The equation $ax + by = 1$ in characteristic p

José Felipe Voloch

Let K be a field and G a finite rank subgroup of the multiplicative group $K^* \times K^*$. Recently there has been a lot of attention devoted to the equation $ax + by = 1$, for given $a, b \in K^*$, where the solutions (x, y) are sought among the elements of G; in the case that K has characteristic zero. In particular, Schlickewei ([S]) proved that the number of solutions could be bounded solely in terms of the rank of G and the current best bound is due to Beukers and Schlickewei ([BS]), where they prove that $ax + by = 1$ has at most 2^{8r+16} solutions $(x, y) \in G$, where r is the rank of G. See also [B].

The purpose of this note is to prove an analogous result in positive characteristic. We note that the techniques of the aforementioned authors do not work in our situation. The first thing to notice is that, contrary to the characteristic zero situation, the equation $ax + by = 1$ can have infinitely many solutions. We will discuss this below. We will give a condition for the number of solutions to be finite and give a bound for the number of solutions in terms of the rank of the group when the condition is satisfied. Part of our approach is similar to that of [AV], [BV] with the geometry being simpler but, since the varieties in question are isotrivial, there are many added complications.

We will study a special case first and then reduce the general case to it. Let K be a field of characteristic $p > 0$ finitely generated over its prime subfield. By a result of Baer [Ba], K has a derivation D such that $\{x \in K \mid Dx = 0\} = K^p$. Let us fix a subgroup G of $K^* \times K^*$. We define two equations $a_i x + b_i y = 1$ to be G-equivalent if $(a_1/a_2, b_1/b_2) \in G$. Clearly there is a bijection between the set of solutions in G of two G-equivalent equations. We will call an equation $ax + by = 1$ G-trivial if $(a, b)^n \in G$ for some integer $n, (n, p) = 1$. If $ax + by = 1$ is G-trivial then it is G-equivalent to $a^q x + b^q y = 1$, for some power q of the characteristic p, that is, $(a^q, b^q) = (a, b)(x_1, y_1), (x_1, y_1) \in G$. If $(x_0, y_0) \in G$ is a solution to $ax + by = 1$ then so is $(x_0^q x_1, y_0^q y_1)$ and, it easy to see that, if $(x_0, y_0) \neq (x_0^q x_1, y_0^q y_1)$ then this process will lead to infinitely many solutions. Let us call an equation G-bad

if it is G-equivalent to an equation with $a, b \in K^p$ and G-good otherwise. Finally, let $H = \{(x, y) \in G \mid Dx = Dy = 0\}$, which is a subgroup of G containing G^p.

Lemma 1. *If $\#G/H = p^r$ and $ax + by = 1$ is G-good, then it has at most p^r solutions $(x, y) \in G$.*

Proof: If $ax + by = 1$ has more than p^r solutions, then it has two solutions $(x_i, y_i), i = 1, 2$ in the same class modulo H and, passing to a G-equivalent equation, we may assume it is the class corresponding to H itself, so $Dx_i = Dy_i = 0, i = 1, 2$. However, if $ax + by = 1$ and $Dx = Dy = 0$, differentiation yields $xDa + yDb = 0$ and the system $ax + by = 1, xDa + yDb = 0$ has at most one solution in $K^* \times K^*$, unless $Da = Db = 0$, in which case the equation is G-bad, proving the lemma.

We only need to study G-bad equations. Suppose $a, b \in K^p$, we will call the equation $ax + by = 1$ G-awful if it has a solution $(x, y) \in G \setminus H$.

Lemma 2. *If $\#G/H = p^r$ then there are at most $(p^r - 1)/(p - 1)$ G-equivalence classes of G-awful equations and that each such equation has at most $(p^r - 1)/(p - 1)$ solutions $(x, y) \in G \setminus H$.*

