EXTENSIONS OF BÜCHI'S HIGHER POWERS PROBLEM TO POSITIVE CHARACTERISTIC

HECTOR PASTEN AND JULIE TZU-YUEH WANG

Abstract. Büchi's n-th power problem on $\mathbb{Q}$ asks whether there exist an integer M such that the *only* monic polynomials $F \in \mathbb{Q}[X]$ of degree n satisfying that $F(1), \dots, F(M)$ are n-th power rational numbers, are precisely of the form $F(X) = (X+c)^n$ for some $c \in \mathbb{Q}$. In this paper, we study analogues of this problem for algebraic function fields of positive characteristic. We formulate and prove an analogue (indeed, such a formulation for $n > 2$ was missing in the literature due to some unexpected phenomena), which we use to derive some definability and undecidability consequences. Moreover, in the case of characteristic zero we extend some known results by improving the bounds for M (from quadratic on n to linear on n).

Contents

1. Introduction and main results

In the seventies, J. R. Büchi investigated the following logic problem on simultaneous representation by quadratic forms, motivated by Matiyasevich's negative solution to Hilbert's tenth problem (after the work of Davis, Putnam and Robinson) [7]:

Büchi's undecidability problem. *Does there exists an algorithm for the following decision problem? Given diagonal quadratic forms $Q_j(x_1, \dots, x_n) = a_{j1}x_1^2 + \dots + a_{jn}x_n^2$ (for $j = 1, \dots, m$) with integer coefficients, and given $\mathbf{b} = (b_j)_j \in \mathbb{Z}^m$, decide whether or not there is a tuple of integers $\mathbf{a} \in \mathbb{Z}^n$ such that for each $j = 1, \dots, m$ we have $Q_j(\mathbf{a}) = b_j$.*

More precisely, Büchi formulated an arithmetic problem which, if true, would imply that the previous problem has a *negative* answer. See [6, 8, 9] for more details. The arithmetic problem formulated by Büchi is the following.

Date: February 16, 2014.
2010 *Mathematics Subject Classification.* Primary 11U05; Secondary 11D41, 14H05.
Key words and phrases. Büchi's problem, n-th powers, positive characteristic, Hilbert's tenth problem.
The first author is partially supported by an Ontario Graduate Scholarship.
The second author is supported in part by Taiwan's NSC grant 101-2115-M-001-001-MY2.

Büchi's squares problem I. *Does there exist an absolute constant M satisfying the following? If $x_1^2, \dots, x_M^2$ is a sequence of integer squares with second differences equal to 2 (i.e. $x_{n+2}^2 - 2x_{n+1}^2 + x_n^2 = 2$ for $1 \leq n \leq M-2$), then there is an integer c such that $x_n^2 = (n+c)^2$ for $1 \leq n \leq M$.*

This problem can be (equivalently) reformulated as follows.

Büchi's squares problem II. *Does there exist a positive integer M such that the* only *monic polynomials of degree two $F \in \mathbb{Z}[X]$ satisfying that $F(1), \dots, F(M)$ are integer squares, are precisely of the form $F(X) = (X+c)^2$ for some $c \in \mathbb{Z}$?*

We stress that the key point of this problem (and in all subsequent variations of it) is the *uniformity* of M: this constant is independent of F. Indeed, it was believed by Büchi that $M = 5$ should work. Since then, Büchi's squares problem has attracted the attention of researchers not only because of the applications in logic, but also as an interesting diophantine problem in its own right. The existence of such an M is still an open question (even Büchi's suggestion $M = 5$ remains open!). However, two conditional results in the direction of a positive solution are known: Vojta [21] proved that the Bombieri-Lang conjecture in the case of surfaces implies a positive solution, and Pasten [11] proved that a version of the *abc* conjecture for number fields implies a positive solution too (and even to a higher exponents generalization).

Analogues of these problems for other structures have also been investigated. An appropriate formulation of Büchi's squares problem for meromorphic functions and function fields in characteristic zero was proposed and solved by Vojta in the same work cited above [21]. An alternative proof for rational functions in characteristic zero, along with a study of the consequences in logic, was given by Pheidas and Vidaux in [14]. Actually, in [14] the case of rational functions in positive characteristic was also considered, but an unfortunate mistake was found by the first author, due to an unexpected behavior in positive characteristic (see [15] for counterexamples and the corrected results). This unexpected phenomenon is explained in more detail in Section 4 below. The statement of Büchi's squares problem for rational functions in positive characteristic turned out to be more subtle, but logic consequences could still be obtained from it as in the case of characteristic zero with only minor adaptations (see [15]). Büchi's squares problem for function fields in positive characteristic was later solved (and used to derive the expected undecidability results) by Shlapentokh and Vidaux in [18].

On the other hand, it is natural to extend Büchi's squares problem to higher exponents and, of course, to investigate the consequences in logic of a positive solution for it. This was first addressed in [16] where *Büchi's n-th powers problem* was formulated for $\mathbb{Z}$, $\mathbb{C}(t)$ and similar rings. The only results on Büchi's n-th powers problem for $\mathbb{Z}$ (and number fields) beyond the case of squares are the results of [11] which are conditional to a version of the *abc* conjecture. However, much (unconditional) progress has been achieved in Büchi's n-th powers problem for meromorphic functions and function fields *in characteristic zero*, and the following result surveys most of what is known (see [1, 10, 11]):

Theorem A (Büchi's problem in characteristic zero)**.** *Let $\mathbf{K}$ be a function field in one variable over an algebraically closed field $\mathbf{k}$ of characteristic zero, or let $\mathbf{K}$ be the field of meromorphic functions over $\mathbf{k} = \mathbb{C}$. Let $n \geq 2$ be an integer. There is a constant M depending only on $\mathbf{K}$ and n such that the following holds:*

Suppose that $F \in \mathbf{K}[X] \smallsetminus \mathbf{k}[X]$ is a monic polynomial of degree n such that $F(b)$ is an n-th power in $\mathbf{K}$ for at least M values of b in $\mathbf{k}$. Then F itself is an n-th power in $\mathbf{K}[X]$.

More precisely, in [1] it is obtained $M \ll_{\mathbf{K}} n^5$ and their proof also works for p-adic meromorphic functions, while in [11] it is obtained $M \ll_{\mathbf{K}} n^2$ using a different approach, and the result also holds for questions on μ-powerful values of F (i.e. all the zeros of $F(b)$ have multiplicity at least μ) rather than n-th powers. The consequences in logic were also investigated in [11]. Let us point out that the first work going beyond the case of squares is [17] where the case of cubes for the ring $\mathbb{C}[t]$ was solved.

However, Büchi's n-th powers problem for one variable function fields in positive characteristic remained mysterious. Even an appropriate *formulation* of it for $n > 2$ was unclear due to several unexpected examples of degree n polynomials taking n-th power values in the spirit of [15], but with unknown general pattern (see [1] for some families of such examples in higher exponents). Note, however, that some partial results are known; power values (with higher exponents) of *quadratic* polynomials have been studied by Garcia-Fritz [4] including the case of positive characteristic, while Hensley's problem in positive characteristic was first considered by Shlapentokh and Vidaux [18], and later solved by Wang [23] when the exponent n is large enough with respect to the genus of the function field (avoiding exceptional examples). (For the convenience of the reader, let us recall that Hensley's problem for n-th powers asks about polynomials of the specific form $F(X) = (X + a)^n - b$ taking n-th power values. For $n = 2$ it is equivalent to Büchi's squares problem, while for $n > 2$ it is weaker than Büchi's n-th powers problem. We remark that the positive characteristic case of Hensley's problem for higher exponent n solved in [23] works for $n \geq (3 + \sqrt{1 + 8\mathfrak{g}})/2$ where $\mathfrak{g}$ is the genus of the function field, and it does not cover the case of squares from [15, 18].)

We prove the following theorem, which provides this missing analogue of Büchi's n-th powers problem for function fields in positive characteristic.

Theorem 1. *Let $\mathbf{K}$ be a function field in one variable of genus $\mathfrak{g}_{\mathbf{K}}$ over an algebraically closed field $\mathbf{k}$ of characteristic $p > 0$. Let $n \geq 2$ be an integer and assume that*

$$p > M := 4n(n-1)\max\{\mathfrak{g}_{\mathbf{K}} - 1, 0\} + 11n^2 - 10n - 3.$$

For any monic polynomial $F \in \mathbf{K}[X] \setminus \mathbf{k}[X]$ of degree n, the following are equivalent:

(i) *$F(1), F(2), \ldots, F(M)$ are n-th powers in $\mathbf{K}$,*
(ii) *$F(m)$ is an n-th power in $\mathbf{K}$ for each $m = 1, 2, \ldots$, and*
(iii) *F is an n-th pseudo-power (see below).*

Again, the crucial point here is the uniformity of M. The notion of an *n-th pseudo-power* will be defined (in a constructive way) in Section 4. For instance, at the end of Section 4 we will show that, if $F \in \mathbf{K}[X]$ is monic of degree two and has some non-constant coefficient, then F is a 2-nd pseudo-power if and only if it factors in the form $F = (X - f)(X - f^{p^r})$ for some $f \in \mathbf{K} \setminus \mathbf{k}$ and $r \geq 0$. From this special case, it follows that Theorem 1 with $n = 2$ implies the main results of [15] and [18] (with different value of M). At the end of Section 4 we will also make Theorem 1 explicit in the cubic case.

The equivalence of the items (ii) and (iii) in Theorem 1, which makes no mention to M, is insufficient for applications in logic (at least in the context of Büchi's problem). However, that (ii) is equivalent to (iii) seems to be of independent arithmetic interest, as it gives a simple characterization of the monic degree n polynomials $F(X) \in \mathbf{K}[X]$ satisfying that $F(1)$, $F(2)$, ... are n-th powers.

The proof of Theorem 1 has two main ingredients. First, we use value distribution (Nevanlinna theory) to prove a general theorem for function fields (Theorem 3 below), which in the case of characteristic zero improves previously known bounds, and in the case of positive characteristic gives a partial result towards Büchi's n-th powers problem under some restrictions (these restrictions precisely avoid

the exceptional examples discussed above). The second ingredient is a detailed study of factorizations of polynomials whose coefficients belong to a function field of positive characteristic (see Section 4), which allows us to characterize in a simple way all the exceptional examples, leading to the notion of n-th pseudo-power. This approach is new, and indeed it gives a new proof of the known case for squares.

In the case of squares, the consequences in logic of [15, 18] in positive characteristic are derived by a slight modification of Büchi's original strategy. Also, in the case of higher exponents and characteristic zero (or even for number fields), analogous consequences in logic are deduced from Büchi's problem by a generalization of Büchi's approach (see [16, 11]). However, deducing consequences in logic from Theorem 1 in the spirit of Büchi's original approach is not straightforward, and a more subtle analysis is necessary. Indeed, these applications in logic lead us to prove a criterion for an n-th pseudo-power to be an actual n-th power (see Porposition 5.7) which can be useful in other contexts. Our main application of Theorem 1 in logic is Theorem 5.1 below (a definability result), but it requires some notational conventions before stating it, see Section 5. However, at this point we can mention the following application to undecidability (which uses work of Pheidas [13]).

Theorem 2. *Let $n \geq 2$ be an integer, let $p > 11n^2 - 10n - 3$ be a prime and let q be a power of p. There is no algorithm for solving the following decision problem:*

Given a system S consisting of equations of the form

$$a_1 X_1^n + a_2 X_2^n + \ldots + a_r X_r^n = b$$

with coefficients a_i, b in $\mathbb{F}_p[t]$, decide whether or not S has a solution in $\mathbb{F}_q(t)$.

This can be stated roughly as follows: *Fix any integer $n \geq 2$ and p, q as above. There is no algorithm for the problem of deciding if a variety defined over the field $\mathbb{F}_p(t)$ by Fermat-like equations of degree n (as above) has an $\mathbb{F}_q(t)$-rational point or not.*

The exceptional examples from [15], that initially seemed to be problematic, turned out to be a useful tool. Indeed, they have applications outside the context of Büchi's problem and constitute one of the main technical tools in [12], yielding *uniform* definability results. We hope that Theorem 1 can be useful in this direction too, since it gives a complete characterization of these exceptional examples for higher exponents, not only squares.

Finally, we state our main theorem for one variable function fields.

