Mathematische Zeitschrift (2022) 302 1–45
https://doi.org/10.1007/s00209-022-03039-y

Mathematische Zeitschrift

Unlikely intersections for curves in products of Carlitz modules

W. D. Brownawell[1] · **D. Masser**[2]

Received: 5 September 2021 / Accepted: 21 April 2022 / Published online: 5 June 2022

Abstract
The conjectures associated with the names of Zilber-Pink greatly generalize results associated with the names of Manin-Mumford and Mordell-Lang, but unlike the latter they are almost exclusively restricted to zero characteristic. Not so long ago the second author made a start on removing this restriction by studying multiplicative groups over positive characteristic, and recently both authors went further for additive groups with extra Frobenius structure. Here we study additive groups with extra structure coming instead from the Carlitz module. We state a conjecture for curves in general dimension and we prove it in three dimensions. The main tool is a new relative version (for cyclotomic fields) of Denis's analogue of Dobrowolski's classical lower bound for heights, as well as a suitable upper bound. We also work out a couple of special cases in two dimensions: for example with respect to prime fields there are exactly 23 Carlitz roots of unity whose reciprocals are also roots of unity.

Keywords Carlitz modules · Unlikely intersections

Mathematics Subject Classification 11G09 · 11G20 · 14G17

1 Introduction

For over two decades now much has been written on the study of what happens when a fixed algebraic variety sitting inside a fixed commutative group variety is intersected with the union of group subvarieties of suitable dimension. When the group variety is the multiplicative group $\mathbf{G}_{\mathrm{m}}^n$, we may refer to the work of Bombieri, Zannier and the second author (for example the early paper [8] on curves, our later paper [11] on varieties of codimension 2, and our paper [12] on planes) and the wide-ranging extension of Habegger to arbitrary varieties (see [31] for example). When the group variety is projectively complete there are the results

✉ D. Masser
David.Masser@unibas.ch

W. D. Brownawell
wdb@math.psu.edu

[1] Department of Mathematics, The Pennsylvania State University, University Park, PA 16802, USA

[2] Departement Mathematik und Informatik, Universität Basel, Spiegelgasse 1, 4051 Basel, Switzerland

of Viada about powers of a fixed elliptic curve (see [58] for example) as well as those of Rémond generalizing to abelian varieties (see [54] for example); see especially the paper [32] of Habegger and Pila. There are also investigations of Zannier and the second author inside varying group varieties such as elliptic and abelian schemes (see [45–47] for example). All this work on "unlikely intersections" takes place over zero characteristic, and one may consult the book [59] of Zannier for a comprehensive survey. The general conjectures are due to Zilber [60] and Pink [51].

Over positive characteristic it is well-known that related simpler problems, such as those associated with the names Manin-Mumford about torsion points, can become false. For example over zero characteristic the equation

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

has only two solutions in roots of unity x and y (involving primitive sixth roots). However over characteristic p there are infinitely many; indeed we can take any $x \neq 0, 1$ in the algebraic closure $\overline{\mathbf{F}_p}$ and then y accordingly.

Another special kind of unlikely intersection occurs when we intersect the variety with a finitely generated group, an area often associated with the names Mordell-Lang. For example over zero characteristic we can ask for solutions of (1) with x a power of 3 and y a power of -2, amounting essentially to the equation $3^a - 2^b = 1$. This has for centuries been known to have only two solutions in integers a, b. However over characteristic p inside the function field $\mathbf{F}_p(t)$, with x a power of t and y a power of $1 - t$, we have infinitely many solutions

$$x = t^q, \quad y = (1-t)^q = 1 - t^q \quad (q = 1, p, p^2, \ldots).$$

For much more see for example the papers [33] of Hrushovski and [49] of Moosa and Scanlon.

And the torsion situation can be combined with the finitely generated situation by allowing finite rank; under this heading see for example the papers [29] of Ghioca and Moosa and [26] of Ghioca.

The second author [43] made a start on Zilber-Pink problems over positive characteristic, formulating a conjecture for curves in $\mathbf{G}_{\mathrm{m}}^n$ and proving it for $\mathbf{G}_{\mathrm{m}}^3$.

Then in [15] we continued the study of such problems, but now for the additive group $\mathbf{G}_{\mathrm{a}}^n$. Over zero characteristic the naive conjectures for $\mathbf{G}_{\mathrm{a}}^n$ become false, because they implicitly involve group subvarieties (of codimension 2), and there are simply far too many of these. For example the union of all of codimension 1 (and even of codimension $n - 1$) is the whole $\mathbf{G}_{\mathrm{a}}^n$.

Over positive characteristic it is well-known that problems of Manin-Mumford or Mordell-Lang type can be formulated for $\mathbf{G}_{\mathrm{a}}^n$ by imposing some extra structure. One immediately thinks of Drinfeld modules (on which the literature is already substantial); but there is an easier way using Frobenius (also see [26], in particular Theorem 2.6 p.3841). It is these "Frobenius modules" or "F-modules" that we recently studied in [15].

In the present paper we go in the direction of Drinfeld, but we restrict ourselves to the simplest and most attractive forerunner, the Carlitz module.

To fix ideas, let us first review the situation for the multiplicative $\mathbf{G}_{\mathrm{m}}^n$ over zero characteristic. The decisive result was obtained by Maurin [48] (see [7] also), and, taking into account [13], we now know the following best possible result.

Theorem A *Let K be an algebraically closed field of characteristic 0, and let C in $\mathbf{G}_{\mathrm{m}}^n$ be an irreducible curve defined over K. Assume for any non-zero $(r_1, \ldots, r_n)$ in $\mathbf{Z}^n$ that the monomial $x_1^{r_1} \cdots x_n^{r_n}$ is not identically 1 on C. Then there are at most finitely many*

$(\xi_1, \dots, \xi_n)$ in $C(K)$ for which there exist linearly independent $(a_1, \dots, a_n)$, $(b_1, \dots, b_n)$ in $\mathbf{Z}^n$ such that

$$\xi_1^{a_1} \cdots \xi_n^{a_n} = \xi_1^{b_1} \cdots \xi_n^{b_n} = 1.$$

It was already pointed out in [43] (p.506) that the naive analogue of this over positive characteristic is false, in that a (stronger) hypothesis about two monomials, not one, is needed. There is an exactly analogous situation in [15] for $\mathbf{G}_{\mathrm{a}}^n$ with Frobenius structure associated with x^p.

Our Carlitz structure is associated with $x^p + tx$ instead, and it may be found surprising that such a small change makes the situation revert back to that of the original multiplicative result in Theorem A above, with only one monomial.

Let us now recall this $\mathbf{G}_{\mathrm{a}}^n$ with Carlitz structure.

We use a distinguished parameter t.

Write $\mathcal{C} = t\mathcal{F}^0 + \mathcal{F}^1$ where $\mathcal{F}^r$ is the Frobenius taking x to x^{p^r}. Thus $\mathcal{C}(x) = tx + x^p$; or we shall usually write just $\mathcal{C}x$. We have the non-twisted $\mathcal{R} = \mathbf{F}_p[\mathcal{C}]$ inside the twisted $\mathcal{K}\{\mathcal{F}^1\}$, where $\mathcal{K}$ is any field of characteristic p. Of course $\mathcal{K}\{\mathcal{F}^1\}$ acts on $\mathbf{G}_{\mathrm{a}}$ by

$$\alpha x = a_0 x + a_1 x^p + a_2 x^{p^2} + \cdots$$

for $\alpha = a_0\mathcal{F}^0 + a_1\mathcal{F}^1 + a_2\mathcal{F}^2 + \cdots$ in $\mathcal{K}\{\mathcal{F}^1\}$. Here we have used the same juxtaposition notation for the module action, as in αx, and the field action, as in ax. In general throughout this paper the first will be used mainly with greek letter coefficents, and the second mainly with roman letter coefficients; and the two actions will rarely be side-by-side.

There is an action on $\mathbf{G}_{\mathrm{a}}^n$ by $\alpha(x_1, \dots, x_n) = (\alpha x_1, \dots, \alpha x_n)$. Any algebraic subgroup of $\mathbf{G}_{\mathrm{a}}^n$ that is an $\mathcal{R}$-module is then defined by several equations of the form

$$\alpha_1 x_1 + \cdots + \alpha_n x_n = 0$$

where $\alpha_1, \dots, \alpha_n$ are in $\mathcal{R}$. The codimension is the rank of the various $(\alpha_1, \dots, \alpha_n)$ in $\mathcal{R}^n$. We believe in the following version of Theorem A.

Conjecture *Let K be an algebraically closed field containing $\mathbf{F}_p(t)$, and let C in $\mathbf{G}_{\mathrm{a}}^n$ be an irreducible curve defined over K. Assume for any non-zero $(\rho_1, \dots, \rho_n)$ in $\mathcal{R}^n$ that the form $\rho_1 x_1 + \cdots + \rho_n x_n$ is not identically zero on C. Then there are at most finitely many $(\xi_1, \dots, \xi_n)$ in $C(K)$ for which there exist linearly independent $(\alpha_1, \dots, \alpha_n)$, $(\beta_1, \dots, \beta_n)$ in $\mathcal{R}^n$ such that*

$$\alpha_1\xi_1 + \cdots + \alpha_n\xi_n = \beta_1\xi_1 + \cdots + \beta_n\xi_n = 0.$$

The case $n = 1$ is empty.

The case $n = 2$ amounts to an analogue of Manin-Mumford. It was proved in the general context of Drinfeld modules by Scanlon [55], using techniques from model theory. Here we will sketch a more elementary method in the Carlitz context.

Already the case $n = 3$, going beyond torsion, is in the sense of Zilber-Pink. The main result of this paper is a proof for $n = 3$.

It is possible that the cases $n = 4, 5$ can be handled by adapting the methods of [10] and the results of Amoroso and David [1].

But for $n \geq 6$ quite different methods will probably be needed, maybe following [7, 48] or [13].

Also for general n it may well be possible to prove a weaker form of the Conjecture under the stronger hypothesis that $\rho_1 x_1 + \cdots + \rho_n x_n$ is not identically constant on C (the analogue of the hypothesis in [8] for zero characteristic $\mathbf{G}_{\mathrm{m}}^n$).

Anyway, we shall prove

Theorem *Let K be an algebraically closed field containing $\mathbf{F}_p(t)$, and let C in $\mathbf{G}_{\mathrm{a}}^3$ be an irreducible curve defined over K. Assume for any non-zero (ρ_1, ρ_2, ρ_3) in $\mathcal{R}^3$ that the form $\rho_1 x_1 + \rho_2 x_2 + \rho_3 x_3$ is not identically zero on C. Then there are at most finitely many (ξ_1, ξ_2, ξ_3) in $C(K)$ for which there exist linearly independent $(\alpha_1, \alpha_2, \alpha_3)$, $(\beta_1, \beta_2, \beta_3)$ in $\mathcal{R}^3$ such that*

$$\alpha_1\xi_1 + \alpha_2\xi_2 + \alpha_3\xi_3 = \beta_1\xi_1 + \beta_2\xi_2 + \beta_3\xi_3 = 0.$$

In fact Theorem A above for $n = 3$ and $K = \overline{\mathbf{Q}}$ was first proved in [8]. There the concept of height was unavoidable, and we needed also results on upper bounds as well as considerably deeper results on lower bounds.

By contrast the proofs in [43] and [15] do not use heights at all.

It turns out that the proof of our Theorem above follows much more closely [8] and in particular we need heights $h(\xi)$ on $\overline{\mathbf{F}_p(t)}$ (see later).

Of course the condition of algebraic closure can be omitted in all the above statements, but its retention is meant to emphasize that we are considering points of unbounded degree (over $\mathbf{F}_p(t)$ for example).

We will prove the following upper bound, a Carlitz analogue of Theorem 1 of [8] (p.1120).

Proposition 1 *For $K = \overline{\mathbf{F}_p(t)}$ let C be an irreducible curve in $\mathbf{G}_{\mathrm{a}}^n$ defined over K. Assume for any non-zero $(\rho_1, \ldots, \rho_n)$ in $\mathcal{R}^n$ that the form $\rho_1 x_1 + \cdots + \rho_n x_n$ is not identically constant on C. Then there is $\mathfrak{B}$ such that*

$$h(\xi_1) + \cdots + h(\xi_n) \le \mathfrak{B}$$

for all $(\xi_1, \ldots, \xi_n)$ on $C(K)$ for which there exists non-zero $(\alpha_1, \ldots, \alpha_n)$ in $\mathcal{R}^n$ with

$$\alpha_1\xi_1 + \cdots + \alpha_n\xi_n = 0.$$

For the lower bound we have to go beyond [8] with a Carlitz analogue of the Néron-Tate height on a elliptic curve. This was constructed by Denis [21] (even for Drinfeld modules), and we shall denote it by $\hat{h}(\xi) \ge 0$ for ξ in $\overline{\mathbf{F}_p(t)}$ (see later). It is well-known that $\hat{h}(\zeta) = 0$ if and only if ζ is torsion in the sense of Carlitz (see later). Then $\mathbf{F}_p(t, \zeta)$ is a cyclotomic (see later) extension of $\mathbf{F}_p(t)$, and every cyclotomic extension F_{c} has this form (see later). They are separable (see later). We have to consider extensions of F_{c} that are not necessarily separable. Thus we will prove the following lower bound, where from now on we write $F_0 = \mathbf{F}_p(t)$.

Proposition 2 *There is a positive constant c depending only on p with the following property. Let F_{c} be a finite cyclotomic extension of F_0 and let F be an extension of F_{c} of degree d. Then for any non-torsion ξ in F we have*

$$\hat{h}(\xi) \ge c^{-1} d^{-1} \frac{(\log 16D)^{-3}}{(\log\log 16D)^{-2}}$$

where $D = [F : F_0]$.

In fact the lower bound can be multiplied by

$$\min\left\{\frac{(\log 16D)^3}{\log\log 16D}, q^2 \log\log 16D, q^2\sqrt{d}\right\} \ge 1,$$

where q is the inseparable degree of F over F_{c}; but this seems such a tiny improvement that we did not bother to include the details.

In the case $F_c = F_0$ (so no cyclotomy) a very slightly stronger result with $(\log\log 16D)^{-3}$ in place of $(\log\log 16D)^{-2}$ was proved by Denis [21] as Théorème 2 (p.218); but only for extensions which are regular (apparently not essential) and separable - a genuine restriction. This restriction was lifted by Demangos [20], even for a class of Drinfeld modules including Carlitz. The lower bound of his Theorem 2 (p.153) involves an extra negative power of the inseparable degree of F over F_0. But it has the advantage that all the constants appearing are explicitly calculated. See also Bosser and Galateau [14] for several simplifications and improvements, especially Theorem 1.8 (p.168).

In the case $F = F_c$ (so only cyclotomy) David and Pacheco [19] have shown in Théorème 1.0.1 (p.1046) that in fact $\hat{h}(\xi) \geq c^{-1}$ (and even the generalizations to abelian extensions and Drinfeld modules). See also Bauchère [4] for further generalizations.

We now describe our proofs.

That of our Conjecture for $n = 2$ follows one of the classical proofs over $\overline{\mathbf{Q}}$.

For $n = 3$, the proof of our Theorem for $K = \overline{F_0}$ follows the general strategy of [8], using height upper and lower bounds.

We prove Proposition 1 by adopting the slightly simplified exposition in [42].

As for Proposition 2, it is an analogue of a result of Amoroso and Zannier [2] over $\overline{\mathbf{Q}}$ (they actually treated abelian extensions). We have at our disposal the proof in [21] for $F_c = F_0$; this is the natural analogue of the classical result of Dobrowolski [23]. But in fact this proof in [21] resembles much more Laurent's analogue [37] for elliptic curves with complex multiplication. It was Ratazzi [53] who extended Laurent's result to abelian extensions. Meanwhile Pontreau [52] gave a simpler proof of the result of [2] restricted to cyclotomic extensions (whose discriminants are known), and it is this that we adapt here to the the Carlitz context. However our choice of parameters in the auxiliary polynomial is rather different from his; for example we differentiate about d times and he only about $\log d$ times.

In much of the early work it was possible in analogues of Proposition 2 to restrict to ξ that are integral in some sense. But already this made some trouble in the elliptic case [37] and in the original Carlitz work [21]. Here we are obliged to distinguish between valuations of small and large ramification and exploit the known ramification properties of cyclotomic extensions (this argument may fail for abelian extensions).

We have mentioned that over positive characteristic there may well be problems with inseparability. In principle this causes trouble here, especially in the use of an analogue of Siegel's Lemma. We overcome this by adapting a version due to Thunder [57]. This version also involves the genus of the function field F_c analogous to well-known results of Bombieri and Vaaler [6] involving the discriminant; for us it is fortunate that the genus of F_c is also known (just as for cyclotomic extensions of $\mathbf{Q}$).

On the other hand inseparability can be an advantage, For example Theorem 6.6 (p.64) of Ghioca [25] about Drinfeld modules implies that $\hat{h}(\xi) \geq p^{-19}$ for any non-torsion ξ in any purely inseparable extension of F_0. So there is no dependence at all on field degrees. In fact here the only possible torsion ξ are 0 when $p > 2$ and $0, 1, t, t+1$ when $p = 2$ (see later).

In the earlier work [43] and [15] over positive characteristic it was relatively easy to extend the results from $\overline{F_0}$ to general K by means of transcendence degree arguments. This does not seem possible here. Instead we follow a specialization strategy of Bombieri, Zannier and the second author in [9]. But the additive situation diverges somewhat from the multiplicative situation and there are several new elements. For example the zeroes and poles of $x_1^{a_1}x_2^{a_2}x_3^{a_3}$ are on an equal footing, but this is not true of the Carlitz analogue $f = \alpha_1 x_1 + \alpha_2 x_2 + \alpha_3 x_3$, whose poles are clear but whose zeroes are far from evident. In [9] we used Mason's *abc* inequality to settle similar problems. We do not yet have a Carlitz *abc* but we can exploit

the underlying differentiating idea by noting that the derivative of $tx + x^p$ is just t. Using certain systems of identities we can in this way show that the degree of f does not drop too much under specialization. Also in [9] we made an innocent-looking appeal to a result in Mumford [50] for counting inverse images of algebraic maps. This result used the concept of topologically unibranch. In the literature we could not find a suitable positive characteristic analogue and so we developed our own substitute.

The rest of our paper is arranged as follows.

In Sect. 2 we consider the case $n = 2$ of the Conjecture. After warming-up with a simple explicit example, we turn to the general proof; but in view of previous work we feel justified with just a sketch. It also follows from our Theorem by "lifting" the curve in $\mathbf{G}_{\mathrm{a}}^2$ to $\mathbf{G}_{\mathrm{a}}^3$ by introducing a sufficiently general constant value of x_3.

Then in Sect. 3 we prove Proposition 1, also after giving an explicit example.

We postpone the comparatively technical proof of Proposition 2, and in section 4 we show that Propositions 1 and 2 imply the Theorem for $K = \overline{F_0}$.

Section 5 contains preliminary material on Siegel's Lemma, and Sect. 6 more preparations for the proof of Proposition 2, which then follows in Sect. 7.

In Sects. 8 and 9 we start some preliminaries for the general K, including the identities mentioned above. After that Sect. 10 contains an extension of Proposition 1. The main specialization arguments follow in Sect. 11.

We are then able almost to complete the proof in Sect. 12, with a final extra argument in Sect. 13 because certain statements of Mordell-Lang type are not quite in the literature.

As a matter of fact inseparability turns out to be not quite such a problem for the application to our Theorem, because we show in an Appendix that the relevant inseparable degree is bounded. Nevertheless, we have to take it into account in the proof of Proposition 2.

It will be clear to the experts that everything in this paper extends immediately from $\mathbf{F}_p$ to arbitrary finite fields.

And it should go without saying that all our results are effective, and indeed we shall make no further reference to such matters.

We have mentioned Drinfeld modules several times already, so there naturally occurs the problem of generalizing this paper to those (or even to t-modules). It is reasonable to expect some sort of analogue of our Conjecture to hold.

We heartily thank Umberto Zannier for valuable correspondence regarding some of the estimates in Sect. 8.

2 Manin-Mumford

As examples we start with an example for the line $x + y = 1$ as in (1) and also the hyperbola $xy = 1$, with K arbitrary as in the Conjecture. In fact we go further and determine the solutions for every p, as Leitner did in [38, 39] with $\mathbf{G}_{\mathrm{m}}^4$ (confirming an expectation of Hrushovski [33] p.669). This leads to the 23 mentioned in the abstract.

Example 1 (a) If $p > 2$ then there are no ξ, η in K with $\xi + \eta = 1$ for which there exist $\alpha \neq 0,\ \beta \neq 0$ in $\mathcal{R}$ with $\alpha\xi = \beta\eta = 0$. If $p = 2$ then there are infinitely many solutions, all obtained by choosing any torsion ξ and taking $\eta = 1 - \xi$.

