ACTA ARITHMETICA
153.4 (2012)

Linear equations over multiplicative groups in positive characteristic

by

Dominik J. Leitner (Basel)

1. Introduction. Let K be a field and let G be a finitely generated subgroup of the multiplicative group K^*. Equations of the form

$$a_1x_1 + \cdots + a_nx_n = 1, \tag{1.1}$$

to be solved with unknowns $x_1, \ldots, x_n$ in G, play a central role in number theory and diophantine geometry. Here $a_1, \ldots, a_n$ are considered constants in K, and for convenience we take them in K^*. To avoid trivialities we can clearly suppose $n \geq 2$.

Much is known about (1.1). Here it is necessary to divide into two cases depending on the characteristic of K.

Suppose that K has zero characteristic. Then it was proved independently by Evertse [E] in 1984 and van der Poorten and Schlickewei [PS] in 1991 that there are at most finitely many solutions of (1.1) which satisfy the subsum restriction $\sum_{i\in I} a_ix_i \neq 0$ for every non-empty subset I of $\{1, \ldots, n\}$. This is a minor restriction because if it fails, then we may use induction to reduce the number of variables. In particular for three-term equations it shows that there are at most finitely many solutions, and for more terms it leads easily to a complete structure.

Now suppose that K has positive characteristic p. The result of Evertse, van der Poorten and Schlickewei then becomes false. The simplest counterexample comes from the equation

$$x + y = 1 \tag{1.2}$$

over the function field $K = \mathbb{F}_p(t)$ with $G = \langle t, 1 - t\rangle$ generated by t and $1 - t$. Namely if $q = 1, p, p^2, \ldots$ then

$$x = t^q, \quad y = 1 - t^q = (1 - t)^q \tag{1.3}$$

clearly supply infinitely many solutions unrestricted by subsums.

2010 *Mathematics Subject Classification*: 11D04, 11D72.
Key words and phrases: linear equations, equations in many variables, S-unit equation.

DOI: 10.4064/aa153-4-1

The situation here was clarified in 1992 by Abramovich and Voloch [AV] who showed (in a much more general context) that such counterexamples can arise only when the equation (1.1) is essentially defined over a finite field; of course (1.2) is literally defined over $\mathbb{F}_p$. See also the work of Voloch [V] for $n = 2$.

A less simple counterexample was observed in 2004 by Masser [M] for the equation

$$x + y - z = 1 \tag{1.4}$$

with the same K and G. Namely there is a doubly infinite family of solutions

$$x = t^{(q-1)q'}, \quad y = (1-t)^{qq'}, \quad z = t^{(q-1)q'}(1-t)^{q'} \tag{1.5}$$

with q, q' ranging independently over $1, p, p^2, \ldots$.

A full structure theorem was found at about the same time by Moosa and Scanlon [MS2], [MS2] (also in the more general context). Independently Derksen and Masser [DM] have given an alternative proof in the present context which is completely effective in the logical sense. As is well-known, this is not yet possible in zero characteristic. But of course there are many effective counting results; for brevity we mention here just [ESS] of Evertse, Schlickewei and Schmidt on (1.1) and (thanks to the referee) the paper [HP] of Hrushovski and Pillay for transcendental points in the more general context.

In the present paper we find all solutions of (1.2) and (1.4) for the above K and G. But first we state one of the main results of [DM] for general K and G, for simplicity in affine rather than projective form. Some preliminary definitions are needed.

We define a *G-automorphism* ψ of K^n by an equation

$$\psi(x_1, \ldots, x_n) = (g_1 x_1, \ldots, g_n x_n) \tag{1.6}$$

with $g_1, \ldots, g_n$ in G.

For a power q of the characteristic p we denote by $\varphi = \varphi_q$ the Frobenius with $\varphi(x) = x^q$. Let $\psi_1, \ldots, \psi_h$ be G-automorphisms. Then we imitate commutator brackets by defining the operator

$$[\psi_1, \ldots, \psi_h] = [\psi_1, \ldots, \psi_h]_q = \bigcup_{e_1=0}^{\infty} \cdots \bigcup_{e_h=0}^{\infty} (\psi_1^{-1}\varphi^{e_1}\psi_1) \cdots (\psi_h^{-1}\varphi^{e_h}\psi_h), \tag{1.7}$$

with of course the identity interpretation if $h = 0$.

For example with $q = p$, $h = 1$ and $\psi_1 = \psi$ as the identity automorphism, we have, for a point Π,

$$[\psi]_p \Pi = \bigcup_{e=0}^{\infty} \varphi^e \Pi = \{\Pi, \Pi^p, \Pi^{p^2}, \ldots\}$$

as in (1.3) with $\Pi = (t, 1-t)$. Or with $q = p, h = 2$ and again ψ_1 as the identity automorphism, $[\psi_1, \psi_2]_p\Pi$ is the union over $q = p^{e_2}$ and $q' = p^{e_1}$ of the $(\psi_2^{-1}(\psi_2\Pi)^q)^{q'}$. For suitable ψ_2, Π (see (1.8) below) this reduces to (1.5).

We will need the radical $\sqrt{G} = \sqrt[K]{G}$. For us this remains in K; thus it is the group of γ in K for which there exists a positive integer s such that γ^s lies in G.

We generalize (1.1) by treating linear varieties V defined by the vanishing of linear polynomials of degree at most 1 in $x_1, \dots, x_n$. Thus our object of study is $V \cap G^n$, which we abbreviate to $V(G)$. A special role is played by the case when all the equations have the form $x_i = a$ or $x_i = ax_j$; these we call linear cosets or just *cosets* for brevity. A point is of course a coset.

THEOREM (Derksen–Masser). *Let K be a field of positive characteristic p, let V be a linear variety defined over K, and suppose that $\sqrt[K]{G}$ is finitely generated. Then there is a power q of p such that $V(G)$ is an effectively computable finite union of sets $[\psi_1, \dots, \psi_h]_q T(G)$ with $\sqrt{G}$-automorphisms $\psi_1, \dots, \psi_h$ $(0 \le h \le n-1)$ and with cosets T contained in V.*

Here are our main results for (1.2) and (1.4), in which ψ_0 denotes the identity automorphism. First (1.2), whose statement (and proof) is relatively simple.

THEOREM 1. *Suppose that $K = \mathbb{F}_p(t)$, $G = \langle t, 1-t\rangle$ and the line L is defined by $x + y = 1$. Then $L(G)$ is $[\psi_0]_p\Pi^+ \cup [\psi_0]_p\Pi^-$ for the points*

$$\Pi^+ = (t, 1-t), \qquad \Pi^- = (1-t, t)$$

provided $p \ge 3$, and is

$$[\psi_0]_p\Pi^+ \cup [\psi_0]_p\Pi^- \cup [\psi_0]_p\Pi_1^+ \cup [\psi_0]_p\Pi_1^- \cup [\psi_0]_p\Pi_2^+ \cup [\psi_0]_p\Pi_2^-$$

for the additional points

$$\Pi_1^+ = \left(\frac{1}{t}, \frac{1-t}{t}\right), \qquad \Pi_1^- = \left(\frac{1-t}{t}, \frac{1}{t}\right),$$
$$\Pi_2^+ = \left(\frac{1}{1-t}, \frac{t}{1-t}\right), \qquad \Pi_2^- = \left(\frac{t}{1-t}, \frac{1}{1-t}\right)$$

when $p = 2$.

Thus for $p \ge 3$ we get not only (1.3) corresponding to Π^+ but also the extra solutions $x = (1-t)^q$, $y = t^q$ corresponding to Π^-. The reason is of course the symmetry of the equation in x, y. For $p = 2$ we get even more solutions, but these can be considered as coming from more symmetry which arises by writing the equation in homogeneous form as $X + Y + Z = 0$.

It is precisely this sort of symmetry which is responsible for the much more complicated situation in (1.4). Define the coset T_x by the equations $x = 1$, $y = z$, and similarly T_y, T_z.

THEOREM 2. *Suppose that $K = \mathbb{F}_p(t)$, $G = \langle t, 1-t\rangle$ and the plane P is defined by $x+y-z=1$. Then $P(G)$ is*

$$T_x(G) \cup T_y(G) \cup \bigcup_{\Pi\in\mathcal{T}} [\psi_0, \psi_\Pi]_p \Pi$$

for a set $\mathcal{T}$ of 40 points Π in G^3 with G-automorphisms ψ_Π provided $p \geq 5$; it is

$$T_x(G) \cup T_y(G) \cup \bigcup_{\Pi\in\mathcal{T}'} [\psi_0]_p \Pi \cup \bigcup_{\Pi\in\mathcal{T}} [\psi_0, \psi_\Pi]_p \Pi$$

for a set $\mathcal{T}$ of 40 points Π in G^3 with G-automorphisms ψ_Π and a set $\mathcal{T}'$ of eight points Π in G^3 when $p=3$; it is

$$T_x(G) \cup T_y(G) \cup T_z(G) \cup \bigcup_{\Pi\in\mathcal{T}'} [\psi_0]_p \Pi \cup \bigcup_{\Pi\in\mathcal{T}} [\psi_0, \psi_\Pi]_p \Pi$$

for a set $\mathcal{T}$ of 216 points Π in G^3 with G-automorphisms ψ_Π and a set $\mathcal{T}'$ of 24 points Π in G^3 when $p=2$.

