On the equation $x + y = 1$ in finitely generated multiplicative groups in positive characteristic

Peter Koymans
Universiteit Leiden

Intercity seminar

Leiden, The Netherlands, March 2017

Introduction

Let K be a number field with unit group $\mathcal{O}_K^*$. For fixed $a, b, c \in K^*$ consider the unit equation

$$ax + by = c$$

to be solved in $x, y \in \mathcal{O}_K^*$. Unit equations frequently show up when solving Diophantine equations. One well known example of such a Diophantine equation is the Thue equation

$$F(x, y) = \delta \text{ in } x, y \in \mathbb{Z}$$

for a given square-free binary form $F(X, Y) \in \mathbb{Z}[X, Y]$ of degree $n \geq 3$ and $\delta \in \mathbb{Z} \setminus \{0\}$.

History of unit equations

Finiteness results have been proved for the following types of unit equations:

Siegel (1921): $ax + by = c$ in $x, y \in \mathcal{O}_K^*$,

Mahler (1933): $ax + by = c$ in $x, y \in \mathbb{Z}_S^*$

$\mathbb{Z}_S := \mathbb{Z}[(p_1 \cdots p_t)^{-1}]$ ($S = \{p_1, \ldots, p_t\}$ finite set of primes)

$\mathbb{Z}_S^* = \{\pm p_1^{e_1} \cdots p_t^{e_t} : e_i \in \mathbb{Z}\}$,

Lang (1960): $ax + by = c$ in $x, y \in A^*$

A arbitrary finitely generated domain

over $\mathbb{Z}$ of characteristic 0.

The above results are all ineffective.

History II

Mahler and Evertse gave explicit upper bounds for the number of solutions of unit equations. The best and most general result is due to Beukers and Schlickewei (1996). They considered

$$x + y = 1 \tag{1}$$

to be solved in $(x, y) \in G$, where G is a multiplicative subgroup of $\mathbb{C}^* \times \mathbb{C}^*$ with $\dim_{\mathbb{Q}}(G \otimes_{\mathbb{Z}} \mathbb{Q}) = r$.

Theorem 1

Equation (1) has at most 2^{8r+8} solutions $(x, y) \in G$.

Characteristic p

A natural question is to prove an analogue in positive characteristic. Let K be a field of characteristic p and let G be a finitely generated multiplicative subgroup of $K^* \times K^*$ with $\dim_{\mathbb{Q}}(G \otimes_{\mathbb{Z}} \mathbb{Q}) = r$. Consider the equation

$$x + y = 1 \tag{2}$$

to be solved in $(x, y) \in G$. Can one still hope to show that there are finitely many solutions?

No, consider for example $K = \mathbb{F}_p(t)$ and $G = \langle (t, 1 - t) \rangle$. Then we have

$$t^{p^k} + (1 - t)^{p^k} = 1$$

for all integers $k \geq 0$, leading to infinitely many solutions of (2).

Our result

In view of the previous one can hope to show finiteness "up to Frobenius".

Theorem 2 (joint work with Carlo Pagano)

Let K be a field of characteristic $p > 0$ and let G be a finitely generated multiplicative subgroup of $K^* \times K^*$ with $\dim_{\mathbb{Q}}(G \otimes_{\mathbb{Z}} \mathbb{Q}) = r$. Then the equation

$$x + y = 1 \text{ in } (x, y) \in G \tag{3}$$

has at most $31 \cdot 19^r$ solutions (x, y) satisfying $(x, y) \notin G^p$. Equivalently, there is a set S of cardinality $|S| \leq 31 \cdot 19^r$ such that any solution of (3) with $x, y \notin \overline{\mathbb{F}_p}$ is of the shape s^{p^k}, where $s \in S$ and $k \in \mathbb{Z}_{\geq 0}$.

This answers a conjecture of Voloch (1998), who had previously shown a bound of the shape p^{Cr} with C an absolute constant.

Proof outline

Our proof is a modified version of the proof due to Beukers and Schlickewei. Their proof consists of roughly four parts:

1. Reduce to the case that G is a finitely generated subgroup of $\overline{\mathbb{Q}}^* \times \overline{\mathbb{Q}}^*$.
2. Prove several height inequalities for the solutions.
3. Map the solutions to a normed vector space V.
4. Transfer the height inequalities to V and deduce finiteness.

We will first discuss the proof of Beukers and Schlickewei in characteristic 0 and then later on the necessary modifications in characteristic $p > 0$.

Heights

Let K be a number field. Denote by M_K the set of places of K, i.e. the equivalence classes of absolute values on K. The height of $x \in K^*$ is defined by

$$H_K(x) := \sum_{v \in M_K} \max(0, \log |x|_v),$$

where we normalize $|\cdot|_v$ in a "nice" way. The normalization can be done in such a way that we have the sum formula

$$\sum_{v \in M_K} \log |x|_v = 0 \text{ for } x \in K^*$$

and moreover, $H_K(x)$ does not depend on the number field K containing x. More generally, we define

$$H_K(x_0, \ldots, x_n) = \sum_{v \in M_K} \max(\log |x_0|_v, \ldots, \log |x_n|_v).$$

By the sum formula this defines a height on the K-rational points of the projective space $\mathbb{P}^n$.