(b) If $p > 3$ then there are no ξ, η in K with $\xi\eta = 1$ for which there exist $\alpha \neq 0,\ \beta \neq 0$ in $\mathcal{R}$ with $\alpha\xi = \beta\eta = 0$. If $p = 3$ there are twelve, corresponding to

$$\xi^4 + t\xi^2 + 1 = 0, \quad \xi^4 + (t+1)\xi^2 + 1 = 0, \quad \xi^4 + (t+2)\xi^2 + 1 = 0. \tag{2}$$

If $p = 2$ there are eleven, corresponding to

$$\xi = 1, \quad \xi^2 + t\xi + 1 = 0, \quad \xi^2 + (t+1)\xi + 1 = 0 \tag{3}$$

and

$$\xi^3 + (t+1)\xi^2 + t\xi + 1 = 0, \quad \xi^3 + t\xi^2 + (t+1)\xi + 1 = 0. \tag{4}$$

Verification. For (a) we start by checking that the line defined by $x + y = 1$ does not lie in a proper Carlitz submodule of $\mathbf{G}_{\mathrm{a}}^2$ defined by say $\rho x + \sigma y = 0$. Otherwise we would get

$$0 = \rho x + \sigma y = \rho x + \sigma(1 - x) = \rho x + \sigma 1 - \sigma x = (\rho - \sigma)x + \sigma 1$$

identically in x. However if $\rho - \sigma \neq 0$ then $(\rho - \sigma)x$ involves x. Thus $\rho - \sigma = 0$ and we are left with $\sigma 1 = 0$. But if $\sigma = S(\mathcal{C})$ for a non-constant polynomial S of degree d then as $p \neq 2$ we can quickly check that $\sigma 1$ is a polynomial in t of degree p^{d-1}. This fails for $p = 2$ and indeed then $\sigma_0 1 = 0$ for $\sigma_0 = \mathcal{C}^2 + \mathcal{C}$ (compare [30] p.61). So here the line lies in $\sigma_0 x + \sigma_0 y = 0$. This even makes the above example (a) fail for $p = 2$: simply take any torsion element ξ then because 1 is torsion so also is $\eta = 1 - \xi$.

In fact this calculation proves (a) at once: if ξ, η are torsion then so is $\xi + \eta = 1$ (it is the Carlitz analogue of say $xy = 2$ in zero characteristic $\mathbf{G}_{\mathrm{m}}^2$).

For (b) it is clear that the hyperbola defined by $xy = 1$ does not lie in a proper Carlitz submodule of $\mathbf{G}_{\mathrm{a}}^2$ defined by $\rho x + \sigma y = 0$, because if $\rho \neq 0$ then ρx has positive degree in x and if $\sigma \neq 0$ then σy for $y = 1/x$ has negative degree in x.

Now we proceed to the main argument, by diophantine approximation as in the classical case of $\mathbf{G}_{\mathrm{m}}^2$ over zero characteristic. There is a monic polynomial N in $\mathbf{F}_p[X]$ of minimal degree, say n, such that the torsion elements ξ, η (if any) satisfy

$$N(\mathcal{C})\xi = N(\mathcal{C})\eta = 0. \tag{5}$$

If ζ is a primitive N-torsion element, then there are polynomials R, S with

$$\xi = R(\mathcal{C})\zeta, \ \ \eta = S(\mathcal{C})\zeta \tag{6}$$

(see for example [30] p.55); further we can assume R, S are of degree at most $n - 1$. We can find polynomials A, B, D not all zero with $AR + BS = DN$; and a simple counting argument shows that we can take A, B, D of degree at most $n/2$. Indeed there are $3(1+[n/2])$ coefficients at our disposal subject to $n + [n/2]$ linear conditions, and the difference is

$$3 + 2\left[\frac{n}{2}\right] - n \geq 3 + 2\left(\frac{n}{2} - \frac{1}{2}\right) - n = 2.$$

It follows that

$$\alpha\xi + \beta\eta = 0, \tag{7}$$

for $\alpha = A(\mathcal{C}), \beta = B(\mathcal{C})$, an equation of total degree at most $p^{n/2}$ in ξ, η. We can apply Bezout to this and $\xi\eta = 1$ by our opening observation. We find that for this particular N there are at most $2p^{n/2}$ pairs (ξ, η).

On the other hand replacing ζ by any of its conjugates over F_0 in (6) gives a conjugate pair also on the same hyperbola. Further these pairs are all different, because by the minimality of N and (5), (6) the polynomials R, S, N can have no common factor; and so we can solve $GR + HS + LN = 1$ giving $\zeta = G(\mathcal{C})\xi + H(\mathcal{C})\eta$.

Now the degree of ζ over F_0 is well-known to be the Carlitz-Euler $\phi(N)$ (see [16] p.173). It follows that

$$2p^{n/2} \geq \phi(N). \tag{8}$$

In terms of a prime factorization $N = \prod_{i=1}^{r} N_i^{e_i}$ it is

$$\phi(N) = \prod_{i=1}^{r} \phi(N_i^{e_i}) = \prod_{i=1}^{r} p^{e_i n_i} \left(1 - \frac{1}{p^{n_i}}\right) = p^n \prod_{i=1}^{r} \left(1 - \frac{1}{p^{n_i}}\right)$$

with n_i as the degree of N_i $(i = 1, \ldots, r)$. Thus

$$\phi(N) \geq p^n \left(1 - \frac{1}{p}\right)^r \geq (p-1)^n. \tag{9}$$

So from (8) and the fact that $2p^{1/2} < p - 1$ for $p > 5$ we reduce to the cases $p = 2, 3, 5$.

If $p = 5$ then $2p^{n/2} \geq (p-1)^n$ forces $n = 1$, so $N = X + a$ $(a = 0, 1, 2, 3, 4)$. But the polynomials $N(\mathcal{C})T = T^5 + tT + a$ for $a \neq 0$ are irreducible and non-reciprocal and none is the reciprocal of another, and for $a = 0$ it is irreducible and non-reciprocal after dividing by T. So by (5) there are no solutions if $p = 5$.

If $p = 3$ then $2p^{n/2} \geq (p-1)^n$ forces $n \leq 4$. Checking each of the 81 possibilities for N we find that

$$N = X^2 + 2, \quad X^2 + 2X, \quad X^2 + X$$

give rise via (5) to the solutions indicated in (2). We find no other ξ.

If $p = 2$ we have to work a bit harder, and we divide into eight cases according to which of the three irreducible polynomials X, $X + 1$, $X^2 + X + 1$ of degree at most 2 divide N.

The worst case is when all three divide N, say as N_1, N_2, N_3 respectively. Thus $r \geq 3$ and

$$\phi(N) = 2^n \left(\frac{1}{2}\right)\left(\frac{1}{2}\right)\left(\frac{3}{4}\right) \prod_{i=4}^{r} \left(1 - \frac{1}{2^{n_i}}\right).$$

In the product $n_i \geq 3$ so it is at least $(7/8)^{r-3}$. But also

$$n = e_1 + e_2 + 2e_3 + \sum_{i=4}^{r} e_i n_i \geq 4 + 3(r-3).$$

So now $n \geq 4$ and we get

$$2(2^{n/2}) \geq 3(2^{n-4}) \left(\frac{7}{8}\right)^{(n-4)/3},$$

which implies $n \leq 7$. Therefore $N = N_1 N_2 N_3 (X^3 + aX^2 + bX + c)$ leading to just eight possibilities for N.

The other seven cases lead to $n \leq 6$ and better. For example when none of N_1, N_2, N_3 divide N, then every $n_i \geq 3$ and so the above product is at least $(7/8)^n$.

Here the factorization of each $N(\mathcal{C})T$ on Maple can take up to two hours. Spotting reciprocal factors or reciprocal pairs by eye is not too easy and so *ad hoc* methods using resultants were developed. Thus if the resultant of $N(\mathcal{C})T$ and its own reciprocal is non-zero, then such factors cannot exist. Already all solutions come from

$$N = X^2 + X, \quad X^3 + X^2, \quad X^3 + X, \quad X^4 + X.$$

The first N gives $\xi = 1$ of course in (3) and the next two give the next two there; but the last N provides the reciprocal pair

$$T^3 + (t+1)T^2 + tT + 1, \quad T^3 + tT^2 + (t+1)T + 1$$

leading to the last two in (4).

This completes the verification of Example 1 (we checked that allowing powers of primes does not yield any additional solutions).

It is not much more difficult to prove the general conjecture above in $\mathbf{G}_{\mathrm{a}}^2$ by these means. We just sketch the details, as it also follows rather easily from our Theorem by "lifting" the curve in $\mathbf{G}_{\mathrm{a}}^2$ to $\mathbf{G}_{\mathrm{a}}^3$ by introducing a sufficiently general constant value of x_3.

It suffices to treat the case $K = \overline{F_0}$. We can assume that C is defined over a finite extension F of F_0, otherwise it contains anyway at most finitely many points algebraic over F_0 and in particular torsion points. We obtain as above (7) and now it is by assumption that we can apply Bezout. This time we are allowed to use the conjugates of ζ only over F, but that hardly affects their number. However as in Example 1 we need a lower bound for $\phi(N)$ better than (9) and at least of the form $c(p^n)^\theta$ for some $\theta > \frac{1}{2}$. The only trouble is at $p = 2$ but in general we can argue

$$\prod_{i=1}^{r}\left(1 - \frac{1}{p^{n_i}}\right) \geq \prod_{j=1}^{n}\left(1 - \frac{1}{p^j}\right)^{r_j}$$

for the number r_j of monic irreducible polynomials of degree j in F_0. Taking logarithms, using $-\log(1-x) \leq 2x$ for $0 \leq x \leq \frac{1}{2}$, and also

$$r_j = \frac{1}{j}\sum_{d|j}\mu(d)p^{j/d} \leq \frac{p^j}{j}\left(1 + \sum_{6\leq d|j}\frac{1}{p^{j(1-1/d)}}\right) \leq \frac{p^j}{j}\left(1 + \frac{j}{2^{5j/6}}\right) \leq 2\frac{p^j}{j} \tag{10}$$

as well as $\sum_{j=1}^{n} 1/j \leq \log n + 1$ we find

$$\phi(N) \geq \frac{p^n}{55n^4} \tag{11}$$

(with n as the degree of N) comfortably of the required form. We shall need both (10) and (11) later.

It will now be seen that $\|N\|$ is a useful notation for p^n.

3 Proof of Proposition 1

We use the standard height function h on $F_0 = \mathbf{F}_{\mathbf{p}}(t)$ normalized to $h(t^d) = d$. Thus if ξ_0 is in F_0 we have

$$h(\xi_0) = \sum_w \log\max\{1, |\xi_0|_w\}$$

taken over all w on F_0 (which correspond to monic irreducible P in F_0, with $|P|_P = e^{-d}$ for d the degree of P, and $|t|_\infty = e$). We extend it in the usual way to the algebraic closure, so that if ξ is in an extension F of degree D over F_0, then

$$h(\xi) = \frac{1}{D}\sum_v D_v \log\max\{1, |\xi|_v\} \tag{12}$$

over all valuations v on F which extend those on F_0 (that is, for all x in F_0 we have $|x|_v = |x|_w$ for some w as above) and $D_v = e_v f_v$ the local degrees. We shall often say that v divides P or ∞.

For example

$$h(t^\theta) = |\theta| \tag{13}$$

when θ is rational.

First here is an illustration in the spirit of Example 1. However we cannot use the line $x + y = 1$ here because $1x + 1y$ is constant on it. And indeed the points with coordinates

$$\xi = \mathcal{C}^m 1, \quad \eta = 1 - \mathcal{C}^m 1 = (1 - \mathcal{C}^m)1$$

have $h(\xi) = p^{m-1}$ going to infinity with m provided $p \neq 2$, in view of our remarks above about the degree of $\sigma 1$ (again the analogue of $xy = 2$ in zero characteristic $\mathbf{G}_{\mathrm{m}}^2$). But the (Carlitz) hyperbola $xy = 1$ is fine.

Example 2 If ξ, η are in $\overline{F_0}$ with $\xi\eta = 1$ for which there exist α, β not both zero in $\mathcal{R}$ with $\alpha\xi = \beta\eta$ then $h(\xi) + h(\eta) \leq 18$.

Verification. We use the canonical height introduced by Denis for Drinfeld modules (for which he proved his lower bound, at least in the Carlitz case). This is defined as

$$\hat{h}(\xi) = \lim_{m\to\infty} \frac{h(\mathcal{C}^m \xi)}{p^m}.$$

Not only do we have the obvious $\hat{h}(\mathcal{C}\xi) = p\hat{h}(\xi)$ but even (in the notation at the end of Sect. 2)

$$\hat{h}(P(\mathcal{C})\xi) = \|P\|\hat{h}(\xi) \tag{14}$$

for any non-zero P in $\mathbf{F}_p[X]$. Also since $\mathcal{C}$ is additive we get

$$\hat{h}(\xi + \eta) \leq \hat{h}(\xi) + \hat{h}(\eta) \tag{15}$$

(but not $\hat{h}(\xi\eta) \leq \hat{h}(\xi) + \hat{h}(\eta)$ or even $\hat{h}(\xi^2) \leq 2\hat{h}(\xi)$, for example ξ can be torsion while ξ^2 is not, as for $\xi = \sqrt{-t}$ with $p = 3$; and even $\hat{h}(1/\xi)$ need not be $\hat{h}(\xi)$, as for $\xi = t$ with $p = 2$).

To compare with (13), it may be shown for $p = 2$ that

$$\hat{h}(t^\theta) = \theta, \quad 0, \quad \frac{1+\theta}{2}, \quad -\theta + 2^{[\theta]-1}(\theta - [\theta] + 1), \quad -\theta$$

when

$$\theta > 1, \quad \theta = 1, \quad 0 < \theta < 1, \quad 0 \geq \theta \notin \mathbf{Z}, \quad 0 \geq \theta \in \mathbf{Z}$$

respectively (we shall not need these values).

Denis showed that $\hat{h}(\xi)$ differs from $h(\xi)$ by a bounded amount, and indeed we now check that

$$|\hat{h}(\xi) - h(\xi)| \leq 3 \tag{16}$$

independently of p.

In the first place we have an upper bound

$$h(\mathcal{C}\xi) = h\left(\xi\left(t+\xi^{p-1}\right)\right) \le h(\xi) + h\left(t+\xi^{p-1}\right)$$
$$\le h(\xi) + 1 + (p-1)h(\xi) = ph(\xi) + 1.$$

For a corresponding lower bound we use the standard Nullstellensatz argument. With $\rho = \xi t^{p+1}$ and

$$\sigma = \xi^{(p-1)p} - t\xi^{(p-1)(p-1)} + t^2\xi^{(p-1)(p-2)} \cdots - t^p = \frac{(\xi^{p-1})^{p+1} - (-t)^{p+1}}{\xi^{p-1} - (-t)}$$

we have

$$\rho + \sigma(\xi^p + t\xi) = \xi^{p^2}.$$

Thus for any ultrametric valuation we deduce

$$\max\left\{1, |\xi|^{p^2}\right\} \le \max\{1, |\mathcal{C}\xi|\} \max\left\{1, |\xi|^{p^2-p}\right\} \max\left\{1, |t|^{p+1}\right\}.$$

Cancelling and taking the product with suitable exponents leads in the usual way to

$$h(\mathcal{C}\xi) \ge ph(\xi) - (p+1).$$

The standard telescoping sum gives

$$|\hat{h}(\xi) - h(\xi)| \le (p+1)\sum_{i=1}^{\infty} p^{-i} = \frac{p+1}{p-1} \le 3.$$

Now take ξ, η as in Example 2. If $\alpha = 0$ then η is torsion so $h(\eta) \le 3$ by (16). Thus $h(\xi) = h(\eta) \le 3$ too, and we are done. Similarly if $\beta = 0$; so we will henceforth assume that $\alpha \ne 0, \beta \ne 0$. Write $\alpha = A(\mathcal{C}), \beta = B(\mathcal{C})$. Then (14) leads to

$$p^l \hat{h}(\xi) = p^m \hat{h}(\eta)$$

with l the degree of A and m the degree of B. As in the situation over $\overline{\mathbf{Q}}$ (see for example Theorem 14.9 of [44] p.176) everything depends on the relation between l and m. We note from (16) that

$$\hat{h}(\eta) = \hat{h}(1/\xi) \le h(1/\xi) + 3 = h(\xi) + 3 \le \hat{h}(\xi) + 6. \tag{17}$$

First suppose $l > m$. Then $p\hat{h}(\xi) \le \hat{h}(\eta)$ which by (17) leads to $\hat{h}(\xi) \le 6$ and so $h(\xi) \le 9$; further $h(\eta) = h(\xi) \le 9$ as well. So

$$h(\xi) + h(\eta) \le 18. \tag{18}$$

By symmetry we get the same result if $l < m$. So it remains only to consider $l = m$.

We can now write

$$\alpha = \alpha_0 + a\mathcal{C}^m, \ \beta = \beta_0 + b\mathcal{C}^m$$

with nonzero a, b in $\mathbf{F}_p$ and α_0, β_0 of degree in $\mathcal{C}$ smaller than m. Then

$$\alpha_0\xi - \beta_0\eta = b\mathcal{C}^m\eta - a\mathcal{C}^m\xi$$

and we proceed to compare the canonical heights.

To begin with the left-hand side, (15) gives

$$\hat{h}(\alpha_0\xi - \beta_0\eta) \le \hat{h}(\alpha_0\xi) + \hat{h}(\beta_0\eta)$$

which is by (14) and (17) at most

$$p^{m-1}\left(\hat{h}(\xi)+\hat{h}(\eta)\right) \le p^{m-1}(2h+6) \tag{19}$$

with $h = h(\xi)$.

To continue with the right-hand side

$$\hat{h}(b\mathcal{C}^m\eta - a\mathcal{C}^m\xi) = p^m\hat{h}(b\eta - a\xi)$$

and

$$\hat{h}(b\eta - a\xi) \ge h(b\eta - a\xi) - 3 = h(b/\xi - a\xi) - 3 = 2h - 3$$

(here we used $ab \ne 0$). Comparison with (19) gives

$$h \le \frac{3p+6}{2p-2} \le 6.$$

The same argument gives the same bound for $h(\eta)$ and by addition we get something stronger than (18). So the verification is complete.

Next we need an analogue of Lemma 2.1 of [42] (p.327).

Lemma 1 *If $\xi_1, \dots, \xi_n$ are in $\overline{F_0}$ for which there exist $\alpha_1, \dots, \alpha_n$ not all zero in $\mathcal{R}$ with $\alpha_1\xi_1 + \cdots + \alpha_n\xi_n = 0$, then for any non-negative integer m there are $\beta_1, \dots, \beta_n$ in $\mathcal{R}$, not all zero, such that*

$$\hat{h}(\beta_1\xi_1 + \cdots + \beta_n\xi_n) \le p^{-m/n}(\hat{h}(\xi_1) + \cdots + \hat{h}(\xi_n))$$

with $\beta_1 = B_1(\mathcal{C}), \dots, \beta_n = B_n(\mathcal{C})$ for

$$\max\{\|B_1\|, \dots, \|B_n\|\} \le p^m.$$

Proof Write $\alpha_i = A_i(\mathcal{C})$ $(i = 1, \dots, n)$ with $p^d = \max\{\|A_1\|, \dots, \|A_n\|\}$. Choose any A in $\mathbf{F}_p[X]$ with degree exactly d. Then the A_i/A $(i = 1, \dots, n)$ are in the completion $\mathbf{F}_p[[1/X]]$ of $\mathbf{F}_p[X]$. It follows that for any positive integer l we can find $Q \ne 0$, $B_1, \dots, B_n$ in $\mathbf{F}_p[X]$ of degree at most m such that

$$Q\frac{A_i}{A} - B_i \text{ is in } X^{-l}\mathbf{F}_p[[1/X]] \qquad (i = 1, \dots, n) \tag{20}$$

as long as $m+1 > n(l-1)$. Thus we can choose $l = [m/n]+1 > m/n$. Now the polynomials $C_i = QA_i - AB_i$ $(i = 1, \dots, n)$ have degrees at most $d - l$.

We now put

$$\beta_i = B_i(\mathcal{C}), \quad \gamma_i = C_i(\mathcal{C}) \quad (i = 1, \dots, n)$$

and note that

$$-A(\mathcal{C})(\beta_1\xi_1 + \cdots + \beta_n\xi_n) = \gamma_1\xi_1 + \cdots + \gamma_n\xi_n. \tag{21}$$

Taking canonical heights gives

$$p^d\hat{h}(\beta_1\xi_1 + \cdots + \beta_n\xi_n) = \hat{h}(\gamma_1\xi_1 + \cdots + \gamma_n\xi_n) \le p^{d-l}\left(\hat{h}(\xi_1) + \cdots + \hat{h}(\xi_n)\right).$$

This is the required result since $l > m/n$. Note that $\beta_1, \dots, \beta_n$ are indeed not all zero otherwise (20) would give a contradiction, because $l > 0$ and some A_i has the same degree as A. □

We can now prove Proposition 1. Suppose that C is defined over a finite extension F of F_0. We fix some m with

$$p^{m/n} \geq 2nd \tag{22}$$

where now d is the degree of the curve C. For any point $P = (\xi_1, \ldots, \xi_n)$ as there we construct $\beta_i = B_i(\mathcal{C})$ $(i = 1, \ldots, n)$ as in Lemma 1. Then the function $y = \beta_1 x_1 + \cdots + \beta_n x_n$ is not constant on C by hypothesis. Thus for any i there is a polynomial $\Phi_i(Y, X)$ in $F[Y, X]$, of positive degree in X, such that $\Phi_i(y, x_i) = 0$ on C; further we can take the degree in Y to be at most d. By specialization there follows $\Phi_i(\eta, \xi_i) = 0$ for $\eta = \beta_1\xi_1 + \cdots + \beta_n\xi_n$. Standard height estimates give now $h(\xi_i) \leq dh(\eta) + c$ for some c independent of P. So for $h = h(\xi_1) + \cdots + h(\xi_n)$ we get by Lemma 1 and (16)

$$\begin{aligned} h &\leq n(dh(\eta) + c) \leq nd\hat{h}(\eta) + c' \\ &\leq ndp^{-m/n}(\hat{h}(\xi_1) + \cdots + \hat{h}(\xi_n)) + c' \leq ndp^{-m/n}h + c'' \end{aligned}$$

for c', c'' also independent of P. The required result $h \leq 2c''$ now follows from (22). This completes the proof of Proposition 1.