For example $\mathcal{T}$ (for every p) includes the point $\Pi = (1, 1-t, 1-t)$, with

$$\psi_\Pi(x,y,z) = \left(tx, y, \frac{t}{1-t}z\right), \tag{1.8}$$

and then $[\psi_0, \psi_\Pi]_p \Pi$ is exactly the set (1.5). But there are in all 40 such classes of solutions when $p \geq 3$, and even 216 when $p = 2$.

As hinted above, the large numbers here arise essentially from the symmetry of the special equation $x+y-z=1$, which in homogeneous form $X+Y=Z+W$ has a natural dihedral D_4-action. When $p=2$ this is even an S_4-action. But in addition the nature of the special group $\langle t, 1-t\rangle$ can be exploited through field automorphisms, which yield an independent S_2-action and for $p=2$ even an S_3-action.

In view of the effectivity of [DM] our own results may not seem too significant, and things are naturally simpler for the special equation. Also the work of [DM] includes explicit estimates for everything appearing, and so at first sight it may seem that only a computer is needed. But in fact the matter is more complicated, for two main reasons.

First, the estimates in [DM] are not very small. For example equation (12.1) there involves an upper bound which in our situation is

$$B = (144.3^{10}(270.5^{15})^7)^{43} p^{86} > 10^{4185} p^{86}.$$

It follows, for example, that each of the g_i in (1.6) is a quotient of polynomials in t of degree at most B. Thus even for $p=2$ a very large computer would be needed.

Second, there is no uniformity in the characteristic p; the coefficients in the polynomials above lie in $\mathbb{F}_p$ and we get no algorithm for treating all p, even if the bound above were independent of p.

Our own results are essentially uniform in p. This is in tune with an existing vague philosophy that the solution set should not depend too much on p. Indeed Hrushovski [H, p. 669], who substantially generalized the work of [AV], has expressed the expectation that "quantifier elimination and elimination of imaginaries hold already in the differential language, without the distinguished basis, and in this language the proof should become entirely uniform with respect to the characteristic." As far as we know, our work is the first confirmation of this uniformity, albeit at an elementary level.

Actually our result for $p = 2$ can also be found in the recent article [ABB] of Arenas-Carmona, Berend and Bergelson. Thus our Theorem 1 for $p = 2$ is essentially their Lemma 5.6 (p. 348); our S_3-symmetry has been factored out. And our Theorem 2 for $p = 2$ is essentially their Proposition 4.1 (p. 345). Here the S_4-symmetry with 24 elements has reduced our set $\mathcal{T}$ with 216 points to nine quadrangles $Q_2, \dots, Q_{10}$ (compare our (F1),...,(F9) in Sections 4, 5, 6) and our $\mathcal{T}'$ to Q_1 (compare our Π_0 in Proposition 3).

Let us now say a few words about the proof. We follow broadly the strategy of [DM], which in general uses differential operators to replace the study of $V(G)$ by that of $W(G)$ for finitely many proper subvarieties W of V. Here we need only d/dt. For Theorem 1 about a line, we get at once points. But for Theorem 2 about a plane we have to cope with lines. Now there is no reason to suppose that these lines will be defined over finite fields, and so one might expect to encounter equations $ax + by = 1$ more general than (1.2). These might easily cause problems. But by carefully estimating we are in fact able to reduce to (1.2) itself.

The paper is arranged as follows. We prove Theorem 1 in Section 2. Then in Section 3 we record some preliminary observations; in particular the proof of Lemma 3.1 contains the crucial uniformity argument and Lemma 3.4 enables us to reduce to (1.2). In Section 4 we prove Theorem 2 for $p \geq 5$. A critical role is played by the field $C = \mathbb{F}_p(t^p)$, which is used to define for each solution (x, y, z) a quantity

$$d = d(x, y, z) = \dim_C(Cx + Cy + Cz).$$

Thus Propositions 1 and 2 treat the cases $d = 3$ and $d = 2$ respectively.

Then in a short Section 5 we prove Theorem 2 for $p = 3$.

Finally in Section 6 we do $p = 2$, which seems to cause quite a lot of complications, even though now the case $d = 3$ cannot occur. We study the case $d = 2$ in Proposition 3. Here it is very reassuring that we are in agreement with [ABB].

2. Proof of Theorem 1. Actually we determine the set $L(\sqrt{G})$ with L as above and $G = \langle t, 1 - t\rangle$ as above in $K = \mathbb{F}_p(t)$; from now on we abbreviate $\sqrt[K]{G}$ to $\sqrt{G}$.

LEMMA 2.1. *The set $\sqrt{G}$ is generated by G together with the elements of $\mathbb{F}_p^*$.*

Proof. For u in $\sqrt{G}$ we have s in $\mathbb{N}$ such that u^s is in G. Let A be any irreducible polynomial of $\mathbb{F}_p[t]$, which is not a constant multiple of $t, 1-t$. Then the corresponding order function satisfies $\operatorname{ord}_A u^s = 0$, which implies $\operatorname{ord}_A u = 0$. Since this holds for all such A, we see that u must lie in the set generated by G and $\mathbb{F}_p^*$. Conversely, we have $G \subset \sqrt{G}$ and further $a^{p-1} = 1$ for a in $\mathbb{F}_p^*$ shows that $\mathbb{F}_p^* \subset \sqrt{G}$, which completes the proof.

We will need to differentiate with respect to t, and we note that the field of differential constants here is $C = \mathbb{F}_p(t^p)$ (cf. [L1, pp. 185–186]).

PROPOSITION. *The set $L(\sqrt{G})$ is*

$$[\psi_0]_p \Pi^+ \cup [\psi_0]_p \Pi^- \cup [\psi_0]_p \tilde{\Pi}_1^+ \cup [\psi_0]_p \tilde{\Pi}_1^- \cup [\psi_0]_p \tilde{\Pi}_2^+ \cup [\psi_0]_p \tilde{\Pi}_2^- \cup \bigcup_{a=2}^{p-1} \Pi^{(a)}$$

for the points

$$\Pi^+ = (t, 1-t), \qquad \Pi^- = (1-t, t), \qquad \Pi^{(a)} = (a, 1-a)$$

and

$$\tilde{\Pi}_1^+ = \left(\frac{1}{t}, -\frac{1-t}{t}\right), \qquad \tilde{\Pi}_1^- = \left(-\frac{1-t}{t}, \frac{1}{t}\right),$$

$$\tilde{\Pi}_2^+ = \left(\frac{1}{1-t}, -\frac{t}{1-t}\right), \qquad \tilde{\Pi}_2^- = \left(-\frac{t}{1-t}, \frac{1}{1-t}\right),$$

where for $p = 2$ the union over a must be omitted.

Proof. We must investigate x and y in $\sqrt{G}$ with $x + y = 1$.

Assume first that the C-vector space $Cx + Cy$ has dimension 2. Using a dot to indicate the derivative with respect to t, we deduce $\dot{y}/y \neq \dot{x}/x$, else y/x would be in C. From $x + y = 1$ and its derivative $(\dot{x}/x)x + (\dot{y}/y)y = 0$ we get in the usual way the identities

$$x = \frac{\dot{y}/y}{\dot{y}/y - \dot{x}/x}, \qquad y = \frac{-\dot{x}/x}{\dot{y}/y - \dot{x}/x}. \tag{2.1}$$

Now if $h = at^r(1-t)^s$ is a typical element of $\sqrt{G}$, then

$$\frac{\dot{h}}{h} = \frac{r}{t} - \frac{s}{1-t} = \frac{r - (s+r)t}{t(1-t)} \tag{2.2}$$

takes just p^2 values, which are also the values of $-\dot{h}/h$. Since $\dot{y}/y - \dot{x}/x$ in (2.1) is $\dot{h}/h$ for $h = y/x$, it follows that x and y are non-zero quotients of these. But with the help of Lemma 2.1 it is easily seen that such a quotient $x = \frac{r' - (s'+r')t}{r-(s+r)t}$ lies in $\sqrt{G}$ if and only if ax belongs to the list

$$1,\ t,\ 1-t,\ \frac{1}{t},\ -\frac{1-t}{t},\ \frac{1}{1-t},\ -\frac{t}{1-t} \tag{2.3}$$

for some a in $\mathbb{F}_p^*$. The first in this list is ruled out because $Cx+Cy = C+Cx$ has dimension 2. And then the fact that $y = 1-x$ also lies in $\sqrt{G}$ restricts $a = 1$. This leads respectively to

$$(x,y) = \Pi^+,\ \Pi^-,\ \tilde{\Pi}_1^+,\ \tilde{\Pi}_1^-,\ \tilde{\Pi}_2^+,\ \tilde{\Pi}_2^- . \tag{2.4}$$

What if $Cx + Cy = C + Cx = C + Cy$ has dimension 1?

