IRREDUCIBLE SPECIALIZATION IN GENUS 0

BRIAN CONRAD, KEITH CONRAD, AND ROBERT GROSS

Abstract. For irreducible $f(T) \in \mathbf{Z}[T]$, a classical conjecture predicts how often f has prime values. The frequency of these prime values is believed to be controlled by local obstructions. We discuss an analogue of this conjecture for irreducible $f(T) \in \kappa[u][T]$, with κ a finite field. Local obstructions are no longer sufficient. When f is inseparable over $\kappa(u)$, there is a new obstruction that is global, and it is quantified and effectively computable through the average of the Möbius function on specializations of $f(T)$.

We build on some results of Swan to prove the surprising fact that the "Möbius average" of $f(g)$ with $g \in \kappa[u]$ of large degree n has periodic behavior in n when f is inseparable over $\kappa(u)$, and that the periodicity is governed by the extrinsic geometry of the plane curve $f = 0$ over κ. We use the periodic Möbius average behavior in two ways: it enables us to show in specific examples that $f(g)$ is not irreducible as often as analogies with the classical case suggest, and we incorporate the Möbius periodicity into a modified conjecture for how often $f(g)$ is irreducible. The modified conjecture matches numerical data well.

1. Introduction

A well-known conjecture going back to Bouniakowsky [6] says that a nonconstant irreducible polynomial in $\mathbf{Z}[T]$ has infinitely many prime values in $\mathbf{Z}$ unless there is a divisibility obstruction, meaning that all values of the polynomial on $\mathbf{Z}$ are divisible by a nontrivial common factor. For example, $3T^2 - T + 2$ is irreducible in $\mathbf{Z}[T]$ but $3n^2 - n + 2$ is always even (and thus hardly ever prime) for $n \in \mathbf{Z}$.

Quantitatively, when $f(T) \in \mathbf{Z}[T]$ is nonconstant and irreducible with no divisibility obstructions, it is expected that

$$\#\{1 \le n \le x : f(n) \text{ prime}\} \stackrel{?}{\sim} \frac{C(f)}{\deg f}\frac{x}{\log x}, \tag{1.1}$$

where the constant $C(f)$ is a certain (nonzero) infinite product whose definition will be recalled in §2. The notation $\stackrel{?}{\sim}$ denotes a conjectural asymptotic relation. It is traditional to assume that f has a positive leading coefficient, but if negative prime values are allowed then this positivity condition on the leading coefficient of $f(T)$ is unnecessary. (The sampling range $1 \le n \le x$ is also traditional. It could be replaced with $|n| \le x$, after making an obvious change on the right side.)

The relation (1.1) is a conjecture due to Hardy and Littlewood [16] in special cases and Bateman and Horn [1, 2] more generally. The only proved case of (1.1) is in degree 1: the prime number theorem is the case $f(T) = T$ and Dirichlet's theorem is the case $f(T) = aT + b$ with a and b nonzero and relatively prime. While (1.1) can be extended to allow for several polynomials, such as twin-prime pairs, no version of the conjecture for several polynomials has been proved, even qualitatively.

Date: 2005-08-06, 10:00 AM.

1991 *Mathematics Subject Classification.* 11N32.

Key words and phrases. Bateman–Horn conjecture, Hardy–Littlewood conjecture, Möbius bias.

In this paper, we discuss an analogue of (1.1) in $\kappa[u][T]$ with κ a finite field. An extension of this work, with $\kappa[u]$ replaced by the coordinate ring of any smooth affine curve over κ with one geometric point at infinity, will be the subject of [10]. The proofs in [10] do not supersede the material here, but rather depend upon it, and in the case of higher genus we shall have to use geometric techniques that are not helpful in the case of genus 0.

The usual dictionary between $\mathbf{Z}$ and $\kappa[u]$ suggests that a polynomial $f(T)$ in $\kappa[u][T]$ that is nonconstant in T should have infinitely many prime (*i.e.*, irreducible) specializations on $\kappa[u]$ if and only if f is irreducible and f has no divisibility obstructions (*i.e.*, values of $f(T)$ on $\kappa[u]$ do not all share a common nontrivial factor). For the rest of this Introduction, it is assumed that $f \in \kappa[u][T]$ satisfies the previous three conditions: it has positive T-degree, it is irreducible in $\kappa[u][T]$, and it has no divisibility obstructions. We will call these the *Bouniakowsky* conditions. Setting $q = \#\kappa$, it is natural to guess that for such f,

$$\#\{g \in \kappa[u] : \deg g = n, f(g) \text{ prime}\} \stackrel{?}{\sim} \frac{C(f)}{\deg_T f} \frac{(q-1)q^n}{\log(q^n)} \tag{1.2}$$

as $n \to \infty$, where the constant $C(f) \neq 0$ is similar to the classical paradigm over $\mathbf{Z}$. (For the definition of $C(f)$, see (2.4).) Note that sampling in (1.2) is over all polynomials in $\kappa[u]$ of degree n, not just monics; this is why $q-1$ occurs in (1.2). Although it is traditional to believe that problems over $\mathbf{Z}$ become more accessible when they are reformulated over $\kappa[u]$, the only proved instance of (1.2) is $\deg_T f = 1$, just as in the classical situation. Counting g in (1.2) with $\deg g \leq n$ (or $(\deg g)|n$) instead of $\deg g = n$ does not simplify matters, and in fact we shall see that counting by separate degrees is essential for a proper understanding of the situation.

Numerical evidence supports (1.2) when f is separable over $\kappa(u)$, *e.g.*, when f is irreducible in $\kappa[T]$. The *raison d'être* of this paper is the discovery that (1.2) can be wrong when f is inseparable over $\kappa(u)$, *e.g.*, when $f(T) = T^p + u$. Thus, we call the right side of (1.2) the *naive estimate*. The rest of this Introduction provides compelling numerical evidence that (1.2) is not generally true and describes both proved counterexamples to (1.2) and our proposed correction to (1.2), relying on new nontrivial theorems about polynomials over finite fields.

Example 1.1. In Table 1.1, we count prime values of $f(g)$, where $f(T) = T^{12}+(u+1)T^6+u^4$ and g runs over polynomials of degree n in $\mathbf{F}_3[u]$, with $9 \leq n \leq 17$. (Here and in later examples, checking the Bouniakowsky conditions for f is left to the reader. All computations in this paper were carried out using PARI, NTL, and MAGMA, with deterministic primality testing.) An estimate for $C(f)$ is 3.52138375. Our range of degrees in Table 1.1 is small, but the sampling sets are substantial; *e.g.*, there are 9,565,938 polynomials of degree 14 in $\mathbf{F}_3[u]$. After each count of prime values in the table, we give the naive estimate for that count according to (1.2) and we give the ratio of these quantities. These data suggest the ratio tends to a number ≈ 1.33 rather than to 1. Incidentally, there is no point in searching for prime values of $f(g)$ in $\mathbf{F}_9[u]$ since $f(T)$ factors non-trivially in $\mathbf{F}_9(u)[T]$ as follows:

$$f(T) = (T^6 - 2i(u+2)T^3 + u^2)(T^6 + 2i(u+2)T^3 + u^2),$$

where $i^2 = -1$. Obviously $f(g)$ is reducible in $\mathbf{F}_9[u]$ for all $g \in \mathbf{F}_9[u]$.

Remark 1.2. To keep the presentation of data in our tables clean and informative, we round naive estimates (that is, the right side of (1.2)) to one digit after the decimal point — as a simple reminder that they are only estimates — and we round ratios between the two sides of (1.2) to three digits after the decimal point. Our policy has been to compute $C(f)$

to high enough accuracy to convince ourselves that we have correctly rounded all estimates presented in the tables; we have not worried about giving rigorous proofs of the correctness of the rounding in these tables, since the data in the tables merely serve to illustrate and motivate theorems and conjectures.

n	Count	Naive Est.	Ratio
9	1624	1168.3	1.390
10	4228	3154.5	1.340
11	11248	8603.2	1.307
12	31202	23658.7	1.319
13	87114	65516.5	1.330
14	244246	182510.2	1.338
15	683408	511028.6	1.337
16	1914254	1437268.0	1.332
17	5409728	4058168.4	1.333

TABLE 1.1. $T^{12} + (u+1)T^6 + u^4$ over $\mathbf{F}_3[u]$

Example 1.3. Let $f(T) = T^9 + (2u^2+u)T^6 + (2u+2)T^3 + u^2 + 2u + 1$ over $\mathbf{F}_3[u]$. Here $C(f) \approx 1.115866$. Table 1.2 suggests (1.2) is wrong if $n \equiv 1$ or $3 \bmod 4$: no irreducible values seem to occur when $n \equiv 1 \bmod 4$, while about twice as many irreducible values seem to occur as predicted by (1.2) when $n \equiv 3 \bmod 4$. The absence of prime $f(g)$ for $\deg g \equiv 1 \bmod 4$ is proved in Example 7.8.

If we look at irreducible values of $f(g)$ as g runs over $\mathbf{F}_9[u]$ instead of $\mathbf{F}_3[u]$, then data suggest the ratio of the two sides in (1.2) falls into a pattern of 2 interlaced convergent sequences rather than 4 as in Table 1.2.

n	Count	Naive Est.	Ratio
5	0	11.0	0
6	28	27.4	1.022
7	146	70.5	2.071
8	173	185.1	0.935
9	0	493.6	0
10	1345	1332.8	1.009
11	7348	3634.9	2.022
12	10138	9996.1	1.014
13	0	27681.4	0
14	77288	77112.5	1.002
15	432417	215915.0	2.003

TABLE 1.2. $T^9 + (2u^2+u)T^6 + (2u+2)T^3 + u^2 + 2u + 1$ over $\mathbf{F}_3[u]$

Example 1.4. Consider $T^8 + u^3$ over $\mathbf{F}_2[u]$. Although (1.2) predicts an exponentially growing number of prime values in each degree, $T^8 + u^3$ has no prime values on $\mathbf{F}_2[u]$! This is a special case of an example of Swan [22, p. 1102] from 1962, but the context of Swan's

work was sufficiently different from questions related to a $\kappa[u]$-analogue of (1.1) that a link between the two was not identified until now. In Example 3.13 we will give an example like Swan's in $\kappa[u]$ for any finite field κ.

Example 1.4 is surprising from a classical point of view, but we regard Example 1.3 as more instructive because it suggests that the ratio of the two sides of (1.2) can have interlaced limiting values as a periodic function of n.

Further numerical work leads to more non-constant polynomials $f(T)$ that do not appear to satisfy (1.2). We observed the following three common features of such polynomials:

- $f(T)$ is a polynomial in T^p, where p is the characteristic of κ.
- The ratio of the two sides in (1.2) appears to have 1, 2, or 4 limiting values as a function of n mod 4 when $n \to \infty$.
- The numbers $\mu(f(g))$, where μ is the Möbius function on $\kappa[u]$ (see Definition 3.1) and g runs over $\kappa[u]$, exhibit unusual statistics. Essentially, this means the nonzero values of $\mu(f(g))$ are not equally often 1 and -1. We call this idea the *Möbius bias*. One of the basic results in this paper is a theorem that lets us rigorously prove such a bias can occur for polynomials in T^p when $p \neq 2$ and in T^4 when $p = 2$.

For an algebraist, it is comforting to find apparent counterexamples to (1.2) only among polynomials in T^p, since irreducible polynomials in T^p are already well-known to exhibit peculiar algebraic properties in characteristic p. These are the irreducible $f \in \kappa[u][T]$ that have positive degree in T and are inseparable over the field $\kappa(u)$. While inseparable irreducibles have no classical analogue, there is no reason to dismiss them from consideration in (1.2). For instance, the nonvanishing of $C(f)$ in (1.2) is unrelated to whether or not $f(T)$ is inseparable in T. Moreover, (1.2) does look good for many inseparable irreducibles. A simple example is $T^p + u^2$ (see Example 3.12 and Table 7.1).

By studying apparent counterexamples to (1.2) in the context of our three observations above, we were led to a new heuristic idea: statistics for irreducible values of $f(g)$ as g varies are influenced by an *appropriate* average value of $\mu(f(g))$ as g varies. Averaging the Möbius bias in the right way enables us to predict the 1, 2, or 4 apparent limits suggested in all numerical examples that we have examined, and moreover these predicted values are effectively computable rational numbers. Whereas Bouniakowsky's divisibility obstruction is of a *local* character (a divisibility obstruction, by its definition, comes from divisibility by a common prime), the consideration of Möbius averages is fundamentally *global*. We are not aware of an explanation of the above phenomena in $\kappa[u]$ by a heuristic use of the circle method in characteristic p, and Möbius bias is not expected to occur for any primitive polynomial in $\mathbf{Z}[T]$ that is separable with positive degree in $\mathbf{Q}[T]$ (nor do we expect it to occur for any primitive polynomial in $\kappa[u][T]$ that is separable with positive T-degree in $\kappa(u)[T]$).

To illustrate our Möbius-bias heuristic for Example 1.1, let $f(T) = T^{12} + (u+1)T^6 + u^4$ over $\mathbf{F}_3[u]$. In Example 4.2 we will show

$$\mu(f(g)) = \left(\frac{g(0)^2(g(1)^2+1)}{3}\right) \tag{1.3}$$

for all g in $\mathbf{F}_3[u]$, where $(\frac{\cdot}{3})$ is the Legendre symbol. (The term $g(0)^2$ should not be omitted from the Legendre symbol, since it could be 0.) As g runs over polynomials of a given degree ≥ 2 in $\mathbf{F}_3[u]$, (1.3) shows that $\mu(f(g))$ is -1 twice as often as it is 1. The average nonzero value of $\mu(f(g))$ in each degree ≥ 2 is therefore $(-1-1+1)/3 = -1/3$ (not

just asymptotically, but exactly). Note that $1-(-1/3) = 4/3 = 1.33\ldots$ seems to fit the deviation from (1.2) in Table 1.1. Such agreement is purely numerical; we have no proof linking $\mu(f(g))$ to the primality statistics of $f(g)$.

Remark 1.5. Since the Möbius bias is a global parity condition on squarefree factorizations (with the squarefreeness of $f(g)$ considered to be a preliminary local condition), it is natural to ask if there are higher-order heuristic global obstructions to primality, such as a mod-3 condition on squarefree factorizations. We have studied many examples over small finite fields (of characteristics 2, 3, 5, and 7) and have found that the Möbius bias leads to a correction factor that gives an excellent numerical fit to all observed deviations from (1.2). Without guidance provided by examples giving evidence to the contrary, the Möbius-bias heuristic provides a satisfactory theory to account for all deviations from (1.2).

To convert our heuristic into a correction term in (1.2), we now describe some new theorems about the Möbius function on $\kappa[u]$. More accurately, our results concern the behavior of $\mu(f(g))$, where $f(T)$ is fixed in $\kappa[u][T^p]$ and g runs through $\kappa[u]$. Our study of $\mu(f(g))$ is inspired by counterexamples to (1.2), but we do not require that $f(T)$ be irreducible in $\kappa[u][T]$: we only need $f(T)$ to be squarefree, and (unlike irreducibility) this is a stable property under extension of the finite constant field ($f(T)$ in Example 1.1 is reducible in $\mathbf{F}_9[u][T]$, but still squarefree). Therefore we now fix $f(T) \in \kappa[u][T^p]$ that is squarefree in $\kappa[u][T]$ and, to avoid trivialities, we assume $f \notin \kappa$.

The key result, to be made precise in Theorem 1.7 below, is that $\mu(f(g))$ is essentially a periodic function of g and we can provide a formula for a modulus of periodicity. When $f(T)$ is monic in T, for instance, a modulus of periodicity is the radical of the $\kappa[u]$-resultant (this means the resultant of polynomials in T with coefficients in $\kappa[u]$) of $f(T)$ and the u-partial derivative $(\partial_u f)(T)$. As an example, let $f(T)$ be the polynomial in Example 1.1. The $\mathbf{F}_3[u]$-resultant of f and $\partial_u f$ is $u^{18}(u-1)^{18}$, whose radical is $u(u-1)$. This is consistent with (1.3), where we see that $\mu(f(g))$ depends on g modulo $u(u-1)$. To give a modulus of periodicity for $\mu(f(g))$ without a T-monicity restriction on f (and to prepare for the viewpoint that works in higher genus), we use geometric language as follows. Let $Z_f = \{f(u,T) = 0\}$ be the affine plane curve corresponding to $f \in \kappa[u,T]$. The projection from Z_f to the T-axis is flat and generically étale, so this projection is non-étale at a finite set of points on Z_f, say at the set B. Projecting B onto the u-axis gives a finite set of points. Define M_f^{geom} to be the monic polynomial in $\kappa[u]$ whose roots are this finite set on the u-axis, each root having multiplicity 1 (that is, M_f^{geom} is squarefree). We label this polynomial M_f^{geom} since it is not affected by replacing κ with a finite extension. In this monic case, this recovers the preceding $\kappa[u]$-resultant construction.

Remark 1.6. Concretely, an element u_0 in an algebraic closure of κ is a root of M_f^{geom} precisely when the specializations $f(u_0,T)$ and $(\partial_u f)(u_0,T)$ have a common T-root. This condition is the same as u_0 being a root of the $\kappa[u]$-resultant of f and $\partial_u f$ *only* when the u_0-specialization of either f or $\partial_u f$ has the same respective T-degree as f or $\partial_u f$. An equivalent description of this latter condition is: u_0 is not a double root of the leading coefficient of f as a polynomial in T.

For example, if $f(T)$ is monic in T then we conclude that M_f^{geom} is indeed the radical of the $\kappa[u]$-resultant of f and $\partial_u f$. For a contrast, let $f = u^2T^p + u + 1$ with $p \neq 2$; note that the leading coefficient of f as a polynomial in T has a double root at $u = 0$. The projection from Z_f to the T-axis is non-étale only at $(u_0, t_0) = (-2, 1/4)$, so $M_f^{\mathrm{geom}} = u+2$. However,

the $\kappa[u]$-resultant of f and $\partial_u f$ is $-u^p(u+2)$, and this has an extra root at 0 in comparison with M_f^{geom}.

The following theorem explains how M_f^{geom} is essentially a modulus of periodicity for $\mu(f(g))$, and that it is a minimal modulus of periodicity after a suitable finite extension of the constant field. In the theorem, the quadratic character of $\kappa^\times$ is denoted χ, with $\chi(0) = 0$. (A more accurate notation than $\mu(f(g))$ and χ is $\mu_{\kappa[u]}(f(g))$ and χ_κ, since the Möbius function and the quadratic character are sensitive to the choice of constant field κ.)

Theorem 1.7. *Let κ have odd characteristic p and $f(T) \in \kappa[u][T^p]$ be squarefree in $\kappa[u][T]$ and not lie in κ.*

There is a nonzero polynomial $M_{f,\kappa}$ in $\kappa[u]$ such that for $g_1 = c_1u^{n_1} + \cdots$ and $g_2 = c_2u^{n_2} + \cdots$ in $\kappa[u]$ with sufficiently large degrees n_1 and n_2,

$$g_1 \equiv g_2 \bmod M_{f,\kappa}, \quad n_1 \equiv n_2 \bmod 4, \quad \chi(c_1) = \chi(c_2) \Longrightarrow \mu(f(g_1)) = \mu(f(g_2)). \tag{1.4}$$

If -1 is a square in κ or $\deg_T f$ is even, the second congruence in (1.4) may be relaxed to $n_1 \equiv n_2 \bmod 2$.

One choice for the modulus $M_{f,\kappa}$ is M_f^{geom}. Using this choice, there is a lower bound on n_1 and n_2 beyond which (1.4) holds when κ is replaced by any finite extension.

The monic modulus $M_{f,\kappa}^{\min}$ of minimal degree in $\kappa[u]$ that makes (1.4) true for large n_1 and n_2 is a factor of any other $M_{f,\kappa}$. Moreover, there is a finite extension κ'/κ such that $M_{f,\kappa''}^{\min} = M_f^{\mathrm{geom}}$ whenever κ'' is a finite extension of κ'.

Motivated by our examples and the technical needs of proofs, throughout the paper we will keep track of the behavior of bounds and other parameters with respect to replacing κ with an arbitrary finite extension κ' while using the same f. In Example 4.9 we will give an f and κ such that $M_{f,\kappa}^{\min} \neq M_f^{\mathrm{geom}}$.

In the proof of Theorem 1.7, the importance of $f(T)$ being a polynomial in T^p is that its T-partial derivative is 0. That implies, for any $g \in \kappa[u]$, the u-derivative of $f(g(u)) \in \kappa[u]$ is $(\partial_u f)(g(u))$. In other words, $\partial_u(f(u, g(u))) = (\partial_u f)(u, g(u))$ if we consider f as a function of two variables u and T. Therefore the u-derivative of $f(g)$ is a polynomial in g with no dependence on $g'(u)$ in such cases.

Remark 1.8. From the geometric point of view, it is surprising to have an implication as in (1.4) that can relate polynomials g_j with different degrees. Since the quadratic nature of -1 in $\kappa^\times$ influences whether or not (1.4) depends on $\deg g \bmod 4$ or on $\deg g \bmod 2$, it seems unlikely that there can be a purely geometric proof of (1.4), although geometric ideas do play a prominent role in our proof.

Example 1.9. Let $f(T) = T^{12} + (u+1)T^6 + u^4$ in $\mathbf{F}_3[u][T]$, as in Example 1.1. Remark 1.6 and an earlier calculation imply that $M_f^{\mathrm{geom}} = u(u-1)$, so Theorem 1.7 says that, for $\deg g \gg 0$, 0 $\mu(f(g))$ depends on $g \bmod u(u-1)$, $\deg g \bmod 4$, and the quadratic character of the leading coefficient of g. This is consistent with (1.3). Viewing (1.3) in the context of Theorem 1.7, note that the mod-4 and quadratic-character conditions in (1.4) are unnecessary and the condition $\deg g \gg 0$ can be made explicit: $\deg g \geq 2$.

Remark 1.10. For both theoretical and numerical purposes, it would be useful to establish an explicit lower bound on n_1 and n_2 beyond which (1.4) holds even if κ is replaced by a finite extension. We do not have any theorems in this direction, but we expect that there should be a sufficient lower bound with order of magnitude $\deg M_f^{\mathrm{geom}}$.

Example 1.11. We return to the polynomial $f(T)$ in Example 1.3, but considered over $\kappa[u]$ for any finite κ of characteristic 3. As preparation for the proof of Theorem 1.7, in Example 4.3 we will show

$$\mu(f(g)) = (-1)^n(\chi(-1))^{n(n-1)/2}\chi(c)^{n+1}\chi(g(1)^2 + g(1) + 2)\chi(g(2)) \tag{1.5}$$

when $g = cu^n + \cdots$, with $n \geq 1$. Therefore $\mu(f(g))$ depends on $g \bmod (u-1)(u-2)$, $\deg g \bmod 4$, and the quadratic character of the leading coefficient c. (One checks that f and $\partial_u f$ have $\kappa[u]$-resultant $-(u-1)^6(u-2)^9$ whose monic radical is $(u-1)(u-2)$, so (1.5) and Remark 1.6 recover Theorem 1.7 in this case.) Formula (1.5) shows that Möbius behavior can change upon extension of the ground field: when -1 is a square in κ, the term $\chi(-1)^{n(n-1)/2}$ drops out, so dependence of $\mu(f(g))$ on $\deg g \bmod 4$ drops to dependence on $\deg g \bmod 2$.

The case of characteristic 2 lies deeper than the case of odd characteristic. Our treatment of characteristic 2 uses liftings to characteristic 0, via Witt vectors. (Readers not interested in characteristic 2 can skip ahead to the paragraph after Remark 1.13.) Here is an analogue of Theorem 1.7 in characteristic 2 for the case of polynomials in T^4; in §6 we will state and prove a more technical theorem that applies to polynomials in T^2.