Proof: Suppose $(x, y) \in G \setminus H$ satisfy $ax + by = 1$, where $a, b \in K^p$. Then $aDx + bDy = 0$, so $a(Dx/x)x + b(Dy/y)y = 0$ and, eliminating by we get $ax(1 - (Dx/x)/(Dy/y)) = 1$. Now we notice that the map $(x, y) \mapsto (Dx/x, Dy/y)$ is a homomorphism from G to $K^+ \times K^+$, whose kernel is H and therefore has p^r elements in its image. Hence the quantity $(Dx/x)/(Dy/y)$ takes at most $(p^r - 1)/(p - 1)$ values as (x, y) varies in $G \setminus H$. But from the equation $ax(1 - (Dx/x)/(Dy/y)) = 1$ and its counterpart $by(1 - (Dy/y)/(Dx/x)) = 1$ we get that $ax + by = 1$ is G-equivalent to $cx + dy = 1$ where $(c, d) = ((1 - (Dx/x)/(Dy/y))^{-1}, (1 - (Dy/y)/(Dx/x))^{-1})$, which proves the first part of the lemma. For the second part, notice that the above calculation shows that (x, y) is determined by $(Dx/x)/(Dy/y)$.

Theorem 1. *If K is a field of characteristic $p > 0$, finitely generated over its prime field, and G is a finitely generated subgroup of $K^* \times K^*$ of rank r such that $\#G/H = p^r$, then*

an equation $ax + by = 1$ has at most $p^r(p^r + p - 2)/(p - 1)$ solutions $(x, y) \in G$ unless it is G-trivial.

Proof: Note that the hypotheses imply that $H = G^p$. If $ax + by = 1$ is G-good, then lemma 1 gives the result. If the equation is G-bad but not G-awful then we may assume $a, b \in K^p$ and the equation has as many solutions as $a^{1/p}x + b^{1/p}y = 1$, by extracting p-th roots. If the new equation is G-good, again we are done. If it is not G awful, we repeat. So either we will eventually reach a G-awful equation or we will continue indefinitely. If we reach a G-awful equation, lemma 2 bounds the number of solutions in $G \setminus G^p$ and for the solutions in G^p, we again extract p-th roots and move on to another equation. We keep repeating this process. If G-awful equations are encountered in $p^r + 1$ steps, then by Lemma 2, two of these equations are G-equivalent and this implies (a, b) equivalent to (a^q, b^q) for some power q of p, which means it is G-trivial. If G-awful equations are encountered in at most p^r steps and a G-good equation is then reached, we get at most $p^r(p^r - 1)/(p - 1) + p^r = p^r(p^r + p - 2)/(p - 1)$ solutions by lemmas 1 and 2.

The remaining possibility is that for all $n \geq 1$, the equation $ax + by = 1$ is G-equivalent to an equation $a_n^{p^n}x + b_n^{p^n}y = 1$. By the above $(a, b) = (a_n^{p^n}, b_n^{p^n})(x_n, y_n)$, for some $(x_n, y_n) \in G$. Lemma 3 below implies that $(a, b)^n$ belongs to G for some $n \geq 1$. If $p|n$, then $(a, b)^n$ is in H hence in G^p, since our hypotheses imply $H = G^p$. So $(a, b)^{n/p}$ also belongs to G. Proceeding in this fashion we may assume that $(n, p) = 1$ and then the equation is G-trivial, completing the proof of the theorem.

Lemma 3. *Let K be a field of characteristic p, finitely generated over a finite field, and G is a finitely generated subgroup of K^*. If $a \in (K^*)^{p^n}G$ for all $n \geq 1$ then there exists $m \geq 1, a^m \in G$.*

Proof: We proceed by induction on the absolute transcendence degree of K, with the case of transcendence degree zero being trivial. There exists a subfield F of K such that K/F is a function field in one variable. Also, we can find a finite set of places S of K/F, for which a and the elements of G are in U_S, the group of S-units of K. Now, U_S/F^* is

finitely generated, therefore so is $B = U_S/(F^*G)$. The image of a in B is, by hypothesis, in B^{p^n} for all $n \geq 1$ and therefore is in the torsion of B, so there exists $m \geq 1, a^m \in F^*G$. Writing $a = a_n^{p^n} x_n, a_n \in K, x_n \in G$, we see that $a_n \in U_S$ and that the image of a_n in B is torsion so, replacing m by a larger integer if necessary, we may assume that a_n^m maps to zero in B, that is, $a_n^m \in F^*G$. Since G is finitely generated, it has a subgroup G_1 such that $F^* \cap G_1$ is finite and $F^*G = F^*G_1$. Let us write $a^m = bc, a_n^m = b_n c_n, x_n = y_n z_n$, where $b, b_n, y_n \in F^*, c, c_n, z_n \in G_1$. If d is the order of $F^* \cap G_1$ we can conclude that $b^d = (b_n^d)^{p^n} y_n^d$. Therefore $b^d \in (F^*)^{p^n}(F^* \cap G)$ for all $n \geq 1$, so by the induction hypothesis applied to F we conclude that some power of b^d belongs to $F^* \cap G$, hence to G. On the other hand we have $a^m = bc$ and $c \in G$, so we conclude that some power of a belongs to G, as desired.