Then x and y lie in C. If they are both in $\mathbb{F}_p^*$, then $p \geq 3$ and we get $(x,y) = \Pi^{(a)}$.

Otherwise by considering degrees we see that there is a greatest power q of p with $x = x'^q$ and $y = y'^q$ for x' and y' in K not both in C. Now $x' + y' = 1$ with x' and y' still in $\sqrt{G}$, and $Cx' + Cy'$ has dimension 2. It follows from the above discussion that (x', y') is one of (2.4). So taking the union over all q gives the Proposition.

Now Theorem 1 follows at once because $L(G) = L(\sqrt{G}) \cap G^2$. Namely if $p \geq 3$ then because of the minus signs only Π^+, Π^- in (2.4) lie in G^2 and a similar assertion holds for the qth powers implicit in the $[\psi_0]_p$; similarly we can omit the $\Pi^{(a)}$. However if $p = 2$ then nothing in (2.4) can be omitted.

3. Preliminaries. Let S be the set of polynomials

$$t^r(1-t)^s \quad (r \geq 0,\ s \geq 0,\ r+s \leq 3) \tag{3.1}$$

in $\mathbb{F}_p[t]$. For A in $\mathbb{F}_p[t]$ let $r(A)$ be the number of (X,Y) in S^2 with $A = X + Y$. The following is the basic reason for our uniformity in p.

LEMMA 3.1. *Suppose $p \geq 5$. Then $r(A) = 0, 1, 2$ apart from $r(A) = 4$ for the following:*

$$A = 1 - t + t^2 = t^2 + (1-t) = (1-t)^2 + t,$$
$$A = t(1 - t + t^2) = t^3 + t(1-t) = t(1-t)^2 + t^2,$$
$$A = (1-t)(1 - t + t^2) = t^2(1-t) + (1-t)^2 = (1-t)^3 + t(1-t).$$

Proof. The analogous assertion for the corresponding set $\tilde{S}$ defined by (3.1) in $\mathbb{Z}[t]$ is readily checked by machine. This means that an equation $\tilde{A} = \tilde{X} + \tilde{Y} = \tilde{Z} + \tilde{W}$ with $\tilde{X}, \tilde{Y}, \tilde{Z}, \tilde{W}$ in $\tilde{S}$ implies $\tilde{Z} = \tilde{X}$ or $\tilde{Z} = \tilde{Y}$ except as indicated when $\tilde{A}$ is the canonical pullback of one of the three A shown above.

But now suppose $A = X + Y = Z + W$ in with X, Y, Z, W in S. Each term $t^r(1-t)^s$ has a canonical pullback $t^r(1-t)^s$ to $\tilde{S}$ with coefficients of absolute values at most 3. Then the polynomial $P = \tilde{X} + \tilde{Y} - \tilde{Z} - \tilde{W}$ lies in $p\mathbb{Z}[t]$ and its coefficients have absolute values at most 12. So if $p \geq 13$ this forces $P = 0$. Now the conclusion for $\tilde{S}$ immediately implies the conclusion for S.

One can cover $p = 11$ by noting that the only element of $\tilde{S}$ with a coefficient of absolute value 3 is $(1-t)^3$. So if P has a coefficient of absolute value greater than 10 then at least three of $\tilde{X}, \tilde{Y}, \tilde{Z}, \tilde{W}$ must be $(1-t)^3$; and this forces $r(A) \leq 2$.

The cases $p = 5, 7$ can be checked by hand.

We note that this equation

$$X + Y = Z + W \tag{3.2}$$

is invariant under the action of the dihedral group D_4 with eight elements acting on the square with vertices X, Z, Y, W in an anti-clockwise direction. This group therefore acts on the solutions of (3.2). Let N be the set of solutions (X, Y, Z, W) in G^4 with

$$\{X, Y\} \neq \{Z, W\}, \quad \dim_C(CX + CY + CZ + CW) \neq 1, \tag{3.3}$$

also stable under this action.

Define an equivalence relation on K^* by two elements having their quotient in C.

LEMMA 3.2. *Suppose $p \geq 3$. Then every D_4-orbit in N contains a point where the equivalence classes in $\{X, Y, Z, W\}$ are described by one of*

(1) $\{X, Y\}, \{Z\}, \{W\}$,
(2) $\{X\}, \{Y\}, \{Z\}, \{W\}$,
(3) $\{Y, W\}, \{X\}, \{Z\}$.

Proof. Take any (X, Y, Z, W) in N, and let h be the number of classes in $\{X, Y, Z, W\}$. Then $h \neq 1$ because of the second condition in (3.3).

If $h = 4$ then we are in case (2) at once.

If $h = 3$ then there must be two singletons and one pair. Under D_4 we can assume that the pair is either $\{X, Y\}$ (opposite points of the square) or $\{Y, W\}$ (adjacent points), leading to cases (1) and (3).

It remains to exclude $h = 2$. This could arise from one singleton and one triplet; but then the equation (3.2) would destroy the singleton. Or we could have two pairs. Under D_4 these could be taken as either $\{X, Y\}, \{Z, W\}$ (opposite points equivalent) or $\{X, Z\}, \{Y, W\}$ (adjacent points). The first means $X = \alpha Y$, $Z = \beta W$ for α, β in C, but then $(1 + \alpha)Y = (1 + \beta)W$, forcing $\alpha = X/Y = -1$ and $\beta = Z/W = -1$, which however are not in G as $p \geq 3$. The second means similarly $X = \alpha Z$, $Y = \beta W$ but then $(1 - \alpha)Z + (1 - \beta)W = 0$, forcing $\alpha = \beta = 1$ and $X = Z$, $Y = W$, contrary to the first condition in (3.3). This completes the proof.

LEMMA 3.3. *Suppose that q is a power of p, and u is in K with u^q in G. Then u is in G.*

Proof. By definition u lies in $\sqrt{G}$, so that as we have seen in Lemma 2.1 $u = ag$ for a in $\mathbb{F}_p^*$ and g in G. But then $a = a^q = u^q/g^q$ lies in G, so $a = 1$.

LEMMA 3.4. *Suppose that u, u_1 are in K with $u_1 \neq 0$ of degree at most 1 and u^p/u_1 in $\sqrt{G}$. Then there is $a \neq 0$ in $\mathbb{F}_p$ such that au_1 is one of* (2.3). *Further if $1 - u_1$ lies in $\sqrt{G}$ then $a = 1$.*

Proof. Let A be any irreducible polynomial of $\mathbb{F}_p[t]$ which is not a constant multiple of $t, 1 - t$. Then the corresponding order function satisfies

$$0 = \operatorname{ord}_A \frac{u^p}{u_1} = p \ \operatorname{ord}_A u - \operatorname{ord}_A u_1;$$

but as $|\operatorname{ord}_A u_1| \leq 1$ this implies $\operatorname{ord}_A u_1 = 0$. Since this holds for all such A we see that u_1 must lie in $\sqrt{G}$, and because it has degree at most 1 we get the list (2.3) as before.

If further $1 - u_1$ lies in $\sqrt{G}$ then $a = 1$ as we saw during the proof of the Proposition.

4. Proof of Theorem 2 for $p \geq 5$. Now we must investigate x, y and z in G with $x + y - z = 1$. This time there are three possibilities for the dimension $d = d(x, y, z)$ of $Cx + Cy + Cz$ and we take these in turn.

We have to exploit the symmetry, which becomes clearer by writing formally

$$x = \frac{X}{W}, \quad y = \frac{Y}{W}, \quad z = \frac{Z}{W} \tag{4.1}$$

so that the equation $x + y - z = 1$ is now just (3.2). There is therefore a D_4-action on $P(G)$.