4 Proof of Theorem for $K = \overline{F_0}$

Here we deduce the Theorem for $K = \overline{F_0}$ from Proposition 1 just proved together with Proposition 2 whose proof will follow in Sect. 7. To avoid logarithmic pedantries we reformulate Proposition 2 as follows.

Assertion. *Given $\varepsilon > 0$, there is a positive constant c depending only on p and ε with the following property. Let F_c be a finite cyclotomic extension of F_0 and let F be a finite extension of F_c. Then for any non-torsion ξ in F we have*

$$\hat{h}(\xi) \geq c^{-1}[F : F_c]^{-1-\varepsilon}[F_c : F_0]^{-\varepsilon}.$$

It will be clear that all we need is any $\varepsilon < 1/3$. But during this section we will use $\ll$ instead of c.

In what follows we shall be relatively brief, as it follows the strategy over $\overline{\mathbf{Q}}$ (see for example [42] p.330).

Given (ξ_1, ξ_2, ξ_3) as in the Theorem, we can find ξ and a torsion point ζ such that

$$F_0(\xi_1, \xi_2, \xi_3) = F_0(\zeta, \xi) \tag{23}$$

and

$$\xi_1 = \sigma_1\xi + \tau_1\zeta, \quad \xi_2 = \sigma_2\xi + \tau_2\zeta, \quad \xi_3 = \sigma_3\xi + \tau_3\zeta \tag{24}$$

for $\sigma_1, \tau_1, \sigma_2, \tau_2, \sigma_3, \tau_3$ in $\mathcal{R}$, somewhat as in (6). Here it is because the $\mathcal{R}$-module generated by ξ_1, ξ_2, ξ_3 in $\overline{F_0}$ has rank at most 1, and so is $\mathcal{R}\xi + \mathcal{Z}$ for a finitely generated torsion module $\mathcal{Z}$, which further has the form $\mathcal{R}\zeta$. And if ζ has order ν then we can take τ_1, τ_2, τ_3 as polynomials in $\mathcal{C}$ of degree less than that of ν.

We now want to find small $\gamma_1, \gamma_2, \gamma_3, \delta$ in $\mathcal{R}$, not all zero, such that

$$\gamma_1\sigma_1 + \gamma_2\sigma_2 + \gamma_3\sigma_3 = 0, \quad \gamma_1\tau_1 + \gamma_2\tau_2 + \gamma_3\tau_3 = \delta\nu.$$

Using any form of Siegel's Lemma over $\mathbf{F}_p[X]$, or by counting as above, we find a solution with

$$\max\{\|\gamma_1\|, \|\gamma_2\|, \|\gamma_3\|\} \ll (\|\nu\| M)^{1/2}, \quad M = \max\{\|\sigma_1\|, \|\sigma_2\|, \|\sigma_3\|\}.$$

It follows from (24) that

$$\gamma_1\xi_1 + \gamma_2\xi_2 + \gamma_2\xi_3 = 0. \tag{25}$$

Clearly $\gamma_1, \gamma_2, \gamma_3$ are not all zero, and we deduce from (25) and Bezout that

$$D = [F_0(\xi_1, \xi_2, \xi_3) : F_0] \ll (\|\nu\| M)^{1/2}. \tag{26}$$

On the other hand (24) gives

$$\hat{h}(\xi_1) = \|\sigma_1\|\hat{h}(\xi), \quad \hat{h}(\xi_2) = \|\sigma_2\|\hat{h}(\xi), \quad \hat{h}(\xi_3) = \|\sigma_3\|\hat{h}(\xi).$$

Let us temporarily assume that no non-trivial $\rho_1 x_1 + \rho_2 x_2 + \rho_3 x_3$ is constant on our curve C, as in Proposition 1. Then summing we get

$$\hat{h}(\xi) \ll M^{-1}.$$

Assuming further that ξ is non-torsion, we are now set up to apply the Assertion, with $F = F_0(\xi_1, \xi_2, \xi_3) = F_c(\xi)$ for $F_c = F_0(\zeta)$. We get

$$\hat{h}(\xi) \gg (D/\phi(\nu))^{-1-\varepsilon}(\phi(\nu))^{-\varepsilon} = D^{-1-\varepsilon}\phi(\nu).$$

By (11) we have $\phi(\nu) \gg \|\nu\|^{1-\varepsilon}$, and comparison gives $\|\nu\|^{1-\varepsilon} M \ll D^{1+\varepsilon}$. But for $\varepsilon < 1/3$ this contradicts (26) or better gives an upper bound for D which implies everything (by the usual Northcott).

If ξ is torsion, then ξ_1, ξ_2, ξ_3 are, and Manin-Mumford on a suitable projection to two dimensions settles the thing.

Finally, what if some non-trivial $\rho_1 x_1 + \rho_2 x_2 + \rho_3 x_3$ is constant on C? We can assume the coefficients are coprime in $\mathcal{R}$ and then use $\mathrm{GL}_3(\mathcal{R})$ to assume that it is x_3 that is some constant, which we can call ξ_3. Then ξ_3 is non-torsion. Now the thing reduces to Mordell-Lang in two dimensions, that is, a group of finite $\mathbf{Q}$-dimension (in fact dimension 1). This was done by Ghioca [27]. But we can too. Namely, we can just eliminate ξ_3 from the two relations between ξ_1, ξ_2, ξ_3 to get a relation between ξ_1, ξ_2 on the projected curve C' in $\mathbf{G}_a^2$. By Proposition 1 for this projection we see that ξ_1, ξ_2 have bounded heights (unless some non-trivial $\phi = \kappa_1 x_1 + \kappa_2 x_2$ is constant on C'); and of course so does ξ_3. We still have (24) and we can argue as before. And really finally, if some non-trivial ϕ is constant on C' then it may as well be $x_2 = \xi_2$. But now ξ_2, ξ_3 must be independent and we cannot have two relations. All this is exactly parallel to the situation over $\overline{\mathbf{Q}}$.

5 Siegel's Lemma

Denis [21] and others use an *ad hoc* version for function fields, rather as in the proof of our Lemma 1. We need a relative version. That for number fields involves discriminants. The correct analogue for function fields involves the genus and was found by Thunder [57].

We stick with our $F_0 = \mathbf{F}_p(t)$, and a finite extension F of F_0. We have a genus $g(F)$ (see [3] for this and much more). In [57] (pp.148,150) a projective absolute height $h_{\mathbf{P}}$ on F^n is

defined; it involves the integer

$$\mathfrak{m}(F, F_0) = \frac{[F : F_0]}{[\mathbf{F} : \mathbf{F}_p]}$$

where $\mathbf{F}$ is the algebraic closure of $\mathbf{F}_p$ in F. It is not hard to see that it coincides with the natural extension of (12) to non-zero vectors by

$$h_{\mathbf{P}}(\xi_1, \ldots, \xi_n) = \frac{1}{D} \sum_v D_v \log \max\{|\xi_1|_v, \ldots, |\xi_n|_v\}. \tag{27}$$

Also it is convenient to define the height of the zero vector as zero.

We have the following extension of Corollary 3 of [57] (p.149), in which a condition about full rank is eliminated. Of course we cannot afford the luxury of a Grassmannian height anymore.

Lemma 2 *Let F_* be a finite extension of F_0 and let F_{s} be a separable extension of F_* of degree r. Let M be a matrix with $m \geq 1$ rows $R_1, \ldots, R_m$ and n columns with entries in F_{s}. If $l = n - rm \geq 1$ then there are linearly independent rows $\mathbf{b} = \mathbf{b}_1, \ldots, \mathbf{b}_l$ in F_*^n with $M\mathbf{b}^t = 0$ that satisfy*

$$\sum_{\nu=1}^{l} h_{\mathbf{P}}(\mathbf{b}_\nu) \ \leq \ rm \max_{\mu=1,\ldots,m} h_{\mathbf{P}}(R_\mu) + l\frac{g + \mathfrak{m}}{\mathfrak{m}}$$

where $g = g(F_*)$ and $\mathfrak{m} = \mathfrak{m}(F_*, F_0)$.

Proof If $M = 0$ the result is obvious (for example with any l standard basis elements of F_*^n), so we assume $M \neq 0$. We follow the argument of Bombieri and Gubler [5] (pp.75,79,80). For that we define $h_{\mathbf{P}}(M')$ for any matrix M' with m' rows and n' columns and $m' < n'$ of rank $s \geq 1$ to be the Grassmannian height $h_{\mathbf{P}}(\hat{M})$ as in [57] (p.151), where $\hat{M}$ consists of any s independent rows of M'. This is the analogue of the definition in [5] (p.75).

Pick a basis $(\lambda_1, \ldots, \lambda_r)$ of F_{s}/F_* and write $M = \lambda_1 M_1 + \cdots + \lambda_r M_r$ with $M_1, \ldots, M_r$ over F_*. Let $\sigma_1, \ldots, \sigma_r$ be the different (here we use separability) embeddings of F_{s}, fixing F_*, into the algebraic closure of F_*. Then we check that $\sigma(M) = \Lambda M_\sigma$ for

$$\sigma(M) = \begin{pmatrix} \sigma_1(M) \\ \vdots \\ \sigma_r(M) \end{pmatrix}, \quad M_\sigma = \begin{pmatrix} M_1 \\ \vdots \\ M_r \end{pmatrix}$$

with invertible Λ.

Let $j \geq l$ be the dimension of the space B of all $\mathbf{b}$ in F_*^n with $M\mathbf{b}^t = 0$. Thus we see that M_σ has rank $n - j > 0$. Thus $\sigma(M)$ too. Let $\hat{M}_\sigma$ be a submatrix of M_σ consisting of $n - j$ independent rows. By Corollary 2 of [57] (p.148) there are linearly independent $\mathbf{b}_1, \ldots, \mathbf{b}_j$ in F_*^n with $\hat{M}_\sigma \mathbf{b}^t = 0$ that satisfy

$$\sum_{\nu=1}^{j} h_{\mathbf{P}}(\mathbf{b}_\nu) \ \leq \ h_{\mathbf{P}}(\hat{M}_\sigma) + jG$$

with $G = (g - 1 + \mathfrak{m})/\mathfrak{m}$. These form of course a basis of B. Also by definition $h_{\mathbf{P}}(\hat{M}_\sigma) = h_{\mathbf{P}}(M_\sigma)$ and it is not difficult to see that $h_{\mathbf{P}}(M_\sigma) = h_{\mathbf{P}}(\sigma(M))$. We can calculate this last height by choosing any $n - j$ independent rows. These have the form $R = \sigma_\rho(R_\mu)$ with

$$h_{\mathbf{P}}(R) = h_{\mathbf{P}}(R_\mu) \leq \max_{\mu=1,\ldots,m} h_{\mathbf{P}}(R_\mu) = H$$

(say), and because $h_{\mathbf{P}}(\hat{M}_\sigma)$ is at most the sum of the heights of its rows we get

$$\sum_{\nu=1}^{j} h_{\mathbf{P}}(\mathbf{b}_\nu) \ \le \ (n-j)H + jG.$$

Ordering by increasing height we deduce

$$\sum_{\nu=1}^{l} h_{\mathbf{P}}(\mathbf{b}_\nu) \ \le \ \frac{l}{j}((n-j)H + jG) = l\frac{n-j}{j}H + lG,$$

and finally recalling $j \ge l$ we get the required result. □

We next drop to a single solution.

Lemma 3 *Let F_* be a finite extension of F_0 and let F_s be a separable extension of F_* of degree r. Let M be a matrix with $m \ge 1$ rows $R_1, \dots, R_m$ and n columns with entries in F_s. If $n > rm$ then there is non-zero $\mathbf{b}$ in F_*^n with $M\mathbf{b}^t = 0$ that satisfies*

$$h_{\mathbf{P}}(\mathbf{b}) \ \le \ \frac{rm}{n-rm} \max_{\mu=1,\dots,m} h_{\mathbf{P}}(R_\mu) + \frac{g+\mathfrak{m}}{\mathfrak{m}}$$

where $g = g(F_*)$ and $\mathfrak{m} = \mathfrak{m}(F_*, F_0)$.

Proof This follows from Lemma 2 after taking the $\mathbf{b}_\nu$ with smallest height. □

Finally we allow inseparable extensions. This seems to be new.

Lemma 4 *Let F_* be a finite extension of F_0 and let F be an extension of F_* of degree d and inseparable degree q. Let M be a matrix with $m \ge 1$ rows $R_1, \dots, R_m$ and n columns with entries in F. If $qn > dm$ then there is non-zero $\mathbf{b}$ in $(F_*^{1/q})^n$ with $M\mathbf{b}^t = 0$ that satisfies*

$$h_{\mathbf{P}}(\mathbf{b}) \ \le \ \frac{dm}{qn-dm} \max_{\mu=1,\dots,m} h_{\mathbf{P}}(R_\mu) + \frac{g+\mathfrak{m}}{\mathfrak{m}}$$

where $g = g(F_*)$ and $\mathfrak{m} = \mathfrak{m}(F_*, F_0)$.

Proof We first solve $M^q\mathbf{c}^t = 0$ using Lemma 3, where M^q is not the q-th power but just has entries the q-th powers of those of M. Now we are over the separable extension F^q of degree $r = d/q$ and so there is a non-zero solution $\mathbf{c}$ in F_*^n with

$$h_{\mathbf{P}}(\mathbf{c}) \ \le \ \frac{rm}{n-rm} \max_{\mu=1,\dots,m} qh_{\mathbf{P}}(R_\mu) + \frac{g+\mathfrak{m}}{\mathfrak{m}}.$$

To finish we take $\mathbf{b} = \mathbf{c}^{1/q}$, so that the height gets divided by q. □

6 More preliminaries

Some of these are analogues of those occurring in the original Dobrowolski proof [23]. We write $\mathcal{O}_0$ for the ring of integers $\mathbf{F}_p[t]$ of $F_0 = \mathbf{F}_p(t)$.

Lemma 5 *For $j \ge 8$ the number r_j of monic irreducible Q in $\mathcal{O}_0$ of degree j satisfies*

$$\frac{1}{2}\frac{p^j}{j} \le r_j \le 2\frac{p^j}{j}.$$

Proof The upper bound is (10) and the lower bound follows by similar arguments. Of course the Prime Number Theorem for $\mathcal{O}_0$ suffices for our purpose. □

From now on we find it more convenient to use the notation ξ^Q instead of $Q(\mathcal{C})\xi$ for our Carlitz $\mathcal{C}$.

Lemma 6 *Let F_* be an extension of F_0 of degree d and let F be an extension of F_*. Suppose $\xi \neq 0$ in F is not torsion.*

(a) As Q runs over all monic irreducible polynomials in $\mathcal{O}_0$, the ξ^Q are all non-conjugate over F_*.

(b) As Q runs over all monic irreducible polynomials in $\mathcal{O}_0$, we have $[F_*(\xi^Q) : F_*] = [F_*(\xi) : F_*]$ with at most $(\log d)/(\log 2)$ exceptions.

Proof This is essentially Lemma 4 of [21] (p.219), where it is merely said that the proof is identical to Dobrowolski's (and was for the separable case). We supply some details.

For (a) suppose ξ^{Q_1}, ξ^{Q_2} are conjugate. Then so are $(\xi^{Q_1})^{Q_1} = \xi^{Q_1^2}$ and $(\xi^{Q_2})^{Q_1} = \xi^{Q_1 Q_2}$, and so are $(\xi^{Q_1})^{Q_2} = \xi^{Q_2 Q_1} = \xi^{Q_1 Q_2}$ and $(\xi^{Q_2})^{Q_2} = \xi^{Q_2^2}$. Iterating we find that the $d+1$ elements $\xi^{Q_1^d}, \xi^{Q_1^{d-1}Q_2}, \ldots, \xi^{Q_2^d}$ are all conjugate over F_*. So two must coincide. As ξ is non-torsion this leads to $Q_1^l = Q_2^l$ for some positive integer l. And as Q_1, Q_2 are monic this leads to $Q_1 = Q_2$.

For (b) suppose $Q_1, \ldots, Q_m$ are different with

$$[F_*(\xi^{Q_i}) : F_*] \neq [F_*(\xi) : F_*] \ \ (i = 1, \ldots, m). \tag{28}$$

Write for brevity $Q = Q_i$ and $R = Q_1 \cdots Q_{i-1}$ (with $R = 1$ if $i = 1$). Then $F_*(\xi^{RQ})$ lies in $F_*(\xi^R)$ and we claim that equality is impossible. Otherwise ξ^R in $F_*(\xi^{RQ})$ is in $F_*(\xi^Q)$. Now there are U, V in $\mathcal{O}_0$ with $UR + VQ = 1$, and then $\xi = (\xi^R)^U + (\xi^Q)^V$ also lies in $F_*(\xi^Q) = F_*(\xi^{Q_i})$. But this would contradict (28).

Taking $i = m, \ldots, 1$ we deduce that the fields

$$F_*(\xi^{Q_1 \cdots Q_m}), \ F_*(\xi^{Q_1 \cdots Q_{m-1}}), \ldots, \ F_*(\xi^{Q_1}), \ F_*(\xi)$$

form a strictly increasing chain. Thus

$$2^m \leq [F_*(\xi) : F_*(\xi^{Q_1 \cdots Q_m})] \leq [F_*(\xi) : F_*] \leq d$$

and now (b) is clear. □

The next result reflects the need in some of the previous literature to distinguish between small and large ramification e_v as in (12).

Lemma 7 *Let F be an extension of F_0 of degree D, let Q be monic irreducible in $\mathcal{O}_0$, let ξ be in F, and let E be the set of v on F dividing Q such that ξ is not v-integral. Then*

$$\hat{h}(\xi) \geq \frac{1}{D} \frac{\log \|Q\|}{\log p} \sum_{v \in E} \frac{D_v}{e_v}.$$

Proof If Q has degree n, the value group of the valuation on F_0 corresponding to Q is generated by $g = e^n = \|Q\|^{1/\log p}$. So that of v by g^{1/e_v}. Thus for each v in E we have $|\xi|_v \geq g^{1/e_v}$. It follows easily that $|\xi^{t^m}|_v \geq g^{p^m/e_v}$ for any positive integer m. Thus $h(\xi^{t^m}) \geq D^{-1} p^m (\log g) \sum_{v \in E} D_v/e_v$. So

$$p^m \hat{h}(\xi) = \hat{h}(\xi^{t^m}) \geq -3 + \frac{1}{D} p^m (\log g) \sum_{v \in E} \frac{D_v}{e_v}$$

by (16). Making m tend to infinity gives the result. □

From now on the intermediate field F_* will be cyclotomic over F_0, so it has the form $F_c = F_0(\zeta)$, where ζ has order N for some monic N in $\mathcal{O}_0$. Its degree over F_0 is $n_c = \phi(N)$. For any monic Q in $\mathcal{O}_0$ prime to N there is a Frobenius automorphism $\sigma = \sigma_Q$ of F_c over F_0 such that $\sigma(\zeta) = \zeta^Q$. Write $\mathcal{O}_c = \mathcal{O}_0[\zeta]$; this is in fact the ring of integers of F_c, that is, the integral closure of $\mathcal{O}_0$ in F_c (see for example [56] p.82).

Lemma 8 *Suppose Q is monic irreducible in $\mathcal{O}_0$ prime to N and Δ is in $\mathcal{O}_c[X]$.*
(a) We have $\Delta^\sigma(X^Q) \equiv \Delta^\sigma(X^{\|Q\|})$ mod Q in $\mathcal{O}_c[X]$
(b) We have $\Delta(X)^{\|Q\|} \equiv \Delta^\sigma(X^{\|Q\|})$ mod Q in $\mathcal{O}_c[X]$.

Proof Part (a) for $\Delta = 1$ is Lemma 3 of [21] (p.219) - and then it holds even in $\mathbf{F}_p[X]$. It follows immediately for general Δ. For part (b) we need to note that every coefficient α of Δ has the form $\Phi(t, \zeta)$ for $\Phi(X, Y)$ in $\mathbf{F}_p[X, Y]$, so $\alpha^\sigma = \Phi(t, \zeta^Q)$ which is congruent to $\Phi(t, \zeta^{\|Q\|})$ mod Q in $\mathcal{O}_c$ by part (a) with $\Delta = \Phi(t, X)$ and then by substituting $X = \zeta$. This in turn is congruent to $\Phi(t^{\|Q\|}, \zeta^{\|Q\|}) = \alpha^{\|Q\|}$ mod Q in $\mathcal{O}_c$ as Q divides $t^{\|Q\|} - t$ in $\mathcal{O}_0$. □

The following is an analogue of an estimate of Amoroso-David [1] (p.157) restricted to a single variable. Again F is a finite extension of F_c. For a polynomial Φ in $F[X]$ or $F[X, Y]$ we write $|\Phi|_v$ for the maximum of $|f|_v$ as f runs over the coefficients (at first sight this could be confused with $|Q|_v$ below - but in fact it is the same notation because Q, even though a polynomial in t, is already in F).

