Diophantine equations

F.Beukers

Spring 2011

4 Three term S-unit equations

4.1 Introduction

Although Thue's method is a brilliant one, approximation of an algebraic number α by rationals is a slight overkill if one wants to prove finiteness of the solution set of Thue's equation. In Thue's work one also finds explicit polynomial constructions which were meant to deal with special values of α. It turns out that these constructions suffice to prove finiteness of the solution set of a large class of Diophantine equations.

Theorem 4.1.1 (Siegel-Mahler) *Let K be an algebraic number field and Γ a finitely generated subgroup of K^*. Let $a, b \in K^*$. Then the equation*

$$ax + by = 1, \quad x, y \in \Gamma$$

has at most finitely many solutions.

Letting r be the rank of Γ, it was shown by J.H.Evertse in 1984 that the number of solutions is bounded by $3 \times 7^{d+2r}$ where $d = [K : \mathbb{Q}]$. In 1996 Beukers-Schlickewei showed that 2^{8r+16} is a bound, independent of K.
As a first application we show finiteness of the number of integer solutions to $F(x, y) = m$ where F is an irreducible form of degree ≥ 3. Write

$$F(x, y) = a_0 x^d + a_1 x^{d-1} y + \cdots + a_{d-1} x y^{d-1} + a_d y^d.$$

Without loss of generality we can assume that $a_0 = 1$. Namely, $a_0^{d-1} F$ can be rewritten in an integer form in $a_0 x$ and y. Call the new form G and consider $G(x, y) = a_0^{d-1} m$ instead. So let us suppose $a_0 = 1$ from now on. Let K be the splittinhg field of $F(x, 1)$ and write $F(x, 1) = \prod_{i=1}^{d} (x - \alpha_i)$ where the α_i are integers in K. Then the equation $F(x, y) = m$ implies that for every i there exist $\mu_i \in \mathcal{O}_K$ and a unit β_i such that $x - \alpha_i y = \mu_i \beta_i$. The elements μ_i can be taken from a finite set of possibilities. Taking the first three, we can combine them linearly to eliminate x, y:

$$(\alpha_2 - \alpha_3)\mu_1\beta_1 + (\alpha_3 - \alpha_1)\mu_2\beta_2 + (\alpha_3 - \alpha_1)\mu_3\beta_3 = 0.$$

Since the group of units in a number field is finitely generated, we can apply the Siegel-Mahler theorem. We conclude that there exist finitely many possibilities for β_1/β_2. Hence there exist finitely many possibilities for $(x-\alpha_1 y)/(x-\alpha_2 y)$ and hence for x/y. This proves the finiteness of the solution set.

4.2 The hyperelliptic equation

Here we deal with a generalization of Mordell's equation, namely

$$y^2 = x^n + a_1 x^{n-1} + \cdots + a_{n-1}x + a_n,$$

where the $a_i \in \mathbb{Z}$ are given and x, y are the unknowns. Let us write $f(x) = x^n + a_1 x^{n-1} + \cdots + a_n$. We have the following theorem.

Theorem 4.2.1 (Siegel) *Suppose that $f(x)$ is squarefree and has at least three distinct complex zeros. Then the number of solutions is finite.*

Proof. Without loss of generality we can assume that f has simple zeros. Let K be the splitting field of f. The equation can be rewritten as $y^2 = (x-\alpha_1)(x-\alpha_2)\cdots(x-\alpha_n)$ where the α_i are integers in K. Let $\wp$ be a prime ideal divisor of $x-\alpha_i$ and $x-\alpha_j$ with $i \neq j$. Then $\wp$ divides $\alpha_i - \alpha_j$. Hence, any ideal $(x-\alpha_i)$ equals $J_i \eta_i^2$ where J_i divides $\prod_{k<l}(\alpha_k - \alpha_l)$. We conclude that there exists a finite set T of triples $(\mu_1, \mu_2, \mu_3) \in K^3$ such that $x - \alpha_i = \mu_i \xi_i^2$, $i = 1, 2, 3$ for some $(\mu_1, \mu_2, \mu_3) \in T$, and where ξ_i are integers in K. Taking all differences,

$$\begin{aligned} \mu_1\xi_1^2 - \mu_2\xi_2^2 &= \alpha_2 - \alpha_1 \neq 0 \\ \mu_2\xi_2^2 - \mu_3\xi_3^2 &= \alpha_3 - \alpha_2 \neq 0 \\ \mu_3\xi_3^2 - \mu_1\xi_1^2 &= \alpha_1 - \alpha_3 \neq 0 \end{aligned}$$

