeta_1\varepsilon_1^{a_1}\cdots\varepsilon_r^{a_r} + \beta\zeta_2\varepsilon_1^{b_1}\cdots\varepsilon_r^{b_r} = 1.$$

We assume without loss of generality that $B := \max(|a_1|, \ldots, |b_r|) = |b_r|$. We estimate from above and below,

$$\Lambda_i := |\sigma_i(\alpha)\sigma(\zeta_1)\sigma_i(\varepsilon_1)^{a_1}\cdots\sigma_i(\varepsilon_r)^{a_r} - 1| = |\sigma_i(\beta)\sigma_i(y)|$$

for a suitable choice of i.

In fact, let $|\sigma_i(y)|$ be the smallest, and $|\sigma_j(y)| = \overline{|y|}$ the largest among $|\sigma_1(y)|, \ldots, |\sigma_d(y)|$. Then by Lemma 5.7,

$$|\sigma_i(y)|^{d-1} \cdot \overline{|y|} \leqslant 1$$

and subsequently by Lemma 5.9,

$$|\sigma_i(y)| \leqslant \overline{|y|}^{-1/(d-1)} \leqslant e^{-B/C(d-1)}.$$

This leads to

$$\Lambda_i \leqslant |\sigma_i(\beta)|e^{-B/C(d-1)}.$$

By Corollary 5.3 we have $|\Lambda_i| \geqslant (eB)^{-C'}$ for some effectively computable number C' depending on α, $\varepsilon_1, \ldots, \varepsilon_r$ and the finitely many roots of unity of K. We infer

$$(eB)^{-C'} \leqslant |\sigma_i(\beta)|e^{-B/C(d-1)}$$

and this leads to an effectively computable upper bound for B. □

Remark. There are practical algorithms to solve equations of the type (5.3) which work well as long as the degree of the field K is not too large, and K has a system of fundamental units whose heights are not too large. These algorithms are based on lower bounds for linear forms in logarithms and the Lenstra-Lenstra-Lovász lattice basis reduction algorithm (LLL-algorithm). For instance, in 2000 Wildanger determined all solutions of the equation $x+y = 1$ in $x, y \in \mathcal{O}_K^*$, with $K = \mathbb{Q}(\cos(2\pi/19))$. The number field K has degree 9 and all its embeddings are real. Thus, the unit group $\mathcal{O}_K^*$ has rank 8.

In general, a *form of degree d in n variables* is a homogenenous polynomial $F(X_1, \ldots, X_n)$ of degree d, i.e., a polynomial consisting of terms $cX_1^{i_1}\cdots X_n^{i_n}$ with

$i_1 + \cdots + i_n = d$. Note that $F(tX_1, \ldots, tX_n) = t^d F(X_1, \ldots, X_n)$. A *binary form of degree* d is a homogeneous polynomial of degree d in two variables, i.e.,

$$F(X, Y) = a_0 X^d + a_1 X^{d-1} Y + \cdots + a_d Y^d.$$

Suppose that F has its coefficients in K. Let a_r be the first non-zero coefficient of F from the left. In some finite extension L of K, we can factor $F(X, 1)$ as $a_r(X - \alpha_1)^{r_1} \cdots (X - \alpha_t)^{r_t}$, where $\alpha_1, \ldots, \alpha_t$ are distinct, and the multiplicities $r_1, \ldots, r_t > 0$. Then we get

$$F(X, Y) = Y^d F(X/Y, 1) = a_r Y^{d-r} (X - \alpha_1 Y)^{r_1} \cdots (X - \alpha_t Y)^{r_t}. \tag{5.4}$$

Thus, a binary form can be factored into linear forms.

A *Thue equation* is an equation of the shape

$$F(x, y) = m \quad \text{in } x, y \in \mathbb{Z}, \tag{5.5}$$

where F is a binary form with coefficients in $\mathbb{Z}$ and m is a non-zero integer.