Lemma 9 *Suppose Ω in $\mathcal{O}_c[X]$ of degree at most M vanishes at some ξ in F to order at least S. Then for any monic irreducible Q in $\mathcal{O}_0$ prime to N and any valuation v on F dividing Q we have*

$$|\Omega^\sigma(\xi^Q)|_v \le |\Omega^\sigma|_v |Q|_v^S \max\{1, |\xi^Q|_v\}^M.$$

Proof Let $\Delta(X) = \alpha_0 X^d + \cdots + \alpha_d$ be a minimal polynomial of ξ over $\mathcal{O}_c$. We use Strong Approximation but not quite as in [1] (whose argument for one variable uses already a special case for two variables). In fact this allows us to assume that $|\Delta|_w = 1$ for all w on F_c dividing Q. Namely for each such w there is $i = i(w)$ with $0 < |\alpha_i|_w = \mu_w = |\Delta|_w \le 1$. Now by the Theorem of [17] (p.67 - see also first paragraph of proof) there is β in F_c with $|\beta - \alpha_i^{-1}|_w < \mu_w^{-1}$ for each w dividing Q and $|\beta|_{w'} \le 1$ for all other w' on F_c not dividing ∞. The first of these imply $|\beta|_w = \mu_w^{-1}$ (in particular $\beta \ne 0$) and so $|\beta\Delta|_w = 1$ for each w dividing Q; and by the second $|\beta\Delta|_{w'} \le 1$ for all other w' not dividing ∞. So $\beta\alpha_0, \ldots, \beta\alpha_d$ are all in $\mathcal{O}_c$. Thus we just have to replace Δ by $\beta\Delta$.

In fact since $|x^\sigma|_w = |x|_{w(\sigma)}$ for any x in F_c and some other $w(\sigma)$, we see that $|\Delta^\sigma|_w = 1$ for all w dividing Q. In particular $|\Delta^\sigma|_v = 1$ for our v as well.

Next $\Omega = \Delta^S \hat{\Omega}$ for $\hat{\Omega}$ in $F_c[X]$ of degree at most $M - dS$ so by Gauss's Lemma we have $|\Omega^\sigma|_v = |\Delta^\sigma|_v^S |\hat{\Omega}^\sigma|_v = |\hat{\Omega}^\sigma|_v$.

Now $\Omega^\sigma(\xi^Q) = \Delta^\sigma(\xi^Q)^S \hat{\Omega}^\sigma(\xi^Q)$ and so

$$|\Omega^\sigma(\xi^Q)|_v \le |\Omega^\sigma|_v |\Delta^\sigma(\xi^Q)|_v^S \max\{1, |\xi^Q|_v\}^{M-dS}. \tag{29}$$

By Lemma 8 we see that $\Delta^\sigma(X^Q) - \Delta(X)^{\|Q\|} = Q\Xi(X)$ for some Ξ in $\mathcal{O}_c[X]$ of degree at most $d\|Q\|$. Putting $X = \xi$ we deduce $\Delta^\sigma(\xi^Q) = Q\Xi(\xi)$ and so

$$|\Delta^\sigma(\xi^Q)|_v \le |Q|_v \max\{1, |\xi|_v\}^{d\|Q\|} = |Q|_v \max\{1, |\xi^Q|_v\}^d.$$

The result now follows from this and (29). □

For the application of Lemma 4 we need information about $g(F_c)$ and $\mathfrak{m}(F_c, F_0)$.

Lemma 10 *(a) We have* $\mathfrak{m}(F_c, F_0) = n_c$.
(b) We have $g(F_c) \le 8000 n_c \log(n_c + 1)$.

Proof For (a) we have by definition $\mathfrak{m}(F_c, F_0) = n_c/[\mathbf{F} : \mathbf{F}_p]$, where $\mathbf{F}$ is the algebraic closure of $\mathbf{F}_p$ in F_c. But according to Gebhardt [24] (p.91) $\mathbf{F} = \mathbf{F}_p$, and the result follows.

For (b) we need a formula given by Keller [34] (see also [24] p.92). Namely

$$g(F_c) = 1 + \frac{1}{2} n_c \left(-2 + \frac{p-2}{p-1} + \sum_{i=1}^{r} \delta_i \left(\frac{e_i q_i - e_i - 1}{q_i - 1} \right) \right)$$

where $N = \prod_{i=1}^{r} N_i^{e_i}$ for distinct monic irreducible $N_1, \ldots, N_r$ of degrees $\delta_1, \ldots, \delta_r$ with $q_1 = p^{\delta_1}, \ldots, q_r = p^{\delta_r}$. Clearly this is at most $n_c \sum_{i=1}^{r} \delta_i e_i = \phi(N) n$ for the degree n of N. Now the result follows without difficulty from (11) above, using $n \le 8000 \log(1 + 2^n/(55n^4))$. □

7 Proof of Proposition 2

We may assume that $F_c(\xi) = F$ because making F smaller decreases d. We may also assume $D = [F : F_0] \ge 16$ from the remarks in [21] (p.218). We continue the notation $F_c = F_0(\zeta)$ for ζ of order N. Thus $D = d\phi(N)$.

We will suppose that

$$\hat{h}(\xi) \le C^{-11} d^{-1} \frac{(\log D)^{-3}}{(\log \log D)^{-2}}. \tag{30}$$

It then suffices to deduce a contradiction if the constant C (which cannot possibly be mistaken for our curve C) is sufficiently large as a function of p. Generally below c will denote various positive quantities depending only on p.

We use the Carlitz exponential function

$$e(z) = \sum_{i=0}^{\infty} \frac{z^{p^i}}{A(i)} \tag{31}$$

where $A(0) = 1$ and

$$A(i) = \prod_{j=1}^{i} (t^{p^i} - t^{p^{i-j}}) = \prod_{j=1}^{i} (t^{p^j} - t)^{p^{i-j}} \quad (i = 1, 2, \ldots) \tag{32}$$

can be taken as the a_i of Lemma 2(ii) of [21] (p.218). We pick any u with $e(u) = \xi$.

It is well-known that F_c is separable over F_0. Actually it follows at once from the identity

$$\frac{\partial}{\partial X} X^N = N \tag{33}$$

which we shall use later.

Let q be the inseparable degree of F over F_c. We fix any monic irreducible Q_0 in $\mathcal{O}_0 = \mathbf{F}_p[t]$ satisfying

$$\|Q_0\| \le C^4 \frac{d \log D}{q \log \log D} < p\|Q_0\| \tag{34}$$

and we define

$$L = \left[C^3 \frac{d \log D}{q \log\log D} \right], \quad T = \left[C^4 \frac{d \log D}{q \log\log D} \right]. \tag{35}$$

Now for a non-zero polynomial Φ in $F[X, Y]$ we define the height

$$h_{\mathbf{P}}(\Phi) = \frac{1}{D} \sum_v D_v \log |\Phi|_v$$

as in (27). Later we will have to be slightly careful about the projectivity. We also use the term hyperzero to remind the reader that we are considering Taylor series expansions rather than actually differentiating.

Lemma 11 *There is a non-zero polynomial $\tilde{\Phi}(X, Y)$ of degree at most L in each of X and Y, such that $\Phi = \tilde{\Phi}^q$ is in $\mathcal{O}_c[X, Y]$ with $h_{\mathbf{P}}(\Phi) \le cC^3 d \log D$, and such that the function*

$$\varphi(z) = \Phi(e(z), e(Q_0 z))$$

has a hyperzero of order at least qT at $z = u$.

Proof With

$$\tilde{\Phi}(X, Y) = \sum_{i=0}^{L} \sum_{j=0}^{L} a_{ij} X^i Y^j, \quad \tilde{\varphi}(z) = \tilde{\Phi}(e(z), e(Q_0 z))$$

and $z = w + u$ we have

$$\tilde{\varphi}(w + u) = \sum_{i=0}^{L} \sum_{j=0}^{L} a_{ij} (e(w) + \xi)^i (e(Q_0 w) + \xi^{Q_0})^j = \sum_{k=0}^{\infty} b_k w^k$$

and we start by solving $b_0 = b_1 = \cdots = b_{T-1} = 0$. These are $m = T$ linear equations in the $n = (L + 1)^2$ unknowns a_{ij} over F, to be solved first in $F_c^{1/q}$ as in Lemma 4. We note from (35) that

$$\frac{dm}{qn} \le cC^{-2} \frac{\log\log D}{\log D} \tag{36}$$

in readiness for an application of this lemma.

Here (31) gives

$$(e(w) + \xi)^i = \sum_{r=0}^{i} \binom{i}{r} \xi^{i-r} \sum_{i_1=0}^{\infty} \cdots \sum_{i_r=0}^{\infty} \frac{w^{p^{i_1} + \cdots + p^{i_r}}}{A(i_1) \cdots A(i_r)}. \tag{37}$$

Similarly

$$(e(Q_0 w) + \xi^{Q_0})^j = \sum_{s=0}^{j} \binom{j}{s} (\xi^{Q_0})^{j-s} \sum_{j_1=0}^{\infty} \cdots \sum_{j_s=0}^{\infty} \frac{(Q_0 w)^{p^{j_1} + \cdots + p^{j_s}}}{A(j_1) \cdots A(j_s)}, \tag{38}$$

so b_l involves an apparent denominator (forgetting ξ, ξ^{Q_0}) the lowest common multiple of all

$$A(i_1) \cdots A(i_r) A(j_1) \cdots A(j_s) \tag{39}$$

subject to

$$p^{i_1} + \cdots + p^{i_r} + p^{j_1} + \cdots + p^{j_s} = l. \tag{40}$$

It is clear from (32) that $A(i_1)\cdots A(i_r)$ for $p^{i_1} + \cdots + p^{i_r} \le p^I$ contains the factor $t^{p^j} - t$ at most $p^{i_1-j} + \cdots + p^{i_r-j} \le p^{I-j}$ times ($j = 1, \ldots, I$), so their lowest common multiple has degree at most $\sum_{j=1}^{I} p^{I-j} p^j = Ip^I$. Thus we get a contribution $cT \log T$ to the height of the rows of the matrix of linear equations in the a_{ij}; here

$$T \log T \le cC^5 \frac{d(\log D)^2}{q \log\log D}. \tag{41}$$

Similarly for the $A(j_1)\cdots A(j_s)$.

Another contribution comes from the ξ in (37). However $h(\xi) \le 3 + \hat{h}(\xi) \le 4$ by (16), so we get only cL from this; here

$$L \le C^3 \frac{d \log D}{q \log\log D}. \tag{42}$$

Similarly in (38) we have

$$h(\xi^{Q_0}) \le 3 + \hat{h}(\xi^{Q_0}) = 3 + ||Q_0||\hat{h}(\xi) \le 4 \tag{43}$$

as well.

Finally the ${Q_0}^{p^{j_1}+\cdots+p^{j_s}}$ in (38) contributes $cT \log \|Q_0\|$; here

$$T \log \|Q_0\| \le cC^5 \frac{d(\log D)^2}{q \log\log D}. \tag{44}$$

Taking into account (41), (42), (44) and not forgetting (36), we get using Lemma 4 (the extra $g/\mathfrak{m}$ there is at most $c \log D$ by Lemma 10) $\tilde{\Phi}$ of degree at most L with coefficients in $F_c^{1/q}$ of projective height at most $cC^3(d/q)\log D$, such that $\tilde{\varphi}$ has a zero of order at least T. The present lemma follows by raising to the power q, a dirty cheap trick which one might well think very wasteful. At first the a_{ij}^q are in $F_c = F_0(\zeta)$ but we can clear denominators to get them into $\mathcal{O}_c = \mathcal{O}_0[\zeta]$. □

We next define

$$T_1 = \left[C^2 d\right];$$

this is quite a bit smaller than qT because later estimates as in the proof of Lemma 11 will not be helped by a small Siegel exponent as in (36). We also fix n satisfying

$$p^n \le C^7 \frac{(\log D)^2}{\log\log D} < p^{n+1}. \tag{45}$$

Lemma 12 *For every monic irreducible Q in $\mathcal{O}_0$ prime to N of degree n and $\sigma = \sigma_Q$ the function $\varphi_\sigma(z) = \Phi^\sigma(e(z), e(Q_0 z))$ has a hyperzero of order at least T_1 at $z = Qu$.*

Proof We show by induction on k that there is a hyperzero of order at least k ($k = 0, 1, \ldots, T_1$).

The case $k = 0$ is empty, so we assume it holds up to $k-1$ for some k with $1 \le k \le T_1$. We can write $\varphi(z) = \Psi(e(z))$ for $\Psi(X)$ in $\mathcal{O}_c[X]$ of degree at most $M = qL + qL\|Q_0\|$ (incidentally it is the second term here that forces our non-integrality considerations). It follows from Lemma 11 that Ψ vanishes at ξ to order at least qT. So the kth hyperderivative

$\Omega = \Psi^{[k]}$ vanishes at ξ to order at least $T' = qT - k \geq qT/2$. Let V be the set of valuations v on F dividing Q such that ξ and so ξ^Q is v-integral. For these we have $|\Psi^\sigma|_v \leq |\Phi^\sigma|_v$ and clearly also $|\Omega^\sigma|_v \leq |\Psi^\sigma|_v$. Thus from Lemma 9 we deduce for $\eta = \Omega^\sigma(\xi^Q)$ the estimate

$$|\eta|_v \leq |\Phi^\sigma|_v |Q|_v^{T'} \leq |\Phi^\sigma|_v |Q|_v^{qT/2} \tag{46}$$

still for v in V.

For v on F not in V we argue analytically, using our induction on k. From $\varphi(z) = \Psi(e(z))$ and the fact that the z-derivative of $e(z)$ is 1, we deduce $\varphi^{[k]}(z) = \Psi^{[k]}(e(z)) + \cdots$, where the missing terms involve lower hyperderivatives of Ψ. Applying σ, putting $z = Qu$ and using our induction we see that η is the k-th Taylor coefficient of $\varphi_\sigma(z) = \Phi^\sigma(e(z), e(Q_0 z))$ at Qu. Estimating as we did in the proof of Lemma 11 with the analogues of (37) and (38) for Φ (instead of $\tilde{\Phi}$) we get

$$|\eta|_v \leq |\Phi^\sigma|_v \mathfrak{M}_v \mathfrak{N}_v \mathfrak{Q}_v \mathfrak{A}_v \tag{47}$$

with

$$\mathfrak{M}_v = \max\{1, |\xi^Q|_v\}^{qL}, \quad \mathfrak{N}_v = \max\{1, |\xi^{Q_0 Q}|_v\}^{qL}, \quad \mathfrak{Q}_v = \max\{1, |Q_0|_v\}^k$$

and

$$\mathfrak{A}_v = \max\left\{1, \max\left\{\left|\frac{1}{A}\right|_v\right\}\right\} \tag{48}$$

the inner maximum running over all A in (39) subject to (40).

We are trying to prove $\eta = 0$. If $\eta \neq 0$ then the sum S of $D^{-1} D_v \log |\eta|_v$ over all v on F should be zero by the Product Formula. We will deduce a contradiction. By (46) and (47)

$$S \leq h_{\mathbf{P}}(\Phi^\sigma) + S_0 + S_{\mathfrak{M}} + S_{\mathfrak{N}} + S_{\mathfrak{Q}} + S_{\mathfrak{A}}$$

where

$$S_0 = \frac{1}{D} \frac{qT}{2} \sum_{v \in V} D_v \log |Q|_v$$

and the last four terms correspond to sums with $\mathfrak{M}_v, \mathfrak{N}_v, \mathfrak{Q}_v, \mathfrak{A}_v$ over all v on F not in V. The first three of the latter are easily estimated. We find

$$S_{\mathfrak{M}} \leq qLh(\xi^Q) \leq 4qL \leq 4C^3 d \log D$$

as in (42) and (43); and even the same for $S_{\mathfrak{N}}$, as $\|Q_0\| \|Q\| \hat{h}(\xi) \leq 1/q \leq 1$. Also

$$S_{\mathfrak{Q}} \leq kh(Q_0) = k \frac{\log \|Q_0\|}{\log p} \leq cC^3 d \log D$$

instead of (44).

It would be tedious to estimate each $\mathfrak{A}_v$ explicitly. But $S_{\mathfrak{A}}$ is at most the height of the corresponding vector of 1 with $1/A$ in (48), so the calculation can be done in F_0, which we already did in (41), now getting $T_1 \log T_1 \leq cC^3 d \log D$ instead.

Thus using Lemma 11 to estimate $h_{\mathbf{P}}(\Phi^\sigma) = h_{\mathbf{P}}(\Phi)$ we get

$$0 = S \leq cC^3 d \log D + S_0. \tag{49}$$

Finally in S_0 we have $|Q|_v = |Q|_Q = \|Q\|^{-1/\log p}$ and V is the complement (among v dividing Q) of the set E in Lemma 7. Now $e_v \leq d$ there because each Q prime to N does not

ramify in the cyclotomic F_c (this step would fail for an arbitrary abelian extension). Thus

$$\sum_{v\in V} D_v = D - \sum_{v\in E} e_v \frac{D_v}{e_v} \geq D - dD\hat{h}(\xi)\frac{\log p}{\log \|Q\|} \geq D - D\frac{\log p}{\log \|Q\|} \geq \frac{D}{2}$$

using just $\hat{h}(\xi) \leq 1/d$ from (30). Therefore

$$S_0 \leq -\frac{qT}{4}\frac{\log \|Q\|}{\log p} \leq -c^{-1}C^4 d \log D$$

and (49) yields our desired contradiction. Thus indeed $\eta = 0$ and this completes the proof of Lemma 12. □

We can now finish the proof of Proposition 2 by showing that the polynomial Ψ defined above by $\Psi(e(z)) = \Phi(e(z), e(Q_0 z))$ has too many hyperzeroes for its degree.

First note that $\Psi \neq 0$ because $\Phi = \tilde{\Phi}^q$ and $e(Q_0 z)$ is a polynomial of degree $\|Q_0\| > L$ in $e(z)$ by (34) and (35). By Lemma 12 its conjugate Ψ^σ for $\sigma = \sigma_Q$ has a hyperzero of order at least T_1 at each ξ^Q. Let τ be any automorphism of $\overline{F}$ over F_0 extending σ^{-1}. Then

$$0 = (\Psi^\sigma(\xi^Q))^\tau = \Psi((\xi^Q)^\tau).$$

By Lemma 6(b) these $(\xi^Q)^\tau = (\xi^\tau)^Q$ are all of degree d over F_c if we exclude at most $2\log d \leq 2\log D$ exceptional Q. This is harmless because by Lemma 5 and (45) the total number of Q at our disposal is at least $M \geq c^{-1}C^6(\log D)^2/(\log\log D)^2$. We should also note that the number of monic irreducible polynomials in $\mathcal{O}_0$ not prime to $N = N_1^{e_1}\cdots N_r^{e_r}$ is g, certainly at most the degree of N which by (11) is at most $c\log\phi(N) \leq c\log||N|| \leq cC\log\log D$.

Now as Q ranges over all those remaining, and τ ranges over all extensions of $\sigma^{-1} = \sigma_Q^{-1}$ from F_c to F we claim that the $(\xi^Q)^\tau$ are all different. In fact an equation $(\xi^Q)^\tau = (\xi^{Q'})^{\tau'}$ would imply that $\xi^Q, \xi^{Q'}$ are conjugate over F_0. Thus by Lemma 6(a) (with $F_* = F_0$) we have $Q = Q'$ and so $\sigma = \sigma'$ for $\sigma' = \sigma_{Q'}$. So $(\xi^\tau)^Q = (\xi^{\tau'})^Q$. To cancel the Q here we note that $F_c(\xi) = F_c(\xi^Q)$ by Lemma 6(b), so that $\xi = R(\xi^Q)$ for R in $F_c(X)$. Now

$$\xi^\tau = R^{\sigma^{-1}}((\xi^Q)^\tau) = R^{\sigma^{-1}}((\xi^Q)^{\tau'}) = R^{\sigma'^{-1}}((\xi^Q)^{\tau'}) = \xi^{\tau'}.$$

As we assumed that ξ generates F over F_c and $\sigma = \sigma'$ we conclude $\tau = \tau'$. This settles the above claim.

Write Δ_Q for the monic minimal polynomial over F_c of ξ^Q. Then Ψ^σ in $F_c[X]$ is divisible by $\Delta_Q^{T_1}$, and so Ψ in $F_c[X]$ is divisible by $(\Delta_Q^{\sigma^{-1}})^{T_1}$. Now the hyperzeroes of $\Delta_Q^{\sigma^{-1}}$ are the $(\xi^Q)^\tau$ (each repeated q times) and so an equation $\Delta_Q^{\sigma^{-1}} = \Delta_{Q'}^{\sigma'^{-1}}$ leads to some $(\xi^Q)^\tau = (\xi^{Q'})^{\tau'}$ as above. Thus $Q = Q'$ and so $\sigma = \sigma'$ and these $\Delta_Q^{\sigma^{-1}}$ in $F_c[X]$ are all different (and irreducible over F_c). Once again by Lemma 6(b) they all have degree d and so we get in all

$$MdT_1 \geq c^{-1}C^8\frac{d^2(\log D)^2}{(\log\log D)^2}$$

hyperzeroes for Ψ. However Ψ has degree at most

$$qL + qL\|Q_0\| \leq cC^7\frac{d^2(\log D)^2}{q(\log\log D)^2}$$

(here we can ignore the q - taking it into account would lead to the tiny improvement mentioned in section 1) and the proof of Proposition 2 is complete.