Let $L = K(\sqrt{\mu_1}, \sqrt{\mu_2}, \sqrt{\mu_3})$. Then

$$\begin{aligned} \xi_1\sqrt{\mu_1} - \xi_2\sqrt{\mu_2} &= \beta_3\epsilon_3 \\ \xi_2\sqrt{\mu_2} - \xi_3\sqrt{\mu_3} &= \beta_1\epsilon_1 \\ \xi_3\sqrt{\mu_3} - \xi_1\sqrt{\mu_1} &= \beta_2\epsilon_2 \end{aligned}$$

where the ϵ_i are integer units in L and the triples $\beta_1, \beta_2, \beta_3 \in L$ belong to a finite set. Hence

$$\beta_1\epsilon_1 + \beta_2\epsilon_2 + \beta_3\epsilon_3 = 0.$$

After division by $\beta_3\epsilon_3$,

$$(\beta_1/\beta_3)(\epsilon_1/\epsilon_3) + (\beta_2/\beta_3)(\epsilon_2/\epsilon_3) + 1 = 0.$$

From the Siegel-Mahler theorem we know that there exist at most finitely many solutions ϵ_1/ϵ_3 and ϵ_2/ϵ_3. Hence there are finitely many possibilities for the ratios ξ_1/ξ_2 and consequently a finite number of possibilities for $(x-\alpha_1)/(x-\alpha_2)$. So the finiteness of the solution set follows.

qed

In a similar way we can deal with the socalled *superelliptic equation*

Theorem 4.2.2 *Let $m \in \mathbb{Z}_{\geq 3}$ and $f \in \mathbb{Z}[x]$. Suppose f is not divisible by an m-th power has at least two distinct zeros. Then the equation*

$$y^m = f(x)$$

has at most a finite number of solutions $x, y \in \mathbb{Z}$.

4.3 Proof of Siegel-Mahler theorem in $\mathbb{Z}$

To avoid technicalities that have to deal with algebraic extension we restrict ourselves to $\mathbb{Q}$ and to so-called S-units. Let S be a finite set of primes in $\mathbb{Z}$. The ring generated by $\mathbb{Z}$ and all inverse $1/p$ with $p \in S$ is denoted by $\mathbb{Z}_S$ (localization of $\mathbb{Z}$ at S). The units in $\mathbb{Z}_S$ are precisely those rational numbers r such that $|r|_p = 1$ for all $p \notin S$. The statement reads as follows.

Theorem 4.3.1 *Let S be a finite set of primes in $\mathbb{Z}$. Then the number of solutions to*

$$x + y = 1$$

in $x, y \in \mathbb{Z}_S^*$ is at most finite.

After multiplying out denominators of possible solutions we get the following corollary.

Corollary 4.3.2 . *Let S be a finite set of primes in $\mathbb{Z}$. Consider the equation*

$$A + B = C \tag{S}$$

in positive integers A, B, C that factor into primes in S, and $\gcd(A, B, C) = 1$. Then the number of solutions is at most finite.

Note that this corollary is equivalent to the original statement. So we just prove the corollary.

Proposition 4.3.3 *Let p, q, r be postive integers and consider the following polynomials in z.*

$$R(z) = \frac{(p+q+r+1)!}{p!q!r!} \int_0^1 x^p (1-x)^q (x-z)^r \, dx$$

$$Q(z) = \frac{(p+q+r+1)!}{p!q!r!} \int_0^1 x^p (1-xz)^q (1-x)^r \, dx$$

$$P(z) = \frac{(p+q+r+1)!}{p!q!r!} \int_0^1 (1-(1-z)x)^p x^q (1-x)^r \, dx.$$

Then $P, Q, R \in \mathbb{Z}[z]$, the degrees of P, Q, R are p, q, r respectively and

$$R(z) = (-1)^r z^{p+r+1} Q(z) + (1-z)^{q+r+1} P(z).$$

Proof. We prove that $R \in \mathbb{Z}[z]$, the proof for P, Q runs similarly. Expanding the integrand of R in powers of z shows that we get terms of the form $\binom{r}{k}\binom{q}{l}x^{p+k+l}z^{r-k}$ with $0 \leq k \leq r, 0 \leq l \leq q$. They are integrated and then multiplied by $\frac{(p+q+r+1)!}{p!q!r!}$. Hence we get coefficients,

$$\frac{(p+q+r+1)!}{p!q!r!} \times \binom{r}{k}\binom{q}{l} \times \frac{1}{p+k+l+1}.$$