We consider some special cases. First suppose that F is linear. Then (5.5) becomes

$$ax + by = m \quad \text{in } x, y \in \mathbb{Z}.$$

As is well-known, this equation has no solution if $\gcd(a, b)$ does not divide m, and infinitely many solutions if $\gcd(a, b)$ does divide m. Next, suppose that F is quadratic. Then (5.5) specializes to

$$ax^2 + bxy + cy^2 = m.$$

If the discriminant $D = b^2 - 4ac < 0$ then this equation describes an ellipsis, and this has only finitely many points $(x, y) \in \mathbb{Z}^2$ on it. In fact, these points may be determined by rewriting the equation as

$$a(x + (b/2a)y)^2 + (|D|/4a)y^2 = m.$$

In case that $D > 0$ the equation may have infinitely many solutions, e.g., the Pell equation $x^2 - dy^2 = 1$ where $d > 1$ is a positive integer, not equal to a square. In fact it can be shown that if $D = b^2 - 4ac > 0$ and D is not a square, then $ax^2 + bxy = cy^2 = m$ has either no, or infinitely many solutions.

Another special case is, where F may have arbitrary degree d but the coefficient a_0 of X^d is 0. Then F is divisible by Y, so if (x, y) is a solution of (5.5), then y

divides m. For each divisor y of m there are at most finitely many integers x with $F(x,y)=m$, which, if they exist, can be determined effectively.

We prove the following.

Theorem 5.13. *Let $F \in \mathbb{Z}[X,Y]$ be a binary form of degree d. Suppose that the coefficient of X^d in F is non-zero and that $F(X,1)$ has at least three distinct zeros in $\mathbb{C}$. Let m be a non-zero integer. Then the equation*

$$F(x,y)=m \quad \text{in } x,y \in \mathbb{Z} \tag{5.5}$$

has only finitely many solutions.

In 1909, the Norwegian mathematician A. Thue proved in an ineffective way that Eq. (5.5) has only finitely many solutions. In 1967, Baker gave an effective proof of this.

Proof. By assumption, $F(X,Y)=a_0X^d+\cdots+a_dY^d$ with $a_0 \neq 0$. We make a reduction to the case $a_0=1$. If (x,y) is a solution of (5.5), then, by multiplying with a_0^{d-1},

$$(a_0x)^d + a_1(a_0x)^{d-1} + a_2a_0(a_0x)^{d-2}y^2 + \cdots + a_da_0^{d-1}y^d = ma_0^{d-1}.$$

Thus, (a_0x,y) satisfies a Thue equation $F'(x',y')=m'$, where the coefficient of X^d in F' is 1.

Henceforth, we assume that the coefficient of X^d in F is 1. Then

$$F(X,Y)=(X-\alpha_1Y)^{r_1}\cdots(X-\alpha_tY)^{r_t}$$

with $\alpha_1,\ldots,\alpha_t \in \mathbb{C}$, $r_1,\ldots,r_t>0$ and $t \geqslant 3$.

We want to reduce (5.5) to a unit equation. The crucial observation here is that the three linear forms in two variables $X-\alpha_iY$ $(i=1,2,3)$ are linearly dependent. More precisely, we have *Siegel's identity*

$$(\alpha_2-\alpha_3)(X-\alpha_1Y)+(\alpha_3-\alpha_1)(X-\alpha_2Y)+(\alpha_1-\alpha_2)(X-\alpha_3Y)=0.$$

This implies that if $(x,y) \in \mathbb{Z}^2$ is a solution of (5.5), then

$$\frac{\alpha_2-\alpha_3}{\alpha_2-\alpha_1}\cdot\frac{x-\alpha_1y}{x-\alpha_3y}+\frac{\alpha_3-\alpha_1}{\alpha_2-\alpha_1}\cdot\frac{x-\alpha_2y}{x-\alpha_3y}=1. \tag{5.6}$$