PROPOSITION 1. *Suppose $p \geq 5$. Then the set $P^*(G)$ of solutions of the equation $x + y - z = 1$ with $d = 3$ is $D_4(\Pi)$ with*

$$\Pi = \left(t, \frac{1-t}{t}, \frac{(1-t)^2}{t}\right).$$

Proof. We write down $x+y-z = 1$ and its derivative $(\dot{x}/x)x+(\dot{y}/y)y-(\dot{z}/z)z = 0$ as well as the second derivative $(\ddot{x}/x)x + (\ddot{y}/y)y - (\ddot{z}/z)z = 0$. There is an associated determinant

$$\Delta = \begin{vmatrix} 1 & 1 & -1 \\ \dot{x}/x & \dot{y}/y & -\dot{z}/z \\ \ddot{x}/x & \ddot{y}/y & -\ddot{z}/z \end{vmatrix},$$

and by multiplying by $-xyz$ we get the Wronskian of x, y, z. Since the latter are linearly independent over our field C of differential constants, we deduce

that $\Delta \neq 0$ (cf. [L2, pp. 174–175]). It follows that

$$x = \frac{\Delta_x}{\Delta}, \quad y = \frac{\Delta_y}{\Delta}, \quad z = \frac{\Delta_z}{\Delta}$$

for

$$\Delta_x = \begin{vmatrix} 1 & 1 & -1 \\ 0 & \dot{y}/y & -\dot{z}/z \\ 0 & \ddot{y}/y & -\ddot{z}/z \end{vmatrix}, \quad \Delta_y = \begin{vmatrix} 1 & 1 & -1 \\ \dot{x}/x & 0 & -\dot{z}/z \\ \ddot{x}/x & 0 & -\ddot{z}/z \end{vmatrix}, \quad \Delta_z = \begin{vmatrix} 1 & 1 & 1 \\ \dot{x}/x & \dot{y}/y & 0 \\ \ddot{x}/x & \ddot{y}/y & 0 \end{vmatrix}.$$

Now using (2.2) for $k = \dot{h}/h$ and also $\ddot{h}/h = \dot{k} + k^2$ we see that each of $\Delta, \Delta_x, \Delta_y, \Delta_z$ has the form

$$\frac{a_0 + a_1 t + a_2 t^2 + a_3 t^3}{t^3(1-t)^3}$$

for a_0, a_1, a_2, a_3 in $\mathbb{F}_p$. Therefore each of x, y, z is a rational function of t of degree at most 3.

In Section 2 it was easy to see when a rational function of degree at most 1 lies in G. To deal with higher degree we note that $t^r(1-t)^s$ has degree $\max\{|r|, |s|, |r+s|\}$. This leads to 37 possibilities for (r, s) in $\mathbb{Z}^2$. So all we have to do is check the $37^3 = 50653$ possibilities for (x, y, z) in $x + y - z = 1$ (not forgetting $d = 3$).

To reduce this work we use again (4.1), now with X, Y, Z, W in $\mathbb{F}_p[t]$ having no common factor. Each can be chosen to have the form $t^r(1-t)^s$ with $r \geq 0$, $s \geq 0$, $r + s \leq 3$. Now there are only ten possibilities for (r, s), so $10^4 = 10000 < 50653$ in all. However Lemma 3.1 implies $Z = X$ or $Z = Y$ or $A = X + Y$ is one of the list of three. But $Z = X$ means $z = x$ contradicting $d = 3$, and similarly $Z \neq Y$. Thus X, Y, Z, W are as in the list. This actually reduces to a single projective (X, Y, Z, W) under the action of D_4, which can be taken as $(t^2, 1-t, (1-t)^2, t)$. So dividing by $W = t$ we get our $(x, y, z) = \Pi$; and for this $d(x, y, z) = 3$ is quickly checked.

Now to the next case $d(x, y, z) = 2$. Here we need a small modification of our notation. Our coset T_x is the set of $(1, y, y)$ in K^3; we define T_x^* as the subset with y not in C. Similarly for T_y^*. Further we define

$$[\psi]_p^* = \bigcup_{e=1}^{\infty} \psi^{-1} \varphi^e \psi$$

as in (1.7) for $h = 1$ but omitting $e = 0$.

Proposition 2. *Suppose $p \geq 3$. Then the set $P^{**}(G)$ of solutions of the equation $x + y - z = 1$ with $d = 2$ is*

$$T_x^*(G) \cup T_y^*(G) \cup \bigcup_{\Pi} D_4([\psi_\Pi]_p^* \Pi)$$

with the following five pairs of points and automorphisms:

$$\Pi = \left(1, \frac{1-t}{t}, \frac{1-t}{t}\right), \qquad \psi_\Pi(x,y,z) = \left(tx, ty, \frac{t}{1-t}z\right),$$

$$\Pi = \left(1, \frac{1}{t}, \frac{1}{t}\right), \qquad \psi_\Pi(x,y,z) = \left(\frac{1-t}{t}x, y, (1-t)z\right),$$

$$\Pi = \left(\frac{t^2}{(1-t)^2}, \frac{1}{1-t}, \frac{t}{(1-t)^2}\right), \quad \psi_\Pi(x,y,z) = \left(\frac{1-t}{t}x, y, (1-t)z\right),$$

$$\Pi = \left(1, \frac{1}{1-t}, \frac{1}{1-t}\right), \qquad \psi_\Pi(x,y,z) = \left(\frac{t}{1-t}x, y, tz\right),$$

$$\Pi = \left(\frac{(1-t)^2}{t^2}, \frac{1}{t}, \frac{1-t}{t^2}\right), \qquad \psi_\Pi(x,y,z) = \left(\frac{t}{1-t}x, y, tz\right).$$

Proof. Let (x,y,z) be in $P^{**}(G)$, and use (4.1) with X, Y, Z, W also in G. The dimension in (3.3) is also $\dim_C(Cx + Cy + Cz + C) = d = 2$, so this part of the condition holds. And $\{X, Y\} = \{Z, W\}$ would mean we are not just in $T_x(G) \cup T_y(G)$ but even in $T_x^*(G) \cup T_y^*(G)$ because $d \neq 1$. Thus we can assume all of (3.3).

So after adjusting by D_4 we can assume by Lemma 3.2 that we are in one of the cases (1), (2), (3). We take each of these in turn.

In case (1) we have $x = \alpha y$ for some α in C. It follows that

$$(1+\alpha)y - z = 1. \tag{4.2}$$

Further α is in G so $\alpha \neq -1$. Since $1+\alpha$ is a differential constant we can easily differentiate, and since $z/y = Z/Y$ is not in C, the arguments of the proof of the Proposition yield

$$(1+\alpha)y = \frac{\dot{z}/z}{\dot{z}/z - \dot{y}/y}, \qquad -z = \frac{-\dot{y}/y}{\dot{z}/z - \dot{y}/y}$$

as in (2.1). In particular from (2.2), $u_1 = (1+\alpha)y$ has degree at most 1. This gives only finitely many possibilities for $z = u_1 - 1$, thus reducing to finitely many lines M on the plane P. In the context of a general variety V, these are the W mentioned at the end of Section 1. However their number may depend on the characteristic p.

To cut down this dependence we note that there is u in K with $u^p = 1+\alpha$, and then $u^p/u_1 = 1/y$ lies in G. So by Lemma 3.4 there is $a \neq 0$ in $\mathbb{F}_p$ such that au_1 lies in (2.3). The first element in this list is ruled out because in our case (1), $y = Y/W$ is not in C. And then the fact that $-z = 1 - u_1$ also lies in $\sqrt{G}$ restricts $a = 1$. But actually $u_1 - 1 = z$ lies in G, which reduces the choice to

$$u_1 = \frac{1}{t} \quad \text{or} \quad u_1 = \frac{1}{1-t}. \tag{4.3}$$

Take first $u_1 = 1/t$. Now temporarily with general coordinates x, y, z, the line M is defined by the equations $x + y - z = 1$ and $z = (1-t)/t$, or equivalently

$$x + y = \frac{1}{t}, \qquad z = \frac{1-t}{t}.$$

So $M' = \psi(M)$ is defined over $\mathbb{F}_p$, where

$$\psi(x, y, z) = \left(tx, ty, \frac{t}{1-t}z\right)$$

is as in the first pair in the list of Proposition 2; in fact if we call this (x', y', z') then the equations become

$$x' + y' = 1, \qquad z' = 1. \tag{4.4}$$

Here we see a copy of the line L, and so M' has the points $(x', y', z') = (t, 1-t, 1)$, $(1-t, t, 1)$. These give rise via ψ^{-1} to points

$$\Pi = \left(1, \frac{1-t}{t}, \frac{1-t}{t}\right), \ \left(\frac{1-t}{t}, 1, \frac{1-t}{t}\right)$$

on $\psi^{-1}(M') = M$, so on P; note that the first Π here is also as in the first pair in the list of Proposition 2.