Wrting out everything into factorials gives us

$$\frac{(p+q+r+1)!(p+k+l)!}{(p+k+l+1)!p!k!l!(r-k)!(q-l)!}.$$

This equals the product of binomial coefficients

$$\binom{p+q+r+1}{q+r-k-l}\binom{p+k+l}{p}\binom{k+l}{k}\binom{q-l+r-k}{r-k} \in \mathbb{Z}.$$

Hence $R(z) \in \mathbb{Z}[z]$.
Denote $G(x,z) = x^p(1-x)^q(x-z)^r$ and observe the trivial identity

$$\int_0^1 G(x,z)\ dx = \int_0^z G(x,z)\ dx + \int_z^1 G(x,z)\ dx.$$

In the first integral on the right we apply the substitution $x \to xz$ to obtain

$$(-1)^r z^{p+r+1} \int_0^1 x^p(1-xz)^q(1-z)^r\ dx.$$

In the second integral on the right we substitute $x \to 1-(1-z)x$ and we obtain

$$(1-z)^{q+r+1} \int_0^1 (1-(1-z)x)^p x^q(1-x)^r\ dx.$$

This proves our identity.

qed

Now choose $n \in \mathbb{N}$ and let P_n, Q_n, R_n be the above polynomials for the choice $q = r = n, p = n-1$ and let $\tilde{P}_n, \tilde{Q}_n, \tilde{R}_n$ be the above polynomials for the choice $p = r = n, q = n-1$.

Proposition 4.3.4 *Let $n \in \mathbb{N}$ be given and let $A(z)$ be any of the six polynomials $(1-z)P_n, Q_n, R_n, \tilde{P}_n, z\tilde{Q}_n, \tilde{R}_n$. Then, for any $\alpha \in \mathbb{R}$ we have*

$$|A(\alpha)| \leq (6\sqrt{3})^n \max(1, |\alpha|^n).$$

Secondly,

$$Q_n\tilde{P}_n - z(1-z)\tilde{Q}_nP_n = c_n \neq 0,$$

where c_n is a constant polynomial.

Proof. We prove the estimate for $R_n(\alpha)$, the others run similarly. Notice

$$\begin{aligned} |R_n(\alpha)| &= \frac{(3n+1)!}{(n!)^3}\left|\int_0^1 x^n(1-x)^n(x-\alpha)^n\,dx\right| \\ &\leq \frac{(3n+1)!}{(n!)^3}\int_0^1 |x^n(1-x)^n(x-\alpha)^n|\,dx. \end{aligned}$$

It is a straightforward exercise to show that

$$\max_{x\in[0,1]} |x(1-x)(x-\alpha)|\max(1,|\alpha|) \leq 6\sqrt{3}$$

and

$$\frac{(3n+1)!}{n!^3} \leq 27^n$$

for all $n \in \mathbb{N}$ (by induction on n). This yields our estimate for $|R_n(\alpha)|$.
Consider the identities

$$\begin{aligned} R_n(z) &= (-1)^n z^{2n} Q_n(z) + (1-z)^{2n+1} P_n(z) \\ \tilde{R}_n(z) &= (-1)^n z^{2n+1} \tilde{Q}_n(z) + (1-z)^{2n} \tilde{P}_n(z) \end{aligned}$$

Subtract $z\tilde{Q}_n$ times the first equality from Q_n times the second equality. We obtain

$$z\tilde{Q}_n(z)R_n(z) - Q_n(z)\tilde{R}_n(z) = (1-z)^{2n}((1-z)zP_n(z)\tilde{Q}_n - Q_n(z)\tilde{P}_n).$$

By construction we know that the degrees of $z\tilde{Q}_n R_n$ and $Q_n\tilde{R}_n$ are $\leq 2n$. Hence, after division by $(1-z)^{2n}$ we find that $(1-z)zP_n(z)\tilde{Q}_n - Q_n(z)\tilde{P}_n)$ is a constant. Since $\tilde{P}_n(0) \neq 0$ and $Q_n(0) \neq 0$ this constant is non-zero.

qed

Corollary 4.3.5 *Let a, b, c be integers with $c = a + b$ and let $n \in \mathbb{N}, \lambda \in \mathbb{R}$. Then there exist integers A_n, B_n, C_n such that*

1. $a^{2n}A_n + b^{2n}B_n = c^{2n}C_n$.
2. $\max(|A_n|, |B_n|, |C_n|) \leq (6\sqrt{3})^n \max(|a|, |b|, |c|)^n$
3. $A_n - \lambda B_n \neq 0$