Theorem 3. *Let $\mathbf{K}$ be a function field of genus $\mathfrak{g}_{\mathbf{K}}$ over an algebraically closed field $\mathbf{k}$, and let C be the associated smooth projective curve. Let $n \geq 2$ and M be positive integers with*

$$M > 4n \max\{\mathfrak{g}_{\mathbf{K}} - 1, 0\} + 11n - 3.$$

Let $F \in \mathbf{K}[X] \smallsetminus \mathbf{k}[X]$ be a monic polynomial of degree n. Write $F = PH$ where $P \in \mathbf{k}[X]$ is monic, $H \in \mathbf{K}[X]$ is monic and H is not divisible by any non-constant polynomial in $\mathbf{k}[X]$. Let $G_1, \ldots, G_\ell \in \mathbf{K}[X]$ be the distinct monic irreducible factors of H and let $k_1, \ldots, k_\ell \geq 1$ be such that $H = \prod_{j=1}^{\ell} G_j^{k_j}$. Furthermore, if the characteristic p of $\mathbf{K}$ is positive, then we assume that for each $1 \leq j \leq \ell$ the polynomial G_j is separable over $\mathbf{K}$ and $G_j \notin \mathbf{K}^p[X]$. Let $\mu \geq \max_j k_j$ be an integer and let $\beta_1, \ldots, \beta_M$ be distinct elements of $\mathbf{k}$.

If for each $1 \leq i \leq M$ the zero multiplicity of $F(\beta_i) \in \mathbf{K}$ at every point $\mathfrak{p} \in C(\mathbf{k})$ is divisible by μ (with the convention that μ divides ∞), then $\mu = k_1 = \cdots = k_\ell$ and $H = G^\mu$, where $G = G_1 \cdots G_\ell$.

Remark. (1) *This result implies Theorem A in the case of function fields, with the improved constant $M \ll_{\mathbf{K}} n$. Indeed, if $F(\beta_i)$ is an n-th power in $\mathbf{K}$ then the zero multiplicity of $F(\beta_i)$ at every point $\mathfrak{p} \in C(\mathbf{k})$ is divisible by μ.*

(2) *The extra condition in the case of positive characteristic cannot be dropped. For instance, say that the characteristic of $\mathbf{K}$ is $p > 0$, let $f \in \mathbf{K}$ be non-constant and let $n \geq 2$ be an integer not divisible by p. For any $r \geq 1$ define the polynomial*

$$F_r(X) = (X - f)^{n-1}(X - f^{p^{r\phi(n)}}) \in \mathbf{K}[X]$$

where ϕ is the Euler function. A straightforward computation using Euler's congruence shows that $F_r(b)$ is an n-th power for p^r values of b in $\mathbf{k}$ (namely, for any b with $b^{p^r} = b$) although all irreducible factors of F_r have multiplicity $< n$. Taking r large we see that no uniform M as in Theorem 3 can work for the polynomials F_r; the problem is that these polynomials have irreducible factors $X - f^{p^{r\phi(n)}}$ in $\mathbf{K}^p[X]$.

The paper is structured as follows. We will recall some definitions and basic results of the Nevanlinna theory (i.e. value distribution) for function fields and prove key Lemmas in Section 2. The proof of Theorem 3 will be given in Section 3. Then we will give a more accurate discussion of the positive characteristic case and prove Theorem 1 in Section 4. Finally, Section 5 is devoted to the consequences in logic of Theorem 1.

2. Preliminaries on value distribution

In this section we present some preliminary results on value distribution in algebraic function fields in one variable. A crucial concept in this context is the notion of height, recalled below. For a more detailed exposition on heights (in a more general context) the reader can consult chapters III and IV of [5].

Let $\mathbf{K}$ be an algebraic function field of one variable over an algebraically closed field $\mathbf{k}$ of characteristic $p \geq 0$. Let C be the smooth projective curve defined over $\mathbf{k}$ associated to $\mathbf{K}$ and write $\mathfrak{g}$ for the genus of $\mathbf{K}$ (or equivalently, of C).

We first define valuations and height functions on the function field $\mathbf{K}$. For each point $\mathfrak{p} \in C(\mathbf{k})$, we may choose a uniformizer $t_{\mathfrak{p}}$ to define a normalized order function $v_{\mathfrak{p}} := \mathrm{ord}_{\mathfrak{p}} : \mathbf{K} \to \mathbb{Z} \cup \{\infty\}$ at $\mathfrak{p}$.

For $f \in \mathbf{K}$, the (relative) height is defined by

$$h_{\mathbf{K}}(f) := \sum_{\mathfrak{p} \in C(\mathbf{k})} - \min\{0, v_{\mathfrak{p}}(f)\}$$

that is, the degree of the pole divisor of f (which is equal to the degree of the zero divisor of f whenever f is not the zero function). For $\mathbf{x} = [x_0{:}...{:}x_n] \in \mathbb{P}^n(\mathbf{K})$ (taking representatives $x_i \in \mathbf{K}$) the projective height is defined by

$$h_{\mathbf{K}}(\mathbf{x}) := \sum_{\mathfrak{p} \in C(\mathbf{k})} - \min\{v_{\mathfrak{p}}(x_0), ..., v_{\mathfrak{p}}(x_n)\}$$

which is independent of the choice of the representative vector $(x_0, ..., x_n) \in \mathbf{K}^{n+1}$. Note that for $f \in \mathbf{K}$ we have $h_{\mathbf{K}}(f) = h_{\mathbf{K}}([1{:}f])$.

For the remaining of this section, we think about $\mathbf{K}$ as fixed, and consider algebraic extensions of it. Denote by $\bar{\mathbf{K}}$ the algebraic closure of $\mathbf{K}$. Let $\alpha \in \bar{\mathbf{K}}$ be a non-constant element. Let L be an algebraic extension of $\mathbf{K}$ containing α, let C_L be a smooth projective curve over $\mathbf{k}$ of genus $\mathfrak{g}_L$ such

that $L = \mathbf{k}(C_L)$ and let $\pi_L : C_L \to C$ be the corresponding morphism. For each $\mathfrak{q} \in C_L(\mathbf{k})$, the normalized order function $v_\mathfrak{q} : L \to \mathbb{Z} \cup \{\infty\}$ satisfies the relation

$$v_\mathfrak{q}(f) = e_\mathfrak{q} v_\mathfrak{p}(f) \quad \text{for} \quad f \in \mathbf{K}, \tag{1}$$

where $\mathfrak{p} = \mathfrak{q}|_\mathbf{K}$, $v_\mathfrak{p} : \mathbf{K} \to \mathbb{Z} \cup \{\infty\}$ is the normalized valuation attached to $\mathfrak{p}$, and $e_\mathfrak{q}$ is the ramification index of $\pi_L : C_L \to C$ at $\mathfrak{q}$. Hence,

$$[L : \mathbf{K}] \cdot v_\mathfrak{p}(f) = \sum_{\mathfrak{q} \in \pi_L^{-1}(\mathfrak{p})} v_\mathfrak{q}(f) \tag{2}$$

for $f \in \mathbf{K}$. The absolute height of the element $\alpha \in \bar{\mathbf{K}}$ is defined by

$$h(\alpha) := \frac{1}{[L : \mathbf{K}]} h_L(\alpha) = \frac{1}{[L : \mathbf{K}]} \sum_{\mathfrak{q} \in C_L(\mathbf{k})} - \min\{0, v_\mathfrak{q}(\alpha)\},$$

where we recall that $L/\mathbf{K}$ is chosen so that it contains α. It is known that the absolute height of α is independent of the choice of L, see for instance [5], Chapter II, Section 1 (this follows from the definition and the relation (1)). Interpreting the height as the degree of the pole divisor (on a suitable curve) gives the following elementary, though fundamental, properties of heights which we will use repeatedly in our computations.

Proposition 2.1. *For $f, g \in \bar{\mathbf{K}}$ we have $h(f + g) \leq h(f) + h(g)$ and $h(fg) \leq h(f) + h(g)$. If $f \neq 0$ we also have $h(1/f) = h(f)$. Moreover, if $b \in \mathbf{k}$ and $f + b \neq 0$ (in particular, for f non-constant) then $h(f + b) = h(f)$.*

We remark that under our notation, $h_\mathbf{K}(f) = h(f)$ for $f \in \mathbf{K}$. For $\mathbf{x} = [\alpha_0, ..., \alpha_n] \in \mathbb{P}^n(\bar{\mathbf{K}})$ the absolute projective height is given by

$$h(\mathbf{x}) := \frac{1}{[L : \mathbf{K}]} h_L([\alpha_0 : ... : \alpha_n])$$

which is independent of the choice of α_i and L, as long as all the α_i belong to L.

Let $A(X) = a_n X^n + a_{n-1} X^{n-1} + \cdots + a_1 X + a_0 \in \mathbf{K}[X]$ with $a_n \neq 0$. Denote by

$$v_\mathfrak{p}(A) := \min\{v_\mathfrak{p}(a_0), ..., v_\mathfrak{p}(a_n)\} \qquad \text{for } \mathfrak{p} \in C(\mathbf{k}).$$

The height of A is defined by

$$h(A) := h_K([a_0 : ... : a_n]) = - \sum_{\mathfrak{p} \in C(\mathbf{k})} v_\mathfrak{p}(A).$$

We recall that Gauss's lemma in this context says that

$$v_\mathfrak{p}(AB) = v_\mathfrak{p}(A) + v_\mathfrak{p}(B), \tag{3}$$

where A and B are in $\mathbf{K}[X]$ and $\mathfrak{p} \in C(\mathbf{k})$. Consequently, we have that

$$h(AB) = h(A) + h(B). \tag{4}$$

If we further assume that A is an irreducible polynomial in $\mathbf{K}[X]$ and let $\alpha \in \bar{\mathbf{K}}$ be a zero of A, then

$$h(A) = \deg A \cdot h(\alpha) \tag{5}$$

because all the roots of A are conjugated over $\mathbf{K}$ and have the same height. Let $\tilde{S}$ be a finite subset of $C_L(\mathbf{k})$ and β be a non-zero element in L. The truncated counting function with respect to $\tilde{S}$ over L is defined by

$$\bar{N}_{L,\tilde{S}}(\beta) := \sum_{\mathfrak{q} \in C_L(\mathbf{k}) \setminus \tilde{S}} \min\{1, v_\mathfrak{q}^+(\beta)\},$$

where $v_{\mathfrak{q}}^+(\beta) := \max\{0, v_{\mathfrak{q}}(\beta)\}$. This function counts the number of zeros (without taking multiplicity into account) of β on C_L, excepting possible zeros in $\tilde{S}$. In particular, since the height of a non-zero β is also equal to the degree of its zero divisor, one has the often useful inequality

$$\bar{N}_{L,\tilde{S}}(\beta) \leq h_L(\beta).$$

The normalized truncated counting function with respect to $\tilde{S}$ is given by

$$\bar{N}_{\tilde{S}}(\beta) := \frac{1}{[L:\mathbf{K}]}\bar{N}_{L,\tilde{S}}(\beta).$$

With this normalization one has $\bar{N}_{\tilde{S}}(\beta) \leq h(\beta)$. We recall the truncated second main theorem for function fields in arbitrary characteristic (this is the special case of Theorem 1 in [22] with $m = 1$, $n = 1$, $f_0 = f$, $f_1 = 1$ and $L_i = X - b_iY$).

Theorem 2.2. *Let $p \geq 0$ be the characteristic of $\mathbf{K}$, S be a set containing a finite number of points of $C(\mathbf{k})$ and $b_1, ..., b_q$ be distinct elements in $\mathbf{k}$. If $f \in \mathbf{K} \setminus \mathbf{k}\mathbf{K}^p$, then*

$$(q-2)h_{\mathbf{K}}(f) \leq \sum_{i=1}^{q} \bar{N}_S(f - b_i) + 2\mathfrak{g}_{\mathbf{K}} - 2 + |S|.$$

In our applications we will fix $\mathbf{K}$ and apply a version of the second main theorem to some α algebraic over $\mathbf{K}$, hence we need to keep track of the genus of the field $\mathbf{K}(\alpha)$. Recall the following from Proposition 3.11.1 in [19].