8 Preliminaries for general K – geometry

To start with we need some purely geometric results; maybe the first three lemmas below are well-known, but as we are not over zero characteristic (where they also hold) we spell out some proof details.

Lemma 13 *Suppose affine B is irreducible of dimension at least 2. Then there are at most finitely many b in B such that the intersection of B with the generic hyperplane through b is reducible.*

Proof If we are in $\mathbf{A}^m$ then Bertini Irreducibility (see for example [35] p.212) gives non-zero (homogeneous) $\Phi(X_0, X_1, \ldots, X_m)$ such that the intersection of B with $\lambda_1 x_1 + \cdots + \lambda_m x_m = \lambda_0$ is irreducible provided $\Phi(\lambda_0, \lambda_1, \ldots, \lambda_m) \neq 0$.

Now the generic hyperplane through $b = (b_1, \ldots, b_m)$ is

$$\mu_1(x_1 - b_1) + \cdots + \mu_m(x_m - b_m) = 0, \tag{50}$$

i.e.

$$\mu_1 x_1 + \cdots + \mu_m x_m = \mu_1 b_1 + \cdots + \mu_m b_m.$$

If the intersection is reducible we must have

$$\Phi(\mu_1 b_1 + \cdots + \mu_m b_m, \mu_1, \ldots, \mu_m) = 0$$

identically in $\mu_1, \ldots, \mu_m$. This means that $b_1 X_1 + \cdots + b_m X_m - X_0$ divides Φ. But that can happen for at most finitely many $(b_1, \ldots, b_m) = b$. □

The example of a quadric cone B in $\mathbf{A}^3$ defined by $uv + vw + wu = 0$ through $(0, 0, 0)$ shows that exceptional b may exist: here the intersection with any hyperplane is a union of two lines.

Lemma 14 *Suppose affine B is irreducible of dimension at least 2. Then for any b in B the intersection of B with the generic hyperplane through b has codimension 1 in B.*

Proof If $l \geq 2$ is the dimension of B, then certainly $\dim(B \cap \Lambda_b) < l$ for Λ_b generic through b. Because if not, then B would be contained in Λ_b. But we can find a bunch of such Λ_b whose intersection is just b, and it would follow that $B = \{b\}$.

Let L_b be a linear polynomial defining Λ_b. It induces a map L_b from B to $\mathbf{A}$. This is dominant. For otherwise L_b would be a constant c on B. As $L_b(b) = 0$ we see that $c = 0$. But then B would be contained in Λ_b, which is excluded above.

We now use the Fibre Dimension Theorem in the version quoted in [11] (p.8); there we were over zero characteristic but it holds too over positive characteristic, as the reference to [18] shows. Part (a) on this L_b from B to $\mathbf{A}$ shows that $L_b^{-1}(0) = B \cap \Lambda_b$ has dimension at least $l - 1$, provided it is non-empty; which it is, because it contains b. □

Lemma 15 *Suppose affine B is irreducible of dimension at least 2. Then for any non-singular b in B not in the finite set of Lemma 13 the point b remains non-singular on the intersection of B with the generic hyperplane through b.*

Proof Let $\Phi_1, \ldots, \Phi_N$ be generators of the ideal of B in $\mathbf{A}^m$, so that the jacobian matrix with rows

$$\left(\frac{\partial \Phi_i}{\partial x_1}(b), \ldots, \frac{\partial \Phi_i}{\partial x_m}(b) \right) \quad (i = 1, \ldots, N)$$

has rank $m - l$, again for l the dimension of B. With L_b as in (50) defining the generic hyperplane, we adjoin the extra row

$$\left(\frac{\partial L_b}{\partial x_1}(b), \ldots, \frac{\partial L_b}{\partial x_m}(b)\right) = (\mu_1, \ldots, \mu_m)$$

and then the rank increases to $m - l + 1 = m - (l - 1)$, because $\mu_1, \ldots, \mu_m$ are generic. Now $\Phi_1, \ldots, \Phi_N$, L_b may not be generators of the ideal of $B \cap \Lambda_b$, but if we extend them to include such generators then the rank will still be at least $m - (l - 1)$. By Lemmas 13 and 14 this irreducible $B \cap \Lambda_b$ has dimension $l - 1$ and so indeed b is non-singular there (see for example [35] p.198). □

Next we record a result of well-known type, although we could not find it precisely in the literature. It is a version of Mumford's (3.25) and (3.26) in [50] (p. 53), which was applied in [9]. As we are over positive characteristic we cannot use his notion of "topologically unibranch". We write π for the projection from affine $\mathbf{A}^n \times \mathbf{A}^m$ to $\mathbf{A}^m$, and for the moment we work over an arbitrary algebraically closed field.

Lemma 16 *Let W be an algebraic set all of whose components have dimension $l \geq 1$ in $\mathbf{A}^n \times \mathbf{A}^m$ and let B be an irreducible variety of dimension l in $\mathbf{A}^m$, with $\pi(W)$ in B. Let T (when $l \geq 2$) be the finite set of Lemma 13 above for B. If b is non-singular on B and not in T (when $l \geq 2$) such that $W \cap \pi^{-1}(b)$ is finite, then its cardinality is at most that of $W \cap \pi^{-1}(\eta)$ for any η generic on B.*

Proof Here it is crucial, as in [9], that the excluded b hardly depend on W.

We will prove the result first under the assumption that the projections from each component of W to B are dominant.

We start with the case of curves $l = 1$ (for which we do not need T or the hypothesis that $W \cap \pi^{-1}(b)$ is finite). We proceed in three stages.

Suppose first that W is irreducible.

Then $W \cap \pi^{-1}(\eta)$ has exactly d elements, where d is the separable degree of π restricted to W. Suppose $W \cap \pi^{-1}(b)$ contains at least $e > d$ points. We can find a linear form λ in the coordinates of $\mathbf{A}^n$ taking e different values at these points. If q is the inseparable degree, then $\mu = \lambda^q$ satisfies an equation

$$\phi_0 \mu^d + \cdots + \phi_d = 0 \tag{51}$$

with $\phi_0, \ldots, \phi_d$ not all zero in the coordinate ring of B.

If we are lucky and $\phi_0, \ldots, \phi_d$ do not all vanish at b, the result is clear: (51) shows that there can be at most d values of μ, so at most d values of $\lambda = \mu^{1/q}$, a contradiction. Here we did not use non-singularity.

If $\phi_0, \ldots, \phi_d$ all vanish at b, we pick $\phi = \phi_i \neq 0$ with $\mathrm{ord}_b \phi_i$ minimal. Here non-singularity is implicit. Now the ϕ_j/ϕ are regular at b and so can be written as ψ_j/ψ for ψ_j, ψ in the coordinate ring of B with $\psi(b) \neq 0$. Multiplying (51) by ψ gives

$$\left(\psi_0 \mu^d + \cdots + \psi_d\right) \phi = 0$$

on W. But $\phi = 0$ on W would contradict dominance. As W is irreducible, it follows that

$$\psi_0 \mu^d + \cdots + \psi_d = 0$$

on W. And this takes us back to the first case, because $\psi_i = \psi$ does not vanish at b.

Next suppose, still for $l = 1$ under the dominance hypothesis, that $W = W_1 \cup \cdots \cup W_r$ for irreducible $W_1, \ldots, W_r$. Then

$$\#\left(W_i \cap \pi^{-1}(b)\right) \leq \#\left(W_i \cap \pi^{-1}(\eta)\right) \quad (i = 1, \ldots, r) \tag{52}$$

for generic η. Thus

$$\#\left(W \cap \pi^{-1}(b)\right) \leq \sum_{i=1}^{r} \#\left(W_i \cap \pi^{-1}(b)\right) \leq \sum_{i=1}^{r} \#\left(W_i \cap \pi^{-1}(\eta)\right).$$

Now any two $W_i \cap \pi^{-1}(\eta)$ are disjoint because any two W_i intersect in a finite set which cannot project to η. Thus the last sum above is indeed $\#\left(W \cap \pi^{-1}(\eta)\right)$.

This settles the case $l = 1$.

We now use induction on l (still assuming dominance). Assuming the result for some dimension $l - 1 \geq 1$, we will deduce it for dimension l.

As above, we do it in stages. We may assume $W \cap \pi^{-1}(b)$ is non-empty and $b \neq \eta$.

First irreducible W.

It is easy to see that a generic hyperplane constricted to pass through b and η is a generic hyperplane constricted only to pass through b. We call it Λ_b. Since b is not in T, the intersection $B \cap \Lambda_b$ is irreducible. We may denote also by Λ_b the product $\mathbf{A}^n \times \Lambda_b$ in $\mathbf{A}^n \times \mathbf{A}^m$ (it is defined by the same equation). Then π induces a map π_b from $W \cap \Lambda_b$ to $B \cap \Lambda_b$.

Now $\dim(W \cap \Lambda_b) \leq l - 1$ else W would be contained in Λ_b. By varying this hyperplane (still through b and η) we would deduce that W is contained in their intersection, which is ($\mathbf{A}^n$ times) the line through b and η. But then $\pi(W)$ would be contained in this line, contradicting dominance.

Let L_b be a linear polynomial defining Λ_b. It induces a map L_b from W to $\mathbf{A}$. This is dominant. For otherwise L_b would be a constant c on W. As $L_b(b) = 0$ and $W \cap \pi^{-1}(b)$ is non-empty we see that $c = 0$. But then W would be contained in Λ_b, which is excluded above.

Now the Fibre Dimension Theorem on this L_b from W to $\mathbf{A}$ shows that every (non-empty) component of $L_b^{-1}(0) = W \cap \Lambda_b$ has dimension at least $l - 1$. Note that $W \cap \Lambda_b$ is non-empty because $W \cap \pi^{-1}(b)$ is.

Thus every component of $W \cap \Lambda_b$ is irreducible of dimension $l - 1$.

Assume for the moment that there is only one component. We try to apply the induction hypothesis to the map π_b from $W \cap \Lambda_b$ to $B \cap \Lambda_b$ also irreducible of dimension $l - 1$. In fact π_b is dominant, otherwise $\pi_b(W \cap \Lambda_b)$ would be of dimension at most $l - 2$ containing b and then the Fibre Dimension Theorem would imply that $W \cap \pi^{-1}(b)$ would be of dimension at least 1, contradicting its assumed finiteness. We see by Lemma 15 that b remains non-singular on $B \cap \Lambda_b$. Thus by induction we have

$$\#\pi_b^{-1}(b) \leq \#\pi_b^{-1}(\eta). \tag{53}$$

But since Λ_b goes through both b and η, it is easy to see that $\pi_b^{-1}(b) = W \cap \pi^{-1}(b)$ and $\pi_b^{-1}(\eta) = W \cap \pi^{-1}(\eta)$.

A similar argument works if there are several different components $Z^{(1)}, \ldots, Z^{(s)}$ (all necessarily of dimension $l - 1$) of $W \cap \Lambda_b$. Then π_b induces projections $\pi^{(1)}, \ldots, \pi^{(s)}$ from $Z^{(1)}, \ldots, Z^{(s)}$ to $B \cap \Lambda_b$. If one of these is not dominant, then again $W \cap \pi^{-1}(b)$ would be infinite. Thus by induction again, $\#(\pi^{(j)})^{-1}(b) \leq \#(\pi^{(j)})^{-1}(\eta)$ for $j = 1, \ldots, s$. Therefore

$$\#\pi_b^{-1}(b) = \sum_{j=1}^{s} \#(\pi^{(j)})^{-1}(b) \leq \sum_{j=1}^{s} \#(\pi^{(j)})^{-1}(\eta).$$

Now any two $(\pi^{(j)})^{-1}(\eta)$ are disjoint because any two $Z^{(j)}$ intersect in something of dimension at most $l-2$ which cannot project to η. Thus the last sum above is just $\#(W \cap \pi_b^{-1}(\eta))$; and we have recovered (53).

Next the reducible case $W = W_1 \cup \cdots \cup W_r$ (still under dominance) follows in a similar way. Namely

$$\#\left(W_i \cap \pi^{-1}(b)\right) \leq \#\left(W_i \cap \pi^{-1}(\eta)\right) \quad (i = 1, \ldots, r). \tag{54}$$

Thus

$$\#\left(W \cap \pi^{-1}(b)\right) \leq \sum_{i=1}^{r} \#\left(W_i \cap \pi^{-1}(b)\right) \leq \sum_{i=1}^{r} \#\left(W_i \cap \pi^{-1}(\eta)\right)$$

and as above any two $W_i \cap \pi^{-1}(\eta)$ are disjoint because any two W_i intersect in something of dimension at most $l-1$ which cannot project to η. Thus the last sum above is indeed $\#\left(W \cap \pi^{-1}(\eta)\right)$.

This settles the lemma under our assumption that the projections from each component of W to B are dominant.

Finally suppose the latter fails for some component W_0 of W. Then b cannot be in $\pi(W_0)$, otherwise the Fibre Dimension Theorem would imply that $W_0 \cap \pi^{-1}(b)$ would have dimension at least $\dim W_0 - \dim \pi(W_0) \geq 1$, contradicting finiteness. Thus $W_0 \cap \pi^{-1}(b)$ is empty; and of course so is $W_0 \cap \pi^{-1}(\eta)$. So these are not seen in the intersections with W. □

Regarding the excluded b, the example of B defined by $v^2 = u^2(u+1)$ in $\mathbf{A}^2$ and W defined by

$$v^2 = u^2(u+1), \quad w^2 = u+1, \quad wu = v$$

in $\mathbf{A}^3$, where $b = (0,0)$ has two inverse images $(0,0,\pm 1)$, shows also that Lemma 16 can be false if b is singular.

The result can become false also if components of dimension less than l are allowed. For example if $l = 1$ and W is the single point (a,b) in $\mathbf{A}^n \times \mathbf{A}^m$ with b non-singular on the curve B, then

$$\#(W \cap \pi^{-1}(b)) = 1 > 0 = \#(W \cap \pi^{-1}(\eta)).$$

9 Preliminaries for general *K* - Carlitz

Now we return to the Carlitz world. The next result concerns the action of a Carlitz polynomial $A(\mathcal{C})$ on a special sort of Laurent polynomial in a variable u. Write

$$A(T) = a_0 + a_1 T + \cdots + a_d T^d$$

with coefficients in $\mathbf{F}_p$. For a positive integer n denote by $S_m^{(n)}$ the mth elementary symmetric polynomial in $-t^p, -t^{p^2}, \ldots, -t^{p^n}$.

For A as above and an integer $n \leq d$, now including $n = 0$, and variables $\lambda_0, \ldots, \lambda_n$, we define $X_n^{(n)}, \ldots, X_{d+n}^{(n)}$ recursively as follows. At level 0 we have

$$X_k^{(0)} = a_k \lambda_0 \quad (k = 0, \ldots, d). \tag{55}$$

Then at level n from level $n-1 \geq 0$ we have first

$$X_{d+n}^{(n)} = a_d \lambda_n, \tag{56}$$

then

$$X_{d+n-1}^{(n)} = \left(X_{d+n-1}^{(n-1)}\right)^p + \left(S_0^{(n)} a_{d-1} + S_1^{(n)} a_d\right) \lambda_n,$$
$$X_{d+n-2}^{(n)} = \left(X_{d+n-2}^{(n-1)}\right)^p + \left(S_0^{(n)} a_{d-2} + S_1^{(n)} a_{d-1} + S_2^{(n)} a_d\right) \lambda_n,$$

and so on, down to

$$X_{d+1}^{(n)} = (X_{d+1}^{(n-1)})^p + (S_0^{(n)} a_{d-n+1} + S_1^{(n)} a_{d-n+2} + \cdots + S_{n-1}^{(n)} a_d)\lambda_n,$$

which define the $X_k^{(n)}$ for $k > d$. And finally for $k = d, d-1, \ldots, n$ we define

$$X_k^{(n)} = \left(X_k^{(n-1)}\right)^p + \left(S_0^{(n)} a_{k-n} + S_1^{(n)} a_{k-n+1} + \cdots + S_n^{(n)} a_k\right) \lambda_n. \tag{57}$$

By induction on n we verify the following

Remark 1 For $k = n, \ldots, d+n$ and $l = \min\{k, d\} \geq n$ we can write $X_k^{(n)}$ as a linear form in $a_l, \ldots, a_{l-n}$ whose coefficients are polynomials over $\mathbf{F}_p$ in $t, \lambda_0, \ldots, \lambda_n$ of degree at most p^{n+1} in each variable.

It will be crucial that for each n, k the number of a_i appearing as well as the polynomial degree are bounded only in terms of n. For example

$$X_k^{(1)} = (a_k \lambda_0)^p + \left(a_{k-1} - t^p a_k\right) \lambda_1 = \left(\lambda_0^p - t^p \lambda_1\right) a_k + \lambda_1 a_{k-1} \quad (k = 1, \ldots, d),$$
$$X_k^{(2)} = \left(\lambda_0^{p^2} - t^{p^2} \lambda_1^p + t^{p+p^2} \lambda_2\right) a_k + \left(\lambda_1^p - (t + t^p)\lambda_2\right) a_{k-1} + \lambda_2 a_{k-2} \quad (k = 2, \ldots, d).$$

We need the "Carnomial coefficients" T_{ij} in $\mathcal{O}_0 = \mathbf{F}_p[t]$ defined by

$$\mathcal{C}^i Z = \sum_{j=0}^{i} T_{ij} Z^{p^j}.$$

Lemma 17 *We have*

$$A(\mathcal{C}) \left(\frac{\lambda_0}{u} + \frac{\lambda_1}{u^p} + \cdots + \frac{\lambda_n}{u^{p^n}} \right) = \sum_{j=0}^{d+n} \frac{P_j^{(n)}}{u^{p^j}} \tag{58}$$

for

$$P_j^{(n)} = \sum_{k=j}^{d+n} T_{kj} (X_k^{(n)})^{p^{j-n}} \quad (j = n, \ldots, d+n). \tag{59}$$

Proof Note that we do not specify $P_j^{(n)}$ for $j = 0, \ldots, n-1$. It is not difficult to do so but we found it is not useful for applications.

From $\mathcal{C}^{i+1} Z = \mathcal{C}^i (\mathcal{C} Z)$ we derive a simple recurrence

$$T_{i\ j-1} = T_{i+1\ j} - t^{p^j} T_{ij},$$

where the T_{ij} are considered zero if $0 \leq j \leq i$ does not hold.

Then iterating n times gives

$$T_{i\ j-1} = \left(S_0^{(n)}\right)^{p^{j-1}} T_{i+n\ j+n-1} + \left(S_1^{(n)}\right)^{p^{j-1}} T_{i+n-1\ j+n-1} + \cdots + \left(S_n^{(n)}\right)^{p^{j-1}} T_{i\ j+n-1}. \tag{60}$$

We use of course induction on n to prove (58).

For $n = 0$ the left-hand side is

$$\sum_{k=0}^{d} a_k \sum_{j=0}^{k} T_{kj} \frac{\lambda_0^{p^j}}{u^{p^j}} = \sum_{j=0}^{d} \frac{1}{u^{p^j}} \sum_{k=j}^{d} a_k T_{kj} \lambda_0^{p^j} = \sum_{j=0}^{d} \frac{1}{u^{p^j}} \sum_{k=j}^{d} T_{kj} (a_k \lambda_0)^{p^j}$$

which is the right-hand side thanks to (55).

Now we assume the thing done for $n - 1 \geq 0$ and we deduce it for $n \geq 1$.

Splitting λ_n / u^{p^n} off the left-hand side of (58), and using the case $n = 0$ with u^{p^n} in place of u we find

$$\sum_{j=0}^{d+n-1} \frac{P_j}{u^{p^j}} + \sum_{j=0}^{d} \frac{Q_j}{u^{p^{j+n}}}$$

where

$$P_j = P_j^{(n-1)} = \sum_{k=j}^{d+n-1} T_{kj} \left(X_k^{(n-1)}\right)^{p^{j-n+1}}, \quad (j = n-1, \ldots, d+n-1)$$

and

$$Q_j = \sum_{k=j}^{d} T_{kj} (a_k \lambda_n)^{p^j} \quad (j = 0, \ldots, d)$$

(with λ_n in place of λ_0). For $j = d + n$ in (58) we get at once

$$P_{d+n}^{(n)} = Q_d = (a_d \lambda_n)^{p^d}$$

as required in (59), thanks to (56).

Also

$$P_j^{(n)} = P_j + Q_{j-n} \quad (j = n, \ldots, d+n-1)$$

which is

$$\sum_{k=j}^{d+n-1} T_{kj} \left(X_k^{(n-1)}\right)^{p^{j-n+1}} + \sum_{k=j-n}^{d} T_{k\ j-n} (a_k \lambda_n)^{p^{j-n}}.$$

We use (60) to see that the second sum is

$$\sum_{k=j-n}^{d} ((S_0^{(n)})^{p^{j-n}} T_{k+n\ j} + (S_1^{(n)})^{p^{j-n}} T_{k+n-1\ j} + \cdots + (S_n^{(n)})^{p^{j-n}} T_{kj}) (a_k \lambda_n)^{p^{j-n}}.$$

Thus

$$P_j^{(n)} = U + U_0 + U_1 + \cdots + U_n \tag{61}$$

with (now adjusting k)

$$U = \sum_{k=j}^{d+n-1} T_{kj}(X_k^{(n-1)})^{p^{j-n+1}}$$

and

$$U_m = (S_m^{(n)})^{p^{j-n}} \sum_{k=j-m}^{d+n-m} T_{kj}(a_{k-n+m}\lambda_n)^{p^{j-n}}.$$

Here in U_m we can restrict the sum from $k = j$.