Remark: The lemma still holds for a field of arbitrary characteristic, finitely generated over its prime field and with p^n replaced by any unbounded sequence p_n of positive integers. The proof is essentially the same.

Theorem 2. *If K is a field of characteristic $p > 0$, and G is a subgroup of $K^* \times K^*$ with $\dim_{\mathbf{Q}} G \otimes \mathbf{Q} = r$ finite, then an equation $ax + by = 1$ has at most $p^r(p^r + p - 2)/(p - 1)$ solutions $(x, y) \in G$ unless $(a, b)^n \in G$ for some $n \geq 1$.*

Proof: Suppose that $ax + by = 1$ has more than $p^r(p^r + p - 2)/(p - 1)$ solutions $(x, y) \in G$, then we can replace G by a finitely generated subgroup of G with the same property. We can also replace K by a subfield, finitely generated over its prime field, containing the coordinates of the new G and a, b. Finally, at the cost of replacing G by a group containing G as a subgroup of finite, p-power index, we can assume that the condition $G^p = H$ is satisfied and we can apply Theorem 1 to get that $ax + by = 1$ is G-trivial for the new G, which implies the theorem. (Because of the last step of the reduction to theorem 1 we cannot guarantee $(n, p) = 1$).

Remarks: (i) If K is a field complete with respect to a discrete valuation, then the principal units of K form a $\mathbf{Z}_p$-module and there is an analogue of Theorem 2 for subgroups of finite $\mathbf{Z}_p$-rank (or even finite $\mathbf{Q}_p$-dimension in the appropriate sense).

(ii) Theorem 2 is a very special case of the Mordell-Lang conjecture (see, e.g. [AV],[H]). It is the first instance in which the full division hull of a finitely generated group in characteristic $p > 0$ is handled, as opposed to the prime-to-p division hull. The key point is that the estimate for the number of points on Theorem 1 depends on the rank only. If, for instance, the estimate of [BV] could be proved for all non-isotrivial curves, instead of just the ones not defined over K^p, the Mordell-Lang conjecture for curves would follow.

(iii) Of course, if G in Theorem 2 is a divisible group then $(a, b)^n \in G$ for some $n \geq 1$ is the same as $(a, b) \in G$.

We wish to consider a couple of special cases, where more can be said.

First suppose $G = \{(t^n, t^m) \mid n, m \in \mathbf{Z}\}$, for some $t \in K^*$, t not a root of unity. So G is a group of rank two and C. Kang (personal communication) has shown that an equation $ax + by = 1$ has at most p^2 solutions $(x, y) \in G$, unless it is G-trivial. This follows from the above and the simple observation that a G-awful equation is G-trivial in this case. However, much more can be said and we proceed to show that in fact the number of solutions is at most two. Indeed, if $(1, 1), (t^n, t^m), (t^r, t^s)$ are solutions of $ax + by = 1$ then

$$\begin{vmatrix} 1 & 1 & 1 \\ 1 & t^n & t^m \\ 1 & t^r & t^s \end{vmatrix} = 0.$$

Therefore $t^{n+s} + t^m + t^r - t^{m+r} - t^n - t^s = 0$. As t is transcendental over $\mathbf{F}_p$ we conclude that the equation in t is identically zero. This can only happen in a finite number of ways. More precisely, six ways, if $p \neq 2$ and twelve ways for $p = 2$. A case-by-case analysis leads to a contradiction each time. For example, if $n + s = n, m = s, r = m + r$, then $s = m = 0$ and the only equation having three solutions $(1, 1), (t^n, 1), (t^r, 1)$ has $a = 0$, which is not being considered.