Proposition 2.3. *Let C_L be a smooth projective curve over $\mathbf{k}$ of genus $\mathfrak{g}_L$ such that its function field, $L := \mathbf{k}(C_L)$, is a finite extension of $\mathbf{K}$. Assume that $\{\alpha_1, ..., \alpha_m\}$ is a basis of $L/\mathbf{K}$ such that all $\alpha_i \in \mathcal{L}(D)$ for some divisor $D \in \mathrm{Div}(C_L)$. Then*

$$\mathfrak{g}_L - 1 \leq [L:\mathbf{K}] \cdot (\mathfrak{g}_{\mathbf{K}} - 1) + \deg D.$$

Proposition 2.4. *Let α be a non-constant algebraic element over $\mathbf{K}$ with $[\mathbf{K}(\alpha) : \mathbf{K}] = m$. Denote by $L = \mathbf{K}(\alpha)$ and let C_L be a smooth projective curve over $\mathbf{k}$ of genus $\mathfrak{g}_L$ such that $L = \mathbf{k}(C_L)$. Then*

$$\mathfrak{g}_L - 1 \leq m(\mathfrak{g}_{\mathbf{K}} - 1) + (m-1)h_L(\alpha).$$

Proof. Since $[\mathbf{K}(\alpha) : \mathbf{K}] = m$, the collection $\{1, \alpha, \alpha^2, ..., \alpha^{m-1}\}$ is a basis of $\mathbf{K}(\alpha)/\mathbf{K}$. Then the divisor D in Proposition 2.3 can be taken as $D := (m-1)(\alpha)_\infty$ where $(\alpha)_\infty$ denotes the divisor of poles of α in C_L. Taking degrees and recalling the definition of the height, we see that

$$\deg D = (m-1)\deg(\alpha)_\infty = (m-1)h_L(\alpha).$$

The assertion then follows directly from Proposition 2.3. □

The previous results lead to the following version of the second main theorem for algebraic functions with bounded degree that is better suited for our applications.

Theorem 2.5. *Let α be a non-constant algebraic over $\mathbf{K}$, let $L = \mathbf{K}(\alpha)$ and assume that α is not a p-th power in L if the characteristic p of $\mathbf{K}$ is positive. Let $m = [L : \mathbf{K}]$, let C_L be the corresponding smooth projective curve and let $\tilde{S}$ be a finite set of points in $C_L(\mathbf{k})$. Let $b_1, ..., b_q$ be distinct elements in $\mathbf{k}$. Then*

$$(q-2m)h(\alpha) \leq \sum_{i=1}^{q} \bar{N}_{\tilde{S}}(\alpha - b_i) + 2(\mathfrak{g}_{\mathbf{K}} - 1) + \frac{1}{m}|\tilde{S}|.$$

Proof. We apply Theorem 2.2 to $f = \alpha \in L$ with $S = \tilde{S}$ (using L instead of $\mathbf{K}$ in the theorem) to get

$$(q-2)h_L(\alpha) \leq \sum_{i=1}^{q} \bar{N}_{L,\tilde{S}}(\alpha - b_i) + 2\mathfrak{g}_L - 2 + |\tilde{S}|.$$

Using Proposition 2.4 we find

$$(q-2)h_L(\alpha) \leq \sum_{i=1}^{q} \bar{N}_{L,\tilde{S}}(\alpha - b_i) + 2m(\mathfrak{g}_{\mathbf{K}} - 1) + 2(m-1)h_L(\alpha) + |\tilde{S}|.$$

Dividing by m we obtain

$$(q-2)h(\alpha) \leq \sum_{i=1}^{q} \bar{N}_{\tilde{S}}(\alpha - b_i) + 2(\mathfrak{g}_{\mathbf{K}} - 1) + 2(m-1)h(\alpha) + \frac{1}{m}|\tilde{S}|,$$

hence the result. □

The condition of $\alpha \notin L^p$ when $p > 0$ can be formulated as follows.

Proposition 2.6. *Suppose that the characteristic p of $\mathbf{K}$ is positive. Let $\alpha \notin \mathbf{k}$ be algebraic over $\mathbf{K}$ and let $L = \mathbf{K}(\alpha)$. Let $P(X) \in \mathbf{K}[X]$ be the monic minimal polynomial of α over $\mathbf{K}$. Then $P(X)$ is not in $\mathbf{K}^p[X]$ if and only if α is not a p-th power in L.*

Proof. Write $P = X^d + c_{d-1}X^{d-1} + \ldots + c_1 X + c_0$ with $c_i \in \mathbf{K}$. If $P \in \mathbf{K}^p[X]$ then for each i we take $b_i \in \mathbf{K}$ with $c_i = b_i^p$. Let $H = X^d + b_{d-1}X^{d-1} + \ldots + b_1 X + b_0 \in \mathbf{K}[X]$ and let $\beta \in \bar{\mathbf{K}}$ be a root of H. Then we have the equation

$$\beta^d + b_{d-1}\beta^{d-1} + \ldots + b_1\beta + b_0 = 0$$

and taking p-th powers we get $P(\beta^p) = 0$. Hence β^p is a Galois conjugate of α. Let σ be a $\mathbf{K}$-automorphism of $\bar{\mathbf{K}}$ taking β^p to α, then $u = \sigma(\beta)$ has the property that $u^p = \alpha$ (hence $\mathbf{K}(\alpha) \subseteq \mathbf{K}(u)$) and moreover $H(u) = 0$ (hence $[\mathbf{K}(u) : \mathbf{K}] \leq \deg H = \deg P = [\mathbf{K}(\alpha) : \mathbf{K}]$). Therefore $\mathbf{K}(\alpha) = \mathbf{K}(u)$ and we see that $\alpha = u^p$ is a p-th power in $\mathbf{K}(\alpha)$.

Conversely, if α is a p-th power in $\mathbf{K}(\alpha)$ then we let $v \in \mathbf{K}(\alpha)$ be such that $\alpha = v^p$. Note that $\mathbf{K}(\alpha) = \mathbf{K}(v)$, so the minimal polynomials of v and α have the same degree d. Let $H = X^d + b_{d-1}X^{d-1} + \ldots + b_1 X + b_0 \in \mathbf{K}[X]$ be the minimal polynomial of v over $\mathbf{K}$ and observe that

$$v^d + b_{d-1}v^{d-1} + \ldots + b_1 v + b_0 = H(v) = 0,$$

thus, taking p-th powers we see that α is a root of the polynomial

$$Q = X^d + b_{d-1}^p X^{d-1} + \ldots + b_1^p X + b_0^p \in \mathbf{K}^p[X].$$

Since Q is monic of degree $d = \deg P$ and $Q(\alpha) = 0$, we conclude that $P = Q$. Therefore $P \in \mathbf{K}^p[X]$. □

Finally, we turn our attention to some lemmas specific to our applications. Let $A \in \mathbf{K}[X]$ be a monic separable (i.e. without repeated roots) polynomial. Consider the factorization $A = A_1 \cdots A_m$ where A_i are distinct monic separable irreducible polynomials in $\mathbf{K}[X]$. Let $u_j \in \bar{\mathbf{K}}$ $(1 \leq j \leq \deg A)$ be the roots of A and note that they are pairwise distinct. Let $\Delta_A(u_i) = \prod_{j \neq i}(u_i - u_j)$. With this notation, we have the following two lemmas.

Lemma 2.7. *With the assumptions listed above, for each $1 \leq j \leq \deg A$ we have*

$$h(\Delta_A(u_j)) \leq (2\deg(A) - 2) \max_{1 \leq i \leq \deg A} \{h(u_i)\}.$$

Proof. Without loss of generality we may assume that $j = 1$. By Proposition 2.1

$$h(\Delta_A(u_1)) = h\left(\prod_{i\neq 1}(u_1 - u_i)\right) \leq \sum_{i=2}^{\deg A} h(u_1 - u_i)$$
$$\leq \sum_{i=2}^{\deg A} (h(u_1) + h(u_i)) \leq (2\deg A - 2) \max_{1\leq i\leq \deg A}\{h(u_i)\}.$$

□

Lemma 2.8. *We keep the notation and assumptions from above, concerning A, the A_i and the u_j. Suppose that $\alpha = u_1$ is a root of A_1 and that $\alpha \notin \mathbf{k}$. Let $L = \mathbf{K}(\alpha)$ and $\pi : C_L \to C$ be the finite ramified covering corresponding to the field extension $\mathbf{K} \subseteq L$. We have the following.*

(i) *$\Delta_A(\alpha)$ belongs to L and it is not zero.*

(ii) *Let $\mathfrak{q}$ be a point in $C_L(\mathbf{k})$ such that $v_\mathfrak{q}(\Delta_A(\alpha)) = 0$, $v_\mathfrak{q}(\alpha) \geq 0$ and $v_\mathfrak{q}(A) \geq 0$. Suppose that $v_\mathfrak{q}(\alpha - \beta) > 0$ for some $\beta \in \mathbf{K}$. Then $v_\mathfrak{q}(A_1(\beta)) = v_\mathfrak{q}(\alpha - \beta)$ and $v_\mathfrak{q}(A_i(\beta)) = 0$ for each $i \neq 1$. Consequently, $v_\mathfrak{q}(\alpha - \beta) \geq v_\mathfrak{p}(A_1(\beta)) > 0$ and $v_\mathfrak{p}(A_i(\beta)) = 0$ for each $i \neq 1$, where $\mathfrak{p} = \pi(\mathfrak{q})$.*

Proof. We can assume that $u_1, \ldots, u_r$ are the roots of A_1 and recall that they are distinct. Evaluating the identity

$$A_1'(X) = \sum_{j=1}^{r} \prod_{i\in\{1,\ldots,r\}\smallsetminus\{j\}} (X - u_i)$$

at $X = u_1$ we find

$$A_1'(u_1) = \prod_{j=2}^{r}(u_1 - u_j).$$

Observe that

$$A_2(u_1)\cdots A_m(u_1) = \prod_{j>r}(u_1 - u_j)$$

(empty products are taken as 1). From the last two displayed equations and recalling the definition of $\Delta_A(\alpha)$ we obtain

$$\Delta_A(\alpha) = A_1'(\alpha)A_2(\alpha)\cdots A_m(\alpha). \tag{6}$$

Since the A_i's are over $\mathbf{K}$, this expression shows that $\Delta_A(\alpha)$ is in L. Moreover, it is not zero as the A_i's are distinct irreducible separable polynomials. Assertion (i) is concluded.

For (ii), we first note that since all the A_i's are monic polynomials and $v_\mathfrak{q}(1) = 0$, we have that $v_\mathfrak{q}(A_i) \leq 0$ for all $\mathfrak{q} \in C_L(\mathbf{k})$. Then the assumption $v_\mathfrak{q}(A) \geq 0$ implies that $v_\mathfrak{q}(A_i) = 0$ for all i by (3). This is equivalent to say that all the coefficients of the A_i's are regular at $\mathfrak{q}$. Therefore, $v_\mathfrak{q}(A_i(\alpha)) \geq 0$ and $v_\mathfrak{q}(A_i'(\alpha)) \geq 0$ for $1 \leq i \leq m$, since $v_\mathfrak{q}(\alpha) \geq 0$. Consequently, as $v_\mathfrak{q}(\Delta_A(\alpha)) = 0$, we have

$$v_\mathfrak{q}(A_1'(\alpha)) = v_\mathfrak{q}(A_2(\alpha)) = \cdots = v_\mathfrak{q}(A_m(\alpha)) = 0 \tag{7}$$

by (6). We also note that $v_\mathfrak{q}(\beta) \geq 0$ since $v_\mathfrak{q}(\alpha) \geq 0$ and $v_\mathfrak{q}(\alpha - \beta) > 0$. Therefore,

$$v_\mathfrak{q}(A_i(\alpha) - A_i(\beta)) \geq v_\mathfrak{q}(\alpha - \beta) > 0 \quad \text{for } 1 \leq i \leq m \tag{8}$$

because $A_i(\alpha) - A_i(\beta)$ is a multiple of $\alpha - \beta$ with an element regular at $\mathfrak{q}$. As $A_i(\beta) = A_i(\alpha) - (A_i(\alpha) - A_i(\beta))$, (7) and (8) imply that $v_\mathfrak{q}(A_i(\beta)) = 0$ for $2 \leq i \leq m$. Similarly, as $A_1(\alpha) = 0$, we can write

$$A_1(\beta) = A_1'(\alpha)(\beta - \alpha) + \frac{A_1''(\alpha)}{2}(\beta - \alpha)^2 + \cdots.$$

Since $v_{\mathfrak{q}}(A_1'(\alpha)) = 0$ by (7), this implies that $v_{\mathfrak{q}}(A_1(\beta)) = v_{\mathfrak{q}}(\alpha - \beta)$. The rest of the result follows from (1). □

3. Proof of main result for function fields

The purpose of this section is to prove Theorem 3.