Theorem 1.12. *Let κ be a finite field with characteristic 2. Fix a nonzero $f(T) \in \kappa[u][T^4]$ that is squarefree in $\kappa[u][T]$ and assume $f \notin \kappa$. There is a nonzero $M_{f,\kappa}$ in $\kappa[u]$ such that for $g \in \kappa[u]$ with sufficiently large degree, $\mu(f(g))$ is determined by $g \bmod M_{f,\kappa}$ and $\deg g \bmod 2$. If $[\kappa : \mathbf{F}_2]$ is even or $\deg_T f \equiv 0 \bmod 8$, then there is no dependence on $\deg g \bmod 2$.*

Let $W(\kappa)$ be the Witt vectors of κ. The modulus $M_{f,\kappa}$ may be chosen to be a polynomial that is the reduction of a certain geometrically-constructed squarefree polynomial in $W(\kappa)[u]$. For this choice of modulus, the "sufficient largeness" on $\deg g$ in the previous paragraph may be chosen uniformly with respect to finite extensions of κ.

An interesting example of Theorem 1.12 is $f(T) = T^8 + (u^3 + u)T^4 + u$ with κ of characteristic 2. For $g \in \kappa[u]$, the proof of Theorem 1.12 implies $M_{f,\kappa} = 1$ and $\mu(f(g)) = 1$ for $\deg g \gg 0$. Thus, (1.2) fails in this example. See Example 6.13 for further information.

It seems likely that the modulus $M_{f,\kappa}$ in Theorem 1.12 need not be squarefree, which is a contrast with Theorem 1.7. For example, when κ has characteristic 2 and

$$f(T) = T^{16} + (u^9 + u^4 + u^2 + u)T^8 + u^5 + u^3 \in \kappa[u][T],$$

then the proof of Theorem 1.12 yields

$$g_1 \equiv g_2 \bmod u^9(u+1)^4 \Rightarrow \mu(f(g_1)) = \mu(f(g_2)) \tag{1.6}$$

when $\deg g_j \geq 2$ (see Example 6.14), and numerical evidence suggests (but we cannot prove) that the modulus in (1.6) cannot be replaced with its radical, even if we restrict attention to $\deg g_j \gg 0$ instead of to $\deg g_j \geq 2$. Over some fields it seems probable that $u^9(u+1)^4$ in (1.6) can be replaced with a proper factor; for example, when $\kappa = \mathbf{F}_2$ the data suggest (but we cannot prove) that $u^3(u+1)$ may be used as a modulus in (1.6) when taking $\deg g_j \geq 2$. In fact, for $\kappa = \mathbf{F}_2$ it appears from the data that we can use $\deg g_j \geq 0$.

Remark 1.13. Numerical examples suggest that $\mu(f(g))$ is not always periodic in g when $f \in \kappa[u][T^2]$. More specifically, in Theorem 6.10 we will provide a general formula for $\mu(h(g^2))$ for $h \in \kappa[u][T]$ such that $h(T^2)$ is squarefree in $\kappa[u][T]$, and the formula involves the second symmetric function of the geometric residues of a certain rational 1-form depending

on g and h. When h is not a polynomial in T^2 (so $h(T^2)$ is not a polynomial in T^4) then this symmetric function of residues is generally nonzero, and for this reason it seems that new ideas are required to generalize Theorem 1.12 beyond the case $f \in \kappa[u][T^4]$.

Since Theorem 1.7 and Theorem 1.12 are rather unexpected from a classical point of view, let us indicate how we are able to prove such periodicity properties for the Möbius function. For simplicity, suppose $p \neq 2$. The starting point is an elementary formula of Swan over finite fields of odd characteristic: $\mu(\kappa[u]/(h)) = (-1)^n \chi(\mathrm{disc}_\kappa(h))$ for any nonzero $h \in \kappa[u]$ with degree n, where χ is the quadratic character on $\kappa^\times$ (vanishing at 0) and $\mathrm{disc}_\kappa(h) \in \kappa$ is the discriminant of the finite κ-algebra $\kappa[u]/(h)$ with respect to the basis $\{1, u, \ldots, u^{n-1}\}$. By expressing the discriminant in terms of the resultant against the derivative (see §4 for our conventions concerning resultants), in the special case $h = f(g)$ with f as in Theorem 1.7 we can exploit the property that $h' = (\partial_u f)(g)$ is also a polynomial in g because it identifies our problem with a special case of the setup in:

Theorem 1.14. *Let F be a perfect with with arbitrary characteristic and let $f_1, f_2 \in F[u,T]$ be nonzero elements whose zero loci Z_{f_1} and Z_{f_2} in the affine plane have finite intersection. For each $x = (u_x, t_x) \in Z_{f_1} \cap Z_{f_2}$, let $i_x(Z_{f_1}, Z_{f_2})$ be the local intersection number.*

There exist $c_0, c_1 \in F^\times$ and $m_0, m_1 \in \mathbf{Z}$ with $m_1 \geq 0$ such that for $g \in F[u]$ with sufficiently large degree n, the resultant $R_F(f_1(u,g), f_2(u,g))$ is given by

$$R_F(f_1(u,g), f_2(u,g)) = c_0 c_1^n (\mathrm{lead}\, g)^{m_0 + m_1 n} \prod_{x \in Z_{f_1} \cap Z_{f_2}} \mathrm{N}_{F(x)/F}(g(u_x) - t_x)^{i_x(Z_{f_1}, Z_{f_2})},$$

where $\mathrm{lead}\, g \in F^\times$ *is the leading coefficient of g.*

We prove this theorem as a universal algebraic identity (with $\deg g$ fixed and large). The intervention of n in this identity is only in two exponents, so upon applying the quadratic character in the case that $F = \kappa$ is finite with odd characteristic we get dependence on only n mod 2 and g mod M with $M = \prod_{u_x} \mathrm{N}_{F(u_x)/F}(u - u_x) \in F[u]$. The comparison between $\mathrm{disc}_F(h)$ and $R_F(h, h')$ involves the sign $(-1)^{(\deg h)(\deg h - 1)/2}$ that depends on $\deg h$ mod 4, and so in this way we are able to deduce Theorem 1.7 from Theorem 1.14. The proof of Theorem 1.14 is a recursive application of elementary properties of resultants and intersection numbers. Allowing the generality in F in Theorem 1.14 is useful: we apply it over 2-adic fields (characteristic 0!) in our study of characteristic 2. The key to generalizing our results to higher genus is to appropriately generalize Theorem 1.14, and even the formulation of such a generalization is not obvious because resultants are no longer available.

Returning to the faulty (1.2), we modify it as follows. Let $f(T)$ satisfy the Bouniakowsky conditions: f has positive T-degree, is irreducible in $\kappa[u][T]$, and has no divisibility obstructions. Assume also that $f(T)$ is a polynomial in T^p when $p \neq 2$ or in T^4 when $p = 2$. Define

$$\Lambda_{\kappa,M}(f;n) := 1 - \frac{\sum_{\deg g = n, \gcd(f(g), M) = 1} \mu(f(g))}{\sum_{\deg g = n, \gcd(f(g), M) = 1} |\mu(f(g))|}, \tag{1.7}$$

where $M \in \kappa[u]$ is any modulus $M_{f,\kappa}$ from Theorem 1.7 or Theorem 1.12; both sums run over g, and the denominator is nonzero for large n by Lemma 7.2. Note $\Lambda_{\kappa,M}(f;n)$ is a rational number in $[0,2]$.

There are two senses in which the sequence $\Lambda_{\kappa,M}(f;n)$ is independent of M:

(1) for any two choices of modulus M, the corresponding sequences $\Lambda_{\kappa,M}(f;n)$ agree for large n (see Theorem 7.5),

(2) in many (but not all!) examples, $\Lambda_{\kappa,M}(f;n) = \Lambda_{\kappa,1}(f;n)$ for large n (that is, the constraint $(f(g), M) = 1$ in (1.7) can be dropped), even when 1 is not a genuine modulus for $g \mapsto \mu(f(g))$.

At the end of Remark 7.11, we will give a general criterion for (2) to hold, which in particular applies to Example 1.1. (We will also explain in that remark why we use the condition $\gcd(f(g), M) = 1$ in the definition of $\Lambda_{\kappa,M}(f;n)$.) Because of (1), we may abbreviate $\Lambda_{\kappa,M}(f;n)$ to $\Lambda_\kappa(f;n)$, provided that the properties that we care about are limited to large n, as they usually are. The independence of the choice of M provides a robustness that makes the definition of (1.7) less sensitive to change in M than it may initially seem to be.

The marvelous fact (Theorem 7.5) is that $\Lambda_\kappa(f;n)$ is *periodic* in n with period 1, 2, or 4 for sufficiently large n; intuitively, this is a consequence of Theorems 1.7 and 1.12, and consequently $\Lambda_\kappa(f;n)$ is far simpler than it at first appears to be. This makes the following proposed correction to (1.2) simple to appreciate: when $f \in \kappa[u][T^p]$ satisfies the Bouniakowsky conditions, with the extra restriction that $f \in \kappa[u][T^4]$ when $p = 2$,

$$\#\{g \in \kappa[u] : \deg g = n, f(g) \text{ prime}\} \stackrel{?}{\sim} \Lambda_\kappa(f;n)\frac{C(f)}{\deg_T f}\frac{(q-1)q^n}{\log(q^n)} \tag{1.8}$$

as $n \to \infty$. If Theorem 1.12 can be generalized to allow $f \in \kappa[u][T^2]$ (see Remark 1.13) then it should be possible to formulate a version of (1.8) in characteristic 2 for any $f \in \kappa[u][T^2]$ that satisfies the Bouniakowsky conditions.

Example 1.15. Let $f_1(T)$ be the polynomial in Example 1.1 and $f_2(T)$ be the polynomial in Example 1.3. We will show in Example 7.7 that $\Lambda_{\mathbf{F}_3}(f_1;n) = 4/3$ for $n \geq 2$ and in Example 7.8 that $\Lambda_{\mathbf{F}_3}(f_2;n) = 0, 1, 2, 1, 0, 1, 2, 1, \ldots$ for $n \geq 1$, Thus, (1.8) appears to fix the discrepancies in the ratio columns for Tables 1.1 and 1.2 when we stay away from (periodic!) $n \gg 0$ such that $\Lambda_\kappa(f;n) = 0$.

Our numerical evidence suggests that the naive estimate (1.2) is correct for many T-inseparable f, and in some loose sense (1.2) seems to be correct much more often than it is incorrect. As a consistency check between (1.2) and (1.8), we have always been able to prove $\Lambda_\kappa(f;n) = 1$ for large n in examples where data suggest that (1.2) holds. Of course, as with any setting where a quantifiable obstruction may often turn out to be trivial, the point is that nontrivial examples really do occur.

The possibility that 0 lies in the period of $\Lambda_\kappa(f;n)$ requires a clarification on the meaning of (1.8) as an asymptotic relation. When 0 is in the period of $\Lambda_\kappa(f;n)$, what does (1.8) mean for the n's where $\Lambda_\kappa(f;n)$ is periodically 0? The vanishing of $\Lambda_\kappa(f;n)$ implies that for all g of degree n, either $\mu(f(g)) = 1$ or $(f(g), M_{f,\kappa}) \neq 1$. When n is large, both cases imply that $f(g)$ is reducible. Therefore, the appearance of 0 in the period for $\Lambda_\kappa(f;n)$ implies that both sides of (1.8) vanish for such n, which proves there is a periodic lack of irreducible specializations $f(g)$. For instance, the mod-4 patterns of 0's in Table 1.2 provably continues for all larger n. For other large degrees n, where $\Lambda_\kappa(f;n) \neq 0$, we only conjecture that (1.8) is a genuine asymptotic relation.

When κ has characteristic $p \neq 2$ and $f(T) \in \kappa[u][T]$ is irreducible with positive T-degree, we believe the correct $\kappa[u]$-variant on Bouniakowsky's (qualitative) conjecture is the following: $f(g)$ is irreducible for infinitely many $g \in \kappa[u]$ except in the following two cases: $f(T)$ has a divisibility obstruction or $f(T)$ is a polynomial in T^p with $\Lambda_\kappa(f;n) = 0$ for $n \gg 0$. Both types of obstructions can be checked with a finite amount of computation. An example which fits the second case but not the first is $f(T) = T^{4p} + u$; for any nonconstant g

in $\kappa[u]$, $g^{4p}+u$ is reducible. For details, see Example 3.11. (We do not make any analogous conjecture in characteristic 2 because the case of characteristic 2 is still not adequately understood when f is a polynomial in T^2 but not a polynomial in T^4.)

In order that our results are not misunderstood, we want to stress that when n runs through a sequence in which $\Lambda_\kappa(f;n)$ does not vanish, we do not prove a connection between $\Lambda_\kappa(f;n)$ and irreducibility counts for $f(g)$ with $\deg g = n$. All we can say is that numerics in those cases suggest that (1.8) holds.

Numerical examples were extremely important in this work. Without them, most of the nontrivial phenomena in this paper would not have been discovered. While examples in the paper concentrate on the finite fields $\mathbf{F}_2$ and $\mathbf{F}_3$, we did testing over $\mathbf{F}_5$ (and some non-prime finite fields) as well. Moreover, nonmonic polynomials played a more prominent role in our numerical work than the examples included here may suggest.

It is natural to ask how often Möbius bias occurs. The answer seems to be that it is rare. We noted above that, as a matter of experience, for generic $f \in \kappa[u,T^p]$ that is squarefree in $\kappa[u,T]$ we have $\Lambda_\kappa(f;n) = 1$ for all large n. It is unclear how to formulate and prove a precise theorem along these lines with a fixed finite field κ and varying f, but we have the following related result in odd characteristic:

Theorem 1.16. *Suppose $p \neq 2$ and let $f \in \kappa[u][T^p]$ be squarefree in $\kappa[u][T]$ with $\deg_T f > 0$. Fix a congruence class $c \in \mathbf{Z}/4\mathbf{Z}$, and for any finite extension κ'/κ let $\lambda_{\kappa'}(f;c)$ be the common value of $\Lambda_{\kappa'}(f;n)$ for large n in the class c. As $[\kappa' : \kappa] \to \infty$, $\lambda_{\kappa'}(f;c)$ has a limiting value of 0, 1, or 2. Moreover, for "generic" f the limiting value is 1.*

Examples 7.7 and 7.8 illustrate this theorem with $\kappa = \mathbf{F}_3$, and the meaning of genericity for f is that the local intersection number at some point in the finite intersection of the zero loci of f and $\partial_u f$ is odd; we shall see in Example 5.11 that, roughly speaking, for generic f each of these intersection numbers is a power of p. We will not prove Theorem 1.16 in this paper, as it is the genus-0 case of a more general result [10, Cor. 8.5] that is proved by using the Lang–Weil estimate and a link between the Λ_κ's and ζ-functions; this link follows from a higher-genus version of Theorem 1.14. In [10, Thm. 11.14] we shall prove an analogue of Theorem 1.16 for $p = 2$ and $f \in \kappa[u][T^4]$, and also for higher-genus coordinate rings in the role of $\kappa[u]$. To summarize, the phenomenon of inseparable counterexamples to (1.2) seems to be non-generic as the constant field grows.

Our work in characteristic p suggests a link between some classical conjectures in analytic number theory. Since (1.1) is not expected to have any counterexamples, and counterexamples to (1.2) in characteristic p appear to be explained by non-vanishing Möbius averages, it seems reasonable to conjecture that if $f(T)$ is irreducible (or more generally is not a perfect square up to sign) in $\mathbf{Z}[T]$ and has no divisibility obstructions then its Möbius average vanishes:

$$\frac{\sum_{n\leq x}\mu(f(n))}{\sum_{n\leq x}|\mu(f(n))|} \to 0 \tag{1.9}$$

as $x \to \infty$. By [14], the *abc*-conjecture implies that (1.9) is equivalent to

$$\sum_{n\leq x}\mu(f(n)) = o(x). \tag{1.10}$$

For linear f, (1.10) is true [21]. Numerical evidence for (1.10) in other cases is encouraging [8, Table 7]. After being led to (1.10) by analogy with our work in characteristic p, we learned that it is a folklore conjecture. The case $f(T) = T^2+1$ is posed in [12, p. 417].

The way that we were just led to (1.10) from our work in characteristic p suggests that any counterexample to (1.10) in $\mathbf{Z}[T]$ is probably a counterexample to (1.1). That is, the truth of (1.1) should imply (1.9) and (1.10). Can such an implication be proved, perhaps assuming some other standard conjectures?

Diophantine consequences of the failure of (1.2) are worth exploring. For instance, the related failure of (1.9) in characteristic p leads to an interesting family of elliptic curves [11].

Here is an outline of the paper. In §2, we discuss the constant in (1.1) and the obvious analogue of (1.1) in $\kappa[u][T]$. This obvious analogue is not true. In §3 and §4 we build on work of Swan to develop an understanding of $\mu(f(g))$ as in Theorem 1.7, which we restate as Theorem 4.7. The proof of Theorem 1.14 is given in §5, where we also deduce Theorem 4.7. In particular, we only begin to prove some non-trivial results in §5*ff* (for odd characteristic); the development in §2–§4 is largely a discussion of examples and some classical facts. Since the phenomena we study are unrelated to any classical ideas concerning prime values of polynomials, we feel that this preliminary discussion will help the reader to understand the nature of the theorems that we prove in §5*ff*.

In §6 we treat characteristic 2, which is much more difficult. Theorem 1.12 appears in a more precise form as Theorem 6.10 and Corollary 6.11. Its proof uses ideas from our treatment of odd characteristic and some considerations with residues of differential forms on the projective lines in characteristic 2 and over 2-adic fields. (The higher genus analogue of this work in characteristic 2 in [10] uses formal and rigid geometry, in addition to algebraic geometry.) Finally, §7 returns to conjectures, discussing the new factor $\Lambda_\kappa(f;n)$ in (1.8). This leads to our modified $\kappa[u]$-conjecture, given as Conjecture 7.9.

Notation and Terminology. Throughout the paper, κ denotes a finite field of size q. For nonzero $g \in \kappa[u]$, we set $\mathrm{N}g = q^{\deg g}$. We often let μ, rather than $\mu_{\kappa[u]}$, denote the Möbius function on $\kappa[u]$, relying on the context to make clear the ring in which we are computing the Möbius function; see Definition 3.1. We likewise often write χ instead of χ_κ to denote the quadratic character on the multiplicative group $\kappa^\times$ of a finite field with odd characteristic.

We write a typical polynomial in $\kappa[u][T]$ as $f(T)$, suppressing the dependence on u in the notation to make analogies to the classical situation more apparent. When, for geometric and other reasons, we want to make the u-dependence explicit, we write $f(T)$ as $f(u,T)$.

For a nonzero polynomial h in one variable, we write the leading coefficient as $\operatorname{lead} h$. For a nonzero polynomial f in two variables u and T over a ring R, the T-degree of f and the leading coefficient of f as a polynomial in T are indicated with a subscript: $\deg_T f \geq 0$ and $\operatorname{lead}_T f \in R[u]$. An element in $R[u]$ is *primitive* when its coefficients generate the unit ideal in R. For a domain K, the discriminant of a one-variable polynomial with coefficients in K is denoted $\operatorname{disc} h$, or $\operatorname{disc}_K h$ for emphasis. Our definition of discriminants does not match the usual definition when the polynomial is not monic; see (3.1) and (3.2). Our notation for resultants is introduced in §4.

All algebras in this paper are assumed to be commutative.

When R is a local ring with residue field k, a *lift* of a polynomial $h \in k[u,T]$ is a polynomial $H \in R[u,T]$ whose reduction to $k[u,T]$ is h.

Acknowledgments. We thank C. Elsholtz, O. Gabber, A. Granville, A. J. de Jong, M. Larsen, B. Poonen, A. Silberstein, and H. Stark for their advice and encouragement. We are also very grateful to an anonymous referee who offered helpful suggestions on an earlier submission of this work. B.C. thanks the NSF and the Alfred P. Sloan Foundation

for support during work on this paper. K.C. thanks the Clay Mathematics Institute and the Number Theory Foundation.

2. The Classical and Naive Conjectures

This section is intended for readers who are unfamiliar with conjectures like (1.1), and it also serves to fix some terminology.

For $h(T) \in \mathbf{Z}[T]$ and prime p, let

$$\pi_h(x) = \#\{1 \le n \le x : h(n) \text{ is prime}\}.$$

and

$$\omega_h(p) := \#\{\overline{n} \in \mathbf{Z}/(p) : h(n) \equiv 0 \bmod p\}.$$

The "probability" that $h(n)$ is not a multiple of p, as n runs over $\mathbf{Z}$, is $1 - \omega_h(p)/p$. When $\omega_h(p) = p$, *i.e.*, the function $h\colon \mathbf{Z} \to \mathbf{Z}/(p)$ is identically zero, we say h has a *local obstruction* at p. (A polynomial h that has no local obstructions must be primitive. For any primitive h, the only primes p at which h can have a local obstruction are those $p \le \deg h$.)

Conjecture 2.1 (Bateman–Horn, Hardy–Littlewood)**.** *Pick $f(T) \in \mathbf{Z}[T]$. Assume the following two conditions:*

1) *$f(T)$ is irreducible in $\mathbf{Q}[T]$.*
2) *$f(T)$ has no local obstructions, i.e., $\omega_f(p) < p$ for all p.*

Then

$$\pi_f(x) \stackrel{?}{\sim} C(f) {\sum_{n \le x}}' \frac{1}{\log|f(n)|} \sim \frac{C(f)}{\deg f} \frac{x}{\log x}. \tag{2.1}$$

where

$$C(f) = \prod_p \frac{1 - \omega_f(p)/p}{1 - 1/p} \tag{2.2}$$

and the $'$ in the summation indicates that we sum only over n large enough so that $|f(n)| > 1$.

The second hypothesis in Conjecture 2.1 is equivalent to f having a pair of relatively prime values, which is how the second hypothesis is checked in practice. The infinite product $C(f)$, taken in order of increasing p, is usually only conditionally convergent.

We turn now to a $\kappa[u]$-analogue of Conjecture 2.1. Pick $f \in \kappa[u][T]$ with $\deg_T f > 0$. Say $f(T)$ has a *local obstruction* at an irreducible $\pi \in \kappa[u]$ when $f(g) \equiv 0 \bmod \pi$ for all $g \in \kappa[u]$. In practice, one checks that $f(T)$ has no local obstructions by finding two specializations of $f(T)$ on $\kappa[u]$ that are relatively prime. Often $T = 0$ and $T = 1$ suffice.

Suppose $f(T) \in \kappa[u][T]$ is irreducible over $\kappa(u)$ and has no local obstructions. Define

$$\pi_f(n) = \#\{g \in \kappa[u] : \deg g = n, f(g) \text{ is irreducible}\}.$$

A conjecture analogous to (2.1) is

$$\pi_f(n) \stackrel{?}{\sim} \frac{C(f)}{\deg_T f} \frac{(q-1)q^n}{\log(q^n)}, \tag{2.3}$$

where $\mathrm{N}h = q^{\deg h}$ and

$$C(f) = \log q \cdot \prod_{(\pi)} \frac{1 - \omega_f(\pi)/\mathrm{N}\pi}{1 - 1/\mathrm{N}\pi}, \quad \omega_f(\pi) = \#\{g \bmod \pi : f(g) \equiv 0 \bmod \pi\}, \tag{2.4}$$

the product running over nonzero prime ideals in $\kappa[u]$. We could write $C(f)$ as $C_{\kappa[u]}(f)$ to emphasize the base ring $\kappa[u]$ (especially the choice of κ).

We call (2.3) the *naive conjecture* over $\kappa[u]$. It is an obvious conjecture to make, but in the Introduction we saw it is wrong: apparent counterexamples were provided. (Incidentally, the standard version of Conjecture 2.1 allows for simultaneous primality of several polynomials, such as twin prime pairs, and it is trivial to adapt that broader conjecture to a multi-polynomial naive conjecture over $\kappa[u]$. This is a false conjecture.)