Now we return to our point (x, y, z) of $P^{**}(G)$, here in $M(G)$. Then $\psi(x, y, z)$ is in $M'(G)$ and from (4.4) and Theorem 1 we see that this is one of

$$(t^q, (1-t)^q, 1), \qquad ((1-t)^q, t^q, 1) \qquad (q = p^e,\ e = 0, 1, 2, \ldots). \tag{4.5}$$

The first of these is, in the notation of Section 1, just $\varphi^e\psi(\Pi)$ with the first ψ above. So we get the family

$$(x, y, z) = \psi^{-1}\varphi^e\psi(\Pi) = \left(t^{q-1}, \frac{(1-t)^q}{t}, \frac{1-t}{t}\right). \tag{F1}$$

Taking the union over all e gives precisely $[\psi]_p(\Pi)$. But in fact $x/y = \alpha$ lies in C, so $q = 1$ is excluded and we end up in $[\psi]_p^*(\Pi)$ as in Proposition 2. The D_4-action (which may however take us out of case (1) of Lemma 3.2) then provides us with the whole $D_4([\psi]_p^*(\Pi))$.

But what if $\psi(x, y, z)$ is the second of (4.5)? Then it is easily seen, in fact through the interchange of x and y, that we get something in the same D_4-orbit.

We can deal with $u_1 = 1/(1-t)$ in (4.3) by noting that $K = \mathbb{F}_p(t)$ has an automorphism ω taking t to $1-t$ which preserves P and G and therefore acts on $P(G)$. It also preserves C and so acts on $P^{**}(G)$. Now if $u_1 = z+1 = 1/(1-t)$ then $\omega(z)+1 = 1/t$ and so for $\omega(x, y, z)$ we are in the case just considered. Therefore $\omega(x, y, z)$ lies in $\mathcal{D} = D_4([\psi]_p^*(\Pi))$; and by applying $\omega^{-1} = \omega$ we see that (x, y, z) lies in $\omega(\mathcal{D})$. However $\omega(\mathcal{D}) = \mathcal{D}$; for example we get from (F1) the point $((1-t)^{q-1}, t^q/(1-t), t/(1-t))$. But in

terms of (3.2) already (F1) says

$$t^q + (1-t)^q = (1-t) + t$$

and so dividing this by $1-t$ gives again the same orbit.

Thus case (1) of Lemma 3.2 leads to only the first pair in the list of Proposition 2.

We next treat the cases (2) and (3). Now x, y are linearly independent over C and because $d = 2$ there are α, β in C with

$$z = \alpha x + \beta y. \tag{4.6}$$

We note that $\alpha \neq 0$ because $z/y = Z/Y$ is not in C; similarly $\beta \neq 0$. Then for $x_z = x/z,\ y_z = y/z$ we get

$$\alpha x_z + \beta y_z = 1. \tag{4.7}$$

We argue as we did for (4.2). Here $y_z/x_z = y/x$ is not in C and so we get

$$\alpha x_z = \frac{\dot{y}_z/y_z}{\dot{y}_z/y_z - \dot{x}_z/x_z}, \qquad \beta y_z = \frac{-\dot{x}_z/x_z}{\dot{y}_z/y_z - \dot{x}_z/x_z} \tag{4.8}$$

as in (2.1). In particular $x_1 = \alpha x_z \neq 0$, $y_1 = \beta y_z \neq 0$ have degree at most 1. Note however that $\alpha = 1$ or $\beta = 1$ are not yet excluded (see below).

Substituting (4.6) into $x + y - z = 1$ gives

$$(1-\alpha)x + (1-\beta)y = 1, \tag{4.9}$$

and the same arguments give

$$(1-\alpha)x = \frac{\dot{y}/y}{\dot{y}/y - \dot{x}/x}, \qquad (1-\beta)y = \frac{-\dot{x}/x}{\dot{y}/y - \dot{x}/x} \tag{4.10}$$

with $u_1 = (1-\alpha)x,\ v_1 = (1-\beta)y$ of degree at most 1.

Now we have

$$u_1 = x - x_1 z, \qquad v_1 = y - y_1 z \tag{4.11}$$

(adding up to our original $1 = x + y - z$), which again reduces to finitely many lines in P. Again we must cut down the dependence on p.

Using Lemma 3.4 for $\alpha/x_1, \beta/y_1$ shows that there are $a \neq 0, b \neq 0$ in $\mathbb{F}_p$ such that ax_1, by_1 are in the list (2.3). The first element in this list is ruled out because in our cases (2), (3), $x_1 = \alpha X/Z$, $y_1 = \beta Y/Z$ are not in C. And then the fact that $y_1 = 1 - x_1$, $x_1 = 1 - y_1$ also lie in $\sqrt{G}$ due to Lemma 2.1 restricts $a = 1$, $b = 1$. Thus x_1, y_1 lie in the sublist

$$t,\ 1-t,\ \frac{1}{t},\ -\frac{1-t}{t},\ \frac{1}{1-t},\ -\frac{t}{1-t} \tag{4.12}$$

of (2.3).

So far we could do both cases (2), (3) of Lemma 3.2 at once. But now we restrict to (2).

Then (4.9) yields $\alpha \neq 1$ and $\beta \neq 1$ because $y = Y/W$ and $x = X/W$ are not in C. So Lemma 3.4 applied to $(1-\alpha)/u_1, (1-\beta)/v_1$ shows that there are $e \neq 0, f \neq 0$ in $\mathbb{F}_p$ such that eu_1, fv_1 are in the list (2.3). Again we can eliminate the first element in the list and restrict to $e = 1$, $f = 1$.

In particular u_1, v_1, as well as x_1, y_1, lie in the sublist (4.12). This eliminates completely the dependence on p in the equations (4.11). Still, the lines do not look like L.

But when we write these equations as

$$\frac{x}{u_1} + \left(-\frac{x_1 z}{u_1}\right) = 1, \qquad \frac{y}{v_1} + \left(-\frac{y_1 z}{v_1}\right) = 1, \tag{4.13}$$

then we do after all observe two points on $L(\sqrt{G})$. It follows from the Proposition that there are powers q_x, q_y of p and points $\Pi_x = (\xi, \zeta_x),\ \Pi_y = (\eta, \zeta_y)$ there such that

$$x = u_1 \xi^{q_x}, \qquad z = -\frac{u_1}{x_1}\zeta_x^{q_x}, \qquad y = v_1 \eta^{q_y}, \qquad z = -\frac{v_1}{y_1}\zeta_y^{q_y}. \tag{4.14}$$

In particular comparing the z-values gives

$$\frac{\zeta_x^{q_x}}{\zeta_y^{q_y}} = \frac{x_1 v_1}{u_1 y_1}. \tag{4.15}$$

The right-hand side w here certainly has degree at most 4; but in fact (4.8) and (4.10) show that

$$w = \frac{\frac{\dot{x}}{x}\left(\frac{\dot{y}}{y} - \frac{\dot{z}}{z}\right)}{\frac{\dot{y}}{y}\left(\frac{\dot{x}}{x} - \frac{\dot{z}}{z}\right)}.$$

So w has degree at most 2. Furthermore it cannot be constant, because a quick scribble shows that when two products $x_1 v_1, u_1 y_1$ of two elements from (4.12) have constant quotient, they are equal. But $w = 1$ leads at once to $(\dot{z}/z)(\dot{x}/x - \dot{y}/y) = 0$, which is ruled out in our current case (2).

Now (4.15) implies that at least one of q_x, q_y must be 1. Otherwise the left-hand side would be non-constant in C and so have degree at least $p > 2$.

Suppose for example $q_x = 1$. Then (4.14) shows that $\zeta_x/\xi = -\alpha$ is in C. Now inspection of the Π in the Proposition shows that this is so only if $\Pi_x = \Pi^{(a)} = (a, 1-a)$ $(a = 2, \ldots, p-1)$. But then $au_1 = x$ lies in G, which by inspection of (4.12) forces $a = p-1$ and $u_1 = -(1-t)/t, -t/(1-t)$ and $x = -u_1$. Then $x_1/2 = x_1/(1-a) = -u_1/z$ also lies in G, which by the same inspection forces $p = 3$ and $x_1 = -(1-t)/t, -t/(1-t)$. But as $z = Z/W$ is not in C and is now u_1/x_1, we must have $x_1 = 1/u_1$ and so $z = u_1^2$. Finally this means $y = 1 - x + z = 1 + u_1 + u_1^2$. If $u_1 = -(1-t)/t$ then $y = 1/t^2$ does indeed lie in G; but the resulting point

$$(x, y, z) = \left(\frac{1-t}{t}, \frac{1}{t^2}, \frac{(1-t)^2}{t^2}\right)$$

has $d = 3$ (see Section 5). Similarly for $u_1 = -t/(1-t)$ using for example ω. And all this means that we find no points of $P^{**}(G)$ in this case (2).