Proof of Theorem 3. We keep the notation from the statement of the theorem. Let $\lambda_i = P(\beta_i) \in \mathbf{k}$ for each $1 \leq i \leq M$ and define $M' = M + 1 - n$. Relabeling the β_i if necessary, we can assume that all the λ_i are non-zero for $1 \leq i \leq M'$ (indeed, $\deg P \leq n-1$ so it can have at most $n-1$ distinct zeros). Let α_i for $1 \leq i \leq \deg G$ be zeros of G; these are pairwise distinct thanks to our assumption that the G_j are separable polynomials. Without loss of generality we assume that $\alpha := \alpha_1$ is a zero of G_1 and $h(\alpha) \geq h(\alpha_i)$ for $1 \leq i \leq \deg G$. Let $L = \mathbf{K}(\alpha)$, $m = [L : \mathbf{K}]$ and let $\pi : C_L \to C$ be the ramified covering associated to the extension $\mathbf{K} \subseteq L$. By Lemma 2.8 applied to $A = G$ (which is allowed since G does not have repeated roots) we see that $\Delta_G(\alpha)$ is a non-zero element of L (hence defines a function on C_L), so we can define the set

$$\tilde{S} = \{\mathfrak{q} \in C_L(\mathbf{k}) : v_{\mathfrak{q}}(\alpha) < 0, \text{ or } v_{\mathfrak{q}}(\Delta_G(\alpha)) > 0, \text{ or } v_{\mathfrak{q}}(G) < 0\}$$

which is finite. Then

$$\frac{1}{m}|\tilde{S}| \leq h(\alpha) + h(\Delta_G(\alpha)) + h(G) \leq (3 \deg G - 1)h(\alpha) \tag{9}$$

where the first inequality uses the interpretation of the height of $h(\Delta_G(\alpha))$ as the (suitably normalized) degree of the divisor of zeros of $\Delta_G(\alpha)$, and the second inequality is due to Lemma 2.7 and to

$$h(G) = \sum_{j=1}^{\ell} h(G_j) \leq \deg G \cdot h(\alpha)$$

which follows from (4), (5), and the assumption that α has maximal height among the zeros of G.

Suppose that $k_1 < \mu$. We first show that if $\mathfrak{q} \in C_L(\mathbf{k}) \smallsetminus \tilde{S}$ and $v_{\mathfrak{q}}(\alpha - \beta_i) > 0$ for some fixed i between 1 and M', then $v_{\mathfrak{q}}(\alpha - \beta_i) \geq 2$. After this is proved, we will derive a contradiction using value distribution. This contradiction will show that $k_1 = \mu$.

By Lemma 2.8 (again with $A = G$), we have that $v_{\mathfrak{p}}(G_j(\beta_i)) = 0$ for $j \neq 1$ and $v_{\mathfrak{q}}(\alpha - \beta_i) \geq v_{\mathfrak{p}}(G_1(\beta_i)) > 0$ where $\mathfrak{p} = \pi(\mathfrak{q})$. Since each of the zeros (in C) of $F(\beta_i) = \lambda_i \prod_{j=1}^{\ell} G_j(\beta_i)^{k_j}$ has multiplicity divisible by μ for each $1 \leq i \leq M$, and since the λ_i are non-zero constants for $1 \leq i \leq M'$, we see that

$$v_{\mathfrak{q}}(\alpha - \beta_i) \geq v_{\mathfrak{p}}(G_1(\beta_i)) = \frac{1}{k_1} v_{\mathfrak{p}}(F(\beta_i)) \geq \frac{\mu}{k_1} > 1.$$

Consequently, $v_{\mathfrak{q}}(\alpha - \beta_i) \geq 2$ whenever $\mathfrak{q} \in C_L(\mathbf{k}) \smallsetminus \tilde{S}$ and $v_{\mathfrak{q}}(\alpha - \beta_i) > 0$, as we claimed. In particular, recalling that the truncated counting function $\bar{N}_{L,\tilde{S}}$ counts zeros on C_L away from S forgetting multiplicities, while the height h_L equals the degree of the zero divisor (counting multiplicities) we find that for each i

$$\bar{N}_{\tilde{S}}(\alpha - \beta_i) = \frac{1}{m} \bar{N}_{L,\tilde{S}}(\alpha - \beta_i) \leq \frac{1}{2m} h_L(\alpha - \beta_i) = \frac{1}{2} h(\alpha - \beta_i).$$

If $p > 0$, our additional assumption $G_1 \notin \mathbf{K}^p[X]$ implies that α is not in L^p, by Proposition 2.6. Then we can use Theorem 2.5 (for any $p \geq 0$), the previous inequality, and (9) to get

$$\begin{aligned}(M' - 2m)h(\alpha) &\leq \sum_{i=1}^{M'} \bar{N}_{\tilde{S}}(\alpha - \beta_i) + 2(\mathfrak{g}_{\mathbf{K}} - 1) + \frac{1}{m}|\tilde{S}| \\ &\leq \frac{1}{2}\sum_{i=1}^{M'} h(\alpha - \beta_i) + 2(\mathfrak{g}_{\mathbf{K}} - 1) + (3\deg G - 1)h(\alpha) \\ &= \left(\frac{M'}{2} + 3\deg G - 1\right) h(\alpha) + 2(\mathfrak{g}_{\mathbf{K}} - 1)\end{aligned}$$

where the last equality is due to Proposition 2.1. Therefore, since $h(\alpha) \geq 1/m$, $n \geq m$ and $\deg G \leq n$ we obtain

$$\left(\frac{M - n + 1}{2} - 5n + 1\right) \leq \frac{2(\mathfrak{g}_{\mathbf{K}} - 1)}{h(\alpha)} \leq 2n \max\{\mathfrak{g}_{\mathbf{K}} - 1, 0\}$$

which implies $M \leq 4n \max\{\mathfrak{g}_{\mathbf{K}} - 1, 0\} + 11n - 3$. This contradicts the actual value of M so we must have $\mu = k_1$.

If $\ell = 1$ then we are done. Otherwise, we now let $F_1 = P\prod_{j=2}^{\ell} G_j^{k_j}$. Then $F = G_1^{\mu} F_1$. Since each of the zeros of $F(\beta_i)$ has multiplicity divisible by μ for each $1 \leq i \leq M$, this expression implies that each of the zeros of $F_1(\beta_i)$ also has multiplicity divisible by μ for each $1 \leq i \leq M$. Therefore, we can repeat the previous argument to conclude that $k_2 = \mu$. This inductive process will lead to the fact that $k_j = \mu$ for each j, as we wanted. □

4. Generalized Büchi's problem in positive characteristic

In this section we use Theorem 3 to solve the higher exponents generalization of Büchi's problem in function fields of positive characteristic. Let us first discuss on the meaning of Büchi's problem in this case.

The classical Büchi's problem over $\mathbb{Z}$ roughly says the following: *For certain uniform M, if a monic polynomial of degree two $F \in \mathbb{Z}[X]$ satisfies that $F(1), \ldots, F(M)$ are squares, then F is a square.*

The analogous statement for a one variable function field $\mathbf{K}$ over an algebraically closed field of characteristic zero (which is a theorem) cannot be stated exactly in this way because if F has constant coefficients then trivially *all* the values $F(1), F(2), \ldots$ are squares, so this exceptional case should be taken into account. This leads to the statement: *For certain uniform M, if a monic polynomial of degree two $F \in \mathbf{K}[X]$ satisfies that $F(1), \ldots, F(M)$ are squares, then F has constant coefficients or F is a square.*

For function fields $\mathbf{K}$ in positive characteristic $p > 0$ (say, p odd) the situation becomes worse. Of course $p > M$ for otherwise $1, \ldots, M$ are not distinct in the field, but there are other non-trivial issues. For example the polynomial $F = (X - f^{p^r})(X - f) \in \mathbf{K}[X]$ (with $f \in \mathbf{K}$ non-constant) has non-constant coefficients and if $r > 0$ then it is not a square. However $F(1), F(2), \ldots$ are squares (provided that $p > 2$) because for $m = 1, 2, \ldots$ one has

$$F(m) = (m - f^{p^r})(m - f) = (m - f)^{p^r + 1}.$$

The point is that there are no other exceptions as proved in [15, 18], so the correct statement is: *For certain uniform M, if a monic polynomial of degree two $F \in \mathbf{K}[X]$ satisfies that $F(1), \ldots, F(M)$ are squares, then F has constant coefficients or $F = (X - f^{p^r})(X - f)$ with $r \geq 0$ and $f \in \mathbf{K}$ non-constant* (note that with $r = 0$ we recover the case when F is a square).

It is not difficult to produce other exceptional cases for the higher exponents analogue of Büchi's problem in positive characteristic (see [1]). With this consideration in mind, it becomes clear that a satisfactory solution in this context should provide an explicit description of all exceptional cases, and this is the purpose of this section.

First we need to introduce some notation. Let $\mathbf{K}$ be a function field of genus $\mathfrak{g}$ over an algebraically closed field $\mathbf{k}$ of characteristic $p > 0$. Given a polynomial $G = a_d X^d + \ldots + a_1 X + a_0 \in \mathbf{K}[X]$ and a positive integer w we define

$$G^{[w]} = a_d^w X^d + \ldots + a_1^w X + a_0^w.$$

If $W = (w_1, w_2, \ldots, w_r)$ is a tuple of positive integers then we define

$$G^{[W]} = G^{[w_1]} G^{[w_2]} \cdots G^{[w_r]}$$

(in practice, we will only use the case when the integers w_i are powers of p). For W a tuple as above we also define its length $|W| = w_1 + \ldots + w_r$. This notation has the following convenient feature.

Proposition 4.1. *Let $G \in \mathbf{K}[X]$ be a polynomial and let $W = (w_1, \ldots, w_r)$ be a tuple of positive integers whose coordinates w_i are powers of p. Then for $m = 1, 2, \ldots$ (i.e. for m in the prime subfield of $\mathbf{k}$) we have*

$$G^{[W]}(m) = G^{|W|}(m)$$

(where $G^{|W|}$ just denotes the polynomial G to the power $|W|$).

Proof. Since w_i are powers of p, and m belongs to the prime subfield of $\mathbf{k}$ we have $m^{w_i} = m$. Hence

$$G^{[W]}(m) = \prod_i G^{[w_i]}(m) = \prod_i G^{[w_i]}(m^{w_i}) = \prod_i G(m)^{w_i} = G(m)^{|W|}.$$

□

We need the following two auxiliary lemmas.

Lemma 4.2. *Let $H \in \mathbf{K}[X]$ be a polynomial of degree n and suppose that $p \nmid n$. Consider the sequence of polynomials*

$$H, H^{[p]}, H^{[p^2]}, H^{[p^3]}, \ldots$$

A polynomial in this sequence is irreducible over $\mathbf{K}$ if and only if all of them are irreducible over $\mathbf{K}$.

Proof. It suffices to show that H is irreducible if and only if $H^{[p]}$ is irreducible. If H is reducible, say $H = H_1 H_2$ with $H_1, H_2 \in \mathbf{K}[X]$ of degree at least 1, then we have $H^{[p]} = H_1^{[p]} H_2^{[p]}$ because taking p-th powers defines a field map $\sigma_p : \mathbf{K} \to \mathbf{K}$ and the coefficients of these polynomials are in $\mathbf{K}$.