Proof. Define the polynomials P_n, Q_n, R_n and $\tilde{P}_n, \tilde{Q}_n, \tilde{R}_n$ as above. Take $A_n = (-1)^n c^n Q_n(a/c),\ B_n = c^{n-1}bP_n(a/c),\ C_n = c^n R_n(a/c)$. The identity $a^{2n}A_n + b^{2n}B_n = c^{2n}C_n$ follows immediately from $R_n(a/c) = (-1)^n(a/c)^{2n}Q_n(a/c) + (b/c)^{2n+1}P_n(a/c)$. Furthermore,

$$|C_n| = |c^n R_n(a/c)| \leq (6\sqrt{3})^n |c|^n |\max(1, |a/c|)^n \leq (6\sqrt{3})^n \max(|a|, |b|, |c|)^n$$

and similarly for $|A_n|$ and $|B_n|$.
Suppose that $A_n - \lambda B_n = 0$. Then we define instead $A_n = (-1)^n c^{n-1} a\tilde{Q}_n(a/c),\ B_n = c^n\tilde{P}_n(a/c),\ C_n = c^n\tilde{R}_n(a/c)$. For this choice of A_n, B_n, C_n we do have $A_n - \lambda B_n \neq 0$.

qed

We use Corollary 4.3.5 to prove another one.

Corollary 4.3.6 *Let $a, b, c, \alpha, \beta, gamma$ be non-zero integers such that $a+b = c$, $\gcd(a, b, c) = 1$ and $\alpha a^{2n} + \beta b^{2n} = \gamma c^{2n}$. Then*

$$\max(|a|, |b, |c|) \leq 6\sqrt{3} \cdot 2^{1/n} \max(|\alpha|, |\beta|, |\gamma|)^{1/n}.$$

Proof. We construct A_n, B_n, C_n satisfying the properties of Corollary 4.3.5 with $\lambda = \alpha/\beta$. Suppose that $|c| \geq \max(|a|, |b|)$. From the equalities $A_n a^{2n} + B_n b^{2n} = C_n c^{2n}, \alpha a^{2n} + \beta b^{2n} = \gamma c^{2n}$ and $\gcd(a, b, c) = 1$ it follows that c^{2n} divides $A_n\beta - B_n\alpha$. Since $A_n\beta - B_n\alpha \neq 0$ we find the estimates

$$|c|^{2n} \leq |\beta A_n| + |\alpha B_n| \leq 2\max(|\alpha|, |\beta|, |\gamma|)(6\sqrt{3})^n \max(|a|, |b|, |c|)^n.$$

Since $|c| = \max(|a|, |b|, |c|)$ it follows that

$$\max(|a|, |b|, |c|)^n \leq 2 \cdot (6\sqrt{3})^n \max(|\alpha|, |\beta|, |\gamma|).$$

Our assertion follows.

qed

Proof of Theorem 4.3.1. We subdivide the solutions of equation (S) according to their exponent vectors. More particularly, write

$$A = \prod_{p \in S} p^{a_p},\ B = \prod_{p \in S} p^{b_p},\ C = \prod_{p \in S} p^{c_p}$$

and interpret the concatenation of sequences $(a_p)_{p\in S}, (b_p)_{p\in S}(c_p)_{p\in S}$ as a vector in $\mathbb{R}^{3s}$ where $s = |S|$. Thus we obtain a mapping ϕ from the set of triples A, B, C to $\mathbb{R}^{3s}$. In particular the solution set of (S) is mapped into $\mathbb{R}^{3s}$. Vectors in $\mathbb{R}^{3s}$ can be denoted as $(\mathbf{x}, \mathbf{y}, \mathbf{z})$ where $\mathbf{x}, \mathbf{y}, \mathbf{z}) \in \mathbb{R}^s$. On $\mathbb{R}^{3s}$ we define the norm

$$||(\mathbf{x}, \mathbf{y}, \mathbf{z})|| = \max(\sum_p |x_p| \log p, \sum_p |y_p| \log p, \sum_p |z_p| \log p).$$

Notice in particular that $||\phi(A, B, C)|| = \log \max(|A|, |B|, |C|)$. We can subdivide $\mathbb{R}^{3s}_{\geq 0}$ into positive cones as follows. Choose $\epsilon > 0$. Then there exists a finite set $V \subset \mathbb{R}^{3s}_{\geq 0}$ of vectors $\mathbf{v}$ with $||\mathbf{v}|| = 1$ having the following property: to any non-zero vector $\mathbf{x} \in \mathbb{R}^{3s}_{\geq 0}$ there exists $\mathbf{v} \in V$ such that $||(\mathbf{v} - \mathbf{x}/||\mathbf{x}||)|| \leq \epsilon$. We say that $\mathbf{x}$ is in the cone defined by $\mathbf{v}$. We will prove that any such cone can contain at most a finite number of points $\phi(A, B, C)$ where (A, B, C) is a solution of (S).