We finally turn to case (3), which we left halfway through the discussion above and which implies that $y = Y/W$ is in C. So by (4.10) we have $\alpha = 1$, $u_1 = 0, v_1 = 1$. Strangely enough it is this somewhat degenerate-looking case which provides most of the points of $P^{**}(G)$.

From (4.11) we see now that $x_1 = x/z$ lies in G. Inspection of (4.12) shows that x_1 must be in the sublist

$$t,\ 1-t,\ \frac{1}{t},\ \frac{1}{1-t} \tag{4.16}$$

of (4.12). We look at each of these in turn.

Suppose first that $x_1 = t$, so that $y_1 = 1 - t$. Thus from (4.11) we are now on the line M defined by the equations

$$x = tz, \quad y - (1-t)z = 1. \tag{4.17}$$

So $M' = \psi(M)$ is defined over $\mathbb{F}_p$, where

$$\psi(x, y, z) = \left(\frac{1-t}{t}x, y, (1-t)z\right) \tag{4.18}$$

is as in the second and third pairs in the list of Proposition 2; in fact also with (x', y', z') the equations become

$$x' = z', \quad y' - z' = 1. \tag{4.19}$$

Here we do not see exactly the line L. However $\tilde{\Pi}_1^+$ and $\tilde{\Pi}_2^+$ lead to the following solutions over G:

$$(x', y', z') = \left(\frac{1-t}{t}, \frac{1}{t}, \frac{1-t}{t}\right),\ \left(\frac{t}{1-t}, \frac{1}{1-t}, \frac{t}{1-t}\right).$$

These give rise via ψ^{-1} to points

$$\Pi = \left(1, \frac{1}{t}, \frac{1}{t}\right),\ \left(\frac{t^2}{(1-t)^2}, \frac{1}{1-t}, \frac{t}{(1-t)^2}\right)$$

on $\psi^{-1}(M') = M$, so on P; note that these are also as in the second and third pairs in the list of Proposition 2.

Now we return to our point (x, y, z) of $P^{**}(G)$. Then $\psi(x, y, z)$ is on $M'(G)$ and from (4.19) and the Proposition we see that this is one of

$$\left(\frac{(1-t)^q}{t^q}, \frac{1}{t^q}, \frac{(1-t)^q}{t^q}\right), \quad \left(\frac{t^q}{(1-t)^q}, \frac{1}{(1-t)^q}, \frac{t^q}{(1-t)^q}\right)$$
$$(q = p^e,\ e = 0, 1, 2, \ldots).$$

Again these are just $\varphi^e\psi(\Pi)$. The first gives

$$(x, y, z) = \psi^{-1}\varphi^e\psi(\Pi) = \left(\frac{(1-t)^{q-1}}{t^{q-1}}, \frac{1}{t^q}, \frac{(1-t)^{q-1}}{t^q}\right), \tag{F2}$$

and the second gives

$$(x,y,z)=\psi^{-1}\varphi^e\psi(\Pi)=\left(\frac{t^{q+1}}{(1-t)^{q+1}},\frac{1}{(1-t)^q},\frac{t^q}{(1-t)^{q+1}}\right). \tag{F3}$$

Taking the union over all e gives again the $[\psi]_p(\Pi)$. But this time y lies in C, so $q\neq 1$ and we end up with the $[\psi_\Pi]_p^*(\Pi)$ as in Proposition 2. The D_4-action (which as before may take us out of case (3) of Lemma 3.2) then provides us with the whole $D_4([\psi_\Pi]_p^*(\Pi))$.

Suppose next that $x_1=1/t$ in (4.16), so that $y_1=-(1-t)/t$. Thus from (4.11) we are now on the line M defined by the equations

$$x=\frac{1}{t}z,\quad y+\frac{1-t}{t}z=1.$$

So $M'=\psi(M)$ is defined over $\mathbb{F}_p$, where

$$\psi(x,y,z)=\left((1-t)x,y,\frac{1-t}{t}z\right)$$

now is not in the list of Proposition 2; anyway with (x',y',z') the equations become

$$x'=z',\quad y'+z'=1.$$

Now we return to our point (x,y,z) of $P^{**}(G)$. Then $\psi(x,y,z)$ is on $M'(G)$ and from the equations immediately above and Theorem 1 we see that this is one of

$$(x',y',z')=(t^q,(1-t)^q,t^q),\ ((1-t)^q,t^q,(1-t)^q)\quad (q=p^e,\ e=0,1,2,\ldots).$$

And via ψ^{-1} they give

$$\left(\frac{t^q}{1-t},(1-t)^q,\frac{t^{q+1}}{1-t}\right),\quad ((1-t)^{q-1},t^q,t(1-t)^{q-1}),$$

again with $q=1$ excluded because y is in C.

However these result in the same D_4-orbits as the second and third respectively above, which is easily seen by considering (F2) and (F3) respectively in terms of (3.2), namely

$$t(1-t)^{q-1}+1=(1-t)^{q-1}+t^q,$$
$$t^{q+1}+(1-t)=t^q+(1-t)^{q+1}.$$

Finally we deal with the remaining $x_1=1-t,1/(1-t)$ in (4.16) simply by applying our automorphism ω, which yields on (F2), (F3)

$$(x,y,z)=\left(\frac{t^{q-1}}{(1-t)^{q-1}},\frac{1}{(1-t)^q},\frac{t^{q-1}}{(1-t)^q}\right), \tag{F4}$$

$$(x,y,z)=\left(\frac{(1-t)^{q+1}}{t^{q+1}},\frac{1}{t^q},\frac{(1-t)^q}{t^{q+1}}\right). \tag{F5}$$

corresponding to the fourth and fifth pairs in Proposition 2. This is thereby proved.

We now prove Theorem 2 for $p \geq 5$. Take a point (x, y, z) in $P(G)$, with as above $d = d(x, y, z) = \dim_C(Cx+Cy+Cz)$. We already treated the cases $d = 3$ and $d = 2$. The case $d = 1$ is treated as at the end of the proof of the Proposition in Section 2. For then x, y, z lie in C. If they are all in $\mathbb{F}_p^*$, then also in $\mathbb{F}_p^* \cap G$, so $x = y = z = 1$ and we are certainly in $T_x(G) \cup T_y(G)$. Otherwise by considering degrees we see that there is a largest power q' of p with $x = x'^{q'}$, $y = y'^{q'}$, $z = z'^{q'}$ for x', y', z' in K not all in C; and by Lemma 3.3, x', y', z' are still in G. Now $x' + y' - z' = 1$ and $d(x', y', z') \geq 2$. It follows from the above discussion that (x', y', z') is as in Proposition 1 or Proposition 2.

Now in Proposition 2 we see T_x^*, T_y^*, which on raising to power q' ($q' = 1, p, p^2, \ldots$) end up in T_x, T_y as in Theorem 2.

We also see various $\delta([\psi_{\Pi}]_p^*\Pi)$ for δ in D_4. But by going back to projective X, Y, Z, W it is not difficult to see that this is $[\psi_{\Pi,\delta}]_p^*\Pi_\delta$ for some $\psi_{\Pi,\delta}$ and $\Pi_\delta = \delta(\Pi)$. This is $[\psi_{\Pi,\delta}]_p\Pi_\delta$ with just Π_δ removed. And the set of q'th powers of elements of $[\psi_{\Pi,\delta}]_p\Pi_\delta$ is nothing else than $[\psi_0, \psi_{\Pi,\delta}]_p\Pi_\delta$. So we get all the $[\psi_0, \psi_\Pi]_p\Pi$ in Theorem 2 except that it seems that the q'th powers of the Π_δ are missing. However these are supplied by Proposition 1, because the Π there has the same D_4-orbit as the third and fifth Π in Proposition 2.

What about the first, second and fourth Π in Proposition 2? These belong anyway in T_x, which we have already taken into account. This completes the proof.

5. Proof of Theorem 2 for $p = 3$. We can follow the arguments of the preceding section, noting that Proposition 2 has been proved for $p = 3$ as well. However Proposition 1 fails because Lemma 3.1 fails. Hand computation yields exactly six further examples with $r(A) = 4$, which come from

$$\begin{aligned}
1 + t(1-t) &= (1-t)^2 + t^2,\\
t + t^2(1-t) &= t(1-t)^2 + t^3,\\
(1-t) + t(1-t)^2 &= (1-t)^3 + t^2(1-t),\\
1 + t^2(1-t) &= (1-t)^3 + t^2,\\
1 + t(1-t)^2 &= (1-t)^2 + t^3,\\
t + (1-t) &= (1-t)^3 + t^3.
\end{aligned}$$

Here the second and third equations give rise to the same projective points as the first, so we may ignore them. Further the fourth, fifth and sixth equations

do not have $d = 3$. The first equation produces the point $(1/t^2, (1-t)/t, (1-t)^2/t^2)$, and its D_4-orbit accounts for the extra set $\mathcal{T}'$ in Theorem 2.