Conversely, suppose that H is irreducible. Let L be the splitting field of H and let G be the Galois group of H (that is, of $L/\mathbf{K}$). Let $\alpha_1, \ldots, \alpha_n \in L$ be the roots of H and note that they are distinct, since H is irreducible and $p \nmid n$. Then the roots of $H^{[p]}$ are $\alpha_1^p, \ldots, \alpha_n^p \in L$ and they are pairwise distinct (indeed, σ_p is injective as it is a field map). Hence, to show that $H^{[p]}$ is irreducible it suffices to show that G acts transitively on the set $\{\alpha_j^p\}_j$. Given indices i, j let $\tau \in G$ be such that $\alpha_i = \tau(\alpha_j)$; this is possible since H is irreducible so that G acts transitively on the roots of H. Taking p-th powers we get $\alpha_i^p = \tau(\alpha_j)^p = \tau(\alpha_j^p)$ because τ is compatible with multiplication. This proves that G acts transitively on the roots of $H^{[p]}$. □

Lemma 4.3. *Given $H \in \mathbf{K}[X]$ with some non-constant coefficient, there exist a unique integer $i \geq 0$ and a unique $G \in \mathbf{K}[X]$ such that $G \notin \mathbf{K}^p[X]$ and $H = G^{[p^i]}$. Moreover, if H is irreducible, then G is irreducible.*

Proof. For the existence, it suffices to show that if $f \in \mathbf{K}$ is non-constant then there is some i and $g \in \mathbf{K} \setminus \mathbf{K}^p$ such that $f = g^{p^i}$ (then we apply this to the coefficients of H). But this follows from the fact that the extension $\mathbf{K}/\mathbf{k}(f)$ is finite, hence, only finitely many terms in the sequence $\{f^{1/p^i}\}_{i\geq 0}$ can belong to $\mathbf{K}$ because $[\mathbf{k}(f^{p^i}) : \mathbf{k}(f)] = p^i$ (alternatively, one can give a proof using valuations). The uniqueness is clear because the map $\mathbf{K} \to \mathbf{K}$ given by $x \mapsto x^p$ is injective. Finally, the claim about irreducibility follows from the identity $(H_1H_2)^{[p]} = H_1^{[p]}H_2^{[p]}$ for $h_i \in \mathbf{K}[X]$ (already used in the first part of the proof of Lemma 4.2). □

The next result gives a way to write polynomials in $\mathbf{K}[X]$ which is convenient for our purposes and can be of independent interest.

Proposition 4.4. *Let $F \in \mathbf{K}[X]$ be a monic polynomial with some non-constant coefficient (that is, $F \notin \mathbf{k}[X]$). Then F can we written as*

$$F = P\prod_{i=1}^{c} G_i^{[W_i]}$$

where

- *$P \in \mathbf{k}[X]$ is monic (possibly $P = 1$),*
- *$G_i \in \mathbf{K}[X]$ are distinct, monic, irreducible polynomials with some coefficient in $\mathbf{K} \setminus \mathbf{K}^p$ (hence, each G_i has some non-constant coefficient), and*
- *W_i are tuples whose coordinates are powers of p with non-negative exponent.*

Moreover, if $\deg F < p$ then this way of writing F is unique up to rearranging the factors $G_i^{[W_i]}$ and permuting the coordinates of each W_i.

Proof. Write $F = PH_1 \cdots H_t$ with $P \in \mathbf{k}[X]$ the monic polynomial of largest degree that divides F, and $H_j \in \mathbf{K}[X]$ monic irreducible for each j. Lemma 4.3 applied to each H_j gives an irreducible polynomial $G_j \in \mathbf{K}[X]$ with some coefficient not in $\mathbf{K}^p$ such that $H_j = G_j^{[p^{n_j}]}$ for certain $n_j \geq 0$. Rearranging terms and relabeling the G_j if necessary, we get the desired factorization (the tuples W_i are obtained by grouping together the factors $G_j^{[p^{n_j}]}$ that have the same G_j).

Finally, we prove uniqueness when $\deg F < p$. Let us compare two such factorizations:

$$P\prod_{i=1}^{c} G_i^{[W_i]} = F = \bar{P}\prod_{j=1}^{\bar{c}} \bar{G}_j^{[\bar{W}_j]}$$

Each G_i and each $\bar{G}_i$ is irreducible with some non-constant coefficient, and P and $\bar{P}$ are monic with constant coefficients, hence $P = \bar{P}$ by unique factorization in $\mathbf{K}[X]$. If w_{ia} is a coordinate of W_i then $G_i^{[w_{ia}]}$ is irreducible by Lemma 4.2 (here we use $\deg F < p$, so that p does not divide the degree of any factor of F) and similarly for factors of the form $\bar{G}_j^{[\bar{w}_{jb}]}$ where $\bar{w}_{jb}$ is a coordinate of $\bar{W}_j$. Therefore each $G_i^{[w_{ia}]}$ is equal to some $\bar{G}_j^{[\bar{w}_{jb}]}$ and conversely, again because $\mathbf{K}[X]$ is UFD. By Lemma 4.3 and the assumption that each G_i and each $\bar{G}_j$ has some coefficient not in $\mathbf{K}^p$, uniqueness follows. □

We refer to the expression $F = P\prod_{i=1}^{c} G_i^{[W_i]}$ provided by the previous proposition as *Frobenius factorization* of F.

For uniqueness, the requirement $\deg F < p$ cannot be dropped. For instance the polynomial $X^p - f^p$ has two different Frobenius factorizations $(X - f)^{[W]} = (X^p - f)^{[(p)]}$ where $W = (1, \ldots, 1)$ has p coordinates (here, $f \in \mathbf{K} \setminus \mathbf{K}^p$ so that $X^p - f$ is irreducible). In our application the condition $\deg F < p$ will hold, thus we will have uniqueness.

Definition 4.5. *Let μ be a positive integer. A monic polynomial $F \in \mathbf{K}[X]$ with some non-constant coefficient is said to be a* μ-th pseudo-power *if for some Frobenius factorization of F, each of the tuples W_i appearing in this factorization has length $|W_i|$ divisible by μ.*

Note that if F is a μ-th pseudo-power it does not need to be true that *all* the Frobenius factorizations satisfy the above condition. For instance, $F = (X-f)^{p+1}$ is a $(p+1)$-st pseudo-power because $F = (X-f)^{[W']}$ where $W' = (1, \dots, 1)$ has $p+1$ coordinates, although $F = (X-f)^{[(1)]}(X^p - f)^{[(p)]}$ (here, $f \in \mathbf{K} \smallsetminus \mathbf{K}^p$). However, in our application we will have uniqueness, so that one can check whether F is a μ-th pseudo-power *or not* just by looking at one Frobenius factorization.

The motivation for the previous definition comes from the following fact.

Proposition 4.6. *Let $F \in \mathbf{K}[X]$ be a monic polynomial with some non-constant coefficient. Then*

1. *If F is a μ-th power then it is also a μ-th pseudo-power.*
2. *If F is a μ-th pseudo-power then $F(m)$ is a μ-th power in $\mathbf{K}$ for each $m = 1, 2, 3, \dots$ (that is, for each m in the prime subfield of $\mathbf{k}$).*

Proof. For Item (1) we have to construct a suitable Frobenius factorization of F. Let $H \in \mathbf{K}[X]$ be such that $F = H^\mu$. Consider any Frobenius factorization $H = P \prod_{i=1}^c G_i^{[W_i]}$. Let W_i' be the concatenation of μ copies of W_i. Then $F = P^\mu \prod_{i=1}^c G_i^{[W_i']}$ is a Frobenius factorization of F and μ divides each $|W_i|$.

For Item (2), we take a Frobenius factorization $F = \prod_{i=1}^c G_i^{[W_i]}$ such that μ divides all the numbers $|W_i|$. Note that Proposition 4.1 gives

$$F(m) = P(m) \prod_i G_i(m)^{|W_i|}.$$

Since $P(m) \in \mathbf{k}$ is always a μ-th power (as $\mathbf{k}$ is algebraically closed) and μ divides each $|W_i|$, we conclude that $F(m)$ is a μ-th power in $\mathbf{K}$. □

The concept of μ-th pseudo-power is relevant in our situation because of Theorem 1, which solves the generalized Büchi problem in positive characteristic. The item (2) of Proposition 4.6 already proves part of the equivalences in Theorem 1, and for the convenience of the reader we recall the part that remains to be proved.

Theorem 4.7. *Let $n \geq 2$ be an integer and assume that $p > M$ where*

$$M = 4n(n-1)\max\{\mathfrak{g} - 1, 0\} + 11n^2 - 10n - 3.$$

For any monic polynomial $F \in \mathbf{K}[X]$ of degree n with some non-constant coefficient, if we have that $F(1), F(2), \dots, F(M)$ are n-th powers in $\mathbf{K}$, then F is an n-th pseudo-power.

Before proceeding to the proof we need an auxiliary result.

Lemma 4.8. *Let $f \in \mathbf{K} \smallsetminus \mathbf{K}^p$. If $\mathfrak{g} = 0$ set $N = 3$ and if $\mathfrak{g} \geq 1$ set $N = 4\mathfrak{g}$. Then there are at most N values of $b \in \mathbf{k}$ such that $b + f \in \mathbf{K}$ has all its zeros with multiplicity strictly larger than 1.*

Proof. Suppose that $b_1, \dots, b_q \in \mathbf{k}$ are distinct and that for each j all the zeros of $b_j + f$ have multiplicity at least 2. Since the truncated counting function $\bar{N}$ counts zeros without considering multiplicity, while the height h takes multiplicity into account, we find for any finite set of points $S \subseteq C(\mathbf{k})$

$$\bar{N}_S(f + b_j) \leq \frac{1}{2} h(f + b_j) = \frac{1}{2} h(f)$$

(the last equality by Proposition 2.1). Taking S empty in Theorem 2.2 we have

$$(q-2)h(f) \leq \sum_{j=1}^{q} \bar{N}(f+b_j) + 2\mathfrak{g} - 2 \leq \frac{q}{2}h(f) + 2\mathfrak{g} - 2$$

which gives $(q-4)h(f) \leq 4(\mathfrak{g}-1)$. Since $h(f) \geq 1$ we get

$$q - 4 \leq \frac{4}{h(f)}(\mathfrak{g}-1) \begin{cases} \leq 4(\mathfrak{g}-1) & \text{if } \mathfrak{g} \geq 1 \\ < 0 & \text{if } \mathfrak{g} = 0 \end{cases}$$

and the result follows. □

Proof of Theorem 4.7. Assume that $F(m)$ are n-th powers in $\mathbf{K}$ for $1 \leq m \leq M$. Consider the Frobenius factorization of F (note that $n < M < p$, hence we have uniqueness)

$$F = P \prod_{i=1}^{c} G_i^{[W_i]}$$

and define the polynomial

$$F_0 = \prod_{i=1}^{c} G_i^{|W_i|^*} \in \mathbf{K}[X]$$

where $|W_i|^* \in \{0, 1, 2, \ldots, n-1\}$ is the residue of $|W_i|$ modulo n, which is a non-negative integer (possibly 0) strictly less than n. Let n_0 be the degree of F_0, then

$$n_0 \leq \sum_i (n-1) \cdot \deg G_i \leq (n-1)n.$$

Observe that the integers $|W_i| - |W_i|^*$ are divisible by n, so we can define the polynomial

$$H = \prod_{i=1}^{c} G_i^{(|W_i| - |W_i|^*)/n} \in \mathbf{K}[X].$$

Note that $H(\lambda) \neq 0$ for all $\lambda \in \mathbf{k}$ because the G_i do not have constant roots. Set $M' = M+1-n$. Since $M < p$, there exist distinct elements $b_1, \ldots, b_{M'}$ in the prime field of $\mathbf{k}$ such that $P(b_1), \ldots, P(b_{M'}) \in \mathbf{k}^\times$. Then Proposition 4.1 shows that for each b_j we have

$$F_0(b_j)H^n(b_j) = \prod_{i=1}^{c} G_i^{|W_i|}(b_j) = \prod_{i=1}^{c} G_i^{[W_i]}(b_j) = \frac{1}{P(b_j)}F(b_j)$$

and it follows that each $F_0(b_j)$ is an n-th power in $\mathbf{K}$, because each $F(b_j)$ is an n-th power by assumption and each $P(b_j)$ is in $\mathbf{k}$.

Suppose that $|W_i|^* \neq 0$ for some i (in particular, n does not divide $|W_i|$). Then $1 \leq n_0 = \deg F_0$ and F_0 is a separable polynomial since $n_0 < M < p$. Moreover, recall that each G_i has some coefficient not in $\mathbf{K}^p$ by definition of Frobenius factorization (see Proposition 4.4). Let us consider two cases: when $n_0 = 1$ and when $n_0 \geq 2$.

In the first case $F_0(X) = X + f$ for some $f \in \mathbf{K} \smallsetminus \mathbf{K}^p$ and each $F_0(b_j) = b_j + f$ is an n-th power, in particular each one of them has all its zeros with multiplicity at least $n \geq 2$. Lemma 4.8 implies that $M' \leq \max\{3, 4\mathfrak{g}\}$.

In the second case $\deg F_0 = n_0 \geq 2$. Recall that $F_0(b_j)$ is an n-th power for each $1 \leq j \leq M'$ which implies that all the zeros of each $F_0(b_j)$ have multiplicity divisible by n. Taking $\mu = n$ in Theorem 3 (which is allowed since $n > |W_i|^*$ for each i) we see that we cannot have $M' > 4n_0 \max\{\mathfrak{g}-1, 0\} + 11n_0 - 3$, for otherwise the theorem would give $n = |W_i|^*$ for each non-zero $|W_i|^*$, a contradiction.