We already checked $P_{d+n}^{(n)}$ in (59) using (56). Now for each $j = n, \ldots, d+n-1$ we pick out the coefficient of the various T_{kj} in (61).

The biggest k is $d+n$ and now we see T_{kj} only in U_0, with coefficient $(S_0^{(n)})^{p^{j-n}}(a_d\lambda_n)^{p^{j-n}}$, which fits with (59) again thanks to (56).

Next for $k = d+n-1$ we see T_{kj} in U as well as in U_0, U_1. This also fits with (59) thanks to the displayed formula just after (56).

We carry on with $k = d+n-2$ down to $k = d+1$ and these also fit, thanks to the later formulae preceding (57).

Then we go further with $k = d, d-1, \ldots, n$ which appear in all of $U, U_0, U_1, \ldots, U_n$ and these fit with (59) because of (57). This completes the proof. □

As mentioned, the $P_j^{(n)}$ for $j = 0, \ldots, n-1$ are not useful for applications. For example one finds

$$P_0^{(n)} = A(t)\lambda_0 = (a_0 + a_1 t + \cdots + a_d t^d)\lambda_0$$

where the number of a_i appearing is not bounded in terms of n as in Remark 1.

Remark 2 Because $T_{kk} = 1$ the system (59) has a triangular nature; for example the equations

$$P_{d+n}^{(n)} = \cdots = P_e^{(n)} = 0$$

for some $e \geq n$ are equivalent to the equations

$$X_{d+n}^{(n)} = \cdots = X_e^{(n)} = 0.$$

It is not difficult to see from Remark 1 that (we will be more precise later) these are equations for $\lambda_0, \ldots, \lambda_n$ again essentially independent of A (as in the examples given just after that Remark).

10 More on curves

From now on $K = \overline{\mathbf{F}_p(t, s_1, \ldots, s_l)}$ for some $l \geq 1$ variables $s_1, \ldots, s_l$ algebraically independent over $F_0 = \mathbf{F}_p(t)$. We define a height h_s on K by regarding it as the closure of $\overline{F_0}(s_1, \ldots, s_l)$. See for example [22] (p.1053) - thus $h_s(s_1) = \cdots = h_s(s_l) = 1$ but $h_s(t) = 0$.

The next result is in the style of Proposition 1.

Lemma 18 *Let C be an irreducible curve in $\mathbf{G}_{\mathrm{a}}^n$ defined over K. Assume for any non-zero $(\rho_1, \ldots, \rho_n)$ in $\mathcal{R}^n$ that the form $\rho_1 x_1 + \cdots + \rho_n x_n$ is not identically constant on C. Then there is $\mathfrak{B}$ such that*

$$h_s(\xi_1) + \cdots + h_s(\xi_n) \leq \mathfrak{B}$$

for all $(\xi_1, \dots, \xi_n)$ *on* $C(K)$ *for which there exists non-zero* $(\alpha_1, \dots, \alpha_n)$ *in* $\mathcal{R}^n$ *and* λ *in* $\overline{F_0}$ *with*

$$\alpha_1\xi_1 + \cdots + \alpha_n\xi_n = \lambda. \tag{62}$$

Proof Our h_s is not a canonical height in the sense of Denis but it does have the same property that $h_s(\xi^P) = \|P\|h_s(\xi)$ as in (14) above. For example $h_s(\xi^p + t\xi) = ph_s(\xi)$ because now t appears as a constant. To see that we can go through the Nullstellensatz argument observing that $|t| = 1$. Or just directly using $\max\{1, |\xi^p + t\xi|\} = \max\{1, |\xi|\}^p$ (that would work better for general P). And of course $h_s(\xi + \eta) \le h_s(\xi) + h_s(\eta)$ as in (15) above.

Now we can follow the proof of Proposition 1 above. Lemma 1 above goes through with h_s instead of $\hat{h}$ because of the remarks above; the extra λ in (62) makes no trouble because it leads to an extra term $Q(\mathcal{C})\lambda$ on the left-hand side of (21), and this has zero height. The subsequent proof also goes through with h_s also instead of h (actually with $c'{=}c''{=}nc$). □

And now an analogue of Proposition 2 of [9] (p. 452).

Lemma 19 *Let* C *be an irreducible curve in* $\mathbf{G}_{\mathrm{a}}^3$ *defined over* K *but not over* $\overline{F_0}$. *Assume for any non-zero* (ρ_1, ρ_2, ρ_3) *in* $\mathcal{R}^3$ *that the form* $\rho_1x_1 + \rho_2x_2 + \rho_3x_3$ *is not identically constant on* C. *Then given any* D *there are at most finitely many* (ξ_1, ξ_2, ξ_3) *on* $C(K)$ *for which there exists non-zero* $(\alpha_1, \alpha_2, \alpha_3)$ *in* $\mathcal{R}^3$ *and* λ *in* $\overline{F_0}$ *with*

$$\alpha_1\xi_1 + \alpha_2\xi_2 + \alpha_3\xi_3 = \lambda \tag{63}$$

and

$$[\overline{F_0}(\xi_1, \xi_2, \xi_3, s_1, \dots, s_l) : \overline{F_0}(s_1, \dots, s_l)] \le D.$$

Proof Assume first that the group of the $(\alpha_1, \alpha_2, \alpha_3)$ for which λ exists in (63) has rank 1. Then much as in (23) and (24) above we can find ξ, ξ' in K with

$$\overline{F_0}(\xi_1, \xi_2, \xi_3) = \overline{F_0}(\xi, \xi') \tag{64}$$

and

$$\xi_1 = \sigma_1\xi + \sigma_1'\xi' + \lambda_1, \quad \xi_2 = \sigma_2\xi + \sigma_2'\xi' + \lambda_2, \quad \xi_3 = \sigma_3\xi + \sigma_3'\xi' + \lambda_3 \tag{65}$$

for $\sigma_1, \sigma_1', \sigma_2, \sigma_2', \sigma_3, \sigma_3'$ in $\mathcal{R}$ and $\lambda_1, \lambda_2, \lambda_3$ in $\overline{F_0}$. Now ξ, ξ' are linearly independent, as are the rows $(\sigma_1, \sigma_2, \sigma_3)$, $(\sigma_1', \sigma_2', \sigma_3')$. With $\sigma_i = S_i(\mathcal{C})$, $\sigma_i' = S_i'(\mathcal{C})$ the rows

$$\mathbf{s} = (S_1, S_2, S_3), \quad \mathbf{s}' = (S_1', S_2', S_3')$$

in $\mathcal{O}_0^3$ (recall $\mathcal{O}_0 = \mathbf{F}_p[t]$ here) satisfy $M\mathbf{s}^t = M\mathbf{s}'^t = 0$ for some 1×3 matrix whose entries are the minors

$$T_1 = S_2S_3' - S_3S_2', \quad T_2 = S_3S_1' - S_1S_3', \quad T_3 = S_1S_2' - S_2S_1'.$$

We shall show (in a fairly familar way) that we may assume that

$$d = \max\{\deg S_1, \deg S_2, \deg S_3\}, \quad d' = \max\{\deg S_1', \deg S_2', \deg S_3'\}$$

and

$$e = \max\{\deg T_1, \deg T_2, \deg T_3\}$$

satisfy

$$d + d' \le e. \tag{66}$$

Namely, by [57] Corollary 2 (p.148) over F_0 there are independent $\tilde{\mathbf{s}}, \tilde{\mathbf{s}}'$ in F_0^3 with $M\tilde{\mathbf{s}}^t = M\tilde{\mathbf{s}}'^t = 0$ and

$$h_{\mathbf{P}}(\tilde{\mathbf{s}}) + h_{\mathbf{P}}(\tilde{\mathbf{s}}') \le h_{\mathbf{P}}(M). \tag{67}$$

We can normalize $\tilde{\mathbf{s}}, \tilde{\mathbf{s}}'$ to lie in $\mathcal{O}_0^3$ and be primitive. Then

$$h_{\mathbf{P}}(\tilde{\mathbf{s}}) = \max\{\deg \tilde{S}_1, \deg \tilde{S}_2, \deg \tilde{S}_3\}, \quad h_{\mathbf{P}}(\tilde{\mathbf{s}}') = \max\{\deg \tilde{S}'_1, \deg \tilde{S}'_2, \deg \tilde{S}'_3\}$$

for the corresponding polynomials in $\mathcal{O}_0$. Also

$$h_{\mathbf{P}}(M) \le \max\{\deg \tilde{T}_1, \deg \tilde{T}_2, \deg \tilde{T}_3\}$$

for the corresponding minors. Thus (66) indeed holds for the new polynomials. And the old $(\sigma_1, \sigma_2, \sigma_3)$, $(\sigma'_1, \sigma'_2, \sigma'_3)$ are linear combinations of the new rows $(\tilde{\sigma}_1, \tilde{\sigma}_2, \tilde{\sigma}_3)$, $(\tilde{\sigma}'_1, \tilde{\sigma}'_2, \tilde{\sigma}'_3)$ with coefficients in F_0. This leads to (65) for new $\tilde{\xi}, \tilde{\xi}'$ with these new rows.

So it suffices to rename new as old, thus achieving (65).

We may suppose $e = \deg T_3$ above. Eliminating ξ' between the first two equations of (65) gives

$$\sigma'_2\xi_1 - \sigma'_1\xi_2 = (\sigma'_2\sigma_1 - \sigma'_1\sigma_2)\xi + \mu = T_3(\mathcal{C})\xi + \mu$$

for μ in $\overline{F_0}$. Taking heights h_s and estimating the left-hand side by Lemma 18, we get the inequality $p^{d'}\mathfrak{B} \ge p^e h_s(\xi)$.

Now ξ is not in $\overline{F_0}$ by (65) and our rank 1 assumption. Thus Lemma 2.1 (p.1053) of [22] shows that

$$h_s(\xi) \ge \left[\overline{F_0}(s_1, \dots, s_l)(\xi) : \overline{F_0}(s_1, \dots, s_l)\right]^{-1}. \tag{68}$$

By (64) the degree here is at most D. It follows that $p^{d'}\mathfrak{B}D \ge p^e$.

A similar argument eliminating ξ in (65) leads to $p^d\mathfrak{B}D \ge p^e$. Multiplying the two inequalities and using (66) we find $p^e \le (\mathfrak{B}D)^2$.

By grassmannian theory this means that there are at most finitely many possibilities for the $\mathcal{O}_0$-module generated by $\mathbf{s}, \mathbf{s}'$ in $\mathcal{O}_0^3$. Thus we can regard $\mathbf{s}, \mathbf{s}'$ as fixed.

Now the original relation (63) implies

$$\alpha_1\sigma_1 + \alpha_2\sigma_2 + \alpha_3\sigma_3 = \alpha_1\sigma'_1 + \alpha_2\sigma'_2 + \alpha_3\sigma'_3 = 0$$

of which we need only one. We can apply $\mathrm{GL}_3(\mathcal{R})$ to assume it is $\alpha_3 = 0$. This reduces the problem to $\mathbf{G}_{\mathrm{a}}^2$.

Then a similar argument brings us down to $\mathbf{G}_{\mathrm{a}}$. But the lemma is then empty, because now $C = \mathbf{G}_{\mathrm{a}}$ is defined over $\overline{F_0}$.

Right at the start of the proof we made an assumption about rank 1. But if the rank is bigger then things only get easier.

For example if the rank of the $(\alpha_1, \alpha_2, \alpha_3)$ in (63) is 2, then we can argue as in (65) without ξ' and there is no longer any need for minors.

And if the rank is 3, then ξ_1, ξ_2, ξ_3 lie in $\overline{F_0}$. But because C is not defined over $\overline{F_0}$, this implies that $C(\overline{F_0})$ is at most finite anyway. This completes the proof (and on the way we proved the analogue in $\mathbf{G}_{\mathrm{a}}^2$). □

11 Completions and specializations

During this section we assume that C satisfies the conditions of Lemma 19. Of course these are more restrictive than the conditions of the Conjecture for $n = 3$, but we will see that this will be no problem. As in [9] we regard the field of definition of C as the function field $\overline{F_0}(s_1, \ldots, s_m)$ of an irreducible variety B, say of dimension $l \geq 1$, in affine $\mathbf{A}^m$ defined over $\overline{F_0}$. We assume as in the previous section that $s_1, \ldots, s_l$ are algebraically independent over F_0. As in [9] we complete C to $\hat{C}$ in projective $\mathbf{P}_3$ and then take a non-singular model $\tilde{C}$.

In [9] (p.463) we informally defined a variety C_B in $\mathbf{A}^3 \times \mathbf{A}^m$; here this amounts to writing the equations of C in $\mathbf{A}^3$ with coefficients in $\overline{F_0}[s_1, \ldots, s_m]$ and adjoining the equations of B. Of course one should more formally define it as the $\overline{F_0}$-Zariski closure of a point (P, η), where η is generic on B and P is generic on C over $\overline{F_0}(\eta)$. This makes it clear that C_B is irreducible of dimension $l + 1$. The natural projection π from $\mathbf{A}^3 \times \mathbf{A}^m$ to $\mathbf{A}^m$ then takes C_B to B. There is also a natural projection γ from $\mathbf{A}^3 \times \mathbf{A}^m$ to $\mathbf{A}^3$.

Then for a point b of B we define the specialization C_b by

$$C_b = \gamma(C_B \cap \pi^{-1}(b))$$

and similarly $\hat{C}_b, \tilde{C}_b$. As in [9] we can assume that these specializations retain enough properties of $C, \hat{C}, \tilde{C}$ that we shall need, at least when b is restricted to a non-empty open subset B_0 of B. And also for $b = \eta$ a generic point; here we shall identify η with $(s_1, \ldots, s_m)$.

For b in B_0 we can regard x_1, x_2, x_3 as functions on C_b but for simplicity we omit any subscript b that may possibly be more precise. Equally we omit the subscript for

$$f = \alpha_1 x_1 + \alpha_2 x_2 + \alpha_3 x_3$$

or also $f - \lambda$ with a constant function λ. But we put it back in the notation $\deg_b(f - \lambda)$ for the degree of $f - \lambda$ on C_b (unless $f - \lambda$ is identically zero on C_b, a possibility that we shall soon essentially discount). Note that by our condition on C, this $f - \lambda$ is certainly non-zero on C_η as long as $\alpha_1, \alpha_2, \alpha_3$ are not all zero.

The next result replaces the simple argument in the last paragraph of [9] (p.463); here we lack any multiplicative structure.

Lemma 20 *There is a non-empty open subset B_{00} of B_0 with the following property. For any λ in $\overline{F_0}$, any b in B_{00} and any $\alpha_1, \alpha_2, \alpha_3$ not all zero the function $f - \lambda$ is not identically zero on C_b and*

$$\deg_b(f - \lambda) \geq \deg_\eta(f - \lambda) - \deg C. \tag{69}$$

Proof We shall first prove the lemma under the assumption that $f - \lambda \neq 0$ on C_b. Then at the end of the proof we shall show that in fact this follows almost automatically.

Counting by poles we have now

$$\deg_b(f - \lambda) = \sum_{\tilde{P} \in \tilde{C}_b} \max\{0, -\mathrm{ord}_{\tilde{P}}(f - \lambda)\}.$$

This holds also for $b = \eta$, but in fact we shall identify C_η with C over $\overline{F_0}[s_1, \ldots, s_m]$. In this generic case we can restrict to $\tilde{P}$ in $\tilde{C}$ with

$$R = R_{\tilde{P}} = \max\{-\mathrm{ord}_{\tilde{P}} x_1, -\mathrm{ord}_{\tilde{P}} x_2, -\mathrm{ord}_{\tilde{P}} x_3\} > 0$$

because f is a polynomial in x_1, x_2, x_3.

With a uniformizer $u = u_{\tilde{P}}$ we have a local expansion

$$x_i = \underline{x_i} + \overline{x_i}$$

where $\underline{x_i}$ involves only non-negative exponents and $\overline{x_i}$ only (finitely many) negative exponents.

We are going to split the negative exponents into subsets of various sets $\{r, rp, rp^2, \ldots\}$ with r prime to p. These sets are disjoint. Of course we will see rp^m only with

$$1 \le r \le R, \quad 0 \le m \le n_r = \left[\frac{\log R/r}{\log p}\right] \le \frac{\log R}{\log p}.$$

Thus we can write

$$\overline{x_i} = \sum_r \sum_{m=0}^{n_r} \frac{\lambda_{irm}}{u^{rp^m}}.$$

At first the λ_{irm} are in the algebraic closure of $\overline{F_0}(s_1, \ldots, s_m)$ but by taking a finite cover of B and introducing more variables we can suppose that they lie in $\overline{F_0}[s_1, \ldots, s_m]$ itself (this is just for painless specialization).

We have corresponding $f = \underline{f} + \overline{f}$ with

$$\overline{f} = \alpha_1 \overline{x_1} + \alpha_2 \overline{x_2} + \alpha_3 \overline{x_3}$$

which is

$$\sum_r \left(\alpha_1 \left(\sum_{m=0}^{n_r} \frac{\lambda_{1rm}}{u^{rp^m}} \right) + \alpha_2 \left(\sum_{m=0}^{n_r} \frac{\lambda_{2rm}}{u^{rp^m}} \right) + \alpha_3 \left(\sum_{m=0}^{n_r} \frac{\lambda_{3rm}}{u^{rp^m}} \right) \right).$$

We calculate these using Lemma 17 with u there replaced by the various u^r and a suitably large d. We find

$$\overline{f} = \sum_r \sum_{j=0}^{d+n_r} \frac{1}{u^{rp^j}} (P_{1rj} + P_{2rj} + P_{3rj}) \tag{70}$$

with

$$P_{1rj} = \sum_{k=j}^{d+n_r} T_{kj} X_{1rk}^{p^{j-n_r}}, \quad P_{2rj} = \sum_{k=j}^{d+n_r} T_{kj} X_{2rk}^{p^{j-n_r}},$$

$$P_{3rj} = \sum_{k=j}^{d+n_r} T_{kj} X_{3rk}^{p^{j-n_r}} \quad (j = n_r, \ldots, d + n_r)$$

and the $X_{1rk}, X_{2rk}, X_{3rk}$ are defined as in (56)-(57) with $n = n_r$ on taking the A_1, A_2, A_3 in F_0 with $\alpha_1 = A_1(\mathcal{C}), \alpha_2 = A_2(\mathcal{C}), \alpha_3 = A_3(\mathcal{C})$. Thus we need d at least $\deg A_1, \deg A_2, \deg A_3$ and $p^d \ge R$. We get

$$\overline{f} = \sum_{s=0}^{S} \frac{P_s}{u^s}$$

for say $S = p^d R^2$.

Up to now we are in the generic situation but soon we shall specialize.

Suppose first the $\omega_{\tilde{P}} = -\mathrm{ord}_{\tilde{P}}(f-\lambda) > 0$. Then it is $-\mathrm{ord}_{\tilde{P}}\overline{f}$ and so has the form $\tilde{s} = \tilde{r}p^{e-1}$ for some unique $\tilde{r}$ and $e \geq 1$. This means of course

$$P_s = 0 \quad (s > \tilde{s}) \tag{71}$$

so in particular

$$P_s = 0 \quad (s = \tilde{r}p^e, \ldots, \tilde{r}p^{d+\tilde{n}}) \tag{72}$$

for $\tilde{n} = n_{\tilde{r}}$, but

$$P_{\tilde{s}} \neq 0. \tag{73}$$

We aim to specialize these to b in B_0 with corresponding $P_s(b)$. Of course (71) and (72) are trivially done, and we must pay attention only to $P_{\tilde{s}}(b)$.

By (70) we have for $s = \tilde{r}p^j$

$$P_s(b) = P_{1\tilde{r}j}(b) + P_{2\tilde{r}j}(b) + P_{3\tilde{r}j}(b) = \sum_{k=j}^{d+\tilde{n}} T_{kj} X_k^{p^{j-\tilde{n}}}(b) \quad (j = \tilde{n}, \ldots, d+\tilde{n})$$

for

$$X_k(b) = X_{1\tilde{r}k} + X_{2\tilde{r}k} + X_{3\tilde{r}k}. \tag{74}$$

Thus at η, the equations (72),(73) together with triangularity as in Remark 2 imply

$$X_k(\eta) = 0 \quad (k = e, \ldots, d+\tilde{n}) \tag{75}$$

provided $e - 1 \geq \tilde{n}$, but

$$X_{e-1}(\eta) \neq 0. \tag{76}$$

Thus for any b in B_0 not a zero of X_{e-1} we can specialize (75),(76); and doing the thing backwards leads to the required specializations of (72),(73). It follows that the specialized $\omega_{\tilde{P}}(b)$ of $f - \lambda$ on C_b is the same as the generic $\omega_{\tilde{P}}$.