First height bound for solutions

We start with an easy lemma.

Lemma 3

Let a, b, c be non-zero elements of K, and let (x_i, y_i, z_i) for $i = 1, 2$ be two K-linearly independent vectors from K^3 such that $ax_i + by_i + cz_i = 0$ for $i = 1, 2$. Then

$$H_K(a, b, c) \leq H_K(x_1, y_1, z_1) + H_K(x_2, y_2, z_2) + \log 2.$$

Corollary 4 (Gap principle)

Let $u, v \in \overline{\mathbb{Q}}^* \times \overline{\mathbb{Q}}^*$ be two solutions of $x + y = 1$ with $u \neq v$. Then we have $H_K(1, u_1, u_2) \leq H_K(1, v_1/u_1, v_2/u_2) + \log 2$.

Some polynomials from Diophantine approximation

Define for $N \in \mathbb{Z}_{>0}$ the binary form $W_N(X, Y) \in \mathbb{Z}[X, Y]$

$$W_N(X, Y) = \sum_{m=0}^{N} \binom{2N-m}{N-m} \binom{N+m}{m} X^{N-m}(-Y)^m.$$

Lemma 5

Put $Z := -X - Y$. Then we have the following identities in $\mathbb{Z}[X, Y]$.

(i) $W_N(Y, X) = (-1)^N W_N(X, Y)$;

(ii) $X^{2N+1} W_N(Y, Z) + Y^{2N+1} W_N(Z, X) + Z^{2N+1} W_N(X, Y) = 0$;

(iii) *there exists a non-zero integer c_N such that*

$$\det \begin{pmatrix} Z^{2N+1} W_N(X, Y) & Y^{2N+1} W_N(Z, X) \\ Z^{2N+3} W_{N+1}(X, Y) & Y^{2N+3} W_{N+1}(Z, X) \end{pmatrix}$$
$$= c_N (XYZ)^{2N+1}(X^2 + XY + Y^2).$$

Second height bound for solutions

We can now make good use of the polynomials $W_N(X, Y)$.

Lemma 6

Let $u, v \in \overline{\mathbb{Q}}^* \times \overline{\mathbb{Q}}^*$ be two solutions of $x + y = 1$ with $u \neq v$. Then for every integer $N > 0$ there exists $M \in \{N, N+1\}$ such that

$$H_K(1, u_1, u_2) \leq \frac{1}{M+1} H_K(1, v_1/u_1^{2M+1}, v_2/u_2^{2M+1}) + \log 8.$$

Proof sketch.

Use the identities

$$u_1^{2M+1} W_M(u_2, -1) + u_2^{2M+1} W_M(-1, u_1) - W_M(u_1, u_2) = 0$$
$$u_1^{2M+1}(v_1 u_1^{-2M-1}) + u_2^{2M+1}(v_2 u_2^{-2M-1}) - 1 = 0$$

and apply Lemma 3. □

A normed vector space

Suppose from now on that G is a finitely generated subgroup of $\overline{\mathbb{Q}}^* \times \overline{\mathbb{Q}}^*$ of rank r. Then there exists an algebraic number field K and a finite set of places S of K containing all infinite places such that

$$G \subseteq \mathcal{O}_S^* \times \mathcal{O}_S^*.$$

Let $[K : \mathbb{Q}] = d$, $|S| = s$. Define a homomorphism $\varphi : G \to \mathbb{R}^{2s}$ by

$$\varphi : (x_1, x_2) \mapsto (\log |x_i|_v : v \in S, i = 1, 2)\,.$$

Let $V \subseteq \mathbb{R}^{2s}$ be the real vector space spanned by $\varphi(G)$. Then V has dimension r. We have

$$H_K(x_1) + H_K(x_2) = ||\varphi(x_1, x_2)||,$$

where $||\cdot||$ is the norm on $\mathbb{R}^{2s}$ defined by

$$||u|| = \frac{1}{2} \sum_{v \in S} \sum_{i=1}^{2} |u_{iv}|.$$

Normed vector space II

Write $\mathcal{S}$ for the image under φ of the set of $(x, y) \in G$ with $x + y = 1$. One can show that φ is at most two to one when restricted to this set.

Lemma 7

The set $\mathcal{S}$ has the following properties:

(i) *for any two distinct $u_1, u_2 \in \mathcal{S}$ we have*

$$||u_1|| \leq 2||u_2 - u_1|| + \log 4;$$

(ii) *for any two distinct $u_1, u_2 \in \mathcal{S}$ and any positive integer N, there is $M \in \{N, N+1\}$ such that*

$$||u_1|| \leq \frac{2}{M+1}||u_2 - (2M+1)u_1|| + \log 64;$$

(iii) *for any three distinct $u_0, u_1, u_2 \in \mathcal{S}$ we have*

$$||u_1 - u_0|| + ||u_2 - u_0|| > 0.09.$$

Completing the proof

The three properties on the previous slide are enough to prove finiteness of $\mathcal{S}$ for any r-dimensional normed vector space V. Furthermore, this upper bound can be made to depend only on r.