In both cases we conclude

$$M' \leq 4n_0 \max\{\mathfrak{g}-1, 0\} + 11n_0 - 3$$

which contradicts the actual value of M in Theorem 4.7 because $n_0 \leq (n-1)n$. Therefore $|W_i|^* = 0$ for each i, hence F is an n-th pseudo-power. □

This also concludes the proof of Theorem 1.

To conclude our discussion on Theorem 1, let us make its statement explicit in the cases $n = 2$ and $n = 3$, as promised in the introduction.

Quadratic polynomials. Suppose that $F \in \mathbf{K}[X]$ is a monic quadratic polynomial with some non-constant coefficient and $p > 2$. Then F is a 2-nd pseudo-power if and only if $F = (X-f)^{[(p^r,p^s)]} = (X - f^{p^r})(X - f^{p^s})$ for some $f \in \mathbf{K} \smallsetminus \mathbf{K}^p$ and some integers $r \geq s \geq 0$. This is the same as saying that $F = (X - h^{p^m})(X - h)$ for some $h \in \mathbf{K} \smallsetminus \mathbf{k}$ and $m \geq 0$ (indeed, take $h = f^{p^s}$ and $m = r - s$). Therefore, in the case of quadratic polynomials Theorem 1 specializes to the main result of [15] and [18], with the following improved value of M:

$$M(\mathfrak{g}) = 8 \max\{\mathfrak{g}-1, 0\} + 21.$$

Cubic polynomials. Let $F \in \mathbf{K}[X]$ be a monic cubic polynomial with some non constant coefficient and suppose $p \neq 3$. One can see (from the definition) that F is a 3-rd pseudo-power if and only if one of the following holds:

- $F = (X-\lambda)(X-f)^{[p^r,p^s]} = (X-\lambda)(X-f^{p^r})(X-f^{p^s})$ with $\lambda \in \mathbf{k}$, $f \in \mathbf{K} \smallsetminus \mathbf{K}^p$ and some integers $r \geq s \geq 0$ satisfying $p^r + p^s \equiv 0 \mod 3$.
- $F = (X-f)^{[p^r,p^s,p^t]} = (X-f^{p^r})(X-f^{p^s})(X-f^{p^t})$ with $f \in \mathbf{K} \smallsetminus \mathbf{K}^p$ and some integers $r \geq s \geq t \geq 0$ satisfying $p^r + p^s + p^t \equiv 0 \mod 3$.

The value of M in Theorem 1 with $n = 3$ is

$$M(\mathfrak{g}) = 24 \max\{\mathfrak{g}-1, 0\} + 66.$$

We remark that no case of Theorem 1 with $n \geq 3$ was previously known.

5. Consequences in logic

Let κ be a field of characteristic $p > 0$ and let $\mathbf{k}$ be its algebraic closure. In the following, t denotes a transcendental element over κ (hence over $\mathbf{k}$). In this section we apply Theorem 1 to the rational function field $\mathbf{k}(t)$, and deduce some consequences in logic for $\kappa(t)$. From now on, we write $\mathbf{K} = \mathbf{k}(t)$.

As usual, the symbol $\models$ means that a structure satisfies a formula. Let $n \geq 2$ be an integer and let P_n be a unary predicate symbol. We interpret P_n in $\kappa(t)$ as follows: For $f \in \kappa(t)$, $P_n(f)$ holds in $\kappa(t)$ (i.e. $\kappa(t) \models P_n(f)$) if and only if f is an n-th power in $\kappa(t)$. We remark that P_n allows us to express that $f \in \kappa(t)$ is an n-th power of *some* element, but we cannot *a priori* express an n-th root of f (hence, we cannot take n-th powers just by using P_n).

Let us consider the language $\mathcal{L}_{t,n} = \{0, 1, +, =, P_n, t\cdot\}$ where $t\cdot$ is a unary function symbol. Then we can make $\kappa(t)$ into an $\mathcal{L}_{t,n}$-structure by interpreting P_n as indicated above, and $t\cdot$ as the multiplication by the indeterminate $t \in \kappa(t)$. Note that multiplication of arbitrary elements of $\kappa(t)$ is *not* part of the language. Our main application in logic of the notion of pseudo-powers and Theorem 1 is the following.

Theorem 5.1. *There is a positive-existential $\mathcal{L}_{t,n}$-formula $\mu_n[x,y,z]$ with free variables x,y,z and depending only on n, such that if the characteristic of κ is*

$$p > M := 11n^2 - 10n - 3$$

then μ_n defines the multiplication in $\kappa(t)$, that is

$$\kappa(t) \models \mu_n[f,g,h] \text{ if and only if } fg = h.$$

Note that the formula is uniform on p. A similar result holds for other function fields of positive characteristic upon choosing a suitable interpretation of the symbol $t\cdot$ as multiplication by a uniformizer at a given prime (but the formula will also depend on the genus) and for integrally closed κ-sub algebras of $\kappa(t)$ containing t. For clarity of the exposition we just consider $\kappa(t)$.

This theorem is the higher exponents generalization of the main logic result of [18] (which was for $n = 2$). However, our proof is different and, indeed, it gives a new proof in the case $n = 2$.

As a remarkable application of Theorem 5.1, we conclude from [13] (see also [20]):

Corollary 5.2. *If $p > M$ and q is a power of p, then the positive existential theory of $\mathbb{F}_q(t)$ over the language $\mathcal{L}_{t,n}$ is undecidable. Hence, there is no algorithm to solve the following decision problem:*

> *Given a system S of finitely many equations of the form*
>
> $$a_1X_1^n + a_2X_2^n + \ldots + a_rX_r^n + c_1Y_1 + \ldots + c_kY_k = b$$
>
> *with coefficients a_i, c_i, b in $\mathbb{F}_p[t]$ and unknowns X_i, Y_i, decide whether or not S has a solution in $\mathbb{F}_q(t)$.*

Proof. In [13] it is shown that the positive existential theory of $\mathbb{F}_q(t)$ over the language $\mathcal{L}_t$ is undecidable, where $\mathcal{L}_t = \{0, 1, t, +, \cdot, =\}$ it the language of rings augmented with a constant symbol t for the transcendental variable $t \in \mathbb{F}_q(t)$. Our language $\mathcal{L}_{t,n}$ can be used to define t in a positive existential way by $t \cdot 1$, and Theorem 5.1 shows that we can also define the multiplication $\cdot$ in a positive existential way over $\mathcal{L}_{t,n}$ provided that $p > M$. Therefore, the positive existential theory of $\mathbb{F}_q(t)$ over $\mathcal{L}_{t,n}$ is undecidable.

Given a positive existential $\mathcal{L}_{t,n}$-formula Θ without free variables, we can write systems of equations $S_1, \ldots, S_r$ of the type given in the statement of this corollary, such that $\mathbb{F}_q(t) \models \Theta$ if and only if at least one of the S_i has a solution in $\mathbb{F}_q(t)$. Indeed, write Θ as a disjunction of positive existential $\mathcal{L}_{t,n}$-formulas where $\vee$ is not used, and replace each variable U under a condition of the type $P_n(U)$ by X^n for a new variable X. After this observation, the consequence about undecidability of systems of equations follows. □

To deduce Theorem 2 from Corollary 5.2 is standard, and we do it at the end of this section.

Similar undecidability consequences can be obtained for other structures. For instance for $\kappa[t]$ (which is an integrally closed κ-sub algebra of $\kappa(t)$ containing t) by using results of Denef [3] – details are left to the reader.

The proof of Theorem 5.1 relies on the understanding of how the Frobenius factorization interacts with the following family of operators defined for integers ℓ:

$$(\cdot)_{(\ell)} : \mathbf{K}[X] \to \mathbf{K}[X], \qquad F \mapsto F_{(\ell)} = t^{\ell \deg F} F(X/t^\ell). \tag{10}$$

Explicitly, the operator acts on coefficients as follows:

$$(a_0X^n + a_1X^{n-1} + \ldots + a_{n-1}X + a_n)_{(\ell)} = a_0X^n + t^\ell a_1X^{n-1} + \ldots + t^{(n-1)\ell}a_{n-1}X + t^{n\ell}a_n.$$

We begin by recording some facts about these operators.

Lemma 5.3. *For each ℓ, the operator $(\cdot)_{(\ell)}$ is an automorphism of $\mathbf{K}[X]$ as multiplicative $\mathbf{K}$-monoid (by which we mean that it fixes $\mathbf{K}$ and preserves multiplication). Moreover it preserves degrees, and the rule $\ell \mapsto (\cdot)_{(\ell)}$ defines an injective group homomorphism $\mathbb{Z} \to \mathrm{Aut}_{\mathbf{K}-\mathrm{Mon}}\mathbf{K}[X]$.*

In particular:

Lemma 5.4. *For each ℓ, we have that $H \in \mathbf{K}[X]$ is irreducible if and only if $H_{(\ell)}$ is irreducible.*

Lemma 5.5. *Let $G_1, \ldots, G_c \in \mathbf{K}[X]$ be monic irreducible polynomials of degree $d_1, \ldots, d_c$ respectively. Let $d = d_1 + \ldots + d_c$ and assume that $d \leq p-2$. Then for some $1 \leq \ell \leq d+1$, we have that each $(G_i)_{(\ell)}$ is irreducible and either it has some coefficient in $\mathbf{K} \smallsetminus \mathbf{K}^p$, or it is the polynomial X.*

Proof. Write $G_i = X^{d_i} + a_{i,1}X^{d_i-1} + \ldots + a_{i,(d_i-1)}X + a_{i,d_i}$ and let $d^+ = \max d_i$. For each $1 \leq j \leq d^+$ define the set

$$R_j = \{v_0(a_{i,j}) : 1 \leq i \leq c \text{ and } a_{i,j} \neq 0\} \subseteq \mathbb{Z}$$

where v_0 denotes the vanishing order at $t = 0$ (recall that $a_{i,j} \in \mathbf{K}$). Consider the set of residue classes

$$\tilde{R} = \bigcup_{j=1}^{d^+} j^{-1} \cdot R_j \mod p \subseteq \mathbb{Z}/p\mathbb{Z}$$

which is well defined because $d^+ < p$, and observe that $\tilde{R}$ contains at most $d = \sum d_i$ elements. Therefore, there is $\ell \in \{1, 2, \ldots, d+1\}$ such that $\ell \mod p$ does not belong to $-\tilde{R}$. Hence, $\ell + \tilde{R}$ does not contain 0, which means that for every $j \geq 1$ we have that $v_0(t^{\ell j}a_{i,j})$ is not divisible by p (in particular $t^{\ell j}a_{i,j} \notin \mathbf{K}^p$), unless $a_{i,j} = 0$. The only polynomial which is monic, irreducible and all of its coefficients but the leading one are zero, is the polynomial X, so the result follows. □

It will be convenient to introduce the following slight variation of the concept of pseudo-power.

Definition 5.6. *We say that a monic polynomial $F \in \mathbf{K}[X]$ is an* n-th almost-power *if it is either an n-th pseudo-power or it belongs to $\mathbf{k}[X]$.*

Proposition 5.7. *Let $F \in \mathbf{K}[X]$ be a monic polynomial of degree $n \leq p-2$. Then $F_{(\ell)}$ is an n-th almost-power for each $1 \leq \ell \leq n+1$ if and only if F is an n-th power (that is, $F = (X+f)^n$ for some $f \in \mathbf{K}$).*

Proof. Apply the previous lemma to the collection of all monic irreducible factors G_j of F. Thus, there is some $1 \leq \ell_0 \leq n+1$ such that each $(G_j)_{(\ell_0)}$ is irreducible and either it is X or it has some coefficient in $\mathbf{K} \smallsetminus \mathbf{K}^p$. Hence, the Frobenius factorization of $F_{(\ell_0)}$ has exponents of the form $W_i = (1, 1, \ldots, 1)$ (if any) and has $P = X^h$ for some $h \geq 0$. Note that this Frobenius factorization is unique because $p > n$. We have two cases: $F_{(\ell_0)}$ has coefficients in $\mathbf{k}$ or it has some coefficient not in $\mathbf{k}$. The first case leads to $F_{(\ell_0)} = X^n$ (because $P = X^h$) while the second case means that $F_{(\ell_0)}$ is an n-th pseudo-power (recall that $F_{(\ell_0)}$ is an n-th almost-power by assumption). The latter case implies that actually there is only one $W_i = (1, \ldots, 1)$ with exactly n components and moreover $P = 1$ (because the degree of $F_{(\ell_0)}$ is n). In either case $F_{(\ell_0)}$ is an n-th power, and applying $(\cdot)_{(-\ell_0)}$ we conclude that F is an n-th power. The converse follows from Item (1) of Proposition 4.6 and the multiplicativity of the operators $(\cdot)_{(\ell)}$. □

With this at hand we can proceed to the proof of Theorem 5.1. First note that we can freely use the following in our positive existential $\mathcal{L}_{t,n}$-formulas, and still obtain formulas of this type that only depend on n:

- Subtraction (indeed, terms with negative signs can be moved to the other side of the equation where they appear).
- Multiplication by specific integers that only depend on n. For instance $3x$ is short-hand for $x+x+x$, and similarly we can use $2^n x$.
- Any constant, relation or function that has been already defined by a positive existential $\mathcal{L}_{t,n}$-formula that only depends on n.