By Remark 1 and (74), this $X_{e-1}(\eta)$ is a linear form in at most $3(\tilde{n}+1)$ of the coefficients of A_1, A_2, A_3, whose coefficients are themselves of total degree at most $p^{\tilde{n}+1}$ in the

$$\lambda_{1\tilde{r}0}, \ldots, \lambda_{1\tilde{r}\tilde{n}}, \lambda_{2\tilde{r}0}, \ldots, \lambda_{2\tilde{r}\tilde{n}}, \lambda_{3\tilde{r}0}, \ldots, \lambda_{3\tilde{r}\tilde{n}}.$$

Thus indeed we can take any b in a set B_{00} independent of $\alpha_1, \alpha_2, \alpha_3$. For example, if we want $X = a_1\Lambda_1 + a_2\Lambda_2 \neq 0$ at b for all non-zero (a_1, a_2) in $\mathbf{F}_p^2$ then it suffices that $Y = \Lambda_1\Lambda_2(\Lambda_1^{p-1} - \Lambda_2^{p-1}) \neq 0$ (related to the Moore determinant in [30] p.8) at b, easy if $Y \neq 0$ already at η.

Above we assumed that $e - 1 \geq \tilde{n}$. If this is not the case, then our counting by poles gives

$$\omega_{\tilde{P}}(b) \geq 0 \geq \tilde{r}p^{e-1} - \tilde{r}p^{\tilde{n}-1} = \omega_{\tilde{P}} - \tilde{r}p^{\tilde{n}-1} \geq \omega_{\tilde{P}} - R.$$

Thus in both cases for e we get

$$\omega_{\tilde{P}}(b) \geq \omega_{\tilde{P}} - R_{\tilde{P}}.$$

So at each $\tilde{P}$ on $\tilde{C}$ where at least one of x_1, x_2, x_3 has a pole, if $\omega_{\tilde{P}} > 0$ we have

$$\max\{0, \omega_{\tilde{P}}(b)\} \geq \max\{0, \omega_{\tilde{P}}\} - \max\{0, -\mathrm{ord}_{\tilde{P}}x_1, -\mathrm{ord}_{\tilde{P}}x_2, -\mathrm{ord}_{\tilde{P}}x_3\};$$

and this holds trivially if $\omega_{\tilde{P}} \leq 0$.

Thus summing over all such $\tilde{P}$ we get

$$\deg_b(f-\lambda) \geq \deg_\eta(f-\lambda) - \sum_{\tilde{P}} \max\{0, -\mathrm{ord}_{\tilde{P}} x_1, -\mathrm{ord}_{\tilde{P}} x_2, -\mathrm{ord}_{\tilde{P}} x_3\}.$$

The sum on the right is the total number of poles (with multiplicity) of a generic linear combination of x_1, x_2, x_3. So the sum is $\deg C$, the total number of zeroes.

As promised we now show that for any b in our present B_{00}, indeed $f-\lambda$ is not identically zero on C_b. The corresponding assertion in the multiplicative situation is proved in [9] at the bottom of page 463.

Suppose on the contrary $f = \lambda$ on C_b. Choose any non-torsion τ in $\overline{F_0}$; then $f \neq \lambda + \tau$ on C_b. Thus $\mathcal{C}^k f \neq \lambda_k = \mathcal{C}^k(\lambda + \tau)$ on C_b for any $k \geq 0$. We apply (69) to $\mathcal{C}^k f - \lambda_k = \mathcal{C}^k\lambda - \lambda_k = -\mathcal{C}^k\tau \neq 0$, getting

$$0 \geq \deg_\eta(\mathcal{C}^k f - \lambda_k) - \deg C. \tag{77}$$

But $f - \lambda - \tau$ is not constant on C_η by our basic hypothesis, so $f - \lambda - \tau$ has at least one pole there. Thus $\mathcal{C}^k(f - \lambda - \tau) = \mathcal{C}^k f - \lambda_k$ has a pole of order at least p^k on C_η. Therefore the first degree on the right-hand side of (77) is at least p^k. Now we obtain a contradiction by making k tend to infinity. □

The next result essentially replaces an argument in the proof of Lemma 6.1 of [9] (p.464); here we lack Mason's *abc* Theorem.

Lemma 21 *For b in $B_{00}(\overline{F_0})$ or $b = \eta$ there is a finite union $\mathcal{E}_b$ (possibly empty) of rank 2 submodules of $\mathcal{R}^3$ with the following property. Suppose the non-zero $(\alpha_1, \alpha_2, \alpha_3)$ in $\mathcal{R}^3$ is not in $\mathcal{E}_b$ (if non-empty). Then if $\mathrm{ord}_{\tilde{P}} f > 0$ for some $\tilde{P}$ in $\tilde{C}_b$, we have*

(a) $\mathrm{ord}_{\tilde{P}} f = 1$ *for any $\tilde{P}$ over a non-singular point P of C_b,*

(b) $\mathrm{ord}_{\tilde{P}} f \leq \deg C$ *for any $\tilde{P}$ over a singular point of C_b,*

(c) $\mathrm{ord}_{\tilde{P}} f \leq (\deg C)(1 + \deg u)$ *for any $\tilde{P}$ over an infinite point of C_b, where u is a corresponding uniformizer.*

Proof We assume first that b is in $B_{00}(\overline{F_0})$.

With as usual $\alpha_i = A_i(\mathcal{C})$ we have by (33)

$$\mathrm{d}f = A_1(t)\mathrm{d}x_1 + A_2(t)\mathrm{d}x_2 + A_3(t)\mathrm{d}x_3$$

on C or C_b. This is the analogue of the multiplicative

$$\frac{\mathrm{d}(x_1^{a_1}x_2^{a_2}x_3^{a_3})}{x_1^{a_1}x_2^{a_2}x_3^{a_3}} = a_1\frac{\mathrm{d}x_1}{x_1} + a_2\frac{\mathrm{d}x_2}{x_2} + a_3\frac{\mathrm{d}x_3}{x_3}.$$

In case (a) at least one of $\mathrm{d}x_1, \mathrm{d}x_2, \mathrm{d}x_3$ must be non-zero at $\tilde{P}$ on C_b. Say it is $\mathrm{d}x \neq 0$. Then

$$\frac{\mathrm{d}f}{\mathrm{d}x} = A_1(t)g_1 + A_2(t)g_2 + A_3(t)g_3 = \phi$$

say, with fixed functions $g_i = \mathrm{d}x_i/\mathrm{d}x$.

If (a) is false we have $\phi(\tilde{P}) = 0$ on C_b. But this implies

$$[F_0(P) : F_0] \leq D_b \tag{78}$$

for some D_b possibly depending on b; unless, that is, ϕ is identically zero on C_b. We deal with this last possibility first.

If ϕ is identically zero on C_b then $(A_1(t), A_2(t), A_3(t))$ in $\mathcal{O}_0^3$ is in the additive relation group (with an obvious extension of the notion in [41] especially section 3) of g_1, g_2, g_3 in $F_0(b)(C_b)$. This group is of course a $\mathcal{O}_0$-module. If it were the full $\mathcal{O}_0^3$ then g_1, g_2, g_3 would all be identically zero on C_b which is absurd, because actually one of them is 1. So there is non-zero $(\varepsilon_1, \varepsilon_2, \varepsilon_3)$ in $\mathcal{R}^3$, possibly depending on b, such that $\varepsilon_1\alpha_1 + \varepsilon_2\alpha_2 + \varepsilon_3\alpha_3 = 0$. We put the corresponding rank 2 submodule into $\mathcal{E}_b$.

Thus indeed we may assume that (78) holds.

Now $f(\tilde{P}) = 0$, and so by Proposition 1 above together with the Northcott property we deduce that there are at most finitely many possibilities for $\tilde{P}$. For each of those $\tilde{P}$ which are over a non-singular point we still have $\phi(\tilde{P}) = 0$, and again this leads to $(\varepsilon_1, \varepsilon_2, \varepsilon_3)$ as above. This completes case (a).

In case (b) for $\tilde{P}$ over a singular point (ξ_1, ξ_2, ξ_3) of C_b we have

$$x_i = \xi_i + \lambda_i u^r + \cdots \tag{79}$$

for a suitably chosen uniformizer u and with $\lambda_1, \lambda_2, \lambda_3$ values, not all zero, of fixed functions on C evaluated at $\tilde{P}$ and specialized at b. Here $r \leq \deg C$ because

$$\operatorname{ord}_{\tilde{P}}(x_i - \xi_i) \leq \deg(x_i - \xi_i) = \deg x_i \leq \deg C.$$

So $f = \mu u^r + \cdots$ for $\mu = A_1(t)\lambda_1 + A_2(t)\lambda_2 + A_3(t)\lambda_3$. This gives the result unless $\mu = 0$; in which case as above this leads again to $(\varepsilon_1, \varepsilon_2, \varepsilon_3)$.

Finally in case (c) we go back to the decomposition $x_i = \underline{x_i} + \overline{x_i}$ in the proof of Lemma 20, where $\underline{x_i}$ involves only non-negative exponents and $\overline{x_i}$ only (finitely many) negative exponents. As $f(\tilde{P}) = 0$ we must have $\alpha_1\overline{x_1} + \alpha_2\overline{x_2} + \alpha_3\overline{x_3} = 0$ identically on C_b. This implies that $\underline{x_1}, \underline{x_2}, \underline{x_3}$ cannot all be zero on C_b, otherwise so would $f = \alpha_1 x_1 + \alpha_2 x_2 + \alpha_3 x_3$ be, already impossible for b in B_{00} thanks to Lemma 20.

So some $\underline{x_i} \neq 0$. If also $\overline{x_i} \neq 0$ then

$$\operatorname{ord}_{\tilde{P}}\underline{x_i} = \operatorname{ord}_{\tilde{P}}(x_i - \overline{x_i}) \leq \deg(x_i - \overline{x_i}) \leq \deg x_i + \deg \overline{x_i}.$$

Here

$$\deg \overline{x_i} \leq |\operatorname{ord}_{\tilde{P}} x_i| \deg u \leq (\deg x_i)(\deg u)$$

and so

$$\operatorname{ord}_{\tilde{P}}\underline{x_i} \leq (\deg C)(1 + \deg u) = s$$

say. And this clearly holds even if $\overline{x_i} = 0$.

So for each i we can write $\underline{x_i} = \lambda_i u^s + \cdots$ as in (79), and as

$$\operatorname{ord}_{\tilde{P}} f = \operatorname{ord}_{\tilde{P}}(\alpha_1\underline{x_1} + \alpha_2\underline{x_2} + \alpha_3\underline{x_3})$$

we can follow the arguments in case (b).

If $b = \eta$ then in (a) the only change is as follows. Now (78) becomes $[F_0(s_1, \ldots, s_l, P) : F_0(s_1, \ldots, s_l)] \leq D_\eta$, so the index in Lemma 19 is no bigger; this lemma shows that there are at most finitely many $\tilde{P}$. The argument for (b) is essentially unchanged. Similarly for (c); here we already know that f is not identically zero on C_η. □

The result would become false without the condition involving $\mathcal{E}_b$. For example with C as the line parametrized by (x, tx, sx) and $\alpha_1 = \mathcal{C}, \alpha_2 = -1, \alpha_3 = 0$ we have $f = x^p$ contradicting (a) at $P = (0, 0, 0)$.

In fact we are not so very far from the classical *abc* Mason result. For example over $\mathbf{C}$ consider $x^{a_1}(x-1)^{a_2}(x-2)^{a_3}-1$ for positive integer exponents. This is a non-zero polynomial of degree $D=a_1+a_2+a_3$; and *abc* implies at once that it has at least $D-2$ zeroes without multiplicity. A Carlitz analogue might be

$$A_1(\mathcal{C})(x)+A_2(\mathcal{C})(tx)+A_3(\mathcal{C})(t^2x)-1$$

for monic A_1, A_2, A_3. If $p \neq 2,3$ it is easy to see that this has degree $D = \max\{||A_1||, ||A_2||, ||A_3||\}$ in x. Using (33) as above we see that its derivative is simply $A_1+tA_2+t^2A_3$. If $p \neq 2,3$ that is non-zero; thus we see that the polynomial now has exactly D zeroes (without multiplicity).

Now we can give an analogue of Lemma 6.1 of [9] (p.464). We write N_{sing} for the number of points of $\tilde{C}$ above singular points of C and S_{inf} for the sum of $1+\deg u$ taken over all points above infinite points of C with u a corresponding uniformizer. Both quantities remain unchanged when we replace C by C_b. For $\alpha_1, \alpha_2, \alpha_3$ as above we define f as usual and then H in $\mathbf{A}^3 \times \mathbf{A}^m$ by the equation $f=0$. It is convenient during this section to assume that $C_B \cap H$ is non-empty.

Lemma 22 *For b in $B_{00}(\overline{F_0})$ or $b=\eta$ suppose the nonzero $(\alpha_1,\alpha_2,\alpha_3)$ in $\mathcal{R}^3$ is not in $\mathcal{E}_b$ (if non-empty). Then $C_B \cap H \cap \pi^{-1}(b)$ is a finite set whose cardinality satisfies*

$$\deg_b f - (\deg C)(N_{\text{sing}}+S_{\text{inf}}) \leq \#(C_B \cap H \cap \pi^{-1}(b)) \leq \deg_b f. \tag{80}$$

Proof We estimate

$$\deg_b f = \sum_{\tilde{P}\in\tilde{C}_b} \max\{0, \operatorname{ord}_{\tilde{P}} f\} \tag{81}$$

from above and below. We need to take into account only zeroes $\tilde{P}$ of f.

By Lemma 21, each $\tilde{P}$ in (81) above a non-singular point of C_b contributes 1. The number of such $\tilde{P}$ is at most the cardinality of $C_b \cap \gamma(H)$ (finite by Lemma 20), which is the same as that of $C_B \cap H \cap \pi^{-1}(b)$ because γ is an injection even on $C_B \cap \pi^{-1}(b)$. Similarly each $\tilde{P}$ above a singular point contributes at most $\deg C$, and each $\tilde{P}$ above an infinite point at most $(\deg C)(1+\deg u)$. The left-hand inequality of (80) follows.

On the other hand (81) is at least the number of zeroes (without multiplicity) of f over finite points of C_b, and this proves the right-hand inequality. □

Next we give an analogue of Lemma 6.2 of [9] (p.464). The Dimension Theorem (see for example [35] p.36) shows that $C_B \cap H$ (here assumed non-empty) has all its components of dimension l unless C_B is in H; but this last possibility is excluded by our original hypothesis on C.

Lemma 23 *For b in $B_{00}(\overline{F_0})$ or $b=\eta$ suppose the nonzero $(\alpha_1,\alpha_2,\alpha_3)$ in $\mathcal{R}^3$ is not in $\mathcal{E}_b$ (if non-empty). Then for any finite union W of components of $C_B \cap H$ we have*

$$\#\left(W \cap \pi^{-1}(b)\right) \geq \#\left(W \cap \pi^{-1}(\eta)\right) - (\deg C)\left(1+N_{\text{sing}}+S_{\text{inf}}\right).$$

Proof Write $s(b), s(\eta)$ for the two cardinalities to be compared. Let W' be the union of the components of $C_B \cap H$ not in the union W, and write $s'(b), s'(\eta)$ analogously. By Lemma 22 we have

$$s(b)+s'(b) \geq \#\left(C_B \cap H \cap \pi^{-1}(b)\right) \geq \deg_b f - (\deg C)\left(N_{\text{sing}}+S_{\text{inf}}\right).$$

Also

$$s(\eta) + s'(\eta) = \#\left(C_B \cap H \cap \pi^{-1}(\eta)\right)$$

since the sets $W \cap \pi^{-1}(\eta)$, $W' \cap \pi^{-1}(\eta)$ are disjoint. This is because $W \cap W'$ has dimension at most $l-1$, so cannot project to η. Also by Lemma 22 we have

$$\#\left(C_B \cap H \cap \pi^{-1}(\eta)\right) \leq \deg_\eta f$$

which by Lemma 20 is at most $\deg_b f + \deg C$.

Comparing these inequalities we see that it suffices now only to verify $s'(b) \leq s'(\eta)$. But this follows from Lemma 16 (with W' not W). We just have to note that $W' \cap \pi^{-1}(b)$ is finite by Lemma 22. □

However the next result has no analogue in [9] which is solidly over zero characteristic and so all inseparable degrees $\deg^{\text{ins}}$ are 1.

Lemma 24 *Suppose the nonzero* $(\alpha_1, \alpha_2, \alpha_3)$ *in* $\mathcal{R}^3$ *is not in* $\mathcal{E}_\eta$ *(if non-empty). Then if*

$$\deg_\eta f > 2(\deg C)(N_{\text{sing}} + S_{\text{inf}})$$

we have

$$\deg^{\text{ins}}_\eta f = 1.$$

Proof By Lemma 22 we have

$$\#(C_B \cap H \cap \pi^{-1}(\eta)) \geq d - (\deg C)(N_{\text{sing}} + S_{\text{inf}})$$

for $d = \deg_\eta f$. Now $d^{\text{sep}} = \deg^{\text{sep}}_\eta f$ is the number of solutions of $f = \lambda$ (without multiplicity) for generic λ. But these solutions are simply translates of the solutions of $f = 0$ by a single point. Thus

$$\#(C_B \cap H \cap \pi^{-1}(\eta)) = d^{\text{sep}} = \frac{d}{\deg^{\text{ins}}_\eta f}$$

and the lemma follows at once. □

12 Almost finishing

Here we prove our Theorem for general K, which as above we can take as $K = \overline{F_0(s_1, \ldots, s_l)}$ with $l \geq 1$, and C defined over $\overline{F_0}(s_1, \ldots, s_m) = \overline{F_0}(\eta)$. We may assume that C is not defined over $\overline{F_0}$, else every point with just one non-trivial relation would already be over $\overline{F_0}$ and so our Theorem for that field suffices.

Throughout this section (as in the previous section) we shall assume that no non-zero $\rho_1 x_1 + \rho_2 x_2 + \rho_3 x_3$ is identically constant on C. In the next section we shall relax this as required.

Now C is defined over some field K_*, finitely generated over $\mathbf{F}_p$, which lies in $\overline{F_0}(\eta)$. So we can find a finite extension F of F_0 with K_* inside $F(\eta)$. Both of these latter fields are finitely generated over F_0 with transcendence degree l and so the index $[F(\eta) : K_*] = e$ is finite.

Fix once and for all some b in $B_{00}(\overline{F_0})$. By Lemma 20 the hypothesis of our Theorem for C_b over $\overline{F_0}$ is satisfied. Thus there are at most finitely many points on $C_b(\overline{F_0})$ satisfying two independent relations. Let $r \geq 0$ be their cardinality.

Let P be any point on C with two independent relations (if it exists at all), and let W be the F-Zariski closure of (P, η) on C_B in $\mathbf{A}^3 \times \mathbf{A}^m$.

Now there is a submodule $\mathcal{M}$ of $\mathcal{R}^3$ of rank at least 2 that kills P. It may be that $\mathcal{M}$ lies in the unions $\mathcal{E}_b, \mathcal{E}_\eta$ (if non-empty) appearing in Lemma 21. But then $\mathcal{M}$ would be in one of the members making up $\mathcal{E}_b \cup \mathcal{E}_\eta$. Using $\mathrm{GL}_3(\mathcal{R})$ as at the end of the proof of Lemma 19, we can assume that $\mathcal{M}$ actually lies in $\mathcal{R}^2$. But now the problem in $\mathbf{G}_a^3$ is reduced to one in $\mathbf{G}_a^2$, an easy torsion problem.

Thus we can assume that $\mathcal{M}$ does not lie in $\mathcal{E}_b \cup \mathcal{E}_\eta$. Pick any $(\alpha_1, \alpha_2, \alpha_3)$ in $\mathcal{M}$ not in $\mathcal{E}_b \cup \mathcal{E}_\eta$. This gives an H as above; but of course there is an element of $\mathcal{M}$ independent of $(\alpha_1, \alpha_2, \alpha_3)$, and this gives analogously some H'.

Thus (P, η) lies in $C_B \cap H \cap H'$ (and in particular $C_B \cap H$ is non-empty as we assumed in the preceding section). The dimension of W is at least l. It follows that W (which is irreducible over F) is a finite union of $\overline{F_0}$-irreducible components of $C_B \cap H$ (all of dimension l as we saw above). Therefore by Lemma 23 we have

$$\#(W \cap \pi^{-1}(b)) \geq \#(W \cap \pi^{-1}(\eta)) - c_1$$

for some c_1 depending only on C.