Let $K = \mathbb{Q}(\alpha_1, \ldots, \alpha_t)$. Then $\alpha_1, \ldots, \alpha_t \in O_K$ since they are zeros of the monic polynomial $F(X, 1) \in \mathbb{Z}[X]$. Let $(x, y) \in \mathbb{Z}^2$ be a solution of (5.5). Then the numbers $x - \alpha_i y$ $(i = 1, 2, 3)$ divide m in O_K. By Corollary 5.11, we have

$$x - \alpha_i y = \mu_i \varepsilon_i,$$

where μ_i belongs to an effectively determinable finite set and $\varepsilon_i \in O_K^*$ for $i = 1, 2, 3$. By substituting this into (5.6) we obtain

$$\left(\frac{\alpha_2 - \alpha_3}{\alpha_2 - \alpha_1} \cdot \frac{\mu_1}{\mu_3}\right) \cdot \frac{\varepsilon_1}{\varepsilon_3} + \left(\frac{\alpha_3 - \alpha_1}{\alpha_2 - \alpha_1} \cdot \frac{\mu_2}{\mu_3}\right) \cdot \frac{\varepsilon_2}{\varepsilon_3} = 1.$$

We may view this as a unit equation with unknowns $\varepsilon_1/\varepsilon_3$, $\varepsilon_2/\varepsilon_3$. We have only finitely many possibilities for each μ_i which can be determined effectively, and by Theorem 5.12, for each choice of μ_1, μ_2, μ_3 we have only finitely possibilities for the pair $(\varepsilon_1/\varepsilon_3, \varepsilon_2/\varepsilon_3)$ which can be determined effectively. Consequently, if (x, y) runs through the solutions of (5.5), then the quotient $(x-\alpha_1 y)/(x-\alpha_3 y) = (\mu_1/\mu_3)(\varepsilon_1/\varepsilon_3)$ runs through a finite set which can be determined effectively. We can compute x/y from $(x - \alpha_1 y)/(x - \alpha_3 y)$ and then x, y from $y^d F(x/y, 1) = F(x, y) = m$. In this way, it follows that (5.5) has only finitely many solutions which can be determined effectively. □

Remark. Today, Thue equations can really be solved in practice, and several packages contain routines to solve Thue equations (KANT, Maple but to my knowledge not yet SAGE). These routines are based on lower bounds for linear forms in logarithms, and the LLL-algorithm.

Exercise 5.2. *Let $F(X, Y) \in \mathbb{Z}[X, Y]$ be a positive definite binary form of degree $d \geqslant 3$, i.e., the coefficient of X^d is > 0 and the zeros of $F(X, 1)$ are all in $\mathbb{C} \setminus \mathbb{R}$. Prove, without using lower bounds for linear forms in logarithms, that for each positive integer m the equation $F(x, y) = m$ has only finitely many solutions in $x, y \in \mathbb{Z}$.*

We finish with stating, without proof, an effective finiteness result for the equation

$$by^n = f(x) \quad \text{in } x, y \in \mathbb{Z}. \tag{5.7}$$

where $n \geqslant 2$, b is a non-zero integer and $f \in \mathbb{Z}[X]$. For $n = 2$ this is called a *hyperelliptic equation* and for $n \geqslant 3$ a *superelliptic equation*. Such equations can be reduced to unit equations or Thue equations.

Theorem 5.14 (Baker, 1968)**.** *Assume that f has no multiple zeros and that f has degree at least 2 if $n \geqslant 3$ and degree at least 3 if $n = 2$. Then (5.7) has only finitely many solutions, and its set of solutions can be determined effectively.*

We consider a special case to illustrate the idea of the proof. Consider the equation

$$y^3 = 2x(x-3) \text{ in } x, y \in \mathbb{Z}. \tag{5.8}$$

Let (x, y) be a solution of (5.8). The gcd of $2x$ and $x-3$ divides 6. So if p is a prime number $\geqslant 5$, then p divides at most one of $2x$, $x-3$ and if it divides one of these numbers, the exponent of p in the unique factorization of that number is divisible by 3. It follows that

$$2x = au^3, \quad x - 3 = bv^3$$

where ab is a third power, and both a and b are composed of primes from $\{2, 3\}$. In fact, we may assume that the exponents on $2, 3$ in a, b are either 0,1, or 2, since powers 2^{3k}, 3^{3l} can be absorbed by u, v. Thus, $a, b \in \{\pm 2^k 3^l : k, l = 0, 1, 2\}$. Considering the solutions (x, y) of (5.8) with fixed a, b, we get a Thue equation

$$au^3 - 2bv^3 = 6.$$

By determining the solutions (u, v) for each of these Thue equations, we can determine the solutions of (5.8).