It will be useful later to record the simple formulas for the degree and leading coefficient of $f(g)$ when $\deg g \gg 0$, and to make the condition $\deg g \gg 0$ effective. Write

$$f(T) = \alpha_d(u)T^d + \alpha_{d-1}(u)T^{d-1} + \cdots + \alpha_0(u),$$

with $\alpha_d(u) \neq 0$ and $d > 0$. For $g \in \kappa[u]$ with sufficiently large degree (depending on f), the degree and leading coefficient of $f(g)$ in $\kappa[u]$ are the same as those for $\alpha_d g^d$:

$$\deg(f(g)) = d \cdot \deg g + \deg \alpha_d = (\deg_T f)n + \deg(\mathrm{lead}_T f), \tag{2.5}$$

$$\mathrm{lead}(f(g)) = (\mathrm{lead}\, \alpha_d)(\mathrm{lead}\, g)^d, \tag{2.6}$$

where $n = \deg g$. In particular, $\deg(f(g))$ is a linear polynomial in $\deg g$ when $\deg g \gg 0$. In Section 5 it will be useful to have an explicit lower bound in terms of f such that (2.5) and (2.6) apply for $\deg g$ above this bound. Such a bound is

$$\nu(f) = \max_{0 \leq i \leq d-1} \frac{\deg \alpha_i - \deg \alpha_d}{d - i}. \tag{2.7}$$

In this maximum, terms with $\alpha_i = 0$ are omitted, or use the convention that $\deg 0 = -\infty$. For completeness, when $f(T) = \alpha(u)T^d$ is a T-monomial, take $\nu(f) = 0$.

Although it won't be used in this paper, we record here the analogue of $C(f) = C_{\kappa[u]}(f)$ for polynomials with S-integer coefficients. Let $\mathscr{O}_{K,S}$ be a ring of S-integers for a global field K, with S containing the set S_∞ of archimedean places in the number-field case. Let $f \in \mathscr{O}_{K,S}[T]$ be irreducible in $K[T]$ with no local obstructions at places on $\mathscr{O}_{K,S}$. (The last condition means that f defines a non-zero function on the residue field of each place outside of S.) Define

$$C(f) = \frac{1}{\mathrm{Res}(\mathscr{O}_{K,S})} \prod_{v \notin S} \frac{1 - \omega_f(v)/\,\mathrm{N}v}{1 - 1/\,\mathrm{N}v}, \tag{2.8}$$

where $\mathrm{Res}(\mathscr{O}_{K,S})$ denotes the residue at $s = 1$ for the zeta-function $\zeta_{K,S}$ of $\mathrm{Spec}(\mathscr{O}_{K,S})$. Such numbers are called *Hardy–Littlewood constants*, and agree with (2.2) and (2.4) for $\mathscr{O}_{K,S} = \mathbf{Z}$ and $\mathscr{O}_{K,S} = \kappa[u]$, *e.g.*, $\mathrm{Res}_{s=1}\, \zeta_{\kappa[u]}(s) = 1/\log q$. The multiplication in (2.8) is carried out according to increasing values of $\mathrm{N}v$, with all $v \notin S$ of a given norm being introduced into the product at the same time. The convergence of the product in (2.8) is absolute if and only if f is linear in characteristic 0 or a linear polynomial in some T^{p^m} in characteristic p.

3. The Möbius function over finite fields

In the Introduction we gave a heuristic explanation of the data in Example 1.1 as an effect of a Möbius bias. We speak of a Möbius bias when $\mu_{\kappa[u]}(f(g))$ does not take its nonzero values 1 and -1 equally often on average as g varies. In this section, we begin the systematic investigation of Möbius fluctuations in characteristic p, with the ultimate goal of using this work to correct the faulty (1.2). The first step in the analysis of $\mu_{\kappa[u]}(f(g))$

as g varies is the description of a formula for $\mu_{\kappa[u]}(h)$ ($h \in \kappa[u]$) other than its definition; the existence of an alternative Möbius formula on $\kappa[u]$ has no analogue in $\mathbf{Z}$. We will then apply the formula to compute $\mu_{\kappa[u]}(f(g))$ for varying g in some examples.

Definition 3.1. Let R be a Dedekind domain. The *Möbius function* on nonzero ideals of R is given by $\mu_R(\mathfrak{p}_1 \cdots \mathfrak{p}_m) = (-1)^m$ for distinct nonzero prime ideals $\mathfrak{p}_j$, $\mu_R((1)) = 1$, and $\mu_R(\mathfrak{b}) = 0$ for any nonzero ideal $\mathfrak{b} \subseteq R$ divisible by the square of a prime. For nonzero $r \in R$, we define $\mu_R(r) = \mu_R(rR)$. If R is understood from context, we write μ rather than μ_R.

When F is a field and h in $F[u]$ is nonconstant of degree d with roots $\gamma_1, \ldots, \gamma_d$ (counted with multiplicity) in a splitting field, we define the discriminant of h to be

$$\operatorname{disc} h := \prod_{i<j} (\gamma_i - \gamma_j)^2 \in F, \tag{3.1}$$

whether or not h is monic. (For nonzero constant h, the empty product is understood to be 1.) In terms of the derivative of h, (3.1) is the same as

$$\operatorname{disc} h = \frac{(-1)^{d(d-1)/2}}{(\operatorname{lead} h)^d} \prod_{i=1}^{d} h'(\gamma_i). \tag{3.2}$$

The factor $(\operatorname{lead} h)^d$ in (3.2) reflects our definition of discriminants of nonmonic polynomials in (3.1). When h is not monic, a variant on (3.1) is often used in the literature to define $\operatorname{disc} h$ (*e.g.*, [17, p. 204]). This variant equals (3.1) multiplied by $(\operatorname{lead} h)^{2d-2}$. In particular, the two competing definitions of the discriminant of a polynomial differ by a nonzero square factor in $F^\times$. We prefer (3.1) for nonmonic h since it agrees with the universally accepted definition of $\operatorname{disc}_F(F[u]/(h)) \in F$ relative to the ordered basis $\{1, u, \ldots, u^{d-1}\}$.

A generalization of the discriminant of a nonzero polynomial over a field F is the discriminant $\operatorname{disc}_F A$ of a finite F-algebra A. Such discriminants are only well-defined up to multiplication by squares in $F^\times$ due to variation in the choice of F-basis of A. We do not define the discriminant of the zero polynomial, just as the discriminant is not defined for an F-algebra with infinite dimension as an F-vector space.

Definition 3.2. Let κ be a finite field. For a finite κ-algebra A, let $\mu(A) = (-1)^{\#\operatorname{Spec} A}$ if A is étale over κ (*i.e.*, reduced) and let $\mu(A) = 0$ otherwise.

Note that $\mu(A)$ only depends on the underlying ring structure of A and not on its κ-algebra structure. If $h \in \kappa[u]$ is nonzero, then $\mu(\kappa[u]/(h)) = \mu_{\kappa[u]}(h)$. The following elementary result extends an observation of Swan.

Theorem 3.3. *Suppose κ is finite with odd characteristic, and let χ_κ be the quadratic character on $\kappa^\times$, with $\chi_\kappa(0) = 0$. For any finite κ-algebra A,*

$$\mu(A) = (-1)^{\dim_\kappa A} \chi_\kappa(\operatorname{disc}_\kappa A). \tag{3.3}$$

Proof. Both sides of (3.3) vanish when A is not étale over κ, so we may assume A is étale over κ. Both sides are multiplicative in A. The case $A = 0$ is trivial, so we reduce to the case when $A = \kappa'$ is a finite extension field of κ, and we want to prove

$$\chi_\kappa(\operatorname{disc}_\kappa \kappa') = (-1)^{d-1} \tag{3.4}$$

in $\mathbf{Z}$, where $d = [\kappa' : \kappa]$. Let γ be a field generator for κ' over κ. Since κ does not have characteristic 2, $\operatorname{disc}_\kappa \kappa'$ is a square in κ precisely when a generator for $\operatorname{Gal}(\kappa'/\kappa)$ acts as an

even permutation on the κ-conjugates of γ. Since this permutation of the roots is a d-cycle, its sign is $(-1)^{d-1}$. ■

Remark 3.4. Theorem 3.3 and its proof carry over *verbatim* to finite algebras over any perfect field k with characteristic not 2 having only cyclic Galois extensions; *e.g.*, we could take $k = \mathbf{C}((X))$. See [20, XIII, Exercise 3] for artificial examples in positive characteristic.

The proof of Theorem 3.3 works for étale algebras A in characteristic 2 if we formulate the result in terms of signs of certain permutations rather than in terms of quadratic characters of certain discriminants. (See [13, p. 237] for an application of this idea.) For our purposes, the role of discriminants is critical and therefore we need an analogue of Theorem 3.3 in characteristic 2 that involves discriminants. This analogue will use a lifting of A into characteristic 0. We shall now formulate a setup for finite κ with arbitrary characteristic (which for odd characteristic will recover a reformulation of Theorem 3.3).

Let κ be any finite field (of characteristic p, say), F the unramified extension of $\mathbf{Q}_p$ with residue field κ, and $W = W(\kappa)$ the valuation ring of F. (In other words, W is the ring of Witt vectors of κ.) We extend Theorem 3.3 to all characteristics by using finite flat liftings of A over W; *i.e.*, finite flat W-algebras $\widetilde{A}$ such that $\widetilde{A}/p\widetilde{A}$ is isomorphic to A as κ-algebras. For instance, a finite flat lifting of $\kappa[u]/(h(u))$ over W is $W[u]/(H(u))$, where $H \in W[u]$ satisfies $H \bmod p = h$ and $\deg H = \deg h$. By Hensel's lemma, if A is étale over κ then $\widetilde{A}$ exists (and is finite étale over W) and is unique up to unique W-isomorphism. If A is not étale over κ, a finite flat lifting of A over W may not exist (see [5, Example 3.2(4)]).

When κ has characteristic 2 and A is étale over κ, $\mathrm{disc}_W \widetilde{A}$ lies in $W^\times/(W^\times)^2$. Writing $W^\times = \kappa^\times \times (1+2W)$ (Teichmüller decomposition), note that the 1-unit part of $\mathrm{disc}_W \widetilde{A}$ lies in $1 + 4W$. (Ambiguity of $\mathrm{disc}_W \widetilde{A}$ up to a unit-square does not affect the meaning of this assertion, since $(1+2w)^2 \in 1+4W$.) Indeed, to prove $\mathrm{disc}_W \widetilde{A}$ has its 1-unit part in $1+4W$ we may make a finite étale local base change on W to split the finite étale W-algebra $\widetilde{A}$, and the discriminant with respect to a primitive idempotent basis is 1.

Here is a Möbius formula using liftings to characteristic 0.

Theorem 3.5. *For any finite κ-algebra A that admits a finite flat lifting $\widetilde{A}$ of A over W,*

$$\mu(A) = (-1)^{\dim_\kappa A} \widetilde{\chi}(\mathrm{disc}_W \widetilde{A}), \tag{3.5}$$

where $\widetilde{\chi}$ is the unique quadratic character on $W^\times/(W^\times)^2 \simeq \kappa^\times/(\kappa^\times)^2$ when κ has odd characteristic and is the unique quadratic character on

$$(\kappa^\times \times (1+4W))/((\kappa^\times \times (1+4W)) \cap (W^\times)^2) \simeq (1+4W)/((1+4W) \cap (W^\times)^2)$$

when κ has characteristic 2. In both cases, $\widetilde{\chi}$ is extended by 0 to pW.

Before we prove Theorem 3.5, we make some remarks on the case $\mathrm{char}(\kappa) = 2$.

Remark 3.6. When κ has characteristic 2, we do not need to extend $\widetilde{\chi}$ to $1 + 2W$ or to all of $W^\times$, and there is no canonical extension anyway. Note that $(1+4W) \cap (W^\times)^2$ is the index-2 kernel of

$$1 + 4W \twoheadrightarrow (1+4W)/(1+8W) \simeq W/2W = \kappa \xrightarrow{\mathrm{Tr}_{\kappa/\mathbf{F}_2}} \mathbf{F}_2,$$

where the middle isomorphism is induced by $1 + 4x \mapsto x$.

Proof. (of Theorem 3.5) The case $A = 0$ is trivial. Since the reduction of $\mathrm{disc}_W \widetilde{A}$ modulo pW is $\mathrm{disc}_\kappa A$, (3.5) is trivial when A is non-étale over κ. (All we need to know about $\widetilde{\chi}$ here is that, by definition, it vanishes on pW.)

When A is étale over κ, the uniqueness of $\widetilde{A}$ lets us assume $A = \kappa'$ is a field, say of degree d over κ, so $\widetilde{A}$ is the valuation ring W_d of an unramified extension of W of degree d and the desired Möbius formula is equivalent to

$$\widetilde{\chi}(\mathrm{disc}_W(W_d)) = (-1)^{d-1}.$$

By the definition of $\widetilde{\chi}$, this formula says that $\mathrm{disc}_W W_d$ is a square in $W^\times$ if and only if d is odd. This criterion for being a square is proved via the argument used to prove (3.4). ■

Remark 3.7. Theorem 3.5 and its proof apply with κ replaced by any perfect field k of positive characteristic such that all finite Galois extensions of k are cyclic. When k has characteristic 2, Artin–Schreier theory ensures that the subgroup $\{x^2 + x \mid x \in k\}$ has index ≤ 2 in k. However, there is no description of this subgroup akin to Remark 3.6 when k is infinite.

Taking $A = \kappa[u]/(h)$ for nonzero $h \in \kappa[u]$, Theorems 3.3 and 3.5 specialize to

$$\mu_{\kappa[u]}(h) = \begin{cases} (-1)^{\deg h}\chi(\mathrm{disc}_\kappa h), & \text{if } \kappa \text{ has odd characteristic,} \\ (-1)^{\deg h}\widetilde{\chi}(\mathrm{disc}_W H), & \text{if } \kappa \text{ has any characteristic,} \end{cases} \tag{3.6}$$

where χ and $\widetilde{\chi}$ are described in Theorems 3.3 and 3.5, and H is a lifting of h into $W[u]$ with $\deg H = \deg h$.

Remark 3.8. The formula in (3.6) for the case of characteristic 2 uses a discriminant in characteristic 0. There is an intrinsic characteristic 2 variant of the discriminant, due to Berlekamp [4] (and developed by later authors, such as Wadsworth [23]), but we have not found this to be useful for our purposes.

Remark 3.9. When $\kappa = \mathbf{F}_p$, (3.6) is a classical formula of Pellet [18] from 1878, with the case when $h \in \mathbf{F}_p[u]$ is separable being related to Stickelberger's formula for the quadratic character of the discriminant of a number field [7, Prop. 4.8.10]. What is crucial for us is not simply the formula (3.6) itself but its interpretation. In the context of Stickelberger's formula, one uses (3.6) with fixed $h \in \mathbf{Z}[u]$ and varying $\kappa = \mathbf{F}_p$. Instead we will use (3.6) with fixed κ and varying $h \in \kappa[u]$. This is an idea that goes back to Swan [22], although he only considered separable h and did not bring out the Möbius aspect of the formula.

Remark 3.10. In numerical work, we used (3.6) to compute Möbius values on $\mathbf{F}_p[u]$ when $p \neq 2$. This is much faster than the definition of the Möbius function, which involves factorizations. However, we computed Möbius values on $\mathbf{F}_2[u]$ directly from the definition, since factoring in $\mathbf{F}_2[u]$ on a computer is much faster than computing discriminants in characteristic 0 and reducing modulo 8.

Example 3.11. Let κ be a finite field with characteristic p, even perhaps $p = 2$. For nonconstant g in $\kappa[u]$, $\mu(g^{4p} + u) = 1$. Indeed, for $p \neq 2$, this follows from (3.6) because $\mathrm{disc}(g^{4p} + u)$ is a square in κ by (3.2). For $p = 2$, let $W = W(\kappa)$. By (3.2),

$$\mathrm{disc}_W(G^8 + u) \in (W^\times)^8 \cdot (1 + 8W) \in (W^\times)^2.$$

Therefore, $\widetilde{\chi}(\mathrm{disc}_W(G^8 + u)) = 1$ when G is a polynomial in $W[u]$ with positive degree and unit leading coefficient. Thus, by (3.6), $\mu(g^{4p} + u) = 1$ for nonconstant $g \in \kappa[u]$ when $p = 2$.

Example 3.12. Let κ be a finite field with characteristic $p \neq 2$. For nonconstant $g = cu^n + \cdots \in \kappa[u]$ we see via (3.6) that

$$\mu(g^p + u) = (-1)^n \chi(c)^n \chi(-1)^{n(n+1)/2}.$$

When n is odd, this equals 1 and -1 equally often as g varies. When n is even, $\mu(g^p + u)$ equals $\chi(-1)^{n/2}$ for all g of degree n.

For instance, when $n \equiv 2 \bmod 4$, $\mu(g^5 + u) = 1$ for all $g \in \mathbf{F}_5[u]$ with degree n, so $g^5 + u$ is reducible. On the other hand, $\mu(g^3 + u) = -1$ for all $g \in \mathbf{F}_3[u]$ with degree $\equiv 2 \bmod 4$.

We similarly find

$$\mu(g^p + u^2) = (-1)^n (\chi(-1))^{n(pn+1)/2} \chi(2)^n \chi(c)^{n+1} \chi(g(0)).$$

In particular, for fixed $n \geq 1$, $\mu(g^p + u^2)$ is equal to 1 as often as it is equal to -1. Therefore there is no Möbius bias, in contrast with $\mu(g^p + u)$ when $\deg g$ is even and $-1 \in \kappa^\times$ is a square.

Example 3.13. Let κ have size q and characteristic p. Choose an integer b such that $1 < b < 4q$ and $(b, p(q-1)) = 1$ (*e.g.*, $b = 2q - 1$). Then the polynomial $f(T) = T^{4q} + u^b$ is irreducible in $\kappa[u][T]$ by [17, p. 297] and has no local obstructions, but $f(g)$ is reducible in $\kappa[u]$ for every $g \in \kappa[u]$. Indeed, this holds when $g = c$ is constant since $u^b + c$ is non-linear and has a root. If g is nonconstant then $f(g)$ has u as a multiple factor if $g(0) = 0$ and (3.6) implies $\mu(f(g)) = 1$ if $g(0) \neq 0$.

When $q = 2$ and $b = 3$, we recover Example 1.4: $T^8 + u^3$ takes no irreducible values on $\mathbf{F}_2[u]$ despite being irreducible with no local obstructions. What happens if we replace $\mathbf{F}_2$ by a larger finite field of characteristic 2? Further work shows that $T^8 + u^3$ takes no irreducible values on $\mathbf{F}_{2^m}[u]$ with m odd, while for m even it takes irreducible values on $\mathbf{F}_{2^m}[u]$ only at constant non-cubes in $\mathbf{F}_{2^m}^\times$. In particular, $T^8 + u^3$ acquires only a finite number of irreducible specializations on any $\mathbf{F}_{2^m}[u]$. In [3], Bender and Wittenberg establish conditions under which a polynomial over $\kappa[u][T]$ is guaranteed to have at least one irreducible specialization after *extending* the constant field κ. The polynomial $T^8 + u^3$ shows such a result is almost optimal without additional constraints.

This completes our discussion on generalities about the Möbius function on $\kappa[u]$. Theorems 3.3 and 3.5 will be important both here and in our higher genus work in [10]. In the present paper, we will prove a refinement of (3.6) when $h = f(g)$ with fixed nonzero $f \in \kappa[u][T^p]$ and varying $g \in \kappa[u]$. Our main results in this direction are Theorems 4.7, 5.7, and 6.10 (and Corollary 6.11).

4. Discriminants and Resultants

For nonconstant $f \in \kappa[u][T^p]$, we wish to understand the behavior of $\mu(f(g))$ as g varies in $\kappa[u]$ with large degree. The formulas (2.5) and (3.6) suggest that we should study $\operatorname{disc}(f(g))$ as an algebraic function of varying g with large but fixed degree. Following Swan [22], we will find it useful to work with resultants and not discriminants. In characteristic 0, where derivative degrees drop by 1, the relation between resultants and discriminants is given by the formula

$$\operatorname{disc} h = \frac{(-1)^{d(d-1)/2} R_{d,d-1}(h, h')}{(\operatorname{lead} h)^{2d-1}}. \tag{4.1}$$

Here $d = \deg h \geq 1$ and $R_{d,d-1}(h, h')$ is the resultant of h and h' using a universal $(d+(d-1))$-dimensional determinant. In all characteristics, if $h' \neq 0$ then the formula is

$$\operatorname{disc} h = \frac{(-1)^{d(d-1)/2} R(h, h')}{(\operatorname{lead} h)^{d+\deg h'}}, \tag{4.2}$$

where $d = \deg h \geq 1$ and $R(h, h')$ is the resultant of h and h' computed with a determinant whose size is based on the actual degree of h' (which might be less than $d-1$). The periodic dependence of $(-1)^{d(d-1)/2}$ as a function of d mod 4 is ultimately where the mod-4 periodicity enters into our work (for an example, see (4.8) in Example 4.3). We now review some of the basic formalism of resultants.

Recall that for an integral domain A, the *resultant* of two nonzero polynomials h_1 and h_2 in $A[u]$, denoted $R_A(h_1, h_2) = R(h_1, h_2)$, is defined to be

$$R(h_1, h_2) = (\operatorname{lead} h_1)^{\deg h_2} \prod_{h_1(\alpha)=0} h_2(\alpha) \tag{4.3}$$

with the product running over the roots of h_1 (counted with multiplicity) in a splitting field over the fraction field of A. In [17, p. 200], an expression for $R(h_1, h_2)$ is given as a *universal* determinant in the coefficients of h_1 and h_2. An essential aspect of this universal formula is that the size of the determinant defining the resultant depends on the degrees of h_1 and h_2. We may write $R_{d_1,d_2}(h_1, h_2)$ to indicate that h_j is being treated as a polynomial of degree d_j for the resultant calculation via a universal determinant. We make the *convention* that when a resultant $R(h_1, h_2)$ appears without degree subscripts then it is defined in terms of the actual degrees of its arguments if h_1 and h_2 are nonzero. We also agree to define $R(h_1, h_2) = 0$ when at least one h_j vanishes. This latter definition is compatible with universal determinants that define resultants (letting the zero polynomial be assigned whatever nonnegative degree we please).

The effect of computing a resultant with a universal formula involving a fake higher degree in the second argument goes as follows. If nonzero h_1 and h_2 have actual degrees d_1 and d_2, then for any $d_3 \geq d_2$,

$$R_{d_1,d_3}(h_1, h_2) = (\operatorname{lead} h_1)^{d_3-d_2} R_{d_1,d_2}(h_1, h_2). \tag{4.4}$$

While the value of the resultant may have changed (although not if h_1 is monic), the property of vanishing or nonvanishing for the resultant does not change. (We work with resultants only over domains, not over arbitrary commutative rings). Though (4.3) is valid as written when h_2 is given a fake higher degree (still denoted $\deg h_2$), it is generally not valid when h_1 is given a fake higher degree; also keep in mind that in general $R(h_1, h_2)$ and $R(h_2, h_1)$ are related by a sign (the precise sign-factor will be recorded shortly in a list of standard algebraic properties of resultants).

Warning. Failure to remember that the construction of resultants is sensitive to degrees can lead to errors when standard universal formulas from characteristic 0 are used in characteristic p.

Example 4.1. Let $f(T) = T^9 + (2u^2 + u)T^6 + (2u + 2)T^3 + u^2 + 2u + 1$ in $\kappa[u][T]$, where κ has characteristic 3 ($\kappa = \mathbf{F}_3$ is Example 1.3). For nonconstant $g = cu^n + \dots$ in $\kappa[u]$ with $c \neq 0$, $f(g)$ has degree $9n$ and $f(g)' = (\partial_u f)(g)$ has degree $6n + 1 < 9n - 1$. The "true" resultant of $f(g)$ and $(\partial_u f)(g)$ is $R_{9n,6n+1}(f(g), f(g)')$, but the resultant needed to compute

$\operatorname{disc} f(g)$ in (4.1) is

$$R_{9n,9n-1}(f(g), f(g)') = (c^9)^{3n-2} R_{9n,6n+1}(f(g), (\partial_u f)(g)). \tag{4.5}$$

Thus

$$\operatorname{disc} f(g) = \frac{(-1)^{n(n-1)/2} R_{9n,6n+1}(f(g), (\partial_u f)(g))}{c^{9(15n+1)}}, \tag{4.6}$$

which also follows directly from (4.2). Using (4.1) instead of (4.2) introduces an erroneous factor of $(c^9)^{3n-2}$. This power of c affects the quadratic nature of the right side of (4.6), so in view of (3.6) such an error would be serious.