But $\#(W \cap \pi^{-1}(\eta))$ is just the degree

$$[F(P, \eta) : F(\eta)]^{\mathrm{sep}} = \frac{[F(P, \eta) : F(\eta)]}{[F(P, \eta) : F(\eta)]^{\mathrm{ins}}}.$$

Now $F(P, \eta)$ lies in the field K_f obtained by adjoining to $F(\eta)$ all solutions of $f = 0$ on C_η. So

$$[F(P, \eta) : F(\eta)]^{\mathrm{ins}} \leq [K_f : F(\eta)]^{\mathrm{ins}} = \deg_\eta^{\mathrm{ins}} f.$$

If $d = \deg_\eta f > c_2$ again for $c_2 \geq 1$ depending only on C, then by Lemma 24 we have $\deg_\eta^{\mathrm{ins}} f = 1$. But otherwise

$$[F(P, \eta) : F(\eta)]^{\mathrm{ins}} \leq \deg_\eta^{\mathrm{ins}} f \leq d \leq c_2.$$

Thus in both cases

$$\#(W \cap \pi^{-1}(\eta)) = [F(P, \eta) : F(\eta)]^{\mathrm{sep}} \geq c_2^{-1}[F(P, \eta) : F(\eta)].$$

And in turn

$$[F(P, \eta) : F(\eta)] = \frac{[F(P, \eta) : K_*(P)][K_*(P) : K_*]}{[F(\eta) : K_*]} \geq \frac{[K_*(P) : K_*]}{e}.$$

Collecting these together, we see that $W \cap \pi^{-1}(b)$ contains at least

$$c_2^{-1} e^{-1} [K_*(P) : K_*] - c_1$$

different points. These project under γ to different points of $C_b \cap \gamma(H) \cap \gamma(H')$, whose cardinality is at most r. Therefore

$$[K_*(P) : K_*] \leq (c_1 + r) c_2 e$$

is bounded above independently of P. Now $K_*(t, s_1, \ldots, s_l)$ contains K_* and $F_0(s_1, \ldots, s_l)$; and all are finitely generated over F_0 with transcendence degrees l. So all indices are finite, and so the index $[F_0(s_1, \ldots, s_l, P) : F_0(s_1, \ldots, s_l)]$ is also bounded above independently of P.

Thus we can use Lemma 19 (in which the index is no bigger) to conclude as desired that there are at most finitely many P. This completes the proof of the Theorem when no non-zero $\rho_1 x_1 + \rho_2 x_2 + \rho_3 x_3$ is identically constant on C.

13 Relaxing the condition

Finally if some non-zero $\rho_1 x_1 + \rho_2 x_2 + \rho_3 x_3$ is identically constant on C, then again via GL_3 we can assume that x_3 is a constant ξ_3 on C. It is then necessarily non-torsion. Now the thing reduces to the analogue of Mordell-Lang on the projection to $\mathbf{G}_{\mathrm{a}}^2$: both ξ_1 and ξ_2 are in the division hull of $\mathcal{R}\xi_3$. In Sect. 4 we had reached a similar stage, but that was over $\overline{F_0}$. Oddly enough the extension to $\overline{F_0(s_1, \ldots, s_l)}$ is not in the literature, although Ghioca comes very close in [27]. In the personal communication [28] he does establish what we need here. But we can also use the ideas of [9] as follows.

Indeed we argue as in Sect. 4.

Namely as in (23) we can find ξ and a torsion ζ such that

$$F_0(s_1, \ldots, s_l)(\xi_1, \xi_2, \xi_3) = F_0(s_1, \ldots, s_l)(\xi, \zeta)$$

and (24). Further if ζ has order ν then τ_1, τ_2, τ_3 are polynomials in $\mathcal{R}$ of degree less than that of ν.

Then we can proceed down to (26), now in the form

$$D = [F_0(s_1, \ldots, s_l)(\xi_1, \xi_2, \xi_3) : F_0(s_1, \ldots, s_l)] \ll (\|\nu\| M)^{1/2} \tag{82}$$

with implied constants depending only on C (and ϵ soon to appear).

On the other hand (24) gives

$$h_s(\xi_1) = \|\sigma_1\| h_s(\xi), \quad h_s(\xi_2) = \|\sigma_2\| h_s(\xi), \quad h_s(\xi_3) = \|\sigma_3\| h_s(\xi) \tag{83}$$

as $h_s = 0$ on all of $\overline{F_0}$.

Now $h_s(\xi_3) \ll 1$ as ξ_3 is fixed. But we claim that also

$$h_s(\xi_1) \ll 1, \quad h_s(\xi_2) \ll 1. \tag{84}$$

To see this, project C down to a curve C' in $\mathbf{G}_{\mathrm{a}}^2$. There is a non-trivial Carlitz relation between ξ_1, ξ_2 and so by Lemma 18 we indeed get (84) unless some non-trivial $\rho_1 x_1 + \rho_2 x_2$ is constant on C'. With GL_2 we could then assume $x_2 = \xi_2$ is constant on C'. But now ξ_2, ξ_3 must be independent and so we could not have had two independent relations among ξ_1, ξ_2, ξ_3.

Now (83) implies

$$h_s(\xi) \ll M^{-1}. \tag{85}$$

We can assume that ξ is not in $\overline{F_0}$, because otherwise by (24) the point (ξ_1, ξ_2, ξ_3) would be in $C(\overline{F_0})$; and because C is not defined over $\overline{F_0}$ this would give the required finiteness at once. Now in (68) we have

$$\left[\overline{F_0}(s_1, \ldots, s_l, \xi) : \overline{F_0}(s_1, \ldots, s_l)\right] = \left[\overline{F_0}(s_1, \ldots, s_l, \xi) : \overline{F_0}(s_1, \ldots, s_l, \zeta)\right]$$

because ζ is in $\overline{F_0}$. This in turn is at most

$$[F_0(s_1, \ldots, s_l, \xi) : F_0(s_1, \ldots, s_l, \zeta)] \leq [F_0(s_1, \ldots, s_l, \xi_1, \xi_2, \xi_3) : F_0(s_1, \ldots, s_l, \zeta)]$$

which is

$$\frac{D}{[F_0(s_1,\dots,s_l,\zeta):F_0(s_1,\dots,s_l)]}.$$

Here the denominator is $[F_0(\zeta):F_0]=\phi(\nu)\gg\|\nu\|^{1-\epsilon}$ for any $\epsilon>0$.

So we can indeed assume

$$h_s(\xi)\gg\frac{\|\nu\|^{1-\epsilon}}{D}$$

a sort of "cyclotomic Lehmer" in the sense of Proposition 2. Comparing with (85) we find $D\gg M\|\nu\|^{1-\epsilon}\geq(M\|\nu\|)^{1-\epsilon}$. We choose $\epsilon<1/2$ to get by (82) $D\ll 1$, and now we conclude with Lemma 19 (which we already noted holds for $\mathbf{G}_{\mathrm{a}}^2$) applied to (ξ_1,ξ_2) on C'.

This completes the proof of the Theorem.

Funding Open access funding provided by University of Basel.

Appendix

Here we show that the inseparable degree $q=[F:F_{\mathrm{c}}]^{\mathrm{ins}}=[F:F_0]^{\mathrm{ins}}$ mentioned just after Proposition 2 is bounded above independently of d and D, at least in the application to the Theorem for $K=\overline{F_0}$. This does not necessarily mean that it can be forgotten about, due to the unfortunate fact that Frobenius does not commute with Carlitz.

Lemma 25 *For $K=\overline{F_0}$ let C in $\mathbf{G}_{\mathrm{a}}^3$ be an irreducible curve defined over K. Assume for any non-zero (ρ_1,ρ_2,ρ_3) in $\mathcal{R}^3$ that the form $\rho_1x_1+\rho_2x_2+\rho_3x_3$ is not identically zero on C. Then there is a constant $\mathfrak{B}$ such that*

$$[F_0(\xi_1,\xi_2,\xi_3):F_0]^{\mathrm{ins}}\leq\mathfrak{B}$$

for all (ξ_1,ξ_2,ξ_3) in $C(K)$ for which there exist linearly independent $(\alpha_1,\alpha_2,\alpha_3),(\beta_1,\beta_2,\beta_3)$ in $\mathcal{R}^3$ such that

$$\alpha_1\xi_1+\alpha_2\xi_2+\alpha_3\xi_3=\beta_1\xi_1+\beta_2\xi_2+\beta_3\xi_3=0. \tag{86}$$

Proof Observe that this lemma is formally implied by our Theorem! But the following proof is self-contained.

There is a non-zero polynomial P in $\mathbf{F}_p[X_1,X_2,X_3]$ such that $P(x_1,x_2,x_3)=0$ on C (that is, as in [43] and [15] we regard C as a surface over $\mathbf{F}_p$). We choose one of minimal degree. From (33) and Proposition 5.3 of [36] (p.371), the equations $P(\xi_1,\xi_2,\xi_3)=0$ with (86) define a separable extension of F_0 (that is, $[F_0(\xi_1,\xi_2,\xi_3):F_0]^{\mathrm{ins}}=1$) unless the Jacobian

$$Q=\begin{vmatrix}P_1 & P_2 & P_3\\ A_1(t) & A_2(t) & A_3(t)\\ B_1(t) & B_2(t) & B_3(t)\end{vmatrix}=0$$

at (ξ_1, ξ_2, ξ_3), where P_1, P_2, P_3 are the partial derivatives and $\alpha_1 = A_1(\mathcal{C})$ etc.

If $Q \neq 0$ on C then by Bezout we get the required bound for the entire degree $[F_0(\xi_1, \xi_2, \xi_3) : F_0]$ and we are done.

Thus we can assume $Q = 0$ on C. We write $Q = U_1(t)P_1 + U_2(t)P_2 + U_3(t)P_3$ for the appropriate minors (not all zero)

$$U_1(t) = (A_2B_3 - A_3B_2)(t), \quad U_2(t) = (A_3B_1 - A_1B_3)(t), \quad U_3(t) = (A_1B_2 - A_2B_1)(t)$$

of the Jacobian matrix. Then with as usual $\mathcal{O}_0 = \mathbf{F}_p[t]$ the $\mathcal{O}_0$-module ∇ of (Q_1, Q_2, Q_3) in $\mathcal{O}_0^3$ such that $Q_1P_1 + Q_2P_2 + Q_3P_3 = 0$ on C has rank r with $r = 1, 2, 3$. We consider each case in turn.

If $r = 3$ then $P_1 = P_2 = P_3 = 0$ on C. As the degree of P is minimal this implies $P_1 = P_2 = P_3 = 0$ identically. But then $P = \tilde{P}^p$ contradicting this very minimality.

If $r = 2$ (the crucial case), then this means there are fixed R_1, R_2, R_3 in $\mathcal{O}_0$, not all zero, such that $R_1Q_1 + R_2Q_2 + R_3Q_3 = 0$ on ∇. In particular

$$R_1(t)U_1(t) + R_2(t)U_2(t) + R_3(t)U_3(t) = 0 \tag{87}$$

which also relates the minors of (86).

Now we could operate with $\mathrm{GL}_3(\mathcal{R})$ in the usual way to produce from (86) a relation $\gamma_3\xi_3 = 0$ on a different curve. But it is more straightforward to assume $U_3 \neq 0$, and with $\rho_1 = R_1(\mathcal{C}), \rho_2 = R_2(\mathcal{C}), \rho_3 = R_3(\mathcal{C})$ then multiply the first relation in (86) by $\rho_1\beta_2 - \rho_2\beta_1$ and the second relation by $\rho_1\alpha_2 - \rho_2\alpha_1$ and subtract. What comes out using (87) is $U_3(\mathcal{C})\xi = 0$ for $\xi = \rho_1\xi_1 + \rho_2\xi_2 + \rho_3\xi_3$. Thus ξ is torsion. The corresponding function $x = \rho_1x_1 + \rho_2x_2 + \rho_3x_3$ on C is by hypothesis not zero on C. It cannot be constant either, because then its value would be torsion also contradicting the hypothesis. Thus $F_0(x_1, x_2, x_3)$ is a fixed finite extension of $F_0(x)$. Similarly for the specializations $F_0(\xi_1, \xi_2, \xi_3)$, $F_0(\xi)$. But we already noted that torsion is separable. It follows that

$$[F_0(\xi_1, \xi_2, \xi_3) : F_0]^{\mathrm{ins}} \leq [F_0(\xi_1, \xi_2, \xi_3) : F_0(\xi)] \leq \mathfrak{B}$$

as desired.

Finally if $r = 1$ then $(U_1(t), U_2(t), U_3(t))$ in projective $\mathbf{P}_2(F_0)$ is uniquely determined by P. But these are the Grassmann coordinates of the space of relations (86). So this space has generators of bounded degree. In other words we can consider each of (86) as a multiple of a fixed relation. Now we can repeat the above argument for just one of them.

It seems an interesting problem to extend the lemma to $\mathbf{G}_{\mathrm{a}}^n$ in the situation of the Conjecture (for $K = \overline{F_0}$ again). □

References

1. Amoroso, F., David, S.: Le problème de Lehmer en dimension supérieure. J. Reine Angew. Math. **513**, 145–179 (1999)
2. Amoroso, F., Zannier, U.: A relative Dobrowolski lower bound over abelian extensions. Ann. Scuola Norm. Sup. Pisa Cl. Sci. **29**, 711–727 (2000)
3. Artin, E.: Algebraic Numbers and Algebraic Functions. Nelson, Nashville (1967)
4. Bauchère, H.: Minoration de la hauteur canonique pour les modules de Drinfeld à multiplications complexes. J. Number Theory **157**, 291–328 (2015)
5. Bombieri, E., Gubler, W.: Heights in Diophantine Geometry. Cambridge University Press, Cambridge (2006)
6. Bombieri, E., Vaaler, J.: On Siegel's Lemma. Invent. Math. **73**, 11–32 (1983)
7. Bombieri, E., Habegger, P., Masser, D., Zannier, U.: A note on Maurin's Theorem. Rend. Lincei Mat. Appl. **21**, 251–260 (2010)

8. Bombieri, E., Masser, D., Zannier, U.: Intersecting a curve with algebraic subgroups of multiplicative groups. Int. Math. Res. Not. **20**, 1119–1140 (1999)
9. Bombieri, E., Masser, D., Zannier, U.: Finiteness results for multiplicatively dependent points on complex curves. Michigan Math. J. **51**, 451–466 (2003)
10. Bombieri, E., Masser, D., Zannier, U.: Intersecting curves and algebraic subgroups: conjectures and more results. Trans. Am. Math. Soc. **358**, 2247–2257 (2006)
11. Bombieri, E., Masser, D., Zannier, U.: Anomalous subvarieties - structure theorems and applications. Int. Math. Res. Not. **2007**, 33 (2007). https://doi.org/10.1093/imrn/rnm057
12. Bombieri, E., Masser, D., Zannier, U.: Intersecting a plane with algebraic subgroups of multiplicative groups. Ann. Scuola Norm. Sup. Pisa Cl. Sci. **VII**, 51–80 (2008)
13. Bombieri, E., Masser, D., Zannier, U.: On unlikely intersections of complex varieties with tori. Acta Arith. **133**, 309–323 (2008)
14. Bosser, V., Galateau, A.: Lower bounds for the canonical height on Drinfeld modules. Int. Math. Res. Not. **2019**, 165–200 (2017)
15. Brownawell, W.D., Masser, D.: Unlikely intersections for curves in additive groups over positive characteristic. Proc. Am. Math. Soc. **145**, 4617–4627 (2017)
16. Carlitz, L.: A class of polynomials. Trans. Am. Math. Soc. **43**, 167–182 (1938)
17. Cassels, J.W.S., Fröhlich, J.: Algebraic Number Theory. Academic Press, Cambridge (1967)
18. Danilov, V.I., Shokurov, V.V.: Algebraic curves. Algebraic manifolds and schemes. In: Encyclopaedia of Mathematical Sciences **23**, Springer, pp.167–297 (1994)
19. David, S., Pacheco, A.: Le problème de Lehmer abélien pour un module de Drinfeld. Int. J. Number Theory **4**, 1043–1067 (2008)
20. Demangos, L.: Lehmer problem and Drinfeld modules. J. Number Theory **189**, 147–185 (2018)
21. Denis, L.: Hauteurs canoniques et modules de Drinfeld. Math. Ann. **294**, 213–223 (1992)
22. Derksen, H., Masser, D.: Linear equations over multiplicative groups, recurrences, and mixing I. Proc. Lond. Math. Soc. **104**, 1045–1083 (2012)
23. Dobrowolski, E.: On a question of Lehmer and the number of irreducible factors of a polynomial. Acta Arith. **34**, 391–401 (1979)
24. Gebhardt, M.: Constructing function fields with many rational places via the Carlitz module. Manuscr. Math. **107**, 89–99 (2002)
25. Ghioca, D.: The Lehmer inequality and the Mordell-Weil theorem for Drinfeld modules. J. Number Theory **122**, 37–68 (2007)
26. Ghioca, D.: The isotrivial case in the Mordell-Lang theorem. Trans. Am. Math. Soc. **360**, 3839–3856 (2008)
27. Ghioca, D.: Towards the full Mordell-Lang conjecture for Drinfeld modules. Canadian Math. Bull. **53**, 95–101 (2010)
28. Ghioca, D.: Personal communication (2020)
29. Ghioca, D., Moosa, R.: Division points on subvarieties of isotrivial semiabelian varieties. Int. Math. Res. Not. **2006**, 23 (2006)
30. Goss, D.: Basic Structures of Function Field Arithmetic. Ergebnisse Math, vol. 35. Springer, Berlin (1996)
31. Habegger, P.: On the bounded height conjecture. Int. Math. Res. Not. **5**, 860–886 (2009)
32. Habegger, P., Pila, J.: O-minimality and certain atypical intersections. Ann. Sci. École Norm. Sup. **49**, 813–858 (2016)
33. Hrushovski, E.: The Mordell-Lang conjecture for function fields. J. Am. Math. Soc. **9**, 667–690 (1996)
34. Keller, A.: Cyclotomic function fields with many rational places. In: Jungnickel, D., Niederreiter, H. (eds.) Finite Fields and Applications. Springer, Berlin, pp. 293–302 (2001)
35. Lang, S.: Introduction to Algebraic Geometry. Addison-Wesley, Boston (1973)
36. Lang, S.: Algebra. Addison-Wesley, Boston (1993)
37. Laurent, M.: Minoration de la hauteur de Néron-Tate, Séminaire de théorie des nombres de Paris 1981-1982, Progress in Math. **38**, Birkhäuser, pp. 137–151 (1983)
38. Leitner, D.J.: Linear equations over multiplicative groups in positive characteristic. Acta Arith. **153**, 325–347 (2012)
39. Leitner, D.J.: Linear equations over multiplicative groups in positive characteristic II. J. Number Theory **180**, 169–194 (2017)
40. Mason, R.C.: Diophantine Equations Over Function Fields. LMS Lecture Notes **96**, Cambridge (1984)
41. Masser, D.: Linear relations on algebraic groups. In: Baker, A. (ed.) New Advances in Transcendence Theory. Cambridge University Press, Cambridge, pp. 248–262 (1988)
42. Masser, D.: Multiplicative dependence of values of algebraic functions. In: Chen, W.W.L., Gowers, W.T., Halberstam, H., Schmidt, W.M., Vaughan, R.C. (eds.) Analytic Number Theory- Essays in Honour of Klaus Roth. Cambridge University press, Cambridge, pp. 324–333 (2009)

43. Masser, D.: Unlikely intersections for curves in multiplicative groups over positive characteristic. Q. J. Math. **65**, 505–515 (2014)
44. Masser, D.: Auxiliary Polynomials in Number Theory. Tracts In Mathematics, vol. 207. Cambridge University Press, Cambridge (2016)
45. Masser, D., Zannier, U.: Torsion points on families of squares of elliptic curves. Math. Ann. **352**, 453–484 (2012)
46. Masser, D., Zannier, U.: Torsion points on families of abelian surfaces and Pell's equation over polynomial rings (with Appendix by V. Flynn). J. Eur. Math. Soc. **17**, 2379–2416 (2015)
47. Masser, D., Zannier, U.: Torsion points, Pell's equation, and integration in elementary terms. Acta Math. **225**, 227–312 (2020)
48. Maurin, G.: Courbes algébriques et équations multiplicatives. Math. Ann. **341**, 789–824 (2008)
49. Moosa, R., Scanlon, T.: F-structures and integral points on semiabelian varieties over finite fields. Am. J. Math. **126**, 473–522 (2004)
50. Mumford, D.: Algebraic Geometry I Complex Projective Varieties. Grundlehren der Math Wiss, vol. 221. Springer, Berlin (1976)
51. Pink, R.: A common generalization of the conjectures of André-Oort, Manin-Mumford, and Mordell-Lang, 13 pages (2005)
52. Pontreau, C.: Minoration effective de la hauteur des points d'une courbe de G_{m}^2 définie sur Q. Acta Arith. **120**, 1–26 (2005)
53. Ratazzi, N.: Théorème de Dobrowolski-Laurent pour les extensions abéliennes sur une courbe elliptique à multiplication complexe. Int. Math. Res. Not. **58**, 3121–3152 (2004)
54. Rémond, G.: Intersection de sous-groupes et de sous-variétés. III. Comment. Math. Helv. **84**, 835–863 (2009)
55. Scanlon, T.: Diophantine geometry of the torsion of a Drinfeld module. J. Number Theory **97**, 10–25 (2002)
56. Thakur, D.: Function Field Arithmetic. World Scientific Publishing, Singapore (2004)
57. Thunder, J.: Siegel's Lemma for function fields. Michigan Math. J. **42**, 147–162 (1995)
58. Viada, E.: The intersection of a curve with algebraic subgroups in a product of elliptic curves. Ann. Scuola Norm. Sup. Pisa Cl. Sci. **5**, 47–75 (2003)
59. Zannier, U.: Some Problems of Unlikely Intersections in Arithmetic and Geometry. Annals of Math. Studies, vol. 181. Princeton University Press, Princeton (2012)
60. Zilber, B.: Exponential sums equations and the Schanuel conjecture. J. Lond. Math. Soc. **65**, 27–44 (2002)