6. Proof of Theorem 2 for $p = 2$. Now Lemma 3.1 fails quite badly, and for example there are A with $r(A) = 10$. But we do not mind this, because in fact there are no points with $d = 3$ for $[K : C] = 2$; so x, y, z must be linearly dependent over C and we can forget about Proposition 1.

We need the following version of Lemma 3.2, which now involves the action of the symmetric group S_4 on four elements which arises by writing (3.2) as

$$X + Y + Z + W = 0. \tag{6.1}$$

This time let N be the set of solutions (X, Y, Z, W) in G^4 with X, Y, Z, W all different and

$$\dim_C(CX + CY + CZ + CW) \neq 1, \tag{6.2}$$

also stable under this action. As before define an equivalence relation on K^* by two elements having their quotient in C.

LEMMA 6.1. *Suppose $p = 2$. Then every S_4-orbit in N contains a point where the equivalence classes in $\{X, Y, Z, W\}$ are described by one of*

(2) $\{X\}, \{Y\}, \{Z\}, \{W\}$,
(3) $\{Y, W\}, \{X\}, \{Z\}$.

Proof. Take any (X, Y, Z, W) in N, and let h be the number of classes in $\{X, Y, Z, W\}$. Then $h \neq 1$ because of (6.2).

If $h = 4$ then we are in case (2) at once.

If $h = 3$ then there must be two singletons and one pair. Under S_4 we can assume that the pair is $\{Y, W\}$ leading to case (3).

It remains only to exclude $h = 2$. This could arise from one singleton and one triplet; but then the equation (6.1) would destroy the singleton. Or we could have two pairs. Under S_4 these could be taken as $\{X, Z\}, \{Y, W\}$. This means $X = \alpha Z$, $Y = \beta W$ for α, β in C, but then $(1 + \alpha)Z + (1 + \beta)W = 0$ forcing $\alpha = \beta = 1$ and $X = Z$, $Y = W$, contrary to the first condition on N. This completes the proof.

PROPOSITION 3. *Suppose $p = 2$. Then the set $P^{**}(G)$ of solutions of the equation $x + y - z = 1$ with $d = 2$ is*

$$T_x^*(G) \cup T_y^*(G) \cup T_z^*(G) \cup S_4(\Pi_0) \cup \bigcup_{\Pi} S_4([\psi_\Pi]_p \Pi)$$

with

$$\Pi_0 = (t^3, (1-t)^3, t(1-t))$$

and the five Π as in Proposition 2 *together with the following four pairs of points and automorphisms*:

$$\Pi = \left(\frac{t}{(1-t)^2}, \frac{t}{1-t}, \frac{1}{(1-t)^2}\right), \quad \psi_\Pi(x,y,z) = \left(\frac{1-t}{t}x, y, (1-t)z\right),$$

$$\Pi = \left(\frac{1-t}{t^2}, \frac{1-t}{t}, \frac{1}{t^2}\right), \quad \psi_\Pi(x,y,z) = \left(\frac{t}{1-t}x, y, tz\right),$$

$$\Pi = (t(1-t), 1-t, t^2), \quad \psi_\Pi(x,y,z) = \left(\frac{1}{1-t}x, y, \frac{1}{t}z\right).$$

$$\Pi = (t(1-t), t, (1-t)^2), \quad \psi_\Pi(x,y,z) = \left(\frac{1}{t}x, y, \frac{1}{1-t}z\right),$$

Proof. Notice that it is now the full $[\psi]$ that appear, not the $[\psi]^*$ as in Proposition 2 for $p \geq 3$. But we will follow the proof of Proposition 2 to obtain a kind of Proposition 3^* with $[\psi]^*$ instead of $[\psi]$. The apparent discrepancy will be explained and eliminated at the end of the proof.

Let (x,y,z) be in $P^{**}(G)$, and use (4.1) with X, Y, Z, W also in G. The dimension in (6.2) is $d = 2$, so this part of the condition holds. And X, Y, Z, W not all different would mean we are not just in $T_x(G) \cup T_y(G) \cup T_z(G)$ but even in $T_x^*(G) \cup T_y^*(G) \cup T_z^*(G)$ because $d \neq 1$. Thus we can assume that we are in N.

So after adjusting by S_4 we can assume by Lemma 6.1 that we are in the cases (2), (3). We can follow quite literally many of the previous arguments of Section 4. Now x, y are linearly independent over C and because $d = 2$ there are α, β in C with (4.6). We note that $\alpha \neq 0$ because $z/y = Z/Y$ is not in C; similarly $\beta \neq 0$. Then for $x_z = x/z,\ y_z = y/z$ we get (4.7) and we argue as we did there to get (4.8). Again $x_1 = \alpha x_z \neq 0,\ y_1 = \beta y_z \neq 0$ have degree at most 1.

Substituting (4.6) into $x + y - z = 1$ gives (4.9), and the same arguments give (4.10) with $u_1 = (1-\alpha)x,\ v_1 = (1-\beta)y$ of degree at most 1.

Now we have (4.11), but of course no more dependence on $p = 2$!

Using Lemma 3.4 on $\alpha/x_1, \beta/y_1$ shows that x_1, y_1 are in the list (2.3) since $p = 2$ and so immediately $a = b = 1$. The first element in this list is ruled out because in our cases (2), (3), $x_1 = \alpha X/Z,\ y_1 = \beta Y/Z$ are not in C. Thus x_1, y_1 lie in the list (4.12).

So far, as in Section 4, we could do both cases (2), (3) at once. But now we restrict to (2).

Then $\alpha \neq 1$ and $\beta \neq 1$ because $y = Y/W$ and $x = X/W$ are not in C. As above we find that u_1, v_1 are in the list (2.3) due to Lemma 3.4 and again we can eliminate the first element in the list.

In particular u_1, v_1, as well as x_1, y_1, lie in (4.12).

Again we have (4.13) leading to two points on $L(\sqrt{G})$. It follows from the Proposition that there are powers q_x, q_y of p and points $\Pi_x = (\xi, \zeta_x)$, $\Pi_y = (\eta, \zeta_y)$ there such that (4.14) holds. In particular comparing the z-values gives (4.15). As in Section 4 the right-hand side w here has degree at most 2. Furthermore it cannot be constant, because $w = 1$ leads at once to $(\dot{z}/z)(\dot{x}/x - \dot{y}/y) = 0$, which is ruled out in our current case (2).

Now (4.15) implies that at least one of q_x, q_y must be 1 or 2. Otherwise the left-hand side would be non-constant in C and so have degree at least $4 > 2$. Suppose first $q_x = 1$. Then (4.14) shows that $\zeta_x/\xi = -\alpha$ is in C. But inspection of the Π in the Proposition shows that this is impossible. We get a similar contradiction from $q_y = 1$.

Next suppose $q_x = 2$. Then (4.14) gives $x = u_1\xi^2$, $z = (-u_1/x_1)(1-\xi)^2$, which are automatically in G; however we cannot say that $y = 1 - x + z$ is in G. There are at most six possibilities for each of u_1, x_1 in (4.12), and also ξ lies in (4.12), giving at most $6^3 = 216$ possibilities for y in all. This number can be reduced by noting that for each ξ in (4.12) there is an automorphism ω_ξ of K taking t to ξ; for example our earlier ω is ω_{1-t}. Further ω_ξ preserves G (we are in characteristic 2) as well as each of the cases (2), (3). We get an automorphism group Σ_3 isomorphic to S_3.

Thus by applying ω_ξ^{-1} we can assume that $\xi = t$. Now a short calculation shows that y is in G in the current case (2) only for $u_1 = t$ with $x_1 = 1-t, -t/(1-t)$ and for $u_1 = -(1-t)/t$ with $x_1 = 1/t$ and for $u_1 = 1/(1-t)$ with $x_1 = -t/(1-t)$. However only one S_4-orbit turns up here, namely that of Π_0.

And if $q_y = 2$ then (4.14) gives $y = v_1\eta^2$, $z = -(v_1/y_1)(1-\eta)^2$, which are automatically in G; then a similar argument with $x = 1 - y + z$ gives again this same S_4-orbit of Π_0. And because we just used ω_ξ^{-1} we should reverse this by noting that this orbit is even Σ_3-invariant.