Resultants have several useful algebraic properties. We summarize six of them without proof, as in [22], and we include (4.4) in the list. In this list, polynomials are nonzero and have coefficients in a domain A.

(1) $R(h_1, h_2) = (-1)^{(\deg h_1)(\deg h_2)} R(h_2, h_1)$.
(2) $R(h_1, h_2)$ is bimultiplicative: $R(h_1 h_3, h_2) = R(h_1, h_2) R(h_3, h_2)$ and $R(h_1, h_2 h_3) = R(h_1, h_2) R(h_1, h_3)$.
(3) $R(u, h) = h(0)$. More generally, $R(u - c, h) = h(c)$ and $R(h, u - c) = (-1)^{\deg h} h(c)$ for $c \in A$.
(4) $R(c, h) = R(h, c) = c^{\deg h}$ for $c \in A$, $h \neq 0$. Thus, $R(c_1, c_2) = 1$ for $c_1, c_2 \neq 0$ in A.
(5) When h_1 has degree d_1, h_2 has degree d_2, and $d_3 \geq d_2$,

$$R_{d_1,d_3}(h_1, h_2) = (\operatorname{lead} h_1)^{d_3 - d_2} R_{d_1,d_2}(h_1, h_2).$$

(6) For nonzero M, h_1, h_2 in $A[u]$,

$$h_1 \equiv h_2 \bmod M \Longrightarrow R(M, h_1) = (\operatorname{lead} M)^{\deg h_1 - \deg h_2} R(M, h_2),$$

where we recall that $\operatorname{lead} M$ denotes the leading coefficient of $M \in A[u]$.

We call property (6) the *quasi-periodicity* of the resultant (in its second argument). When M is monic, $R(M, h)$ is genuinely periodic in h, with period (M). More generally (and of greater relevance to our work), for monic M in $A[u]$ and any $b(T) \in A[u][T]$, $R(M, b(h))$ is genuinely periodic in h. Swan's definition of $R(h_1, h_2)$ in [22] is what we call $R(h_2, h_1)$, so property (1) warns us that any comparison with [22] must keep this distinction in mind.

The following two examples use the resultant to compute a formula for $\mu(f(g))$ as a function of g:

Example 4.2. Let $f(T) = T^{12} + (u+1)T^6 + u^4 \in \kappa[u][T]$ with finite κ of characteristic 3. (Example 1.1 is $\kappa = \mathbf{F}_3$.) Let q denote the size of κ and let χ be the quadratic character on $\kappa^\times$, with $\chi(0) = 0$. We shall compute $\mu(f(g))$ when $n = \deg g \geq 1$.

Since $4 | \deg(f(g))$ and $\operatorname{lead}(f(g))$ is a square, (3.6) and (4.2) with $h = f(g)$ give

$$\begin{aligned} \mu(f(g)) &= \chi(\operatorname{disc} f(g)) \\ &= \chi(R_{12n,12n-1}(f(g), (f(g))')) \\ &= \chi(R_{12n,12n-1}(f(g), (g^2+u)^3)) \\ &= \chi(R(g^2+u, f(g))). \end{aligned}$$

Since $f(g) \equiv u^6 - u^3 \bmod g^2 + u$ and the leading coefficient of $g^2 + u$ is a square, quasi-periodicity of the resultant gives (using c_g to denote $(\operatorname{lead} g)^{12 \deg g - 6}$)

$$\begin{aligned} R(g^2+u, f(g)) &= c_g^2 R(g^2+u, u^6-u^3) \\ &= c_g^2 R(g^2+u, u)^3 R(g^2+u, u-1)^3 \\ &= c_g^2 g(0)^6 (g(1)^2+1)^3, \end{aligned}$$

so

$$\mu(f(g)) = \chi(\operatorname{disc} f(g)) = \chi(g(0))^2 \chi(g(1)^2+1). \tag{4.7}$$

(This calculation also shows that $\operatorname{disc} f(g)$ is a constant square multiple of $g(0)^{18}(g(1)^2+1)^9$, with the constant multiplier depending on $\operatorname{lead} g$ and $\deg g$.) As g runs over all polynomials of a given degree $n \geq 2$ in $\kappa[u]$, $g(0)$ and $g(1)$ can be "independently assigned" (think about $g \bmod u(u-1)$). So, for instance, if -1 is not a square in κ, we see that $\mu(f(g))$ vanishes $1/q$ of the time (when $g(0) = 0$), and is -1 twice as often as it is 1.

Example 4.3. Let κ be a finite field with characteristic 3, and χ the quadratic character on $\kappa^\times$. Let

$$f(T) = T^9 + (2u^2+u)T^6 + (2u+2)T^3 + u^2 + 2u + 1$$

in $\kappa[u][T]$. This polynomial was already met in Example 4.1. We will compute a formula for $\mu(f(g))$ as g runs over nonconstant polynomials in $\kappa[u]$. The argument is long compared to Example 4.2, but at the same time it is more indicative of the general case, and thus is more *instructive*.

For nonconstant $g(u) = cu^n + \cdots$ with degree $n \geq 1$, we have $\deg(f(g)) = 9n$ and $\deg(f(g)') = 6n+1$, so $\mu(f(g)) = (-1)^{9n}\chi(\operatorname{disc} f(g))$ by (3.6). By (4.6),

$$\mu(f(g)) = (-1)^n(\chi(-1))^{n(n-1)/2}\chi(c)^{n+1}\chi(R(f(g), (\partial_u f)(g))). \tag{4.8}$$

We now compute a universal formula for $R(f(g), (\partial_u f)(g))$ in five steps, working over any field (or even domain) of characteristic 3. The formula is given in (4.13) as an algebraic identity, so for the purposes of the following calculation we may take g to be the universal polynomial of degree n over a field of characteristic 3 (so g has coefficients in a rational function field of transcendence degree $n+1$ over $\mathbf{F}_3$). In particular, the operation of division by $g(2)$ in Step 1 is not problematic.

Step 1. Explicitly,

$$f(g) = g^9 + (2u^2+u)g^6 + (2u+2)g^3 + u^2 + 2u + 1, \quad (\partial_u f)(g) = (u+1)g^6 + 2g^3 + 2u + 2. \tag{4.9}$$

Using (4.9), write $R(f(g), (\partial_u f)(g)) = (-1)^n R((\partial_u f)(g), f(g))$ to make the lower-degree term $(\partial_u f)(g)$ appear as the first argument. We want to simplify the resultant by quasi-periodicity, but the leading terms in (4.9) suggest it is easier to reduce $(u+1)f(g)$, rather than $f(g)$, modulo $(\partial_u f)(g)$. Apply bimultiplicativity to introduce a factor of $u+1$:

$$R(f(g), (\partial_u f)(g)) = \frac{(-1)^n R((\partial_u f)(g), (u+1)f(g))}{R((\partial_u f)(g), u+1)} = \frac{(-1)^n R((\partial_u f)(g), (u+1)f(g))}{g(2)^3}. \tag{4.10}$$

Treating g as if it is generic (so $g(2)$ is a unit) ensures that (4.10) is a meaningful (and correct) algebraic formula. Our derivation of (4.10) used bimultiplicativity to create convenient leading terms for quasi-periodicity. This idea will be used again in Step 3.

Step 2. Since $(u+1)f(g) = (\partial_u f)(g)(g^3+2u^2+u) + g^6 + u^2g^3 + u + 1$, quasi-periodicity of the resultant implies (recall $c = \operatorname{lead} g$)

$$\begin{aligned} R((\partial_u f)(g), (u+1)f(g)) &= (c^6)^{9n+1-6n} R((\partial_u f)(g), g^6 + u^2g^3 + u + 1) \\ &= (c^6)^{3n+1} R(g^6 + u^2g^3 + u + 1, (\partial_u f)(g)). \end{aligned}$$

The nonzero constant in front will disappear when we apply χ as part of (4.8).

Step 3. Since $(\partial_u f)(g) \equiv 2(u+2)(u^2+2u+2)g^3 + 2(u+1)(u+2) \bmod g^6+u^2g^3+u+1$, quasi-periodicity implies $R(g^6+u^2g^3+u+1, (\partial_u f)(g))$ is the product of $(c^6)^{6n+1-(3n+3)} = (c^6)^{3n-2}$ and $R(g^6 + u^2g^3 + u + 1, 2(u+2)(u^2+2u+2)g^3 + 2(u+1)(u+2))$. Writing the second argument of this resultant as a product $2(u+2)((u^2+2u+2)g^3 + u + 1)$, this resultant is a product of $2^{6n} = 1$, $(g(1)^2 + g(1) + 2)^3$, and $R((u^2+2u+2)g^3 + u + 1, g^6 + u^2g^3 + u + 1)$. To simplify this last resultant, we again use bimultiplicativity to make leading terms more compatible. This resultant equals the ratio

$$\frac{R((u^2+2u+2)g^3 + u + 1, (u^2+2u+2)(g^6 + u^2g^3 + u + 1))}{R((u^2+2u+2)g^3 + u + 1, u^2 + 2u + 2)}. \tag{4.11}$$

Step 4. The denominator in (4.11) is 1 by quasi-periodicity (switch the two terms, which introduces no sign, and then reduce mod $u^2 + 2u + 2$). As for the numerator,

$$(u^2+2u+2)(g^6 + u^2g^3 + u + 1) \equiv (2u+2)g^3 + 2u^2 + u + 2 \bmod (u^2+2u+2)g^3 + u + 1,$$

so the numerator of (4.11) is $(c^3)^{3n+1} R((u^2+2u+2)g^3 + u + 1, (2u+2)g^3 + 2u^2 + u + 2)$. The resultant factor equals

$$R(2(u+1)(g^3+u+1), (u^2+2u+2)g^3+u+1) = (-1)^n g(2)^3 R(g^3+u+1, (u^2+2u+2)g^3+u+1).$$

Putting everything together into (4.10), we have a cancellation of $g(2)^3$ and obtain

$$R(f(g), (\partial_u f)(g)) = c^{45n-3}(g(1)^2 + g(1) + 2)^3 R(g^3 + u + 1, (u^2+2u+2)g^3 + u + 1). \tag{4.12}$$

Step 5. Finally, $(u^2+2u+2)g^3 + u + 1 \equiv 2(u+1)^3 \bmod g^3 + u + 1$, so

$$\begin{aligned} R(g^3 + u + 1, (u^2+2u+2)g^3 + u + 1) &= (c^3)^{3n-1}(-1)^n R(2(u+1)^3, g^3 + u + 1) \\ &= (c^3)^{3n-1} g(2)^9. \end{aligned}$$

Feeding this into (4.12) gives the resultant formula

$$R(f(g), (\partial_u f)(g)) = c^{54n-6}(g(1)^2 + g(1) + 2)^3 g(2)^9. \tag{4.13}$$

(The reader may check that the projection from $\{f = 0\}$ onto the T-axis is non-étale at precisely $(2, 0)$ and the two geometric points $(1, t)$ with $t^2 + t + 2 = 0$, with the branch scheme having respective lengths 9 and 3. The relation with the factors and exponents in (4.13) will be explained in §5.)

Inserting (4.13) into (4.8), we find our Möbius formula:

$$\mu(f(g)) = (-1)^n \chi(-1)^{n(n-1)/2} \chi(c)^{n+1} \chi(g(1)^2 + g(1) + 2)) \chi(g(2)) \tag{4.14}$$

for nonconstant g in $\kappa[u]$. This depends on $g \bmod (u-1)(u-2)$, $\deg g \bmod 4$, and the quadratic character of the leading coefficient of g. Taking $\kappa = \mathbf{F}_3$, we will show in Example 7.8 that (4.14) is numerically compatible with the statistics in Table 1.2.

Motivated by the goal of making patterns in $\mu(f(g))$ provable when $f(T)$ is irreducible and inseparable, as in Examples 4.2 and 4.3, we discovered that the function $g \mapsto \mu(f(g))$ admits a periodicity in g when f is squarefree with irreducible factors that are inseparable (in T). Before stating our periodicity theorem, we need a lemma.

Lemma 4.4. *Let F be perfect of characteristic $p > 0$.*

1) *Choose a nonzero $f \in F[u][T^p]$ such that f is squarefree in $F[u,T]$. Then f and $\partial_u f$ have no nonconstant common factor in $F[u,T]$, or equivalently the zero loci $\{f = 0\}$ and $\{\partial_u f = 0\}$ in the affine plane $\mathbf{A}^2_F$ intersect at finitely many points.*

2) *The same conclusion holds if $f \in F[u,T]$ is nonzero and $f(T^p)$ is squarefree in $F[u,T]$ (so f is squarefree in $F[u,T]$).*

Note that if $f \notin F$ then $f(T)$ cannot lie in $F[u^p, T]$ under either hypothesis in the lemma, so $\partial_u f \neq 0$ in such cases. It may happen that $\partial_u f$ is constant; *e.g.*, $f = u^p T^p + u + 1$ (or $f = u$). The second case in Lemma 4.4 will be used only when $p = 2$.

Proof. The case $f \in F^\times$ is trivial, so we may assume $f \notin F$. In particular, $\partial_u f \neq 0$. Let Z_f and $Z_{\partial_u f}$ be the respective zero loci of f and $\partial_u f$ in the affine plane (so the latter locus may be empty). Since F is perfect, extending scalars to an algebraic closure of F preserves the property of being squarefree and hence we may assume F is algebraically closed. The hypothesis on f in case (1) (resp. case (2)) implies that $f(T)$ (resp. $f(T^p)$) is a squarefree element in $F(T^p)[u]$ with nonzero u-derivative, so the projection of $Z_f \cap Z_{\partial_u f}$ onto the T-axis does not contain the generic point and hence is F-finite. To conclude the finiteness of $Z_f \cap Z_{\partial_u f}$ it therefore suffices (since F is algebraically closed) to prove that Z_f contains no lines $T = c$ for $c \in F$. But if Z_f contains such a line then the squarefree element $f \in F[u, T^p]$ (resp. $f(T^p) \in F[u, T^p]$) is divisible by $T^p - c^p = (T-c)^p$ (resp. $(T - c^{1/p})^p$), contrary to the squarefreeness hypothesis. ∎

Definition 4.5. If $f_1, f_2 \in F[u,T]$ are two nonzero polynomials over a perfect field F such that their zero loci Z_{f_1} and Z_{f_2} in $\mathbf{A}^2_F$ have finite intersection, $M^{\mathrm{geom}}_{f_1,f_2} \in F[u]$ is the monic separable polynomial whose zero locus is the projection of $Z_{f_1} \cap Z_{f_2}$ onto the u-axis (so $M^{\mathrm{geom}}_{f_1,f_2} = 1$ if $Z_{f_1} \cap Z_{f_2}$ is empty, such as when some f_j lies in $F^\times$). When $f \in F[u,T]$ is nonconstant, define $M^{\mathrm{geom}}_f := M^{\mathrm{geom}}_{f,\partial_u f}$ when this makes sense (*i.e.*, when $\partial_u f \neq 0$ and $Z_f \cap Z_{\partial_u f}$ is finite).

Note that the formation of $M^{\mathrm{geom}}_{f_1,f_2}$ commutes with extension of the perfect ground field. When $\mathrm{lead}_T f$ is separable, $M^{\mathrm{geom}}_{f_1,f_2}$ is the radical of the resultant $R_{F[u]}(f_1, f_2)$. (We saw in Remark 1.6 that this need not hold when $\mathrm{lead}_T f$ is not separable.)

For $f \in F[u,T]$ with $f \notin F$, Lemma 4.4 gives some sufficient conditions for M^{geom}_f to be defined when F has positive characteristic. The next lemma gives a general geometric criterion in any characteristic.

Lemma 4.6. *If F is perfect with arbitrary characteristic and $f \in F[u,T]$ is not in F, then the zero loci of f and $\partial_u f$ in $\mathbf{A}^2_F$ have finite intersection if and only if f is squarefree in $F[u,T]$ with no irreducible factors in $F[T]$ and the projection*

$$\mathrm{pr}_T : Z_f = \mathrm{Spec}\, F[u,T]/(f) \to \mathrm{Spec}\, F[T] = \mathbf{A}^1_F$$

onto the T-axis is generically étale on Z_f. When this happens, the non-étale locus of pr_T is finite and its projection onto the u-axis is the zero locus of M^{geom}_f in $\mathbf{A}^1_F$.

The generically-étale property is always satisfied for squarefree nonzero $f \in F[u,T]$ in characteristic 0 since pr_T is *a priori* quasi-finite and flat. We will apply this lemma over a 2-adic field in our later study of Möbius bias in characteristic 2.

Proof. Necessity of the conditions that f be squarefree and have no irreducible factors in $F[T]$ is clear. Granting these conditions, the plane curve Z_f is reduced (hence geometrically reduced since F is perfect) and its projection to the T-axis is quasi-finite and hence flat. Thus, the property of pr_T being étale at a point of Z_f may be checked on the geometric fibers of pr_T. Extending scalars to an algebraic closure of F, we thereby see that the non-étale locus for pr_T is where Z_f meets $Z_{\partial_u f}$ in $\mathbf{A}^2_F$. This completes the proof of the desired equivalence, and also yields the asserted relationship between M_f^{geom} and the non-étale locus of pr_T. ■

Here is our main result in odd characteristic. The proof will be given in §5, using Theorem 5.7.

Theorem 4.7. *Let κ be a finite field with odd characteristic p, and let χ be the quadratic character of $\kappa^\times$. Fix a nonzero $f(T) \in \kappa[u][T^p]$ that is squarefree in $\kappa[u][T]$. Assume $f \notin \kappa$.*

For $g_1 = c_1 u^{n_1} + \dots$ and $g_2 = c_2 u^{n_2} + \dots$ in $\kappa[u]$ with sufficiently large degrees n_1 and n_2 (depending on f), we have the implication

$$g_1 \equiv g_2 \bmod M_f^{\mathrm{geom}}, \quad n_1 \equiv n_2 \bmod 4, \quad \chi(c_1) = \chi(c_2) \Longrightarrow \mu(f(g_1)) = \mu(f(g_2)). \tag{4.15}$$

The largeness of degrees n_j can be chosen uniformly with respect to finite extensions of κ.

If -1 is a square in κ or $\deg_T f$ is even, the second congruence in (4.15) may be relaxed to $n_1 \equiv n_2 \bmod 2$.

If $M_{f,\kappa}^{\min} \in \kappa[u]$ is the monic polynomial M of least degree such that

$$g_1 \equiv g_2 \bmod M, \quad n_1 \equiv n_2 \bmod 4, \quad \chi(c_1) = \chi(c_2) \Longrightarrow \mu(f(g_1)) = \mu(f(g_2))$$

for all $g_j = c_j u^{n_j} + \dots$ with sufficiently large degrees n_1 and n_2 then $M_{f,\kappa}^{\min}$ is a factor of any other nonzero polynomial $M \in \kappa[u]$ with the same property (so $M_{f,\kappa} | M_f^{\mathrm{geom}}$). For some finite extension κ'/κ we have $M_{f,\kappa''}^{\min} = M_f^{\mathrm{geom}}$ for any finite extension κ'' of κ'.

Remark 4.8. The finiteness of κ in Theorem 4.7 may be relaxed in odd characteristic exactly as in Remark 3.4 without changing the proof, but we do not know any interesting examples of this generalized theorem with infinite κ.

Although Theorem 4.7 does not say that $g \mapsto \mu(f(g))$ is genuinely periodic in g, we will refer to any nonzero M satisfying the role of M_f^{geom} in (4.15) as a *modulus* for $\mu(f(g))$. Since any congruence class in $\kappa[u]/(M)$ may be represented by a polynomial of any large degree with any desired leading coefficient, it is a trivial exercise with the Chinese remainder theorem to check that for any two moduli M_1 and M_2 for $\mu(f(g))$, $\gcd(M_1, M_2)$ is also a modulus. It therefore follows trivially from (4.15) that $M_{f,\kappa}^{\min}$ divides all other moduli for $\mu(f(g))$. The fact that $M_{f,\kappa''}^{\min} = M_f^{\mathrm{geom}}$ for all finite extensions κ''/κ containing some sufficiently large finite extension κ'/κ rests on an understanding of how we prove (4.15).

Examples 4.2 and 4.3 illustrated some techniques that will be used in the proof of Theorem 4.7. The following example focuses only on explicit Möbius formulas, illustrating the conclusions of Theorem 4.7.

Example 4.9. The variation of $M_{f,\kappa}^{\min}$ as κ grows is interesting. Since $M_{f,\kappa}^{\min} | M_f^{\mathrm{geom}}$, there are only finitely many possibilities for $M_{f,\kappa'}^{\min}$ as κ' varies over finite extensions of κ. We now give an example where $M_{f,\kappa}^{\min} \neq M_f^{\mathrm{geom}}$.

Let $f(T) = T^{12} + (2u^4 + u^3 + u^2 + 2)T^6 + 2u^3 + 1$ in $\kappa[u][T]$, where κ has characteristic 3. For nonconstant g in $\kappa[u]$, the proof of Theorem 4.7 shows

$$\mu(f(g)) = \chi(g(0)^2 + 1)^2\chi(g(1))\chi(R(u^2 + 1, f(g))).$$

Note that $\chi(g(0)^2 + 1)^2$ is not always 1 because it may vanish. This Möbius formula, like (4.7), has no dependence on $\deg g$ mod 4 or on the quadratic character of the leading coefficient of g. Since $R(u^2+1, f(g))$ only depends on g modulo u^2+1 (by quasi-periodicity of resultants), we see that $\mu(f(g))$ only depends on g modulo $u(u-1)(u^2+1)$. (Since $R_{\kappa[u]}(f, \partial_u f) = u^{12}(u-1)^{18}(u^2+1)^{12}$, we have $M_f^{\mathrm{geom}} = u(u-1)(u^2+1)$.) If $[\kappa : \mathbf{F}_3]$ is odd then $g(0)^2+1$ is nonzero, so $\mu(f(g))$ only depends on g modulo $(u-1)(u^2+1)$ for such κ; hence, $M_{f,\kappa}^{\min} = (u-1)(u^2+1) \neq M_f^{\mathrm{geom}}$. This illustrates that the minimal modulus in Theorem 4.7 can be sensitive to a change in the base field κ.

5. A resultant formula

We will obtain Theorem 4.7 from a periodicity property for resultants over arbitrary perfect fields. We indulge in the following notational device: for a field F and a nonzero $M \in F[u]$, we write $F[u]/(M)$ to denote the vector-scheme of remainders upon long division by M over F-algebras A. That is, $F[u]/(M)$ is viewed as an affine space of dimension $\deg M$, whose coordinates arise from coefficients of u^i for $0 \leq i < \deg M$. (This space is $\operatorname{Spec} F$ when $\deg M = 0$.) Such abuse of notation is standard for vector-schemes in the theory of algebraic groups. The context will indicate whether $F[u]/(M)$ denotes an affine space over $\operatorname{Spec} F$ or its set of F-valued points, the "usual" F-vector space $F[u]/(M)$.

We will also work with the scheme

$$\mathrm{Poly}_{n/F} = \mathbf{A}^n \times_F \mathbf{G}_m = \operatorname{Spec} F[a_0, \ldots, a_n, 1/a_n]$$

of polynomials of exact degree $n \geq 0$, as well as the scheme

$$\mathrm{Poly}_{\leq n/F}\mathbf{A}_F^{n+1} = \operatorname{Spec} F[a_0, \ldots, a_n]$$

of polynomials of degree $\leq n$. The coordinates $(a_0, \ldots, a_n)$ correspond to $\sum_{i \leq n} a_i u^i$, with $\mathrm{Poly}_{n/F}$ the locus in $\mathrm{Poly}_{\leq n/F}$ where a_n is a unit. For example, given nonconstant $M \in F[u]$ and any $n \geq \deg M$, formation of remainders under long division by M defines an algebraic morphism

$$\rho_{n,M} : \mathrm{Poly}_{n/F} \to F[u]/(M) \simeq \mathrm{Poly}_{\leq (\deg M - 1)/F} \tag{5.1}$$

of smooth F-schemes and this is a smooth surjection (it is a trivial $\mathrm{Poly}_{d/F}$-bundle with $d = n - \deg M$, by the division algorithm). When $M \in F^\times$, the map

$$\rho_{n,M} : \mathrm{Poly}_{n/F} \to \operatorname{Spec} F \tag{5.2}$$

is the structure map to a point.