The basic *strategy* is the same as the one originally devised by Büchi in the case of squares for $\mathbb{Z}$:

Step 1. Reduce the problem to find a positive existential $\mathcal{L}_{t,n}$-formula $\pi_n[x,y]$ (depending only on n) that defines the n-th power map $f \mapsto f^n$ in $\kappa(t)$. That is, $\kappa(t) \models \pi_n[f,g]$ if and only if $g = f^n$ in $\kappa(t)$.

Step 2. Use the solution to Büchi's problem to find such formula $\pi_n[x,y]$. Note that $\pi_n[x,y]$ is required to define the *graph* of the n-th power map, not just its image (unlike the predicate P_n); here is the main complication of the proof.

The first step is fairly standard (see for instance [11]): if we have such a formula $\pi_n[x,y]$ then we can take n-th powers in our positive existential $\mathcal{L}_{t,n}$-formulas. Note that there are integers $a \neq 0, a_0, \ldots, a_n$ depending only on n such that the following formal identity holds

$$aT^2 = a_0 T^n + a_1 (T+1)^n + \ldots + a_n (T+n)^n$$

because the $(T+i)^n$ $(0 \leq i \leq n)$ form a $\mathbb{Z}$-linearly independent set inside the $\mathbb{Z}$-module of polynomials in T with integers coefficients and degree at most n. Indeed, computing the Wronskian determinant (which in this case is essentially a Vandermonde determinant) one sees that the polynomials $(T+i)^n$ $(0 \leq i \leq n)$ form a basis of the $\mathbb{Q}$-vector space of polynomials in $\mathbb{Q}[T]$ of degree at most n.

It follows that we can use the squaring map in our positive existential $\mathcal{L}_{t,n}$-formulas, and now the formula $\mu_n[x,y,z]$ can be constructed by observing that in any domain of characteristic different from 2 one has

$$(x+y)^2 = x^2 + 2z + y^2 \text{ if and only if } z = xy.$$

On the other hand, the second step (finding $\pi_n[x,y]$) requires a different technique and has not been done in the literature for $n > 2$ in positive characteristic. The reason is that it was not clear what was the correct formulation of Büchi's problem in this setting, as explained in the introduction. Now we proceed to construct $\pi_n[x,y]$, by first constructing several intermediate formulas and studying their properties.

Consider the formula

$$p[u_1, \ldots, u_n, w_1, \ldots, w_n] : \qquad \bigwedge_{i=1}^{n} (u_i = i^n + i^{n-1} w_1 + \ldots + i w_{n-1} + w_n). \tag{11}$$

This formula satisfies:

Claim 5.8. *Given $f_1, \ldots, f_n, a_1, \ldots, a_n \in \kappa(t)$ we have that $\kappa(t) \models p[f_1, \ldots, f_n, a_1, \ldots, a_n]$ if and only if $f_i = F(i)$ for each $1 \leq i \leq n$, where $F(X) = X^n + a_1 X^{n-1} + \ldots + a_n \in \kappa(t)[X]$. Moreover, fixing the f_i uniquely determines the a_i and conversely.*

This follows from the fact that the Vandermonde determinant is invertible and $n < p$.

Let $\Delta^{(j)}$ be the j-th forward difference operator on sequences of length at least $j+1$ (with coordinates in an abelian group), which is inductively defined as follows: if $\mathbf{s}$ is a sequence of length at least 1 we

set $\Delta^{(0)}\mathbf{s} = \mathbf{s}$, and if $\mathbf{s} = (s_1, \dots, s_n)$ is a sequence of length at least $j+2$ (with $j \geq 0$) we set

$$\Delta^{(j+1)}\mathbf{s} = \left((\Delta^{(j)}\mathbf{s})_2 - (\Delta^{(j)}\mathbf{s})_1, \dots, (\Delta^{(j)}\mathbf{s})_{n-j} - (\Delta^{(j)}\mathbf{s})_{n-j-1}\right)$$

where $(\Delta^{(j)}\mathbf{s})_i$ denotes the i-th coordinate of $\Delta^{(j)}\mathbf{s}$. For instance, if $\mathbf{s} = (s_1, \dots s_4)$ we have $\Delta^{(1)}\mathbf{s} = (s_2 - s_1, s_3 - s_2, s_4 - s_3)$ and $\Delta^{(2)}\mathbf{s} = (s_3 - 2s_2 + s_1, s_4 - 2s_3 + s_2)$.

We will need the following basic properties of difference operators. Although these properties are well-known we could not find references, hence, we include proofs.

Lemma 5.9. *Let $k \leq d < N$ be positive integers. Let A be a commutative ring with unit. If $f(X) \in A[X]$ has degree d and leading coefficient $a \in A$, then there is a polynomial $g(X) \in A[X]$ of exact degree $d-k$ and leading coefficient*

$$d(d-1)\cdots(d-k+1)a$$

satisfying

$$\Delta^{(k)}(f(1), \dots, f(N)) = (g(1), \dots, g(N-k)).$$

Proof. Note that if $f(X) \in A[X]$ is a polynomial of degree $d \geq 1$ with leading coefficient a, then $f(X+1) - f(X)$ is a polynomial of degree $d-1$ with leading coefficient da. We also observe that applying $\Delta^{(k)}$ to a sequence is the same as applying $\Delta^{(1)}$ a total of k times. These two remarks give the result. □

Proposition 5.10. *Let A be an integral domain, and if* $\mathrm{char} A = 0$ *we assume that A contains $\mathbb{Q}$. Let $j \geq 0$ and $N \geq j+1$ be integers. Let $s_1, \dots, s_N$ be elements in A. Assume that there is a polynomial $Q(X) \in A[X]$ of exact degree d such that*

$$\Delta^{(j)}(s_1, \dots, s_N) = (Q(1), \dots, Q(N-j)).$$

We further assume that $(d+j)! \neq 0$ in A. Then there is a polynomial $P(X) \in A[X]$ of exact degree $d+j$ such that $s_i = P(i)$ for each $1 \leq i \leq N$. Moreover, if p and q are the leading coefficients of P and Q, then

$$q = (d+j)(d+j-1)\cdots(d+1)p$$

(for $j = 0$ this product is empty and we set it as 1).

Proof. For $j = 0$ there is nothing to prove. Let us first prove the result for $j = 1$. Write $Q(X) = qX^d + q_{d-1}X^{d-1} + \ldots + q_0$ and note that for $1 \leq i \leq N$

$$s_i - s_1 = \sum_{m=1}^{i-1} Q(m) = \sum_{m=1}^{i-1} (qm^d + q_{d-1}m^{d-1} + \ldots + q_0) = \frac{q}{d+1} i^{d+1} + h(i)$$

where $h(X)$ is a polynomial of degree at most d which depends only on Q, and satisfies $h(1) = -q/(d+1)$ (this condition is obtained by setting $i = 1$, and note that by assumption $d+1$ is invertible). Here, we used the well-known formulas for power sums

$$\sum_{m=1}^{i} m^r = \frac{i^{r+1}}{r+1} + (\text{lower degree terms})$$

(see for instance Theorem 3, p.384 [2]) that can be applied thanks to our assumption $(d+j)! \neq 0$ in A. The result for $j = 1$ follows, taking

$$P(X) = \frac{q}{d+1} X^{d+1} + h(X) + s_1. \tag{12}$$

The general case is proved by induction. Assume that the result is proved for $j = \ell$. Let $(s_i)_{i=1}^N$ be a sequence satisfying all the hypotheses of the result for $j = \ell + 1$. Consider the new sequence $s'_i = s_{i+1} - s_i$ for $1 \leq i \leq N' = N - 1$. We have $N' \geq \ell + 1$,

$$\Delta^{(\ell)}(s'_i)_i = \Delta^{(\ell+1)}(s_i)_i = (Q(1), \ldots, Q(N - (\ell + 1))) = (Q(1), \ldots, Q(N' - \ell))$$

and $(d + \ell)! \neq 0$ (because $(d + \ell + 1)! \neq 0$). Hence, by induction hypothesis there is a polynomial Q' of degree $d + \ell$ with $s'_i = Q'(i)$. Now we use the case $j = 1$ (already proved) applied to the sequence $(s_i)_i$ which satisfies

$$\Delta^{(1)}(s_i)_i = (s'_i)_i = (Q'(1), \ldots, Q'(N')) = (Q'(1), \ldots, Q'(N - 1)).$$

We obtain a polynomial P of degree $\deg(Q') + 1 = d + \ell + 1$ such that $s_i = P(i)$ for $1 \leq i \leq N$. The relation between leading coefficients also holds (using the induction hypothesis and (12)), proving the result. □

If we write an equality between sequences in our formulas, we mean the conjunction of all the coordinatewise equalities. With this notation we can define the positive existential $\mathcal{L}_{t,n}$-formula

$$b[u_1, \ldots, u_M] : \qquad \left(\bigwedge_{i=1}^{M} P_n(u_i) \right) \wedge \Delta^{(n)}(u_1, \ldots, u_M) = (n!, \ldots, n!)$$

with M as in the statement of Theorem 5.1. Note that this value of M is the same as in Theorem 1 with $\mathfrak{g} = 0$.

Claim 5.11. *We have that $\kappa(t) \models b[f_1, \ldots, f_M]$ if and only if all the f_i are n-th powers in $\kappa(t)$ and there is a monic polynomial $F \in \kappa(t)[X]$ of degree n such that $f_i = F(i)$ for each $1 \leq i \leq M$. The polynomial F is uniquely determined by the conditions $f_i = F(i)$ for $1 \leq i \leq M$ and it is an n-th almost-power in $\mathbf{K}[X]$.*

Proof. First assume that $\kappa(t) \models b[f_1, \ldots, f_M]$. The existence of $F \in \kappa(t)[X]$ monic of degree n with $F(i) = f_i$ follows from the previous proposition by taking $A = \kappa(t)$, $s_i = f_i$, $j = n$, $d = 0$ and $Q(X) = n!$ (note that $(d + j)! = n! \neq 0$ because $p > M > n$) because $\kappa(t) \models \Delta^{(n)}(f_1, \ldots, f_M) = (n!, \ldots, n!)$. The polynomial F is unique because $M > n$. Additionally, we have that each f_i is an n-th power in $\kappa(t)$ because $\kappa(t) \models \bigwedge_{i=1}^{M} P_n(f_i)$.

We still have to show that F is an n-th almost-power. Note that all the f_i are n-th powers in $\mathbf{K}$ because they are n-th powers in $\kappa(t)$. Hence, Theorem 1 implies that either $F \in \mathbf{k}[X]$ or F is an n-th pseudo-power in $\mathbf{K}[X]$. In either case we get that F is an n-th almost-power in $\mathbf{K}[X]$ (see Definition 5.6).