We finally turn to case (3) with $y = Y/W$ in C. So by (4.10) we have $\alpha = 1$, $u_1 = 0$, $v_1 = 1$, $x_1 = x/z$. It is this case which provides the remaining points of $P^{**}(G)$.

Thanks to the Σ_3-action we can assume that $x_1 = t$, so that $y_1 = 1 - t$ and then (4.11) shows that we are now on the line M defined by (4.17). So $M' = \psi(M)$ is defined over $\mathbb{F}_p$, with ψ as in (4.18), which is as in the first list of Proposition 3, and (4.19) holds.

Since $p = 2$ we can use Theorem 1 and the points Π_1^+, Π_2^+ give rise to the families (F2), (F3) as in Section 4. Further the points Π^+, Π^- lead again to (F2) and (F3). But Π_2^-, Π_1^- lead to

$$(x', y', z') = \left(\frac{1}{1-t}, \frac{t}{1-t}, \frac{1}{1-t}\right), \ \left(\frac{1}{t}, \frac{1-t}{t}, \frac{1}{t}\right)$$

and via ψ^{-1} to points

$$\Pi = \left(\frac{t}{(1-t)^2}, \frac{t}{1-t}, \frac{1}{(1-t)^2}\right), \left(\frac{1}{1-t}, \frac{1-t}{t}, \frac{1}{t(1-t)}\right)$$

on $\psi^{-1}(M') = M$, so on P; note that the first of these is as in the first pair in the list of Proposition 3.

Now our point (x, y, z) of $P^{**}(G)$ gives rise to

$$(x, y, z) = \left(\frac{t}{(1-t)^{q+1}}, \frac{t^q}{(1-t)^q}, \frac{1}{(1-t)^{q+1}}\right) \tag{F6}$$

and

$$(x, y, z) = \left(\frac{1}{t^{q-1}(1-t)}, \frac{(1-t)^q}{t^q}, \frac{1}{t^q(1-t)}\right);$$

however these are in the same S_4-orbit, which is easily seen by considering both in terms of (6.1).

As $q = p, p^2, \ldots$ still we get from (F6) the first of the sets $[\psi]_p^*(\Pi)$ in the list in our modified Proposition 3*. The S_4-action then provides us with the whole $S_4([\psi]_p^*(\Pi))$.

Thus we have ended up with the S_4-orbits of (F2), (F3), (F6). But where have (F1), (F4), (F5) gone, and where are the other three sets in the list of Proposition 3? Again we have to reverse the Σ_3-action. For convenience we start with (F1).

We find that ω_{1-t} and of course ω_t take (F1) into a point in the same S_4-orbit, that $\omega_{1/t}$ and $\omega_{(1-t)/t}$ take (F1) into a point in the S_4-orbit of (F4), and that $\omega_{1/(1-t)}$ and $\omega_{t/(1-t)}$ take (F1) into a point in the S_4-orbit of (F2). Thus Σ_3 takes the S_4-orbit of (F2) also into the S_4-orbits of (F1), (F2), (F4), and similarly for the S_4-orbit of (F4). This at least accounts for the missing (F1) and (F4).

Next let us calculate the Σ_3-action on (F3). It yields points in the S_4-orbit of (F5) and (F6) by applying ω_{1-t} and $\omega_{1/t}$ respectively. Further $\omega_{1/(1-t)}$ yields what we get by applying our original ω_{1-t} to (F6), namely

$$\left(\frac{1-t}{t^{q+1}}, \frac{(1-t)^q}{t^q}, \frac{1}{t^{q+1}}\right), \tag{F7}$$

which corresponds to the second pair in Proposition 3*. And applying $\omega_{t/(1-t)}$ to (F3) we find something in the S_4-orbit of

$$(t^q(1-t), (1-t)^q, t^{q+1}) \tag{F8}$$

which corresponds to the third pair in Proposition 3*. Finally $\omega_{(1-t)/t}$ yields what we get by applying ω_{1-t} to (F8), which is

$$(t(1-t)^q, t^q, (1-t)^{q+1}) \tag{F9}$$

corresponding to the fourth pair in Proposition 3*.

Thus we find that the Σ_3-action on the orbits of (F3), (F6) provides the missing (F5), (F7), (F8), (F9). So indeed everything in Proposition 3* has turned up.

But why is this the same as the original Proposition 3? The discrepancy lies only in $q = 1$. We find that the points (F1), (F2), (F4) with $q = 1$ lie in the coset T_x. We also find (up to the S_4-action) that the points (F3), (F5) with $q = 1$ reduce to the point (F1) with $q = 2$. And (F7), (F8) with $q = 1$ reduce to (F2) with $q = 2$. And finally (F6), (F9) with $q = 1$ reduce to (F4) with $q = 2$, which completes the proof of Proposition 3.

We now prove Theorem 2 for $p = 2$. Take a point (x, y, z) in $P(G)$, with as above $d = d(x, y, z) = \dim_C(Cx + Cy + Cz)$. As noted, $d \neq 3$; and we already treated $d = 2$. The case $d = 1$ is treated as in Section 4. For then x, y, z lie in C. If they are all in $\mathbb{F}_p^*$, then also in $\mathbb{F}_p^* \cap G$, so $x = y = z = 1$ and we are certainly in $T_x(G) \cup T_y(G) \cup T_z(G)$. Otherwise by considering degrees we see that there is a greatest power q' of p with $x = x'^{q'},\ y = y'^{q'},\ z = z'^{q'}$ for x', y', z' in K not all in C; and by Lemma 3.3, x', y', z' are still in G. Now $x' + y' - z' = 1$ and $d(x', y', z') \geq 2$. It follows from the above discussion that (x', y', z') is as in Proposition 3.

Now in Proposition 3 we see T_x^*, T_y^*, T_z^*, which on raising to power q' end up in T_x, T_y, T_z as in Theorem 2. We also see various $\sigma([\psi_{\Pi}]_p \Pi)$ for σ in S_4. But as before this is $[\psi_{\Pi,\sigma}]_p \Pi_\sigma$ for some $\psi_{\Pi,\sigma}$ and $\Pi_\sigma = \sigma(\Pi)$, so the proof is complete.

Acknowledgments. I wish to thank my research supervisor David Masser for suggesting the problems treated here and for advice on the presentation of the material.

References

[ABB] L. Arenas-Carmona, D. Berend and V. Bergelson, *Ledrappier's system is almost mixing of all orders*, Ergodic Theory Dynam. Systems 28 (2008), 339–365.

[AV] D. Abramovich and F. Voloch, *Towards a proof of the Mordell–Lang conjecture in characteristic p*, Int. Math. Res. Notices 1992, no. 5, 103–115.

[DM] H. Derksen and D. Masser, *Linear equations over multiplicative groups, recurrences, and mixing I*, Proc. London Math. Soc., to appear.

[E] J.-H. Evertse, *On sums of S-units and linear recurrences*, Compos. Math. 53 (1984), 225–244.

[ESS] J.-H. Evertse, H. P. Schlickewei and W. M. Schmidt, *Linear equations in variables which lie in a multiplicative group*, Ann. of Math. 155 (2002), 807–836.

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

[HP] E. Hrushovski and A. Pillay, *Effective bounds for the number of transcendental points on subvarieties of semi-abelian varieties*, Amer. J. Math. 122 (2000), 439–450.

[L1] S. Lang, *Introduction to Algebraic Geometry*, Addison-Wesley, 1972.

[L2] —, *Fundamentals of Diophantine Geometry*, Springer, 1983.

[M] D. Masser, *Mixing and linear equations over groups in positive characteristic*, Israel J. Math. 142 (2004), 189–204.

[MS1] R. Moosa and T. Scanlon, *The Mordell–Lang conjecture in positive characteristic revisited*, in: Model Theory and Applications, Quaderni Mat. 11, L. Belair et al. (eds.), Dipartimento di Matematica, Seconda Università di Napoli, 2002, 273–296.

[MS2] —, —, *F-structures and integral points on semiabelian varieties over finite fields*, Amer. J. Math. 126 (2004), 473–522.

[PS] A. J. van der Poorten and H. P. Schlickewei, *Additive relations in fields*, J. Austral. Math. Soc. Ser. A 51 (1991), 154–170.

[V] J. F. Voloch, *The equation $ax + by = 1$ in characteristic p*, J. Number Theory 73 (1998), 195–200.

Dominik J. Leitner
Mathematisches Institut
Universität Basel
Rheinsprung 21
CH-4051 Basel, Switzerland
E-mail: dominik.leitner@unibas.ch

Received on 4.11.2010
and in revised form on 26.7.2011 (6540)