Since $\deg(f(g))$ is determined by $n = \deg g$ for g of large degree (depending on f, as in (2.5)), there is a well-posed algebraic function

$$\mathrm{disc} \circ f : \mathrm{Poly}_{n/F} \to \mathbf{A}_F^1 \tag{5.3}$$

defined by $g \mapsto \mathrm{disc}(f(g))$ when n is sufficiently large; note that (5.3) does *not* extend to an algebraic function on $\mathrm{Poly}_{\leq n/F}$ (*cf.* Remark 1.8). Our aim is to understand the structure of the algebraic function (5.3) for f as in Lemma 4.6, and in particular the extent to which it factors through some remainder morphism $\rho_{n,M}$ for some nonzero $M \in F[u]$.

To exploit inductive arguments, it is convenient to re-interpret our discriminant problem as the study of the resultant $R(f(g),(\partial_u f)(g))$ for varying g of large (fixed) degree; the utility of this point of view is that it allows us to consider the more general algebraic function $\mathrm{Poly}_{n/F} \to \mathbf{A}^1_F$ defined by

$$g \mapsto R(f_1(g), f_2(g))$$

for large n, with fixed nonzero relatively prime $f_1, f_2 \in F[u,T]$ (a condition satisfied for $f_1 = f$ and $f_2 = \partial_u f$ under either hypothesis in Lemma 4.4 when $f \notin F$). The merit of this generality is that we may separately vary f_1 and f_2. Restricting attention to finite or perfect F of positive characteristic is not adequate: our later work in characteristic 2 will use the present considerations with a 2-adic field F.

Let us now fix a pair of nonzero relatively prime elements $f_1, f_2 \in F[u,T]$, so the zero loci $Z_{f_1} = \{f_1 = 0\}$ and $Z_{f_2} = \{f_2 = 0\}$ are (possibly empty) curves in $\mathbf{A}^2_F$ with no common irreducible components. For $g \in F[u]$ of degree n, (2.5) gives the degree of $f_j(g) \in F[u]$ when $n \gg 0$. We give this formula the label $d_{j,n}$. That is,

$$d_{j,n} := (\deg_T f_j)n + \deg(\mathrm{lead}_T f_j). \tag{5.4}$$

The largeness of $n = \deg g$ that makes (2.5) hold for both f_1 and f_2 depends only on $\deg_T f_1$, $\deg_T f_2$, and the u-degrees of the coefficients of f_1 and f_2 when the f_j's are viewed as polynomials in T. See (2.7) for an explicit universal lower bound on n that makes (2.5) valid when g is a point of $\mathrm{Poly}_{n/F}$ with values in any F-algebra domain.

Fixing such large n, let

$$G = a_0 + a_1 u + \cdots + a_n u^n \in F[a_0, \ldots, a_n][u]$$

denote the universal polynomial over the scheme $\mathrm{Poly}_{\leq n/F} = \mathrm{Spec}\, F[a_0, \ldots, a_n]$ of polynomials of degree $\leq n$ over F-algebras; we are not requiring a_n to be a unit. We wish to study the following universal polynomial depending on f_1 and f_2:

$$R_n(G) := R_{F[a_0,\ldots,a_n]}(f_1(G), f_2(G)) \in F[a_0, \ldots, a_n], \tag{5.5}$$

where the resultant is computed by viewing $f_j(G)$ as having u-degree $d_{j,n}$; since n is large, $d_{j,n}$ is also the u-degree of the specialization of $f_j(G)$ at all field-valued points of the open subscheme $\mathrm{Poly}_{n/F} \subseteq \mathrm{Poly}_{\leq n/F}$ where a_n is a unit.

Lemma 5.1. *$R_n(G) \neq 0$.*

Proof. We need to prove that the nonzero $f_1(G)$ and $f_2(G)$ have no common factor in $F(a_0, \ldots, a_n)[u]$. We first show that the $f_j(G)$'s in $F[a_0, \ldots, a_n][u]$ have no non-trivial common factor that lies in $F[u]$. We may assume F is algebraically closed, so it suffices to prove that for each $c \in F$, $f_1(c, G(c))$ and $f_2(c, G(c))$ do not both vanish in $F[a_0, \ldots, a_n]$. Since some $f_{j_c}(u,T)$ is not divisible by $u - c$, as $f_1(u,T)$ and $f_2(u,T)$ cannot both be divisible by $u - c$, so $f_{j_c}(c,T) \neq 0$, clearly $f_{j_c}(c, G(c)) \neq 0$ since $G(c)$ is transcendental over F.

Since f_1 and f_2 have no common factor in $F[u,T]$, and hence no common factor in $F[u]$, we may assume that f_1 and f_2 are not divisible by nonunits in $F[u]$. In particular, if some f_j has T-degree equal to 0 then that f_j lies in $F^\times$. Hence, we may assume both $\deg_T f_j$'s are positive. The relative primality of f_1 and f_2 ensures that we can find $q_1, q_2 \in F[u,T]$ such that

$$q_1 f_1 + q_2 f_2 = h(u) \in F[u] - \{0\},$$

so if $f_1(G)$ and $f_2(G)$ have a non-trivial common monic factor in $F(a_0,\dots,a_n)[u]$ then such a factor must divide $h(u)$ and so must lie in $F[u]$. Thus, there is no such factor. ■

We want to understand the structure of $R_n(G)$ as an algebraic function in the a_j's. For each of the finitely many intersection points $x = (u_x, t_x)$ of Z_{f_1} and Z_{f_2} in $\mathbf{A}^2_F$, the finite extension $F(x)/F$ is generated over F by the subextensions $F(u_x)$ and $F(t_x)$.

Definition 5.2. For $n \geq 1$, define $P_{x,n}(a_0,\dots,a_n)$ to be the norm-form polynomial

$$\mathrm{N}_{F(x)[a_0,\dots,a_n]/F[a_0,\dots,a_n]}(a_0 + a_1u_x + \cdots + a_nu_x^n - t_x) \in F[a_0,\dots,a_n].$$

For any F-algebra F' and any $g \in \mathrm{Poly}_{\leq n/F}(F')$, we have

$$P_{x,n}(g) = \mathrm{N}_{(F(x)\otimes_F F')/F'}(g(u_x) - t_x \otimes 1) \in F'.$$

Lemma 5.3. *Assume $n \geq 1$. For each $x \in Z_{f_1} \cap Z_{f_2}$ such that $F(x)/F$ is separable, $P_{x,n}$ is irreducible in the coordinate ring of $\mathrm{Poly}_{\leq n/F}$. If x and x' are two such distinct points, then $P_{x,n}$ and $P_{x',n}$ are not unit multiples of each other in this coordinate ring.*

If we do not assume $F(x)/F$ to be separable, then $P_{x,n}$ need not be irreducible. For example, if F has characteristic $p > 0$ and $F(x)$ is a purely inseparable extension of F with degree p^2 such that the fields $F(u_x)$ and $F(t_x)$ have degree p over F, then $P_{x,n}$ is a pth power.

Proof. Since the extension $F(x)/F$ is finite separable and $P_{x,n}$ is a norm-form of a polynomial in $F(x)[a_0,\dots,a_n]$ whose coefficients generate $F(x)$ over F (since $n \geq 1$), the irreducibility is obvious. If L/F is a finite Galois extension into which $F(x)$ admits an F-embedding, then over L we see that $P_{x,n}$ factors as a product of linear forms $P_{x_i,n}$ defined by the L-points x_i of $\mathbf{A}^2_F$ that lie over x. Thus, if x' is another point on $Z_{f_1} \cap Z_{f_2}$ such that $F(x')/F$ is separable then the geometric zero locus of $P_{x,n}$ is distinct from that of $P_{x',n}$. Hence, $P_{x,n}$ and $P_{x',n}$ are not unit multiples of each other. ■

Now assume F is perfect, so Lemma 5.3 applies to all $x \in Z_{f_1} \cap Z_{f_2}$. By Definition 4.5 we have

$$M^{\mathrm{geom}}_{f_1,f_2}(u) = \prod_{u_x} \mathrm{N}_{F(u_x)/F}(u - u_x) \in F[u] - \{0\}, \tag{5.6}$$

where u_x runs over the distinct images of the x's on the u-axis. In particular, $M^{\mathrm{geom}}_{f_1,f_2} = 1$ if Z_{f_1} and Z_{f_2} are disjoint. If $g_1, g_2 \in F[u]$ have respective large degrees n_1 and n_2 then from (5.6) and the definition $P_{x,n}(g) = \mathrm{N}_{F(x)/F}(g(u_x) - t_x)$ for $n \geq \deg g$ we see

$$g_1 \equiv g_2 \bmod M^{\mathrm{geom}}_{f_1,f_2} \Longrightarrow P_{x,n_1}(g_1) = P_{x,n_2}(g_2)$$

where $n_j = \deg g_j$.

For $M := M^{\mathrm{geom}}_{f_1,f_2} \neq 0$, consider the division-algorithm morphism $\rho_{n,M}$ as in (5.1) and (5.2). Assume $Z_{f_1} \cap Z_{f_2}$ is nonempty, so $M \notin F$. Choose $x \in Z_{f_1} \cap Z_{f_2}$, so $M(u_x) = 0$. Clearly $P_{x,n} = P_{x,M} \circ \rho_{n,M}$ for the algebraic function $P_{x,M}$ on $\mathrm{Poly}_{\leq(\deg M - 1)/F}$ given by the norm construction $g \mapsto \mathrm{N}_{(F(x)\otimes_F F')/F'}(g(u_x) - t_x)$ for F-algebras F' and $g \in F'[u]$ with degree at most $\deg M - 1$.

Lemma 5.4. *Let $f_1, f_2 \in F[u,T]$ be nonzero and relatively prime, with zero loci Z_{f_1} and Z_{f_2} in $\mathbf{A}^2_F$. Assume that F is perfect. For n sufficiently large there exists a unique $b_n \in F^\times$*

and integers $e_n \geq 0$ *and* $e_{x,n} > 0$ *for all* $x \in Z_{f_1} \cap Z_{f_2}$ *such that*

$$R_n(G) = b_n a_n^{e_n} \cdot \prod_x P_{x,n}^{e_{x,n}} = b_n a_n^{e_n} \cdot \prod_x P_{x,M}^{e_{x,n}} \circ \rho_{n,M} \tag{5.7}$$

as algebraic functions on $\mathrm{Poly}_{n/F}$*, where* $M = M_{f_1,f_2}^{\mathrm{geom}}$*. The exponent* e_n *is positive if and only if* $\deg_T f_1, \deg_T f_2 > 0$*.*

Remark 5.5. The algebraic functions in (5.7) are all polynomial functions on $\mathrm{Poly}_{\leq n/F}$ (*i.e.*, there is no intervention of $1/a_n$).

The functorial construction of $R_n(G)$ as a universal resultant for large n (an alternative to the explicit definition (5.5)) only makes sense over $\mathrm{Poly}_{n/F}$ and not over $\mathrm{Poly}_{\leq n/F}$, so it does not seem possible to use geometric methods alone to determine how the discrete parameters e_n and b_n depend on n (though clearly b_n is generally sign-dependent on the ordering of the pair f_1 and f_2). A geometric interpretation of the $e_{x,n}$'s is given in Theorem 5.7; the product over $x \in Z_{f_1} \cap Z_{f_2}$ in (5.7) is understood to be 1 if $Z_{f_1} \cap Z_{f_2}$ is empty.

Proof. Let us first show $a_n | R_n(G)$ in $F[a_0, \ldots, a_n]$ if and only if both $\deg_T f_j$'s are positive. Specializing $R_n(G)$ into a field in which a_n vanishes causes $R_n(G)$ to specialize to 0 if both $\deg_T f_j$'s are positive and n is large (as then the $f_j(G)$'s have leading coefficients divisible by a_n). If some $\deg_T f_j$ vanishes, say $\deg_T f_1 = 0$, then specializing a_n to zero causes $R_n(G)$ to have non-vanishing specialization because $f_1(u)$ must be relatively prime to $f_2(u, a_0 + a_1 u + \cdots + a_{n-1}u^{n-1})$ (as $f_1(u)$ is relatively prime to $f_2(u,T)$). Thus, the geometric zero locus for $R_n(G) \in F[a_0, \ldots, a_n]$ on $\mathrm{Poly}_{\leq n/F} \simeq \mathbf{A}_F^{n+1}$ contains the hyperplane $a_n = 0$ when both $\deg_T f_j$'s are positive and otherwise it does not contain this hyperplane.

Since the irreducible $P_{x,n}$'s are not scalar multiples of a_n, to establish (5.7) it remains (by the Nullstellensatz) to show that the restriction of $R_n(G)$ to $\mathrm{Poly}_{n/F}$ has geometric zero locus equal to the union of the geometric zero loci of the $P_{x,n}$'s. If $\overline{F}/F$ is an algebraic closure then by separability of $F(x)/F$ the irreducible factorization of $P_{x,n}$ in $\overline{F}[a_0, \ldots, a_n]$ is the product of the $P_{x_i,n}$'s for the $\overline{F}$-points x_i of $\mathbf{A}_F^2$ over the physical point x. Thus, we may assume F is algebraically closed and we wish to prove that if $g \in F[u]$ has large exact degree n then the resultant of $f_1(u, g(u))$ and $f_2(u, g(u))$ vanishes if and only if $g(u_x) = t_x$ for some x in the intersection of the zero-loci Z_{f_j}. But this is obvious since the vanishing of the resultant says that $f_1(u, g(u))$ and $f_2(u, g(u))$ have a common root $u_0 \in F$, and then $x = (u_0, g(u_0))$ lies on both zero-loci Z_{f_j}. ∎

Corollary 5.6. *Let* F *be a perfect field with positive characteristic* p *and* $f(T) \in F[u,T]$ *a nonzero squarefree element.*

1) *If* f *lies in* $F[u][T^p]$ *then, for* g *of sufficiently large degree, the property of* $f(g)$ *being separable in* $F[u]$ *is determined by* $g \bmod M_f^{\mathrm{geom}}$*.*

2) *If* $f(T^p)$ *is squarefree in* $F[u,T]$*, then for* g *of sufficiently large degree, the property of* $f(g^p)$ *being separable in* $F[u]$ *is determined by* $g \bmod M_f^{\mathrm{geom}}$*.*

The "sufficient largeness" of $\deg g$ *may be chosen uniformly with respect to arbitrary extensions of* F*.*

For the study of $p = 2$ we will need the second case in this corollary.

Proof. The case $f \in F^\times$ is trivial, so we may assume $f \notin F$. Thus, in either case, Lemma 4.4 assures us that $\partial_u f \neq 0$ and that f and $\partial_u f$ have no nonconstant common factor in $F[u,T]$ (so M_f^{geom} makes sense). Hence, we may apply Lemma 5.4 with $f_1 = f$ and $f_2 = \partial_u f$ to

conclude that for g with large degree, the vanishing of the resultant of $f(g)$ and $(\partial_u f)(g)$ only depends on $g \bmod M_f^{\mathrm{geom}}$. Also, $f(g)$ is inseparable in $F[u]$ precisely when it has a common geometric root with its derivative $f(g)'$.

In case (1), $f(g)' = (\partial_u f)(g)$ has a common geometric root with $f(g)$ if and only if the resultant of $f(g)$ and $(\partial_u f)(g)$ vanishes. Since $(f(g^p))' = (\partial_u f)(g^p)$, in case (2) we see that separability of $f(g^p)$ only depends on $g^p \bmod M_f^{\mathrm{geom}}$ for $\deg(g) \gg 0$. ■

A defect in Lemma 5.4 is that it does not provide a description of how b_n, e_n, and $e_{x,n}$ depend on large n. These deficiencies are settled by:

Theorem 5.7. *Let F be a perfect field and $f_1, f_2 \in F[u,T]$ nonzero and relatively prime, with $Z_{f_j} \subseteq \mathbf{A}^2_F$ the zero locus of f_j. Let b_n, e_n, and $e_{x,n}$ be as in Lemma 5.4 for the ordered pair (f_1, f_2). For large n and $x \in Z_{f_1} \cap Z_{f_2}$, $e_{x,n}$ is equal to the intersection number $i_x(Z_{f_1}, Z_{f_2})$ at x. Also, there exist unique $\beta_0, \beta_1 \in F^\times$ such that $b_n = \beta_0 \beta_1^n$ for large n, and if $\deg_T f_1, \deg_T f_2 > 0$ then e_n is a polynomial in n with degree ≤ 1 and $\mathbf{Z}$-coefficients for large n.*

In particular, for the fixed choice of ordered pair (f_1, f_2), there exist $c \in F^\times$, integers m_0 and m_1 with $m_1 \geq 0$, and an algebraic function $L_{f_1,f_2} : F[u]/(M) \to \mathbf{A}^1_F$ for some nonzero $M \in F[u]$ such that for large n there is an equality of algebraic functions

$$R_n(G) = c^n a_n^{m_0 + m_1 n} \cdot (L_{f_1,f_2} \circ \rho_{n,M}) \tag{5.8}$$

on $\mathrm{Poly}_{n/F}$, with $\rho_{n,M}$ as in (5.1) or (5.2). If $\deg_T f_1, \deg_T f_2 > 0$ then $m_1 > 0$.

For f as in Lemma 4.4, the pair $f_1 = f$ and $f_2 = \partial_u f$ satisfies the hypotheses in Theorem 5.7. A local calculation shows that in this case $B = Z_{f_1} \cap Z_{f_2}$ is the non-étale locus for projection from $\{f = 0\}$ onto the T-axis, and $i_x(Z_{f_1}, Z_{f_2})$ is the length of B at x. As an illustration (via Example 4.2), for f as in Example 1.1 the projection from the plane curve $\{f = 0\}$ onto the T-axis is non-étale at precisely the geometric points $(0,0)$ and $(1,t)$ with $t^2+1 = 0$, and the branch scheme has respective lengths 18 and 9 at these points. Theorem 5.7 thereby explains why $\mu(f(g))$ has the form given in (1.3), since we can equivalently write it as the quadratic symbol for $g(0)^{18}(g(1)^2+1)^9$.

Before we prove Theorem 5.7, we use it to prove Theorem 4.7.

Proof. (of Theorem 4.7) For g in $\kappa[u]$ of sufficiently large degree, $f(g)$ is nonzero and (3.6) and (4.2) yield

$$\begin{aligned} \mu(f(g)) &= (-1)^d \chi(\mathrm{disc}\, f(g)) \\ &= (-1)^d \chi(\mathrm{lead}\, f(g))^{d + \deg f(g)'} (\chi(-1))^{d(d-1)/2} \chi(R(f(g), f(g)')), \end{aligned}$$

with $d = \deg f(g)$. Note $f(g)' = (\partial_u f)(g)$. Since f is squarefree and $f \notin \kappa$, so $(\partial_u f)(T) \neq 0$ by Lemma 4.4, we have $(\partial_u f)(g) \neq 0$ when $\deg g \gg 0$.

When $\deg g \gg 0$, both $d = \deg f(g)$ and $\deg((\partial_u f)(g))$ are linear in $\deg g$ by (2.5). Using (2.5) and (2.6) to compute $\deg f(g)$ and $\mathrm{lead}\, f(g)$ in terms of $\deg g$ and $\mathrm{lead}\, g$ for $\deg g \gg 0$, we have by Theorem 5.7 that there exist $\varepsilon_0, \varepsilon_1 \in \{\pm 1\}$ and integers m_0 and m_1 such that for $\deg g \gg 0$,

$$\mu(f(g)) = \varepsilon_0 \varepsilon_1^{\deg g} (\chi(-1))^{(\deg f(g))(\deg f(g)-1)/2} \chi(\mathrm{lead}\, g)^{m_0 + m_1 \deg g} \chi(L(g)) \tag{5.9}$$

where L is an algebraic function on the affine space $\kappa[u]/(M_f^{\mathrm{geom}})$ over κ. This formula depends on $\deg g$ modulo 4. If -1 is a square in κ or $\deg_T f$ is a multiple of 4 then the formula (5.9) depends on $\deg g$ modulo 2.

Now let us establish the final part of Theorem 4.7 concerning the behavior of $M^{\min}_{f,\kappa'}$ for sufficiently large finite extensions κ' of κ. Let κ'/κ be a finite extension such that all points in the finite set $Z_f \cap Z_{\partial_u f} \subseteq \mathbf{A}^2_\kappa$ are κ'-rational, and so in particular M^{geom}_f splits into linear factors in $\kappa'[T]$. This rationality property is inherited by all finite extensions of κ'.

We claim that no proper factor of M^{geom}_f can serve as a modulus for $\mu_{\kappa''[u]}(f(g))$ with κ'' any finite extension of κ'. Since $M^{\min}_{f,\kappa''} | M^{\mathrm{geom}}_f$, we can assume M^{geom}_f is nonconstant.

Choose a monic linear factor of M^{geom}_f in $\kappa'[u]$, so it has the form $h = u - u_x$ for some (κ'-rational) point $x = (u_x, t_x) \in Z_f \cap Z_{\partial_u f}$. We can find polynomials g_1 and g_2 with any large degree n and a common leading coefficient such that $g_1(u_x) = t_x \neq g_2(u_x)$ and $g_1(u_{x'}) = g_2(u_{x'}) \neq t_{x'}$ for all $x' \in Z_f \cap Z_{\partial_u f}$ with $x' \neq x$, in which case (5.7) and the positivity of the exponents $e_{x'}$ ensure that for sufficiently large n we have the vanishing of the resultant of $f(g_1)$ and $(\partial_u f)(g_1) = f(g_1)'$ and the non-vanishing of the resultant of $f(g_2)$ and $f(g_2)'$; that is, $\mu_{\kappa'[u]}(f(g_1)) = 0$ and $\mu_{\kappa'[u]}(f(g_2)) \neq 0$. The same properties persist after replacing κ' with any finite extension κ''. Since g_1 and g_2 are clearly congruent modulo $M^{\mathrm{geom}}_f/(u - u_x)$, we conclude that this divisor of M^{geom}_f cannot be a modulus for $\mu_{\kappa''[u]}(f(g))$ and so cannot be divisible by $M^{\min}_{f,\kappa''}$. Thus, the monic factor $M^{\min}_{f,\kappa''}$ of the monic M^{geom}_f must equal M^{geom}_f. ∎

Let us now prepare for the proof of Theorem 5.7. We first establish a key point: if $e_{x,n} = i_x(Z_{f_1}, Z_{f_2})$ for sufficiently large n and all $x \in Z_{f_1} \cap Z_{f_2}$ (or more generally, if $e_{x,n}$ is independent of n for $n \gg 0$ and all such x) then a formula of the shape (5.8) holds for some nonzero M if and only if e_n is a $\mathbf{Q}$-polynomial of degree ≤ 1 in n for large n and $b_n = \beta_0 \beta_1^n$ for some $\beta_0, \beta_1 \in F^\times$ for large n. Sufficiency is obvious by Theorem 5.4, and for necessity we may replace M with $M M^{\mathrm{geom}}_{f_1,f_2}$ to get to the case where $M^{\mathrm{geom}}_{f_1,f_2} | M$, so with $e_x := i_x(Z_{f_1}, Z_{f_2}) > 0$ we have formulas

$$R_n(G) = b_n a_n^{e_n} \prod_x P^{e_x}_{x,M} \circ \rho_{n,M}$$

and

$$R_n(G) = c^n a_n^{m_0 + m_1 n} \cdot L_{f_1,f_2} \circ \rho_{n,M}$$

on $\mathrm{Poly}_{n/F}$ for large n. Thus, for large n the rational function

$$g \mapsto b_n c^{-n} a_n(g)^{e_n - (m_0 + m_1 n)}$$

on $\mathrm{Poly}_{n/F}$ factors through $\rho_{n,M}$, or equivalently for generic (or universal) g it only depends on g mod M. This forces $e_n = m_0 + m_1 n$ for large n, so

$$(b_n c^{-n}) \prod_x P^{e_x}_{x,M} \circ \rho_{n,M} = L_{f_1,f_2} \circ \rho_{n,M}$$

for large n. We can assume $\deg M > 0$, so for large n we have

$$b_n c^{-n} \prod_x P^{e_x}_{x,M} = L_{f_1,f_2}$$

on $\mathrm{Poly}_{\leq (\deg M - 1)/F}$. Since L_{f_1,f_2} and the $P^{e_x}_{x,M}$'s do not depend on n, we conclude that $b_n c^{-n} \in F^\times$ is equal to a constant c' that does not depend on large n. Thus, $b_n = c' c^n$ for $c, c' \in F^\times$ and large n, as desired.