Let $\mathbf{v} \in V$ and suppose its cone contains infinitely many solutions. Let (a, b, c) and (A, B, C) be two such solutions of (S). We may suppose that $\max(A, B, C) > \max(a, b, c)$. Then, by definition of a cone,

$$\left|\left| \frac{\phi(A, B, C)}{||\phi(A, B, C)||} - \frac{\phi(a, b, c)}{||\phi(a, b, c)||} \right|\right| \leq 2\epsilon.$$

Hence

$$\left|\left| \phi(A, B, C) - \frac{||\phi(A, B, C)||}{||\phi(a, b, c)||} \phi(a, b, c) \right|\right| \leq 2\epsilon ||\phi(A, B, C)||.$$

Now let n be the integer such that the difference between $2n$ and $\frac{||\phi(A,B,C)||}{||\phi(a,b,c)||}$ is minimal. Then

$$|\phi(A,B,C) - 2n\phi(a,b,c)| \leq 2\epsilon||\phi(A,B,C)|| + ||\phi(a,b,c)||.$$

This inequality implies that $A/a^{2n}, B/b^{2n}, C/c^{2n}$ are rational numbers whose numerators and denominators are bounded by $\max(A,B,C)^{2\epsilon}\max(a,b,c)$. From the construction of n we know that $\max(A,B,C) \leq \max(a,b,c)^{2n+1}$. Hence, the numerator and denominator of $A/a^{2n}, B/b^{2n}, C/c^{2n}$ are bounded by $\max(a,b,c)^{(4n+2)\epsilon+1}$. Let D be the common denominator of these numbers. We apply Corollary 4.3.6 with $\alpha = DA/a^{2n}, \beta = DB/b^{2n}, \gamma = DC/c^{2n}$. We obtain

$$\max(a,b,c) \leq 6\sqrt{3} \cdot 2^{1/n}(\max(a,b,c)^{((16n+8)\epsilon+8)/n}.$$

Since we have assumed infinitely solutions in V we can take n as large as we like and we conclude that

$$\max(a,b,c) \leq 6\sqrt{3}\max(a,b,c)^{16\epsilon}.$$

Take $\epsilon < 1/16$ and taking $\max(a,b,c)$ large enough we get a contradiction.
Thus the cone defined by $\mathbf{v}$ contains at most finitely many solutions.

qed

4.4 Exercises

1. Solve the diophantine equation $y^2 = x(x-1)(x-2)$ in $x, y \in \mathbb{Z}$.

2. Solve the diophantine equation $y^2 = x(x-1)(x-2)(x-3)$ in $x, y \in \mathbb{Z}$. You can do it in the usual way, but you can also notice that $x(x-1)(x-2)(x-3) = (x^2-3x+1)^2 - 1$.

3. Solve the diophantine equation $y^2 = x(x-1)(x-2)(x-3)(x-4)(x-5)$ in $x, y \in \mathbb{Z}$. Hint: you can do it the usual way, but you can also use the identity

$$x(x-1)(x-2)(x-3)(x-4)(x-5)$$

$$= (x^3 - 15x^2/2 + 115x/8 - 75/16)^2 - 189x^2/64 + 945x/64 - 5625/256.$$

4. Solve the diophantine equation $1 + 2^x = 3^y$ in integers $x, y \geq 0$. Also solve $1 + 3^x = 2^y$.

5. Let a be a positive integer > 20. We will solve $1 + a^x = (a+1)^y$ in integers $x, y \geq 0$ using the polynomial method sketched in the previous section and show that $x = y = 1$ is the only solution. Assume that we are given a solution x, y.

   (a) Show that $x \geq y \geq 0.9x$.

   (b) Show that x is odd.

   (c) Let $x = 2n+1$. Using Corollary 4.2.5 construct integers A, B, C such that $A + a^{2n}B = (a+1)^{2n}C$ and $A \neq B$ and $|A|, |B|, |C| \leq (6\sqrt{3})^n(a+1)^n$.

   (d) Replace in the above equality a^{2n} by $(a+1)^y - 1$ and show that we get a contradiction.