Now let us analyse the case of rank one, i.e., $G = \{(u^n, v^n) \mid n \in \mathbf{Z}\}$. From lemma 2 we know that there is at most one G-awful equation and this equation has one solution in $G \setminus H$, if it exists. So we need only study G-good equations and may assume $H = G^p$, so a G-good equation will have at most p solutions, one in each coset of H in G. We can

actually show that a G-good equation will have at most three solutions, so an equation $ax+by=1$ with $(a,b)^n \notin G$ for any $n \geq 1$, will have at most four solutions. Let us sketch the argument. Denote by $F_{nm}(X,Y)$, for $0<n<m$ integers, the determinant

$$\begin{vmatrix} 1 & 1 & 1 \\ 1 & X^n & Y^n \\ 1 & X^m & Y^m \end{vmatrix}.$$

Suppose $ax+by=1$ has solutions $(1,1),(u^n,v^n),(u^m,v^m),(u^r,v^r)$, where $0<n<m<r$, and $0,n,m,r$ are distinct modulo p, then u,v satisfy $F_{nm}(u,v)=F_{nr}(u,v)=0$. Since G has rank one, $\mathbf{F}_p(u,v)$ is transcendental over $\mathbf{F}_p$, but if the above equations are satisfied, the transcendence degree of $\mathbf{F}_p(u,v)$ over $\mathbf{F}_p$ is one. hence it enough to show that the plane curves given by $F_{nm}(X,Y)=0, F_{nr}(X,Y)=0$ cannot have a component in common besides $X=Y, X=1, Y=1$. The reader will verify easily that the curve $F_{nm}(X,Y)=0$ has $n+m-3$ branches at infinity, besides $X=Y, X=1, Y=1$, with expansions

$$Y=\alpha X+\beta X^{-(n-1)}+\cdots, \alpha^{m-n}=1, \alpha\neq 1, \beta=(1-\alpha^n)\alpha^{1-n}/(m-n),$$

$$Y=\alpha+\beta X^{-(m-n)}+\cdots, \alpha^n=1, \alpha\neq 1, \beta=\alpha(\alpha^m-1)/n,$$

$$X=\alpha+\beta Y^{-(m-n)}+\cdots, \alpha^n=1, \alpha\neq 1, \beta=\alpha(\alpha^m-1)/n.$$

If $F_{nm}(X,Y)=0, F_{nr}(X,Y)=0$ have a component of the first kind above in common, then α and β will be the same and, looking at β we get $m-n\equiv r-n(\bmod p)$, contradicting the hypothesis. If the component is of the other two kinds, then looking at the exponent of the second term of the expansion gives $m-n=r-n$, again a contradiction. The cases where the exponent in (u,v) of a solution is negative can be handled similarly. Alternatively, the reader may notice that, if the equation has at least six solutions, then replacing (u,v) by (u^{-1},v^{-1}) if necessary, it can be assumed that three of the exponents are positive. Be as it may, we get a small absolute bound for the number of solutions. It is tempting to ask if a bound, depending on r but not on p can be given in the general case, like in characteristic zero.

Acknowledgements: The author would like to thank C. Kang for the calculations mentioned above and D. Saltman for pointing out Baer's result. The author would also like to thank the NSA (grant MDA904-97-1-0037) for financial support.

References.

[AV] D. Abramovich and J. F. Voloch, *Toward a proof of the Mordell- Lang conjecture in characteristic p*, International Math. Research Notices No. 5 (1992) 103-115.

[Ba] R. Baer, *Algebraische Theorie der differentiierbaren Funktionenkorper I*, Sitzungberichte, Heidelberger Akademie, 1927, pp. 15-32.

[B] A. Buium, *Uniform bounds for generic points of curves in tori*, Crelle, **469** (1995) 211-219.

[BS] F. Beukers, H. P. Schlikewei, *The equation $x + y = 1$ in finitely generated groups*, Acta Arithmetica, **78** (1996/97), 189-199.

[BV] A. Buium, J. F. Voloch, *Mordell's conjecture in characteristic p: an explicit bound*, Compositio Math., 103 (1996) 1-6.

[H] E. Hrushovski, *The Mordell-Lang conjecture for function fields*, J. Amer. Math. Soc., **9** (1996) 667-690.

[S] H. P. Schlikewei, *Equations $ax + by = 1$*, Ann. of Math., to appear.

Dept. of Mathematics, Univ. of Texas, Austin, TX 78712, USA

e-mail: voloch@math.utexas.edu