We shall now aim to prove $e_{x,n} = i_x(Z_{f_1}, Z_{f_2})$ for all $x \in Z_{f_1} \cap Z_{f_2}$ and large n, as well as an identity of the form (5.8), by means of induction on the ordered pair (f_1, f_2). The

flexibility in the choice of M will be essential for the success of the induction. For example, the preceding argument shows that if this goal is satisfied for a particular pair (f_1, f_2) then upon replacing M with a nonzero multiple so that it is divisible by $M_{f_1,f_2}^{\mathrm{geom}}$ we must have

$$L_{f_1,f_2} = c_0 \prod_x P_{x,M}^{i_x(Z_{f_1},Z_{f_2})}$$

for some $c_0 \in F^\times$. In what follows we will work with a generic field-valued point g of the geometrically integral F-variety $\mathrm{Poly}_{n/F}$ for large n, though one can instead work throughout in the universal case with g having a unit leading coefficient and large degree n. Since local intersection numbers for plane curves enjoy properties analogous to the properties of resultants that were summarized below Example 4.1, our inductive manipulations with resultants below will be well-behaved with respect to the desired equality $e_{x,n} = i_x(Z_{f_1}, Z_{f_2})$ for large n and all $x \in Z_{f_1} \cap Z_{f_2}$.

Note that although $R(f_1(g), f_2(g))$ generally depends on the ordering of f_1 and f_2, the existence of an identity as in (5.8) does not depend on this ordering (nor do the intersection numbers between Z_{f_1} and Z_{f_2}). Indeed, for g of sufficiently large degree, say $\deg g > \nu(f_1), \nu(f_2)$ (see (2.7)),

$$\begin{aligned} R(f_1(g), f_2(g)) &= (-1)^{(\deg f_1(g))(\deg f_2(g))} R(f_2(g), f_1(g)) \\ &= (-1)^{e_0}(-1)^{e_1 \deg g} R(f_2(g), f_1(g)), \end{aligned}$$

where $e_0 = (\deg \alpha_{1,d_1})(\deg \alpha_{2,d_2})$ and $e_1 = d_1 \deg \alpha_{2,d_2} + d_2 \deg \alpha_{2,d_2} + d_1 d_2$, with $d_j = \deg_T f_j$ and $f_j = \sum \alpha_{j,i} T^i$. Thus, we need not be concerned with sign-changes in resultants when $f_1(g)$ and $f_2(g)$ are interchanged. We will use this repeatedly.

Our proof of Theorem 5.7 will roughly be a series of identities

$$R(f_1(g), f_2(g)) = c_0 c_1^{\deg g} (\mathrm{lead}\, g)^{m_0 + m_1 \deg g} R(f_3(g), f_4(g)), \quad Z_{f_1} \cap Z_{f_2} = Z_{f_3} \cap Z_{f_4}$$

for generic g of large degree (or universal g with a unit leading coefficient and large degree), where $c_0, c_1 \in F^\times$ and $m_0, m_1 \in \mathbf{Z}$, and the ordered pair (f_3, f_4) of nonzero relatively prime polynomials in $F[u, T]$ is in some sense smaller than (f_1, f_2). (There is more than one sense that we use, depending on the stage of our argument.) In this way, induction will establish (5.8).

To get started, the case when $f_1(T)$ has T-degree 0, say $f_1(T) = a(u) \in F[u]$, is trivial: writing $a(u) = c a_1(u)$ with $c \in F^\times$ and $a_1(u)$ monic,

$$R(a(u), f_2(g)) = R(c, f_2(g)) R(a_1(u), f_2(g)) c^{\deg f_2(g)} R(a_1(u), f_2(g)). \tag{5.10}$$

For $\deg g > \nu(f_2)$, $c^{\deg f_2(g)} = c_0 c_1^{\deg g}$ for suitable c_0 and c_1 in $F^\times$ that are independent of g. The factor $R(a_1(u), f_2(g))$ is an algebraic function of g modulo $a_1(u)$, since $a_1(u)$ is monic. This proves (5.8) in the present case for large n, and so we next have to relate $e_{x,n}$ to an intersection number at x for large n in this case.

For each $x \in Z_{f_1} \cap Z_{f_2} = Z_{a_1} \cap Z_{f_2}$, since $F(x)/F$ is separable it is clear that if F'/F is any extension and $\{x'_j\}$ is the finite set of points over x in $\mathbf{A}^2_{F'}$ then $P_{x,n} = \prod_j P_{x'_j,n}$ over F' and $i_x(Z_{f_1}, Z_{f_2}) = i_{x'_j}(Z_{f_1/F'}, Z_{f_2/F'})$. Hence, to identify each $e_{x,n}$ with the intersection number at x for large n we may assume F is algebraically closed. The monicity of a_1 and the bimultiplicativity of resultants and local intersection numbers reduce us to the case $a_1 = u - u_0$ for some $u_0 \in F$. By long division against $u - u_0$ we can assume $f_2 \in F[T]$, so since F is algebraically closed we can use bimultiplicativity to reduce to the case $f_2 = c(T - t_0)^{e_0}$ for some $c \in F^\times$, $t_0 \in F$, and $e_0 \geq 0$. The case $e_0 = 0$ is trivial, so we

can assume $c = 1$ and $e_0 = 1$. Since $R(u - u_0, g - t_0) = g(u_0) - t_0 = P_{(u_0,t_0)}(g)$, this case is settled.

To prove Theorem 5.7 in general, we can assume that the coefficients of f_1 as a polynomial in T have no common factor in $F[u]$, and similarly for f_2. Indeed, if $f_1(T) = a(u)h(T)$ for $a(u)$ in $F[u]$ (so f_2 is relatively prime to both $a(u)$ and $h(T)$ in $F[u, T]$), then

$$R(f_1(g), f_2(g)) = R(a(u), f_2(g))R(h(g), f_2(g)), \tag{5.11}$$

with the first factor on the right side satisfying the inductive hypothesis by the preceding discussion. In view of the bimultiplicativity of local intersection numbers, if the second factor on the right side in (5.11) satisfies the inductive hypothesis (for large n) then we will indeed be done. Removing a common factor from the coefficients of f_2 as a polynomial in T is also compatible with Theorem 5.7.

We will prove Theorem 5.7 by two inductions: on the maximum of $\deg_T f_1$ and $\deg_T f_2$ when these degrees are distinct, and for f_1 and f_2 of equal T-degree we will induct on the minimum u-degree of their leading coefficients as polynomials in T.

Lemma 5.8. *Let $h_1(T)$ and $h_2(T)$ in $F[u][T]$ have common T-degree $d \geq 1$:*

$$h_1(T) = \alpha(u)T^d + \cdots, \quad h_2(T) = \beta(u)T^d + \cdots.$$

Assume $\alpha \nmid \beta$ and $\beta \nmid \alpha$ (so $\alpha, \beta \notin F$). There exist $c \in F^\times$, $\varepsilon = \pm 1$, $m \in \mathbf{Z}$, and a second pair of polynomials $\widetilde{h}_1(T)$ and $\widetilde{h}_2(T)$ in $F[u][T]$ with T-degree d whose leading coefficients as polynomials in T, $\widetilde{\alpha}(u)$ and $\widetilde{\beta}(u)$, satisfy

$$\min(\deg \widetilde{\alpha}, \deg \widetilde{\beta}) < \min(\deg \alpha, \deg \beta) \tag{5.12}$$

such that for all extensions F'/F and all g in $F'[u]$ with sufficiently large degree (depending on h_1 and h_2, and uniform with respect to F')

$$R(h_1(g), h_2(g))c\varepsilon^{\deg g}(\operatorname{lead} g)^m R(\widetilde{h}_1(g), \widetilde{h}_2(g)). \tag{5.13}$$

If the h_j's are relatively prime in $F[u, T]$ then the $\widetilde{h}_j$'s must be relatively prime in $F[u, T]$.

Proof. We will prove the lemma when $\deg \alpha \leq \deg \beta$. (When $\deg \alpha > \deg \beta$, we can reduce to the other case by interchanging h_1 and h_2, at the cost of changing c and ε in the conclusion.) In $F[u]$, write $\beta(u) = \alpha(u)q(u) + r(u)$, where $r \neq 0$ and $\deg r < \deg \alpha$. Since $r \neq 0$, $k(T) := h_2(T) - q(u)h_1(T)$ has leading term $r(u)T^d$ as a polynomial in T with coefficients in $F[u]$. For all g, clearly $h_2(g) \equiv k(g) \bmod h_1(g)$. When $\deg g$ exceeds $\nu(h_1)$, $\nu(h_2)$, and $\nu(k)$ (see (2.7)), quasi-periodicity gives

$$\begin{aligned} R(h_1(g), h_2(g)) &= (\operatorname{lead} h_1(g))^{\deg h_2(g) - \deg k(g)} R(h_1(g), k(g)) \\ &= c(\operatorname{lead} g)^m R(h_1(g), k(g)), \end{aligned}$$

where $c = (\operatorname{lead} \alpha)^{\deg \beta - \deg r}$ and $m = d(\deg \beta - \deg r)$. Let $\widetilde{h}_1 = h_1$ and $\widetilde{h}_2 = k$, or $\widetilde{h}_1 = k$ and $\widetilde{h}_2 = h_1$. By Lemma 5.1, the identity (5.13) forces relative primality of the $\widetilde{h}_j$'s when the h_j's are relatively prime. ∎

Now we modify the hypothesis in the previous lemma. Rather than assuming the leading coefficients $\alpha(u)$ and $\beta(u)$ do not divide each other, we assume $h_1(T)$ and $h_2(T)$ are relatively prime as polynomials in T.

Lemma 5.9. *Let $h_1(T)$ and $h_2(T)$ in $F[u][T]$ have common T-degree $d \geq 1$:*

$$h_1(T) = \alpha(u)T^d + \cdots, \quad h_2(T) = \beta(u)T^d + \cdots.$$

Assume the h_j's are relatively prime in $F[u,T]$. There exist $c \in F^\times$, $\varepsilon = \pm 1$, $m \in \mathbf{Z}$, and a second pair of nonzero relatively prime polynomials $\widetilde{h}_1(T)$ and $\widetilde{h}_2(T)$ in $F[u][T]$ with $\deg_T \widetilde{h}_1 < \deg_T \widetilde{h}_2 = d$ such that for all extensions F'/F and all g in $F'[u]$ with sufficiently large degree (uniform with respect to F'),

$$R(h_1(g), h_2(g)) c\varepsilon^{\deg g}(\operatorname{lead} g)^m R(\widetilde{h}_1(g), \widetilde{h}_2(g)).$$

Proof. If neither α nor β divides the other in $F[u]$, apply Lemma 5.8 to get a second pair of polynomials in $F[u][T]$ with T-degree d. Repeat this process if again neither leading coefficient as a polynomial in T divides the other. (Note that terms like $c\varepsilon^{\deg g}(\operatorname{lead} g)^m$ behave well under multiplication: the c's and ε's are multiplicative, while the m's are additive.) The condition (5.12) ensures that we eventually reach the case where $\alpha(u)|\beta(u)$ or $\beta(u)|\alpha(u)$. Thus, we may interchange h_1 and h_2 if necessary to suppose $\alpha(u)|\beta(u)$. Write $\beta(u) = \alpha(u)q(u)$. The polynomial $k(T) := h_2(T) - q(u)h_1(T)$ has T-degree less than d. This polynomial is nonzero and is relatively prime to h_1 since $\gcd(h_1, h_2) = 1$. Proceed as in the proof of Lemma 5.8, taking $\widetilde{h}_1 = k$ and $\widetilde{h}_2 = h_1$. ∎

In the two preceding lemmas, if the h_j's are relatively prime then for the $\widetilde{h}_j$'s constructed in the proofs we have $Z_{h_1} \cap Z_{h_2} = Z_{\widetilde{h}_1} \cap Z_{\widetilde{h}_2}$ scheme-theoretically in $\mathbf{A}^2_F$. In particular, local intersection numbers are unaffected by replacing the pair (h_1, h_2) with the pair $(\widetilde{h}_1, \widetilde{h}_2)$. This observation will be used without comment below.

We are finally ready to prove Theorem 5.7:

Proof. (of Theorem 5.7). We argue by induction on $\max(\deg_T f_1, \deg_T f_2)$. Set $d_1 = \deg_T f_1$ and $d_2 = \deg_T f_2$. We can assume both d_1 and d_2 are positive, since the cases $d_1 = 0$ or $d_2 = 0$ have been settled via (5.10). Remove any nontrivial common factor from the $F[u]$-coefficients of $f_1(T)$ as a polynomial in T, using (5.11), so $f_1(T)$ is primitive over $F[u]$. Similarly make f_2 primitive. By Lemma 5.9 and induction, we may assume $d_1 \neq d_2$, and without loss of generality $0 < d_1 < d_2$. Writing

$$f_1(T) = \alpha(u)T^{d_1} + \ldots, \quad f_2(T) = \beta(u)T^{d_2} + \ldots, \tag{5.14}$$

we wish to reduce to the case $\deg \beta < \deg \alpha$ (at the expense of possibly losing the primitivity condition for f_2 but not for f_1).

Write $\beta(u) = \alpha(u)q(u) + r(u)$, where $r = 0$ or $\deg r < \deg \alpha$. The polynomial $k(T) = f_2(T) - q(u)T^{d_2-d_1}f_1(T)$ is nonzero and relatively prime to f_1. If r is nonzero, then $k(T)$ has leading term $r(u)T^{d_2}$. If $r = 0$, then $\deg_T k < d_2$. In either case, $f_2(g) \equiv k(g) \bmod f_1(g)$ for all field-valued points g of $\mathrm{Poly}_{n/F}$. When $n = \deg g$ is sufficiently large,

$$R(f_1(g), f_2(g)) = (\operatorname{lead} f_1(g))^{\deg f_2(g) - \deg k(g)} R(f_1(g), k(g)).$$

The power of $\operatorname{lead} f_1(g)$ can be written in the form $c_0 c_1^{\deg g}(\operatorname{lead} g)^{m_0 + m_1 \deg g}$ for suitable c_0, c_1 in $F^\times$ and integers m_0 and m_1 that do not depend on g. (The number m_1 is nonzero when $\deg_T k < d_2$.) By construction $Z_{f_1} \cap Z_{f_2} = Z_{f_1} \cap Z_k$ scheme-theoretically, so we are now reduced to proving Theorem 5.7 with f_2 replaced by k.

Either $\deg_T k = d_2$ and the leading coefficient of k as a polynomial in T has smaller degree than $\deg \alpha$, or $\deg_T k < d_2$. In the latter case, $\max(\deg_T f_1, \deg_T k) < d_2$, so Theorem 5.7 with f_1 and k has already been proved by the inductive hypothesis. Thus, it remains to treat

the case (5.14) with $\deg\beta < \deg\alpha$; observe that this reduction step preserves primitivity for f_1 but possibly loses it for f_2.

Our resultant now looks like $R(f_1(g), f_2(g)) = R(\alpha(u)g^{d_1} + \cdots, \beta(u)g^{d_2} + \cdots)$. Since $d_1 < d_2$, it is natural to want to reduce $f_2(g)$ modulo $f_1(g)$ and use quasi-periodicity, hoping to lower the maximum T-degree of the pair f_1, f_2 in our resultants. However, $\deg\beta < \deg\alpha$, so there is no progress through a division algorithm on the leading coefficients as in the proof of Lemma 5.8.

We now apply a generalization of the trick with $u+1$ in (4.10). Consider the universal identity

$$R(f_1(g), \alpha(u))R(f_1(g), f_2(g)) = R(f_1(g), \alpha(u)f_2(g)) \tag{5.15}$$

with g the universal polynomial of large degree n with a unit leading coefficient. The first term in (5.15) is nonzero, since primitivity of f_1 forces $\gcd(f_1(g), \alpha(u)) = 1$. Since all three resultants admit expressions as in Theorem 5.4 for a common modulus M, and since we know that if the $e_{x,n}$'s for large n have been proved to be independent of n then an identity as in (5.8) is equivalent to e_n being a $\mathbf{Z}$-polynomial of degree ≤ 1 in n and b_n having the form $\beta_0\beta_1^n$ for large n (for some $\beta_0, \beta_1 \in F^\times$), it is obvious (again with the help of bimultiplicativity of local intersection numbers) that if the theorem is proved for two of the three pairs (f_1, α), (f_1, f_2), and $(f_1, \alpha f_2)$ then it follows for the third. Since the case of a polynomial of T-degree zero has already been settled, it suffices to prove (5.8) for the ordered pair $(f_1, \alpha(u)f_2)$.

The right side of (5.15) has the form $R(\alpha(u)g^{d_1} + \cdots, \alpha(u)\beta(u)g^{d_2} + \cdots)$. Let $h(T) = \alpha(u)f_2(T) - \beta(u)f_1(T)T^{d_2-d_1}$. Since $\gcd(f_1, f_2) = 1$ and f_1 is primitive over $F[u]$, and we may assume $\deg_T f_1 > 0$, it follows that h is nonzero and satisfies $\deg_T h < d_2$ and $\gcd(f_1, h) = 1$. Since $h(g) \equiv \alpha(u)f_2(g) \bmod f_1(g)$ for all g, when $\deg g \gg 0$ the right side of (5.15) is

$$\begin{aligned} R(f_1(g), \alpha(u)f_2(g)) &= (\operatorname{lead} f_1(g))^{\deg\alpha + \deg f_2(g) - \deg h(g)} R(f_1(g), h(g)) \\ &= c_0 c_1^{\deg g} (\operatorname{lead} g)^{m_0 + m_1 \deg g} R(f_1(g), h(g)) \end{aligned}$$

for suitable c_0, c_1 in $F^\times$ and integers m_0 and m_1. (For instance, $m_1 = d_2 - \deg_T h$.) Since $\deg_T f_1$ and $\deg_T h$ are both less than d_2, the theorem holds for the pair (f_1, h) by induction on the maximum T-degree. The scheme-theoretic equality $Z_{f_1} \cap Z_{\alpha f_2} = Z_{f_1} \cap Z_h$ therefore allows us to infer the desired result for the pair $(f_1, \alpha f_2)$. ∎

Corollary 5.10. *Let F be a perfect field of characteristic $p > 0$ and let $f_1, f_2 \in F[u, T]$ be nonzero and relatively prime. Assume $f_j = h_j(u, T^{p^m})$ with $m \geq 0$. For each $x \in Z_{f_1} \cap Z_{f_2}$, the multiplicity e_x of $P_{x,n}$ as a factor of the algebraic function $g \mapsto R(f_1(g), f_2(g))$ on $\mathrm{Poly}_{n/F}$ for sufficiently large n is equal to $p^m \cdot i_{(1\times\phi^m)(x)}(Z_{h_1}, Z_{h_2})$, with ϕ the relative Frobenius on the T-line over F.*

Proof. By Theorem 5.7, we just have to prove

$$i_x(Z_{f_1}, Z_{f_2}) = p^m \cdot i_{(1\times\phi^m)(x)}(Z_{h_1}, Z_{h_2}).$$

Since F is perfect and the relative Frobenius map commutes with extension of the base field we can assume F is algebraically closed. By a linear translation we can assume $x = (0, 0)$. Thus, the finite F-algebra $F[\![u, T]\!]/(f_1, f_2)$ is identified with the scalar extension of $F[\![u, T]\!]/(h_1, h_2)$ via the local F-algebra map $F[\![T]\!] \to F[\![T]\!]$ given by $T \mapsto T^{p^m}$. This extension of scalars is finite flat of degree p^m, so it multiplies the vector-space dimension by p^m. ∎

Example 5.11. The case of most interest in Corollary 5.10 is $f_1 = f$ and $f_2 = \partial_u f$ with f as in Corollary 5.6(1) and $p \neq 2$. For "generic" such f and maximal m the associated polynomials h_1 and h_2 are T-separable, so the intersection numbers between Z_{h_1} and Z_{h_2} are all equal to 1. Hence, for "generic" such f the exponents e_x are a power of p (and hence odd if $p \neq 2$). Also, if F is a p-adic field then for a "generic" f in $F[u, T^p]$ the intersection numbers between Z_f and $Z_{\partial_u f}$ are all equal to 1; this latter situation will be relevant for our work with $p = 2$.

Although we have given a geometric interpretation to the exponents $e_{x,n}$ for large n, as we noted after Remark 5.5, we lack a good understanding of the constants $\beta_0, \beta_1 \in F^\times$ (with $b_n = \beta_0 \beta_1^n$ for large n) and the $\mathbf{Z}$-polynomial function e_n of degree ≤ 1. For example, if we write $e_n = m_0 + m_1 n$ for large n then we do not know a conceptual interpretation of the parity of the m_j's; such parities influence Möbius bias, as we shall see later. Due to our poor understanding of β_0, β_1, and e_n, to compute Möbius formulas in large degrees for specific examples it seems unavoidable to essentially carry out the recursive algebraic procedure in the proof of Theorem 5.7.

6. Characteristic 2

The analogue of Theorem 4.7 in characteristic 2 is subtle because (3.5) in characteristic 2 requires liftings into characteristic 0. Fix a perfect field k of characteristic 2, and let $W = W(k)$ (the Witt vectors of k) and $F = \mathrm{Frac}(W)$.

Hypothesis. Our running convention throughout this section is that h denotes a polynomial in $k[u, T]$ such that $h \notin k$ and $h(T^2)$ is squarefree in $k[u, T]$.

This hypothesis forces h to be squarefree in $k[u, T]$ and not to have any irreducible factors in $k[T]$, and also forces $h(g^2) \neq 0$ for all $g \in k[u]$. We are interested in studying specializations of $h(T^2)$ on $k[u]$ for finite k, but we will initially focus on $h(T)$ for any perfect k with characteristic 2.

Since $h \notin k$, Lemma 4.4(2) ensures $\partial_u h \neq 0$ and that there is no common irreducible factor of h and $\partial_u h$ in $k[u, T]$. Thus, $R_{k[u]}(h, \partial_u h) \neq 0$ and we may define M_h^{geom} as in Definition 4.5. We emphasize that M_h^{geom} is not to be confused with $M_{h(T^2)}^{\mathrm{geom}}$; in our study of Möbius bias for specializations of $f(T) = h(T^2)$ in characteristic 2, it is M_h^{geom} that will turn out to be of more interest than M_f^{geom}. Corollary 5.6(2) ensures that the separability property of $h(g^2)$ in $k[u]$ only depends on g mod M_h^{geom} provided that $\deg g$ is sufficiently large, with largeness that depends on h and is uniform with respect to all perfect extensions of k.

Since $h(T^2)$ is squarefree in $k[u, T]$ and $h \notin k$, we can find $g \in k[u]$ of any sufficiently large degree such that $h(g^2)$ is nonconstant and separable in $k[u]$: use [19, Theorem 3.1] if k is finite, and use Lemma 5.1 and the Zariski-denseness of the locus of k-rational points in an affine space over k if k is infinite. In particular, $(\partial_u h)(g^2) = h(g^2)'$ is nonzero and $R_k(h(g^2), h(g^2)')$ is nonzero. Fix such a choice of g; concretely, g is a representative of some member of a (nonempty) collection of residue classes modulo M_h^{geom}.