A similar approach can be followed for equations $y^n = f(x)$ with $n \geqslant 3$ if f is not reducible over $\mathbb{Z}$. Then f factorizes over some algebraic number field K, and we have to make a reduction to Thue equations of which the unknowns are taken from O_K instead of $\mathbb{Z}$. For such equations one has a finiteness result similar to Theorem 5.13. In the case $n = 2$, one can make a reduction only to Thue equations where the involved binary form has degree 2, and in this case, Thue's theorem is not applicable. Then one needs a more complicated argument.

Exercise 5.3. *Let a_i, b_i, c_i $(i = 1, 2)$ be integers with*

$$a_1, b_1, c_1, \ a_2, b_2, c_2, \ a_1b_2 - a_2b_1, \ b_1c_2 - b_2c_1 \neq 0.$$

Prove that there are only finitely many triples $(x, y, z) \in \mathbb{Z}^3$ satisfying the system of equations

$$a_1x^2 - b_1z^2 = c_1, \quad a_2y^2 - b_2z^2 = c_2. \tag{5.9}$$

Hint. *Let* $K = \mathbb{Q}(\sqrt{a_1}, \sqrt{b_1}, \sqrt{a_2}, \sqrt{b_2})$. *Apply Theorem 5.12 and the ideas in the proof of Theorem 5.13 to the identities*

$$\sqrt{b_2}(x\sqrt{a_1} + z\sqrt{b_1}) - \sqrt{b_1}(y\sqrt{a_2} + z\sqrt{b_2}) = x\sqrt{a_1 b_2} - y\sqrt{a_2 b_1},$$
$$\sqrt{b_2}(x\sqrt{a_1} - z\sqrt{b_1}) - \sqrt{b_1}(y\sqrt{a_2} - z\sqrt{b_2}) = x\sqrt{a_1 b_2} - y\sqrt{a_2 b_1}.$$

Then conclude that if (x, y, z) *runs through the solutions of* (5.9) *then* $(x\sqrt{a_1} + z\sqrt{b_1})/(x\sqrt{a_1} - z\sqrt{b_1})$ *runs through a finite set.*

Exercise 5.4. *Use Exercise 5.3 to prove that the equation*

$$y^2 = x(2x-3)(4x-5) \quad \text{in } x, y \in \mathbb{Z}$$

has only finitely many solutions.

In 1976, Schinzel and Tijdeman obtained the surprising result that Eq. (5.7) has no solutions with $y \neq 0, \pm 1$ if n is too large.

Theorem 5.15 (Schinzel-Tijdeman, 1976)**.** *Let* b *be a non-zero integer and* $f(X) \in \mathbb{Z}[X]$ *a polynomial of degree at least* 2 *without multiple zeros. Then there is an effectively computable number* C *depending on* f *such that if* $by^n = f(x)$ *is solvable in* $x, y \in \mathbb{Z}$ *with* $y \neq 0, \pm 1$*, then* $n \leqslant C$*.*

Exercise 5.5. *(i) Prove that the equation*

$$x^n - 2y^n = 1 \quad \text{in } x, y \in \mathbb{Z} \text{ with } x \geqslant 2, y \geqslant 2$$

has no solutions if $n > 15000$*.*

Hint. *Applying the estimate of Laurent-Mignotte-Nesterenko from Exercise 5.1 to an appropriate linear form in two logarithms you will get a lower estimate depending on* n *and* x, y*. But you can derive also an upper estimate which depends on* n, x, y*. Comparing the two estimates leads to an upper bound for* n *independent of* x, y*.*

(ii) Let a, b, c *be positive integers. Prove that there is a number* C*, effectively computable in terms of* a, b, c*, such that the equation*

$$ax^n - by^n = c$$

has no solutions if $n > C$*. In the case* $a = b$ *you may give an elementary proof, without using the result of Laurent-Mignotte-Nesterenko.*

(iii) Prove that the equation

$$y^z = \binom{x}{3} \quad \textit{in } x, y, z \in \mathbb{Z} \textit{ with } x \geqslant 4,\ y \geqslant 2,\ z \geqslant 3$$

has only finitely many solutions.