Conversely, suppose that all the f_i are n-th powers in $\kappa(t)$ and there is a monic polynomial $F \in \kappa(t)[X]$ of degree n such that $f_i = F(i)$ for each $1 \leq i \leq M$. The clause $\bigwedge_{i=1}^{M} P_n(f_i)$ is satisfied because the f_i are n-th powers. The clause $\Delta^{(n)}(f_1, \ldots, f_M) = (n!, \ldots, n!)$ is satisfied by Lemma 5.9 because $f_i = F(i)$ and F is a monic polynomial of degree n. □

Recall that $M > n$. Now we see that the formula (see (11))

$$B[u_1, \ldots, u_M, w_1, \ldots, w_n] : \qquad b[u_1, \ldots, u_M] \wedge p[u_1, \ldots, u_n, w_1, \ldots, w_n]$$

has the following property: if $\kappa(t) \models B[f_1, \ldots, f_M, a_1, \ldots, a_n]$ then the polynomial $F = X^n + a_1 X^{n-1} + \ldots + a_n \in \kappa(t)[X]$ is an n-th almost-power in $\mathbf{K}[X]$ and $f_i = F(i)$ for each $1 \leq i \leq M$. In particular, the formula

$$A[w_1, \ldots, w_n] : \qquad \exists u_1 \cdots \exists u_M B[u_1, \ldots, u_M, w_1, \ldots, w_n]$$

satisfies the following: if $\kappa(t) \models A[a_1, \dots, a_n]$ then the polynomial $F = X^n + a_1 X^{n-1} + \dots + a_n \in \kappa(t)[X]$ is an n-th almost-power in $\mathbf{K}[X]$. Now consider the formula

$$A_\ell[w_1, \dots, w_n] : \qquad A[t^\ell \cdot w_1, t^{2\ell} \cdot w_2, \dots, t^{n\ell} \cdot w_n]$$

defined for ℓ a positive integer, where $t^r \cdot$ denotes the symbol $t\cdot$ repeated r times. It follows that if $\kappa(t) \models A_\ell[a_1, \dots, a_n]$ then $F_{(\ell)}$ (defined in (10)) is an n-th almost-power in $\mathbf{K}[X]$, where $F = X^n + a_1 X^{n-1} + \dots + a_n \in \kappa(t)[X]$.

Consider the positive existential $\mathcal{L}_{t,n}$-formula

$$P[w_1, \dots, w_n] : \qquad \bigwedge_{\ell=1}^{n+1} A_\ell[w_1, \dots, w_n] \tag{13}$$

Claim 5.12. *Let $a_1, \dots, a_n \in \kappa(t)$. We have that $\kappa(t) \models P[a_1, \dots, a_n]$ if and only if the polynomial $F = X^n + a_1 X^{n-1} + \dots + a_n \in \kappa(t)[X]$ is of the form $F = (X + \nu)^n$ for some $\nu \in \kappa(t)$. This ν is given by $a_1 = n\nu$.*

Proof. If $\kappa(t) \models P[a_1, \dots, a_n]$ then by our discussion on A_ℓ we see that each $F_{(\ell)}$ is an n-th almost-power in $\mathbf{K}[X]$, for each $1 \leq \ell \leq n+1$. By Proposition 5.7 we conclude that there is $\nu \in \mathbf{K}$ such that $F = (X + \nu)^n$. Since

$$X^n + a_1 X^{n-1} + \dots + a_n = F = (X + \nu)^n = X^n + n\nu X^{n-1} + \dots + \nu^n$$

we see that $a_1 = n\nu$, hence, $\nu \in \kappa(t)$.

Conversely, if $X^n + a_1 X^{n-1} + \dots + a_n = F = (X + \nu)^n$ with $\nu \in \kappa(t)$ then each $F_{(\ell)}$ is monic and it is an n-th power in $\kappa(t)[X]$; more concretely

$$F_{(\ell)} = (X + t^\ell \nu)^n.$$

Hence $F_{(\ell)}(i)$ is an n-th power in $\kappa(t)$ for each ℓ and each i. In particular (see Claim 5.11) we get that $\kappa(t) \models P[a_1, \dots, a_n]$. □

After all this analysis, we can use the formulas p and P (see (11) and (13)) to finally define the positive existential $\mathcal{L}_{t,n}$-formula

$$\pi_n[x, y] : \quad \exists w_1 \cdots \exists w_n \exists u_1 \cdots \exists u_n P[w_1, \dots, w_n] \wedge p[u_1, \dots, u_n, w_1, \dots, w_n] \wedge y = u_1 \wedge w_1 + n = nx.$$

Claim 5.13. *Let $f, g \in \kappa(t)$. We have $\kappa(t) \models \pi_n[f, g]$ if and only if $g = f^n$.*

Proof. Suppose that $\kappa(t) \models \pi_n[f, g]$. We see that there is a polynomial $F = (X + \nu)^n = X^n + w_1 X^{n-1} + \dots + w_n \in \kappa(t)[X]$ with $\nu \in \kappa(t)$ (by Claim 5.12) such that the variables u_i appearing in $\pi_n[f, g]$ satisfy $F(i) = u_i$ for $1 \leq i \leq n$ (by Claim 5.8). In particular, since the clauses $g = u_1$ and $w_1 + n = nf$ hold, we see that $g = F(1) = (1 + \nu)^n$ and $n\nu + n = nf$, from which we obtain $g = f^n$.

Conversely, if $g = f^n$, then we define $\nu = f - 1$ and $F = (X + \nu)^n$. Let $a_i \in \kappa(t)$ (for $1 \leq i \leq n$) be defined by $F = X^n + a_1 X^{n-1} + \dots + a_n$. With this notation, $\pi_n[f, g]$ holds in $\kappa(t)$ by choosing $w_i = a_i$ and $u_i = F(i)$, for $1 \leq i \leq n$. The verification is immediate from Claim 5.8 and Claim 5.12. □

As explained before (see the discussion after Proposition 5.7), this claim concludes the proof of Theorem 5.1.

Finally, we present a proof of Theorem 2, using Corollary 5.2.

Proof of Theorem 2. Consider a system of equations S as in Corollary 5.2. For each i one can introduce new variables $X_{i,1}, X_{i,2}, \dots, X_{i,n}$ and define the diagonal form of degree n

$$D_i = \frac{1}{n!}\Delta^{(n-1)}(X_{i,1}^n, X_{i,2}^n, \dots, X_{i,n}^n) \in \mathbb{F}_p[X_{i,1}, X_{i,2}, \dots, X_{i,n}]$$

which is well-defined because $p > M > n$ (here is a slight abuse of notation: strictly speaking we wrote a sequence of length 1 whose only term is our form D_i). We claim that D_i is universal in $\mathbb{F}_q(t)$, i.e. it represents all the elements in $\mathbb{F}_q(t)$. Indeed, by Lemma 5.9 applied to the polynomial X^n (with $d = n$, $a = 1$ and $k = n - 1$) we have the identity (here, X is a variable)

$$D_i(X, X + 1, \dots, X + n) = X + b$$

for certain $b \in \mathbb{F}_p$. Hence, given $f \in \mathbb{F}_q(t)$ we have

$$D_i(f - b, f - b + 1, \dots, f - b + n) = f$$

which proves that D_i is universal in $\mathbb{F}_q(t)$.

Let us substitute, in each equation of S, the variable Y_i by the form D_i and call S' the new system. Since D_i is universal in $\mathbb{F}_q(t)$, we see that S' has a solution if and only if S has a solution (there is nothing special about D_i; any universal diagonal form of degree n would work).

If we have an algorithm $\mathcal{A}$ for the decision problem stated in Theorem 2 then we could apply it to S' and decide whether or not S' has solutions in $\mathbb{F}_q(t)$. Hence we would be able to decide whether or not S has solutions in $\mathbb{F}_q(t)$. After Corollary 5.2 this is not possible for all such systems S, and we conclude that the algorithm $\mathcal{A}$ cannot exist. This proves Theorem 2. □

Acknowledgments

We would like to thank Xavier Vidaux for carefully reading part of this manuscript and making helpful comments. We also want to express our deep gratitude to the anonymous referees for their excellent work in correcting several issues in a previous version of this manuscript, and for their detailed suggestions that improved the presentation of this work.

References

[1] T. T. H. An, H.-L. Huang, J. T.-Y. Wang; *Generalized Büchi's problem for algebraic functions and meromorphic functions*, Mathematische Zeitschrift **273** (2013), 95-122

[2] A. Borevich, I. Shafarevich; *Number theory.* Translated from the Russian by Newcomb Greenleaf. Pure and Applied Mathematics, Vol. 20 Academic Press, New York-London 1966 x+435 pp.

[3] J. Denef; *The diophantine problem for polynomial rings of positive characteristic*, Logic Colloquium **78**, M. Boffa, D. van Dalen, K. McAloon (eds.), North Holland, 131-145 (1979).

[4] N. Garcia-Fritz; *Representation of powers by polynomials and the language of powers*, J. Lond. Math. Soc. (2) 87 (2013), no. 2, 347-364.

[5] S. Lang, *Diophantine geometry*, Interscience Tracts in Pure and Applied Mathematics, No. 11 Interscience Publishers (a division of John Wiley & Sons), New York-London 1962 x+ 170 pp.

[6] L. Lipshitz; *Quadratic forms, the five square problem, and diophantine equations, The collected works of J. Richard Büchi* (S. MacLane and Dirk Siefkes, eds.) Springer, 677-680, (1990).

[7] Y. Matiyasevich, *Enumerable sets are diophantine*, Dokladii Akademii Nauk SSSR, 191 (1970), 279-282; English translation. Soviet Mathematics Doklady 11, 354-358 (1970).

[8] B. Mazur; *Questions of decidability and undecidability in number theory*, J. of Symbolic Logic **59-2**, 353-371 (1994).

[9] H. Pasten, T. Pheidas, X. Vidaux; *A survey on Büchi's problem: new presentations and open problems.* Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI) 377 (2010), Issledovaniya po Teorii Chisel. 10, 111–140, 243; translation in J. Math. Sci. (N. Y.) **171** (2010), no. 6, 765-781

[10] H. Pasten; *Representation of powers by polynomials over function fields and a problem of Logic,* unpublished (2011), available at http://arxiv.org/abs/1107.4019

[11] H. Pasten; *Powerful values of polynomials and a conjecture of Vojta*, J. Number Theory 133 (2013), no. 9, 2964-2998.

[12] H. Pasten; T. Pheidas; X. Vidaux; *Uniform existential interpretation of arithmetic in rings of functions of positive characteristic*, Invent. Math. (2013) DOI 10.1007/s00222-013-0472-1.
[13] T. Pheidas; *Hilbert's tenth problem for fields of rational functions over finite fields*, Invent. Math. **103**, 1-8 (1991).
[14] T. Pheidas, X. Vidaux; *The analogue of Büchi's problem for rational functions.* J. London Math. Soc. (2) **74** (2006), no. 3, 545-565.
[15] T. Pheidas, X. Vidaux; *Corrigendum: The analogue of Büchi's problem for rational functions.* J. London Math. Soc. (2) **82** (2010), no. 1, 273-278.
[16] T. Pheidas, X. Vidaux; *Extensions of Büchi's problem: questions of decidability for addition and kth powers.* Fund. Math. **185** (2005), no. 2, 171-194.
[17] T. Pheidas, X. Vidaux; *The analogue of Büchi's problem for cubes in rings of polynomials.* Pacific J. Math. **238** (2008), no. 2, 349-366.
[18] A. Shlapentokh, X. Vidaux; *The analogue of Büchi's problem for function fields.* J. Algebra **330** (2011), 482-506.
[19] H. Stichtenoth; *Algebraic function fields and codes.* Second edition. Graduate Texts in Mathematics, 254. Springer-Verlag, Berlin, 2009. xiv+355 pp. ISBN: 978-3-540-76877-7
[20] C. Videla; *Hilbert's tenth problem for rational function fields in characteristic* 2, Proceedings of the American Mathematical Society **120** (1994), no. 11, 249-253.
[21] P. Vojta; *Diagonal quadratic forms and Hilbert's tenth problem.* (English summary) Hilbert's tenth problem: relations with arithmetic and algebraic geometry (Ghent, 1999), 261-274, Contemp. Math., 270, Amer. Math. Soc., Providence, RI, 2000.
[22] J. T.-Y. Wang; *The truncated second main theorem of function fields.* J. Number Theory **58** (1996), no. 1, 139-157.
[23] J. T.-Y. Wang; *Hensley's problem for function fields.* Int. J. Number Theory Vol. 8, No. 2 (2012) 507-524.

Department of Mathematics and Statistics, Queen's University
Jeffery Hall, University ave.
Kingston, ON Canada, K7L 3N6
E-mail address, H. Pasten: hpasten@gmail.com

Institute of Mathematics, Academia Sinica
6F, Astronomy-Mathematics Building, No. 1, Sec. 4,
Roosevelt Road, Taipei10617, Taiwan
E-mail address, J. Wang: jwang@math.sinica.edu.tw