The key idea is to subdivide V into C^r cones where C is some absolute constant. Then one can show that there can only be a bounded number B of points from $\mathcal{S}$ inside each cone by using the three properties.

This leads to an upper bound of the shape $B \cdot C^r$.

Positive characteristic

Most of the machinery from the Beukers and Schlickewei proof carries through to positive characteristic. There are two obvious issues:

(i) Recall that there exists a non-zero integer c_N such that

$$\det \begin{pmatrix} Z^{2N+1} W_N(X,Y) & Y^{2N+1} W_N(Z,X) \\ Z^{2N+3} W_{N+1}(X,Y) & Y^{2N+3} W_{N+1}(Z,X) \end{pmatrix} = c_N (XYZ)^{2N+1} (X^2 + XY + Y^2).$$

But now we would like $c_N \not\equiv 0 \bmod p$, which imposes some restrictions on N.

(ii) The Beukers and Schlickewei method shows finiteness and this is no longer true in positive characteristic.

A formula for c_N

So far no explicit formula was known for c_N in the literature. We were able to derive an explicit formula for c_N.

Lemma 8

We have

$$W_N(2,-1) = \sum_{i=0}^{N} \binom{2N-i}{N}\binom{N+i}{N} 2^{-i} = 4^N \binom{\frac{3}{2}N}{N}.$$

Proof.

This follows from some identities for hypergeometric functions. □

Corollary 9

Let p be an odd prime number and let N be a positive integer with $N < \frac{p}{3} - 2$. Then $c_N \not\equiv 0 \bmod p$.

Proof.

Use the previous lemma to give an explicit formula for c_N. □

Height bounds for solutions in positive characteristic

Let K be a finitely generated field of characteristic $p > 0$. We can endow K with a set of discrete valuations satisfying a sum formula and with this we can define heights in a similar way as for number fields.

Note that there are no archimedean places in positive characteristic. This enables us to give slightly better height inequalities.

We get a similar homomorphism $\varphi : G \to \mathbb{R}^{2s}$ from G into a finite dimensional vector space. Denote by $\mathcal{S}$ the image under φ of all $(x, y) \in G$ satisfying $x + y = 1$ and $x, y \notin \overline{\mathbb{F}_p}$. In this case the restriction of φ to the set just defined is injective.

Properties of $\mathcal{S}$

In our new setting $\mathcal{S}$ has the following properties.

Lemma 10

(i) *for any two distinct* $u, v \in \mathcal{S}$ *we have*

$$||u|| \leq 2||v-u||;$$

(ii) *for any two distinct* $u, v \in \mathcal{S}$ *and any positive integer* N *such that* $N < \frac{p}{3} - 2$*, there is* $M \in \{N, N+1\}$ *such that*

$$||u|| \leq \frac{2}{M+1}||v-(2M+1)u||;$$

(iii) $p\mathcal{S} \subseteq \mathcal{S}$.

How to finish the proof

Just as before we divide our vector space into cones. The “gap principle” still holds, hence two points in $\mathcal{S}$ can not be too close inside the same cone. But how are we going to show that points in $\mathcal{S}$ can not be too far apart?

Idea: we want the RHS of Lemma 10(ii) to be small. In the Beukers and Schlickewei proof Lemma 10(ii) was true for all integers $N > 0$, but in our case it is only true for $N < \frac{p}{3} - 2$. So Lemma 10(ii) does not give much information if $||v||$ is much greater than $||u||$.

The solution

Let u and v be distinct points in $\mathcal{S}$ lying in the same cone with $||u|| < ||v||$. Then we can apply Frobenius a number of times to get a new point $u' \in \mathcal{S}$ such that

$$1 \leq \frac{||u'||}{||v||} \leq \sqrt{p} \text{ or } 1 \leq \frac{||v||}{||u'||} \leq \sqrt{p},$$

where $u' = p^k u$ for some integer $k \geq 0$. To see this, we can construct k explicitly as follows: define $\alpha \in \mathbb{R}$ such that $\frac{||u||}{||v||} = p^\alpha$. Then k is an integer such that $|\alpha - k| \leq \frac{1}{2}$.

Now we are in the position to apply Lemma 10(ii). This implies that

$$1 \leq \frac{||u'||}{||v||} \leq 100 \text{ or } 1 \leq \frac{||v||}{||u'||} \leq 100.$$

The solution II

Conclusion: for every family $\{p^k u\}_{k\geq 0}$ of points in $\mathcal{S}$ in a given cone we can pick a special member u' in such a way that all the u' chosen this way are close in the sense that:

$$1 \leq \frac{||u'||}{||v||} \leq 100 \text{ or } 1 \leq \frac{||v||}{||u'||} \leq 100.$$

Now apply the gap principle to conclude that there are a bounded number of families in each cone. This completes the proof.

The end

Questions?