5.4 p-adic analogues

The results mentioned in Section 5.1 have so-called p-adic analogues. We give one example.

Recall that each non-zero rational number can be expressed uniquely as a product of prime powers. We may express this as

$$a = \pm \prod_{p \in \mathcal{P}} p^{\mathrm{ord_p(a)}},$$

where $\mathcal{P}$ is the set of prime numbers, and the exponents $\mathrm{ord}_p(a)$ are integers, at most finitely many of which are non-zero. We define the *p-adic absolute value* of a by

$$|a|_p := p^{-\mathrm{ord}_p(a)} \text{ for } a \in \mathbb{Q}^*, \quad |0|_p := 0.$$

For instance, $-72/343 = -2^3 \cdot 3^2 \cdot 7^{-3}$, hence

$$|-72/343|_2 = 2^{-3},\ |-72/343|_3 = 3^{-2},\ |-72/343|_7 = 7^3.$$

Notice that for any prime number p we have

$$|ab|_p = |a|_p |b|_p, \quad |a+b|_p \leqslant \max(|a|_p, |b|_p) \ \text{ for } a, b \in \mathbb{Q}.$$

The last inequality is called the *strong triangle inequality* or *ultrametric inequality*. In general, if $a_1, \ldots, a_r$ are rational numbers such that $|a_1|_p > |a_i|_p$ for $i = 2, \ldots, r$, then

$$|a_1 + \cdots + a_r|_p = |a_1|_p. \tag{5.10}$$

The strong triangle inequality implies that the p-adic absolute value $|\cdot|_p$ defines a metric d_p on $\mathbb{Q}$, given by $d_p(x-y) := |x-y|_p$. Two numbers $x, y \in \mathbb{Q}$ are p-adically

close, if $d_p(x-y)$ is small, which means that $x-y=a/b$ where a,b are coprime integers and a is divisible by a high power of p. From topology it is known how to complete a metrical space, by adding to this space the limits of all its Cauchy sequences. The metrical completion of $\mathbb{Q}$ with metric d_p is denoted $\mathbb{Q}_p$. As it turns out, addition and multiplication on $\mathbb{Q}$ can be extended to $\mathbb{Q}_p$, and this makes $\mathbb{Q}_p$ into a field, the *field of p-adic numbers.* In Diophantine approximation, $|\cdot|_p$ and $\mathbb{Q}_p$ have the same 'status' as the ordinary absolute value and $\mathbb{R}$, and many results in Diophantine approximation and transcendence theory have analogues in the p-adic setting. We will not go into this.

To gave a flavour, we give an analogue of Corollary 5.3 in the case that $\alpha_1,\ldots,\alpha_m$ are rational numbers. There is a more general version for algebraic $\alpha_1,\ldots,\alpha_m$ but it requires knowledge of algebraic number theory to state this.

Theorem 5.16. (Yu, 1986) *Let p be a prime number, let $a_1,\ldots,a_m$ be non-zero rational numbers with $|a_i|_p=1$ for $i=1,\ldots,m$. Further, let $b_1,\ldots,b_m$ be integers such that*

$$a_1^{b_1}\cdots a_m^{b_m}\neq 1.$$

Put $B:=\max(|b_1|,\ldots,|b_m|)$. Then

$$|a_1^{b_1}\cdots a_m^{b_m}-1|_p\geqslant (eB)^{-C}$$

where C is an effectively computable number depending on p, m and $a_1,\ldots,a_m$.

For $m=1$ there is a sharper result which can be proved by elementary means (exercise below). But for $m\geqslant 2$ the proof is very difficult. One may define p-adic logarithms and translate the theorem into a lower bound for the p-adic absolute value of a linear form in p-adic logarithms. We do not work this out.