Let H be a lift of h in $W[u][T] = W[u, T]$ such that $\deg_T H = \deg_T h$ and $\mathrm{lead}_T(H) \in W[u]$ has the same u-degree as $\mathrm{lead}_T(h) \in k[u]$, so $\mathrm{lead}_T(H) \in W[u]$ has unit leading coefficient and reduces to $\mathrm{lead}_T(h) \in k[u]$. Let $G \in W[u]$ be a lift of g with unit leading coefficient (so $\deg G = \deg g$). Assume $\deg g$ is sufficiently large so that the degree of $h(g^2) \in k[u]$ is given by the generic formula as in (2.5), and likewise for the degree of $H(G^2)$.

Note that $H(G^2) \in W[u]$ has unit leading coefficient (and hence the same degree as $h(g^2)$), so $W[u]/(H(G^2))$ is a finite flat W-algebra that lifts the finite étale k-algebra $k[u]/(h(g^2))$. By (3.6), we need to understand how the unit discriminant $\mathrm{disc}_W(H(G^2))$ mod $8W$ depends on G.

Though $H(G^2)' \neq (\partial_u H)(G^2)$ in characteristic 0, the mod-2 reductions agree. Thus, the F-resultants

$$R_F(H(G^2), H(G^2)'), \;\; R_F(H(G^2), (\partial_u H)(G^2)) \tag{6.1}$$

lie in W and have reductions in k that are $k^\times$-multiples of each other (see (4.4) and the Warning above Example 4.1). Both reductions therefore lie in $k^\times$ since $h(g^2)$ is separable, so both terms in (6.1) lie in $W^\times$. The quadratic nature of the first resultant in (6.1) intervenes in the study of $\mathrm{disc}_W(H(G^2))$, and the second resultant in (6.1) is a form to which Theorem 5.4 and Theorem 5.7 may be applied (over the field F of characteristic zero). We are going to show that the unit ratio of the resultants in (6.1) can be made explicit in $(W/8W)^\times$ modulo unit-square factors, so we will be able to use Theorems 5.4 and 5.7 to study the quadratic nature of $\mathrm{disc}_W(H(G^2))$.

The leading coefficient of $H(G^2)$ is a unit and the reduction $h(g^2)$ is separable, so the roots of $H(G^2)$ in an algebraic closure $\overline{F}$ are integral, lie in an unramified extension of F, and have pairwise-distinct reductions. Let $\{\alpha\}$ be the (nonempty) set of roots of $H(G^2)$ in $\overline{F}$, with $\overline{\alpha}$ denoting the reduction of α, so $(\partial_u h)(g^2)(\overline{\alpha}) = (h(g^2))'(\overline{\alpha})$ is nonzero and hence $(\partial_u H)(G^2)(\alpha)$ is an integral unit for all α.

Since $H(G^2)' = (\partial_u H)(G^2) + 2(\partial_T H)(G^2)GG'$, the classical formula (4.3) for resultants in terms of products over geometric roots gives

$$\frac{R_F(H(G^2), H(G^2)')}{R_F(H(G^2), (\partial_u H)(G^2))} = \mathrm{lead}(H(G^2))^{d_G} \prod_\alpha \left(1 + 2 \cdot \frac{(\partial_T H)(G^2)GG'}{(\partial_u H)(G^2)}\bigg|_\alpha\right), \tag{6.2}$$

where $d_G = \deg(H(G^2)') - \deg((\partial_u H)(G^2))$ is a "universal" polynomial of degree at most 1 in $\deg G = \deg g$ when $\deg g$ is large. The largeness depends on H but is uniform with respect to perfect extensions of k.

Remark 6.1. For $\deg g$ large, $d_G = 0$ if $\mathrm{lead}_T H \in W[u]$ is nonconstant (or equivalently, if $\mathrm{lead}_T h \in k[u]$ is nonconstant). If $\mathrm{lead}_T H \in W^\times$, then for $\deg g$ large we have

$$\begin{aligned} d_G &= 2\deg_T h \deg g - 1 - \deg(\mathrm{lead}_T \partial_u H) - 2(\deg_T \partial_u H)\deg G. \\ &= 2(\deg_T h - \deg_T \partial_u H)\deg g - (1 + \deg(\mathrm{lead}_T \partial_u H)). \end{aligned}$$

We need to understand the product in (6.2) modulo $8W$. The remarkable surprise is that there is a very simple formula for this product mod $8W$ (see (6.4)), and the formula only depends on g and h (not on G or H). This formula uses residues of a certain differential form. We need to make two definitions before we can state the formula of interest.

Definition 6.2. For any perfect field K and any rational differential form ω on $\mathbf{P}^1_K$, set

$$s_2(\omega) := \sum_{\{y_1, y_2\}} \mathrm{Res}_{y_1}\omega \cdot \mathrm{Res}_{y_2}\omega \in K, \tag{6.3}$$

where the sum runs over unordered pairs of distinct geometric poles of ω on $\mathbf{P}^1_K$.

In words, $s_2(\omega)$ is the second symmetric function of the geometric residues of ω. Our interest in $s_2(\omega)$ will be restricted largely to cases when ω has simple poles. We are grateful to Gabber for pointing out to us that, for ω varying with only simple geometric poles, $s_2(\omega)$

is not algebraic in ω if we do not fix the number of simple geometric poles of ω. For example, let

$$\omega = b \cdot \frac{\mathrm{d}u}{u} + \frac{\mathrm{d}u}{u-a},$$

with $b, b+1 \neq 0$. This has three simple poles when $a \neq 0$ and two simple poles when $a = 0$. If $a \neq 0$ then $s_2(\omega) = -b(b+1) - 1$, but if $a = 0$ then $s_2(\omega) = -(b+1)^2$. This non-algebraicity is analogous to the fact that (5.3) does not extend to an algebraic function on $\mathrm{Poly}_{\leq n/F}$.

Definition 6.3. For $\gamma \in k[u]$, define

$$\omega_{h,\gamma} := \frac{(\partial_T h)(\gamma^2)\gamma}{h(\gamma^2)}\,\mathrm{d}\gamma;$$

the initial hypotheses on $h \in k[u][T]$ in this section ensure that $h(\gamma^2) \neq 0$.

When γ is a square in $k[u]$ (so $\mathrm{d}\gamma = 0$) or h is a polynomial in T^2 (so $\partial_T h = 0$), clearly $\omega_{h,\gamma} = 0$. For $g \in k[u]$ with large degree such that $h(g^2)$ is separable, we may write

$$\omega_{h,g} = \frac{(\partial_T h)(g^2)g^2}{h(g^2)} \cdot \frac{\mathrm{d}g}{g},$$

so this rational differential form on $\mathbf{P}^1_k$ has simple poles. We will see in Theorem 6.10 that $s_2(\omega_{h,\gamma})$ intervenes in the behavior of $\mu(h(\gamma^2))$ when k is finite. The vanishing of $s_2(\omega_{h,\gamma^2})$ will therefore make the behavior of $\mu(h(\gamma^4))$ quite tractable for finite k.

Theorem 6.4. *Choose $H \in W[u,T]$ reducing to $h \in k[u,T]$ such that $\mathrm{lead}_T(H) \in W[u]$ has unit leading coefficient in W. For $g \in k[u]$ of large degree with $h(g^2)$ separable and $G \in W[u]$ lifting g with $\mathrm{lead}(G) \in W^\times$,*

$$\prod_\alpha \left(1 + 2 \cdot \left.\frac{(\partial_T H)(G^2)GG'}{(\partial_u H)(G^2)}\right|_{u=\alpha}\right) \equiv 1 + 2\deg g \deg_T h + 4s_2(\omega_{h,g}) \bmod 8W, \tag{6.4}$$

where α runs over the geometric roots of $H(G^2)$. The largeness of $\deg g$ depends on H and may be chosen uniformly with respect to perfect extensions of k.

Proof. Let $P = H(G^2)$. Since P has simple zeros at each of its roots α, and hence serves as a local coordinate there, we get the residue description

$$\left.\frac{(\partial_T H)(G^2)GG'}{(\partial_u H)(G^2)}\right|_{u=\alpha} = \mathrm{Res}_\alpha\left(\frac{(\partial_T H)(G^2)GG'}{(\partial_u H)(G^2)} \cdot \frac{\mathrm{d}P}{P}\right). \tag{6.5}$$

We will first show

$$2\left.\frac{(\partial_T H)(G^2)GG'}{(\partial_u H)(G^2)}\right|_{u=\alpha} \equiv 2\mathrm{Res}_\alpha\omega_{H,G} + 4(\mathrm{Res}_\alpha\omega_{H,G})^2 \bmod 8\overline{W}, \tag{6.6}$$

where $\overline{W}$ is the integral closure of W in an algebraic closure $\overline{F}$ of F. Note that we can replace the residue in the final term in the mod-8 equation (6.6) with a residue in characteristic 2, namely $\mathrm{Res}_{\overline{\alpha}}(\omega_{h,g})$ with $\overline{\alpha}$ the reduction of α.

Since $(H(G^2))' \equiv (\partial_u H)(G^2) \bmod 2W[u]$ with $H(G^2)'(\alpha) \in \overline{W}^\times$, we have

$$\mathrm{Res}_\alpha\left(\frac{((\partial_T H)(G^2)GG')^2}{(\partial_u H)(G^2)H(G^2)}\,\mathrm{d}u\right) \equiv \mathrm{Res}_\alpha\left(\frac{(\partial_T H)(G^2)GG'}{(\partial_u H)(G^2)}\right)^2 \frac{\mathrm{d}H(G^2)}{H(G^2)} \bmod 2\overline{W}.$$

However,

$$\begin{aligned}\frac{(\partial_T H)(G^2)GG'}{(\partial_u H)(G^2)}\cdot\frac{\mathrm{d}P}{P} &= \frac{(\partial_T H)(G^2)GG'((\partial_u H)(G^2)+2(\partial_T H)(G^2)GG')}{(\partial_u H)(G^2)H(G^2)}\,\mathrm{d}u\\ &= \frac{(\partial_T H)(G^2)G}{H(G^2)}\,\mathrm{d}G+2\frac{((\partial_T H)(G^2)(GG'))^2}{(\partial_u H)(G^2)H(G^2)}\,\mathrm{d}u\end{aligned}$$

and $P=H(G^2)$, so by (6.5) we conclude that in $\overline{W}/8\overline{W}$

$$2\left.\frac{(\partial_T H)(G^2)GG'}{(\partial_u H)(G^2)}\right|_{u=\alpha}=2\cdot\mathrm{Res}_\alpha\frac{(\partial_T H)(G^2)G}{H(G^2)}\,\mathrm{d}G+4\cdot\mathrm{Res}_\alpha\left(\left(\frac{(\partial_T H)(G^2)GG'}{(\partial_u H)(G^2)}\right)^2\frac{\mathrm{d}P}{P}\right).$$

The first residue on the right side is $\mathrm{Res}_\alpha\omega_{H,G}$. The second residue only matters modulo 2. Reducing it modulo 2 gives the square of the residue at $\overline{\alpha}$ of

$$\frac{(\partial_T h)(g^2)gg'}{(\partial_u h)(g^2)}\cdot\frac{\mathrm{d}(h(g^2))}{h(g^2)}\frac{(\partial_T h)(g^2)g^2}{h(g^2)}\cdot\frac{\mathrm{d}g}{g}\omega_{h,g}$$

since $\mathrm{Res}_x(s^p\mathrm{d}r/r)=\mathrm{Res}_x(s\mathrm{d}r/r)^p$ in characteristic $p>0$. This establishes (6.6).

Using (6.6), expanding the product on the left side of (6.4) modulo 8 gives

$$1+2\sum_\alpha\mathrm{Res}_\alpha\omega_{H,G}+4\sum_{\alpha_1\neq\alpha_2}\mathrm{Res}_{\overline{\alpha}_1}\omega_{h,g}\mathrm{Res}_{\overline{\alpha}_2}\omega_{h,g}+4\sum_\alpha\mathrm{Res}_{\overline{\alpha}}(\omega_{h,g})^2, \tag{6.7}$$

where α_1 and α_2 in the second sum run over unordered pairs of distinct $\overline{F}$-roots of $H(G^2)$. By the residue theorem in characteristic 0, the first sum over the zeros α of $H(G^2)$ in (6.7) is equal to

$$-\mathrm{Res}_\infty\left(\frac{(\partial_T H)(G^2)G^2}{H(G^2)}\cdot\frac{\mathrm{d}G}{G}\right)=\deg G\deg_T H=\deg g\deg_T h$$

since $(\partial_T H)(G^2)G^2$ and $H(G^2)$ have the same degree and have leading coefficients with ratio $\deg_T H$.

Since (6.7) is being considered in $W/8W$, the final sum in (6.7) only matters in $W/2W$, where it can be computed to be

$$\left(\sum_{\overline{\alpha}}\mathrm{Res}_{\overline{\alpha}}(\omega_{h,g})\right)^2=\mathrm{Res}_\infty(\omega_{h,g})^2\mathrm{Res}_\infty(\omega_{h,g})\cdot\sum_{\overline{\alpha}}\mathrm{Res}_{\overline{\alpha}}(\omega_{h,g})$$

by the residue theorem in characteristic 2. The second and third sums in (6.7) therefore combine to give $4s_2(\omega_{h,g})$ in (6.4). ■

By (4.1), (6.2), and Theorem 6.4, if $\deg g\gg 0$ and $h(g^2)$ is *separable* then the discriminant $\mathrm{disc}_W(H(G^2))$ is congruent modulo $8W$ to

$$\frac{(-1)^{\delta_g(\delta_g-1)/2}}{(\mathrm{lead}\,H(G^2))^{2\delta_g-1-d_G}}R_W(H(G^2),(\partial_u H)(G^2))(1+2\deg g\deg_T h+4s_2(\omega_{h,g})), \tag{6.8}$$

where

$$\delta_g=\deg(h(g^2))=\deg(\mathrm{lead}_T h)+2\deg g\deg_T h$$

and d_G is given by Remark 6.1; the exponent $2\delta_g-1-d_G$ of $\mathrm{lead}\,H(G^2)$ in (6.8) is linear in $\deg g=\deg G$ when $\deg g$ is large. Since $-4\equiv 4 \bmod 8$, $\mathrm{disc}_W(H(G^2))$ mod $8W$ is therefore

equal to

$$\frac{R_W(H(G^2),(\partial_u H)(G^2))}{(\operatorname{lead} H(G^2))^{2\delta_g-1-d_G}}((-1)^{\delta_g(\delta_g-1)/2}(1+2\deg g\deg_T h)+4s_2(\omega_{h,g})).$$

Write $\delta_g = c+2ab$, with $c=\deg(\operatorname{lead}_T h)$, $a=\deg g$, and $b=\deg_T h$, so

$$\frac{\delta_g(\delta_g-1)}{2} \equiv \frac{c(c-1)}{2}+ab \bmod 2$$

and (by checking cases for ab modulo 4)

$$(-1)^{ab}(1+2ab) \equiv 1+4\left\lfloor \frac{1+ab}{2}\right\rfloor \bmod 8;$$

here, $\lfloor\cdot\rfloor$ denotes the greatest-integer function. Thus, separability of $h(g^2)$ implies that $\operatorname{disc}_W(H(G^2)) \bmod 8W$ is equal to

$$\frac{R_W(H(G^2),(\partial_u H)(G^2))}{(\operatorname{lead} H(G^2))^{2\delta_g-1-d_G}}(-1)^{\deg(\operatorname{lead}_T h)(\deg(\operatorname{lead}_T h)-1)/2}(1+4(m_g+s_2(\omega_{h,g}))), \tag{6.9}$$

where $m_g = \lfloor (1+(\deg g)(\deg_T h))/2\rfloor$ and $\deg g \gg 0$.

If we had instead chosen g of large degree such that $h(g^2)$ is *not* separable and $G\in W[u]$ is a lift of g with $\operatorname{lead}(G)\in W^\times$, then since $H(G^2)$ has the same degree as its reduction $h(g^2)$ we see via (4.4) that $R_W(H(G^2),(\partial_u H)(G^2))$ has reduction that is a $k^\times$-multiple (depending on G) of

$$R_k(h(g^2),(\partial_u h)(g^2)) = R_k(h(g^2),h(g^2)') = 0.$$

Thus, $R_W(H(G^2),(\partial_u H)(G^2))\in 2W$ in such cases, so although $\operatorname{disc}_W(H(G^2))$ may not be congruent modulo 8 to (6.9) when $h(g^2)$ is not separable, the expression (6.9) *always makes sense* in W and is a non-unit precisely when $\operatorname{disc}_W(H(G^2))$ is a non-unit. Thus, we can use the resultant $R_W(H(G^2),(\partial_u H)(G^2))$ from characteristic 0 to study $\operatorname{disc}_W(H(G^2)) \bmod 8W$ even though usually $(\partial_u H)(G^2)\neq H(G^2)'$ in characteristic 0.

Since $\operatorname{lead}_T H\in W[u]$ has leading coefficient in $W^\times$ and $h = H \bmod 2 \in k[u,T]$ is not in k and has no irreducible factors in $k[T]$ (as $h(T^2)$ is squarefree), we conclude that H is not in W and H has no irreducible factors in $W[T]$. Moreover, since h is squarefree in $k[u,T]$ we see that its lift H is squarefree in $W[u,T]$. The same therefore holds using F-coefficients, so $\partial_u H\neq 0$ and the zero loci $Z_H=\{H=0\}$ and $Z_{\partial_u H}=\{\partial_u H=0\}$ in $\mathbf{A}^2_F$ have finite intersection by Lemma 4.6. In particular,

$$R_H := \operatorname{Res}_{W[u]}(H,\partial_u H)\in W[u]$$

is *nonzero* and we may form the monic squarefree polynomial $M_H^{\mathrm{geom}}\in F[u]$ as in Definition 4.5, where the geometric roots of M_H^{geom} are the u-coordinates of intersection points of Z_H and $Z_{\partial_u H}$ in $\mathbf{A}^2_F$.

We may use Theorem 5.7 to obtain the identity of algebraic functions

$$R_F(H(G),(\partial_u H)(G)) = \beta_0\beta_1^n\cdot \operatorname{lead}(G)^{m_0+m_1 n}\cdot \prod_x P_{x,n}(G)^{e_x} \tag{6.10}$$

on $\mathrm{Poly}_{n/F}$ for large n, where the integers $m_0,m_1\in\mathbf{Z}$ and the scalars $\beta_0,\beta_1\in F^\times$ are independent of n, the indexing set $\{x\}$ is the set of intersection points of Z_H and $Z_{\partial_u H}$ in $\mathbf{A}^2_F$, e_x is the intersection number of Z_H and $Z_{\partial_u H}$ at x, and $P_{x,n}(G)=\mathrm{N}_{F(x)/F}(G(u_x)-t_x)$ where (u_x,t_x) are the coordinates of $x\in\mathbf{A}^2_F$. Of course, all of the parameters in (6.10) may depend on the fixed choice of H lifting h (subject to the conditions that $\deg_T H=\deg_T h$

and $\deg_u(\mathrm{lead}_T(H)) = \deg_u(\mathrm{lead}_T(h)))$. When $G \in W[u]$, the left side of (6.10) is a resultant over W. We now show that the identity (6.10) over F can be factored in a manner that is well-behaved with respect to W.

Lemma 6.5. *For large n (uniform with respect to perfect extensions of k), the algebraic maps*

$$\beta_0 \cdot \prod_{|u_x|\leq 1, |t_x|>1} P_{x,n}^{e_x}(\cdot), \quad \beta_1^n \cdot \prod_{|u_x|>1} P_{x,n}^{e_x}(\cdot) : \mathrm{Poly}_{\leq n/F} \to \mathbf{A}^1_F \tag{6.11}$$

extend uniquely to W-maps $\mathrm{Poly}_{\leq n/W} \to \mathbf{A}^1_W$ with nonzero reduction. That is, these polynomial functions in $a_0, \ldots, a_n$ have W-coefficients and have nonzero reduction.

Proof. When $|u_x| \leq 1$ and $|t_x| > 1$, we have an identity

$$P_{x,n}(G) = \mathrm{N}_{F(x)/F}(G(u_x) - t_x) = \mathrm{N}_{F(x)/F}(t_x) \cdot \mathrm{N}_{F(x)/F}(t_x^{-1}G(u_x) - 1) \tag{6.12}$$

as algebraic functions of $G \in \mathrm{Poly}_{\leq n/F}$. Likewise, if we let G^* denote the polynomial of (possibly fake) degree n obtained by reversing the order of the coefficients of G, then for $|u_x| > 1$ we have an identity

$$P_{x,n}(G) = \mathrm{N}_{F(x)/F}(G(u_x) - t_x) = \mathrm{N}_{F(x)/F}(u_x)^n \cdot \mathrm{N}_{F(x)/F}(G^*(1/u_x) - u_x^{-n}t_x) \tag{6.13}$$

with $|u_x^{-n}t_x| \ll 1$ for large n. Hence, to see that (6.11) extends over W it is enough to show that the elements

$$b_0 := \beta_0 \cdot \prod_{|u_x|\leq 1, |t_x|>1} \mathrm{N}_{F(x)/F}(t_x)^{e_x}, \quad b_1 := \beta_1 \cdot \prod_{|u_x|>1} \mathrm{N}_{F(x)/F}(u_x)^{e_x} \tag{6.14}$$

in F are integral. We shall prove these are in fact units in W. It then follows trivially that the first map in (6.11) extends over W and has constant reduction $\overline{b}_0 \in k^\times$. Likewise, the second map in (6.11) then extends over W and has reduction

$$g \mapsto \overline{b}_1 \cdot a_n(g)^{\sum_{|u_x|>1}[F(x):F]e_x}$$

for $g = \sum_{i\leq n} a_i(g)u^i$, since $G \in \mathrm{Poly}_{\leq n/F}(\overline{F}) = \overline{F}^{n+1}$ has coefficients in $\overline{W}$ and $G^*(1/u_x)$ has the same reduction as $G^*(0) = a_n(G)$ when $|u_x| > 1$.