Exercise 5.6. *Let a be an integer, and p a prime, such that $|a|_p\leqslant p^{-1}$ if $p>2$ and $|a|_2\leqslant 2^{-2}$ if $p=2$. Prove that for any positive integer b we have*

$$|(1+a)^b-1|_p=|ab|_p\geqslant 1/ab.$$

Hint. *First prove this for b with $|b|_p=1$ and for $b=p$. Then prove it for $b=up^t$ where u is an integer not divisible by p and $t\geqslant 0$.*

We give an application. Let $S=\{p_1,\ldots,p_t\}$ be a finite set of primes numbers and define the multiplicative group

$$U_S:=\{\pm p_1^{z_1}\cdots p_t^{z_t}:\ z_1,\ldots,z_t\in\mathbb{Z}\}.$$

Theorem 5.17. *The equation*

$$x + y = 1 \quad \text{in } x, y \in U_S \tag{5.11}$$

has only finitely many solutions, and these can be determined effectively.

Proof. Let (x, y) be a solution of (5.11). We may write $x = u/w$, $y = v/w$ where u, v, w are integers with $\gcd(u, v, w) = 1$. Then

$$u + v = w. \tag{5.12}$$

The integers u, v, w are composed of primes from S, and moreover, no prime divides two numbers among u, v, w since u, v, w are coprime. After reordering the primes $p_1, \ldots, p_t$, we may assume that

$$u = \pm p_1^{b_1} \cdots p_r^{b_r}, \quad v = \pm p_{r+1}^{b_{r+1}} \cdots p_s^{b_s}, \quad w = \pm p_{s+1}^{b_{s+1}} \cdots p_t^{b_t},$$

where $0 \leqslant r \leqslant s \leqslant t$ and the b_i are non-negative integers (empty products are equal to 1; for instance if $r = 0$ then $u = \pm 1$). We have to prove that $B := \max(b_1, \ldots, b_t)$ is bounded above by an effectively computable number depending only on $p_1, \ldots, p_t$. By symmetry, we may assume that $B = b_t$. Then using $-(u/v) - 1 = -(w/v)$ we obtain

$$0 < | \pm p_1^{b_1} \cdots p_r^{b_r} p_{r+1}^{-b_{r+1}} \cdots p_s^{-b_s} - 1|_{p_t} = |w/v|_{p_t} = p_t^{-b_t} = p_t^{-B}.$$

From Theorem 5.17 we obtain that $|\cdots|_{p_t} \geqslant (eB)^{-C}$, where C is effectively computable in terms of $p_1, \ldots, p_t$. Hence

$$(eB)^{-C} \leqslant p_t^{-B}.$$

So indeed, B is bounded above by an effectively computable number depending on $p_1, \ldots, p_t$. □

Remark. In his PhD-thesis from 1988, de Weger gave a practical algorithm, based on strong linear forms in logarithms estimates and the LLL-basis reduction algorithm, to solve equations of the type (5.11). As a consequence, he showed that the equation $x + y = 1$ has precisely 545 solutions in positive integers $x, y \in U_S$ with $0 < x \leqslant y$, where $S = \{2, 3, 5, 7, 11, 13\}$.

Let K be an algebraic number field and let Γ be a finitely generated, multiplicative subgroup of K^*, i.e., there are $\gamma_1, \ldots, \gamma_t \in \Gamma$ such that every element of Γ can be expressed as

$$\zeta \gamma_1^{z_1} \cdots \gamma_t^{z_t}$$

where ζ is a root of unity in K, and $z_1, \ldots, z_t$ are integers. Further, let a, b be non-zero elements from K and consider the equation

$$ax + by = 1 \quad \text{in } x, y \in \Gamma. \tag{5.13}$$

The following result is a common generalization of both Theorems 5.12 and 5.17.

Theorem 5.18 (Győry, 1979)**.** *Equation* (5.13) *has only finitely many solutions, and these can be determined effectively.*

In 1960, Lang gave an ineffective proof of this result, by combining earlier work of Siegel (1921), Mahler (1933) and Parry (1950). Győry's proof is based on Corollary 5.3 and a generalization of Theorem 5.16 for algebraic numbers.