We have seen (in the beginning of this section) that for all large n there exists $g_n \in k[u]$ of degree n such that

$$R_k(h(g_n), (\partial_u h)(g_n)) \neq 0.$$

For $G_n \in W[u]$ lifting any such g_n with $\mathrm{lead}(G_n) \in W^\times$, clearly the W-resultant of $H(G_n)$ and $(\partial_u H)(G_n)$ is a unit in W. Thus, the left side of (6.10) is a unit in W when evaluated at G_n. Now consider the right side of (6.10) when evaluated at G_n. The contribution of $\mathrm{lead}(G_n)$ is an integral unit, so we conclude

$$\beta_0\beta_1^n \prod_x P_{x,n}(G_n)^{e_x} \in W^\times.$$

By the norm-scaling calculations (6.12) and (6.13), we thereby obtain

$$(\beta_0 \cdot \prod_{|u_x|\leq 1, |t_x|>1} \mathrm{N}_{F(x)/F}(t_x)^{e_x})(\beta_1 \cdot \prod_{|u_x|>1} \mathrm{N}_{F(x)/F}(u_x)^{e_x})^n \cdot \prod_{|u_x|, |t_x|\leq 1} P_{x,n}(G_n)^{e_x} \in W^\times,$$

or equivalently

$$b_0 b_1^n \cdot \prod_{|u_x|, |t_x|\leq 1} P_{x,n}(G_n)^{e_x} \in W^\times.$$

Obviously a $\overline{W}$-point $x = (u_x, t_x)$ in the zero loci of H and $\partial_u H$ reduces to a geometric point in the zero loci of h and $\partial_u h$. Thus, for such x we conclude via Lemma 5.4 that the reduction of $P_{x,n}(G_n) \in W$ must be nonzero, since the resultant of $h(g_n)$ and $(\partial_u h)(g_n)$ is nonzero. Hence, $P_{x,n}(G_n) \in W^\times$ for such x, so $b_0 b_1^n \in W^\times$ for all large n. This forces $b_0, b_1 \in W^\times$. ■

In the study of (6.10) on G^2 for $G \in W[u]$ with unit leading coefficient, we will be able to ignore x's with $|u_x| > 1$ due to:

Theorem 6.6. *For $G \in W[u]$ with unit leading coefficient and large degree n (uniform with respect to perfect extensions of k),*

$$\beta_1^{2n} \cdot \prod_{|u_x|>1} P_{x,2n}(G^2)^{e_x} \in (W^\times)^2.$$

Proof. By Lemma 6.5, the square $\beta_1^{2n} \cdot \prod_{|u_x|>1} \mathrm{N}_{F(x)/F}(u_x)^{2ne_x} = b_1^{2n}$ is a unit, so we may divide by this without harm. This leaves us with

$$\prod_{|u_x|>1} \mathrm{N}_{F(x)/F}(G^*(1/u_x)^2 - u_x^{-2n} t_x)^{e_x}, \tag{6.15}$$

where G^* is the polynomial of (possibly fake) degree n obtained by reversing the order of the coefficients of G. Note that the square $G^*(1/u_x)^2$ is a unit when $|u_x| > 1$, as its reduction is $\mathrm{lead}(g)^2 \neq 0$. Since $u_x^{-2n} t_x \to 0$ as $n \to \infty$, for large n we see that $G^*(1/u_x)^2 - u_x^{-2n} t_x$ is very close to a unit square in the valuation ring $W(x)$ of $F(x)$. Hence, depending *just* on the amount of ramification in $F(x)$ (bounded by $[F(x) : F]$), we can make n large enough, uniformly with respect to perfect extensions of k, such that $G^*(1/u_x)^2 - u_x^{-2n} t_x$ is a square in $W(x)^\times$. Passing to n so uniformly large for all finitely many x's such that $|u_x| > 1$, the norm-product (6.15) is a unit square in W. ■

To emphasize that $b_0 \in W^\times$ in (6.14) depends on H, we now rename it: define

$$\eta_H = \beta_0 \cdot \prod_{|u_x|\leq 1, |t_x|>1} \mathrm{N}_{F(x)/F}(t_x)^{e_x} \in W^\times,$$

so η_H depends on H since the algebraic factorization on the right side of (6.10) depends on H. Using Lemma 6.5 and Theorem 6.6, together with the obvious fact that $\mathrm{lead}(G^2)$ is a unit square when $G \in W[u]$ has unit leading coefficient, the identity (6.10) yields an identity

$$R_H(G) \in \eta_H \cdot \prod_{|u_x|\leq 1, |t_x|>1} \mathrm{N}_{W(x)/W}(t_x^{-1} G(u_x)^2 - 1)^{e_x} \cdot \prod_{|u_x|, |t_x|\leq 1} P_{x,2n}(G^2)^{e_x} \cdot (W^\times)^2 \tag{6.16}$$

when $G \in W[u]$ with $\mathrm{lead}(G) \in W^\times$ and $\deg G \gg 0$, where

$$R_H(G) := R_W(H(G^2), (\partial_u H)(G^2)).$$

Since $\eta_H \in W^\times$ and all terms in the products in (6.16) are visibly integral, the resultant $R_H(G)$ is a unit in W if and only if each of the terms in the products in (6.16) is a unit, in which case the image of $R_H(G)$ in $W^\times/(W^\times)^2$ is represented by the expression in (6.16).

Define

$$\widetilde{\eta}_H = (-1)^{\deg(\mathrm{lead}_T h)(\deg(\mathrm{lead}_T h)-1)/2} \cdot \mathrm{lead}(\mathrm{lead}_T H)^{e_H} \cdot \eta_H \in W^\times$$

where $e_H = 1$ if $\mathrm{lead}_T H \notin W^\times$ and $e_H = \deg(\mathrm{lead}_T \partial_u H)$ if $\mathrm{lead}_T H \in W^\times$; $\widetilde{\eta}_H$ absorbs both the constant sign-factor and (by Remark 6.1) the odd-exponent power of the unit

$\mathrm{lead}(H(G^2))$ in (6.9) modulo $(W^\times)^2$. Choose $g \in k[u]$ with large degree and choose $G \in W[u]$ lifting g with $\deg G = \deg g$. When $h(g^2)$ is *separable* it follows from (6.9) that $\mathrm{disc}_W(H(G^2)) \in W^\times$ is a unit-square multiple of the visibly integral

(6.17)
$$\widetilde{\eta}_H \cdot (1 + 4(m_g + \widetilde{s}_2(\omega_{h,g}))) \cdot \prod_{|u_x|\leq 1, |t_x|>1} \mathrm{N}_{W(x)/W}(t_x^{-1}G(u_x)^2 - 1)^{e_x} \cdot \prod_{|u_x|, |t_x|\leq 1} P_{x,2n}(G^2)^{e_x},$$

with $m_g = \lfloor (1 + \deg g \deg_T h)/2 \rfloor$ and $\widetilde{s}_2(\omega_{h,g}) \in W$ any lift of $s_2(\omega_{h,g}) \in k$ (see (6.3)). On the other hand, if $h(g^2)$ is not separable then (6.16) implies that one of the terms $P_{x,2n}(G^2)$ with $|u_x|, |t_x| \leq 1$ is in the maximal ideal of W, so (6.17) is also in the maximal ideal of W in such cases.

Motivated by (6.17), consider the W-scheme map $L_{H,n} : \mathrm{Poly}_{\leq n/W} \to \mathbf{A}^1_W$ defined by

$$L_{H,n} : G = \sum_{i \leq n} a_i u^i \mapsto \widetilde{\eta}_H \cdot \prod_{|u_x|\leq 1, |t_x|>1} \mathrm{N}_{W(x)/W}(t_x^{-1}G(u_x)^2 - 1)^{e_x} \cdot \prod_{|u_x|, |t_x|\leq 1} P_{x,2n}(G^2)^{e_x}.$$

Each term on the right, viewed as an algebraic function of G, factors through the division-algorithm morphism

$$\widetilde{\rho}_{n,H} := \rho_{n,(M_H^{\mathrm{geom}})^{\leq 1}} : \mathrm{Poly}_{\leq n/W} \to W[u]/((M_H^{\mathrm{geom}})^{\leq 1}) \tag{6.18}$$

to the affine W-scheme of remainders modulo the F-separable monic polynomial

$$(M_H^{\mathrm{geom}})^{\leq 1} := \prod_{|u_x|\leq 1} (u - u_x) \in W[u].$$

Here, we are viewing $W[u]/((M_H^{\mathrm{geom}})^{\leq 1})$ as an affine space over $\mathrm{Spec}\, W$. Since $\widetilde{\rho}_{n,H}$ is smooth and surjective, it follows by Yoneda's lemma (or a direct construction with norms) that $L_{H,n} = L_H \circ \widetilde{\rho}_{n,H}$ for a unique W-scheme map $L_H : W[u]/((M_H^{\mathrm{geom}})^{\leq 1}) \to \mathbf{A}^1_W$ that is independent of n

Summarizing the conclusions of the above efforts, for any $g \in k[u]$ with large degree and any $G \in W[u]$ lifting g with $\deg G = \deg g$, we have

(6.19)
$$\mathrm{disc}_W(H(G^2)) \equiv (1 + 4(\lfloor (1 + \deg g \deg_T h)/2 \rfloor + s_2(\omega_{h,g}))) \cdot L_H(\widetilde{\rho}_{n,H}(G)) \cdot (W^\times)^2 \bmod 8$$

when $h(g^2)$ is separable, and otherwise the right side lies in $2W/8W$.

We will use the quadratic nature of (6.19) to investigate $\mu(h(g^2))$ in the case of finite k, but before passing to the finite case we need to study the relationship between $(M_H^{\mathrm{geom}})^{\leq 1}$ and M_h^{geom}. We may factor the separable monic M_H^{geom} in $F[u]$ into monic polynomials

$$M_H^{\mathrm{geom}} = (M_H^{\mathrm{geom}})^{\leq 1}(M_H^{\mathrm{geom}})^{>1},$$

where the roots of $(M_H^{\mathrm{geom}})^{\leq 1}$ are the roots of M_H^{geom} in $\overline{W}$ and $(M_H^{\mathrm{geom}})^{>1}$ contains the other roots. Each root of the squarefree *monic* polynomial $(M_H^{\mathrm{geom}})^{\leq 1} \in W[u]$ is an integral root of the resultant

$$\mathscr{R}_H = R_{W[u]}(H, \partial_u H) \in W[u] - \{0\},$$

so $\mathscr{R}_H$ is divisible by $(M_H^{\mathrm{geom}})^{\leq 1}$ in $W[u]$.

Definition 6.7. The reduction of $(M_H^{\mathrm{geom}})^{\leq 1}$ is denoted $\overline{M}_H^{\mathrm{geom}}$.

Up to $k^\times$-multiple, $\overline{M}_H^{\mathrm{geom}}$ is the mod-2 reduction of a primitively-scaled multiple of M_H^{geom} in $W[u]$. By reduction of divisibility over W we conclude that $\overline{M}_H^{\mathrm{geom}}$ divides $R_{k[u]}(h, \partial_u h)$; note that $\overline{M}_H^{\mathrm{geom}}$ need not be squarefree (see Example 6.14).

Remark 6.8. Obviously M_h^{geom} divides the radical of $R_{k[u]}(h, \partial_u h)$. One can have proper divisibility here if the nonzero $\mathrm{lead}_T h \in k[u]$ has a double root at some c, since the resultant $R_{k[u]}(h, \partial_u h)$ vanishes at such c for determinantal reasons but the specializations $h(c, T)$ and $(\partial_u h)(c, T)$ might not have a common geometric root; *cf.* Remark 1.6.

The general relationship between M_h^{geom} and the radical of $\overline{M}_H^{\mathrm{geom}}$ is:

Lemma 6.9. *For all lifts $H \in W[u][T]$ of $h \in k[u][T]$ such that $\deg_T H = \deg_T h$ and $\mathrm{lead}_T(H) \in W[u]$ has the same u-degree as $\mathrm{lead}_T(h) \in k[u]$, $M_h^{\mathrm{geom}} | \overline{M}_H^{\mathrm{geom}}$; in particular, the property of $h(g^2)$ being squarefree is determined by $g \bmod \overline{M}_H^{\mathrm{geom}}$. If $\mathrm{lead}_T h$ is separable (e.g., h is monic in T), then M_h^{geom} is the radical of $\overline{M}_H^{\mathrm{geom}}$.*

Proof. Recall that by Corollary 5.6(2), $g \bmod M_h^{\mathrm{geom}}$ determines whether or not $h(g^2)$ is squarefree. Since M_h^{geom} is squarefree, clearly $M_h^{\mathrm{geom}} | \overline{M}_H^{\mathrm{geom}}$ if and only if each root of M_h^{geom} is the reduction of an integral root of M_H^{geom}. We will prove this root-lifting property by using the structure theorem for quasi-finite separated morphisms.

We know h is not a unit in $k[u, T]$, and $\partial_u h$ is not a zero divisor in $k[u, T]/(h)$ since no irreducible factor of h divides $\partial_u h$ (by Lemma 4.4(2)). Thus, $k[u, T]/(h, \partial_u h)$ is a finite k-algebra. Moreover, since $W[u, T]$ is W-flat, it follows from the local flatness criterion that $\partial_u H$ is nowhere a zero divisor on $\operatorname{Spec} W[u, T]/(H)$ at points over the closed point of $\operatorname{Spec} W$ and that $\operatorname{Spec} W[u, T]/(H, \partial_u H)$ is W-flat at points over the closed point of $\operatorname{Spec} W$. On the generic fiber over $\operatorname{Spec} F$, $F[u, T]/(H, \partial_u H)$ is a finite (flat) F-algebra since $\{H = 0\}$ meets $\{\partial_u H = 0\}$ at only finitely many points in $\mathbf{A}_F^2$. To summarize, the finite-type separated morphism $\operatorname{Spec} W[u, T]/(H, \partial_u H) \to \operatorname{Spec} W$ is quasi-finite and flat.

By the structure theorem for quasi-finite separated schemes over a henselian local base [15, 18.5.11], it follows that $W[u, T]/(H, \partial_u H) = R^{\mathrm{f}} \times R'$, where R^{f} is a finite product of finite local W-algebras and R' is a quasi-finite (hence finite) F-algebra. Moreover, R^{f} must be W-flat. The image of the map

$$\operatorname{Spec} R^{\mathrm{f}} \coprod \operatorname{Spec} R' = \operatorname{Spec} W[u, T]/(H, \partial_u H) \to \operatorname{Spec} W[u] = \mathbf{A}_W^1$$

is topologically a union of a closed subscheme that is finite flat over W (the image of $\operatorname{Spec} R^{\mathrm{f}}$) and an F-finite closed subscheme of the generic fiber (the image of $\operatorname{Spec} R'$). The geometric points of this image in the closed and generic geometric fibers of $\mathbf{A}_W^1$ over $\operatorname{Spec} W$ are the roots of M_h^{geom} and M_H^{geom} respectively. Thus, the problem of identifying roots of M_h^{geom} with reductions of integral roots of M_H^{geom} is brought down to the problem of realizing each geometric closed point of a finite flat W-scheme (specifically, $\operatorname{Spec} R^{\mathrm{f}}$) as the specialization of an integral generic-fiber geometric point. For this we may reduce ourselves to the consideration of a finite flat local W-scheme S that is irreducible and reduced. We can replace S with its normalization, so $S = \operatorname{Spec} B$ where B is the integral closure of W in a finite extension of F. This situation is trivial to handle.

To prove that M_h^{geom} is the radical of $\overline{M}_H^{\mathrm{geom}}$ when $\mathrm{lead}_T h$ is separable, we check that if (c, t) is a geometric point in the common zero locus of H and $\partial_u H$, where c is integral (such c's are the roots of $(M_H^{\mathrm{geom}})^{\leq 1}$), then t is also integral. It suffices to show that $H(c, T)$ or $(\partial_u H)(c, T)$ has unit leading coefficient. That is, if $(\mathrm{lead}_T h)(c) = 0$ then we want $(\mathrm{lead}_T h)'(c) \neq 0$. Since $\mathrm{lead}_T h$ is separable, we are done. ∎

Now let $g \in k[u]$ be arbitrary with large degree. By Lemma 6.9, whether or not $h(g^2)$ is separable is determined by $g \bmod \overline{M}_H^{\mathrm{geom}}$, and even by g modulo the radical of $\overline{M}_H^{\mathrm{geom}}$. Thus,

the monic $\overline{M}_H^{\mathrm{geom}}$ constructed by reduction from characteristic 0 controls the separability of $h(g^2)$ in characteristic 2 when $\deg g$ is large.

Let us now specialize to the case of a *finite* field $k = \kappa$ of characteristic 2. We fix nonconstant $h \in \kappa[u,T]$ such that $h(T^2)$ is squarefree. Choose a lift H of h as in Lemma 6.9. Pick $g \in \kappa[u]$ of large degree, and choose a lift $G \in W[u]$ of g with the same degree (*i.e.*, with unit leading coefficient). Hence, $H(G^2)$ is a lift of $h(g^2)$ with the same degree, and $\mathrm{disc}_W(H(G^2))$ is a unit precisely when $h(g^2)$ is separable. Recall also (as we explained above Theorem 3.5) that if $\mathrm{disc}_W(H(G^2)) \in W^\times$ then it lies in $\kappa^\times \times (1+4W)$; that is, its 1-unit part lies in $1+4W$, not merely in $1+2W$, when it is a unit in W.

By Theorem 3.5 and Remark 3.6,

$$\mu(h(g^2)) = (-1)^{\deg(\mathrm{lead}_T\, h)}\widetilde{\chi}(\mathrm{disc}_W(H(G^2))), \tag{6.20}$$

where $\widetilde{\chi}$ is defined to vanish on $2W$ and is defined on $(\kappa^\times \times (1+4W))/(W^\times)^2$ by

$$\widetilde{\chi}(c\cdot(1+4w)) = (-1)^{\mathrm{Tr}_{\kappa/\mathbf{F}_2}(w \bmod 2)}. \tag{6.21}$$

We can now prove an analogue of (5.9) in characteristic 2:

Theorem 6.10. *Let κ be finite of characteristic 2, and $h \in \kappa[u,T]$ be such that $h \notin \kappa$ and $h(T^2)$ is squarefree in $\kappa[u,T]$. Fix $H \in W[u,T]$ lifting h such that $\deg_T(H) = \deg_T(h)$ and $\mathrm{lead}_T(H) \in W[u]$ has unit leading coefficient (so $\deg_u(\mathrm{lead}_T(H)) = \deg_u(\mathrm{lead}_T(h))$).*

For g of sufficiently large degree n,

$$\mu(h(g^2)) = (-1)^{\deg \mathrm{lead}_T(h) + [\kappa:\mathbf{F}_2]\lfloor (1+n\deg_T h)/2\rfloor + \mathrm{Tr}_{\kappa/\mathbf{F}_2}(s_2(\omega_{h,g}))}\cdot \widetilde{\chi}(L_H(\widetilde{\rho}_{n,H}(G))), \tag{6.22}$$

where $G \in W[u]$ is any lift of g with degree n. Here, $s_2(\omega_{h,g})$ is defined by (6.3)*, $\widetilde{\rho}_{n,H}$ is defined by* (6.18)*, and L_H is defined below (6.18). The "sufficient largeness" for $\deg g$ may be chosen uniformly with respect to finite extensions of κ.*

In particular, if $g_1, g_2 \in \kappa[u]$ have sufficiently large degrees, $\deg g_1 \equiv \deg g_2 \bmod 4$, and $g_1 \equiv g_2 \bmod \overline{M}_H^{\mathrm{geom}}$, then

$$(-1)^{\mathrm{Tr}_{\kappa/\mathbf{F}_2}(s_2(\omega_{h,g_1}))}\mu(h(g_1^2)) = (-1)^{\mathrm{Tr}_{\kappa/\mathbf{F}_2}(s_2(\omega_{h,g_2}))}\mu(h(g_2^2)). \tag{6.23}$$

The "sufficient largeness" for $\deg g_1$ and $\deg g_2$ may be chosen uniformly with respect to finite extensions of κ.

If $\deg_T h$ is even, the congruence on $\deg g_j$'s need only be taken modulo 2, and if $4 | \deg_T h$ or if $[\kappa:\mathbf{F}_2]$ is even then no congruence is necessary on the $\deg g_j$'s.

Proof. The preceding calculations ensure that $L_H(\widetilde{\rho}_{n,H}(G)) \in W$ lies in $\kappa^\times \times (1+4W)$ when it is a unit (because the same is true for both $\mathrm{disc}_W(H(G^2))$ and squares in $W^\times$). Thus, the asserted formula (6.22) for $\mu(h(g^2))$ makes sense and is immediate from (6.20), (6.21), and (6.19). Since any two elements $g_1, g_2 \in \kappa[u]$ that are congruent modulo the reduction $\overline{M}_H^{\mathrm{geom}}$ of the *monic* $(M_H^{\mathrm{geom}})^{\le 1}$ may be respectively lifted to $G_1, G_2 \in W[u]$ with unit leading coefficients such that $G_1 \equiv G_2 \bmod (M_H^{\mathrm{geom}})^{\le 1}$ (so $\widetilde{\rho}_{n_1,H}(G_1) = \widetilde{\rho}_{n_2,H}(G_2)$ with $n_j = \deg G_j = \deg g_j$), we conclude via (6.22) that the indicated congruence conditions on g_j's and $\deg g_j$'s are enough to imply (6.23). ∎

An easy argument with the Chinese remainder theorem shows that Theorem 6.10 remains true with $\overline{M}_H^{\mathrm{geom}} \in \kappa[u]$ replaced by the greatest common divisor, say $\widetilde{M}_{h,\kappa}$, of all $\overline{M}_H^{\mathrm{geom}}$'s as H runs over all lifts of h to $W[u,T]$ with the same T-degree and with $\mathrm{lead}_T(H) \in W[u]$ having unit leading coefficient in W. Though $\widetilde{M}_{h,\kappa}$ is a multiple of M_h^{geom} (by Lemma 6.9)

and is a factor of $R_{\kappa[u]}(h, \partial_u h)$, it probably can fail to be squarefree (see Example 6.14 below). We do not know if $\widetilde{M}_{h,\kappa}$ is the "minimal modulus" for $g \mapsto (-1)^{\mathrm{Tr}_{\kappa/\mathbf{F}_2}(s_2(\omega_{h,g}))}\mu(h(g^2))$ when specializing at elements $g \in \kappa[u]$ with large degree (but see Remark 7.4).

Corollary 6.11. *Let κ be finite of characteristic 2, and $h \in \kappa[u,T]$ be such that $h \notin \kappa$ and $h(T^2)$ is squarefree in $\kappa[u,T]$. Fix $H \in W[u,T]$ lifting h as in Theorem 6.10.*

For g of sufficiently large degree n,

$$\mu(h(g^4)) = (-1)^{\deg \mathrm{lead}_T h + [\kappa:\mathbf{F}_2](\deg_T h)\cdot n} \cdot \widetilde{\chi}(L_H(\widetilde{\rho}_{n,H}(G))),$$

where $G \in W[u]$ is any lift of g with degree n.

In particular, for $g_1, g_2 \in \kappa[u]$ of sufficiently large degrees,

$$g_1 \equiv g_2 \bmod \overline{M}_H^{\mathrm{geom}},\ \deg g_1 \equiv \deg g_2 \bmod 2 \Rightarrow \mu(h(g_1^4)) = \mu(h(g_2^4)). \tag{6.24}$$

The "sufficient largeness" for $\deg g_j$'s may be chosen uniformly with respect to finite extensions of κ. There is no dependence on $\deg g \bmod 2$ if $[\kappa : \mathbf{F}_2]$ or $\deg_T h$ is even.

We now give some Möbius calculations in characteristic 2, using Corollary 6.11 (and omitting further tables of data). Our second and third example will justify what we said above Remark 1.13 in the Introduction.

Example 6.12. Let $f(T) = T^4 + u$. Take $H(T) = T + u \in W[u][T]$ as a lift of $h(T) = T + u$ from $\kappa[u][T]$. Clearly $\overline{M}_H^{\mathrm{geom}} = 1$ in $\kappa[u]$, so $\mu(f(g)) = (-1)^{[\kappa:\mathbf{F}_2]\deg g}$ when $\deg g \gg 0$. The treatment of $\mu(g^2 + u)$ in [9] (replacing g with g^2 and taking simplifications into account) lets us make the condition "$\deg g \gg 0$" effective: $\deg g \geq 1$. It follows that the conjecture for $T^4 + u$ in (2.3) fails in even degrees when $[\kappa : \mathbf{F}_2]$ is odd, and in all degrees when $[\kappa : \mathbf{F}_2]$ is even. See [8, Table 1] for data on irreducible values of $T^4 + u$ over $\mathbf{F}_2[u]$.

Example 6.13. Let $f(T) = T^8 + (u^3+u)T^4 + u$ in $\kappa[u][T]$. Take $H(T) = T^2 + (u^3+u)T + u$. A calculation shows $M_H^{\mathrm{geom}} = 6u^5 + 2u^3 + 1$, so $\overline{M}_H^{\mathrm{geom}} = 1$ and $\deg_T H$ is even. Thus, $\mu(f(g)) = 1$ for $\deg g \gg 0$. A closer analysis, carried out in [9], shows that $\mu(f(g)) = 1$ for $\deg g \geq 3$ and $\mu(f(cu^2)) = -1$ for some $c \in \kappa^\times$, so the lower bound on $\deg g$ is sharp.

Example 6.14. In $\kappa[u][T]$, let $f(T) = T^{16} + (u^9 + u^4 + u^2 + u)T^8 + u^5 + u^3$. Using the proof of Theorem 6.10 to make sufficient largeness explicit, for g_1 and g_2 with degree at least 2 we have

$$g_1 \equiv g_2 \bmod u^9(u+1)^4 \Longrightarrow \mu(f(g_1)) = \mu(f(g_2)).$$

Numerical evidence suggests that we can use $u^3(u+1)$ instead of $u^9(u+1)^4$ when $\kappa = \mathbf{F}_2$, and it seems likely that the minimal modulus is not squarefree for any κ. Unfortunately, we do not have proofs for these two assertions.

7. Conjectures over $\kappa[u]$

We now correct the naive conjecture (2.3) over $\kappa[u]$. Numerical testing supports the belief that (2.3) is correct when f is separable (in any characteristic). We have seen that (2.3) is not always true for inseparable f. To define a correction factor we begin with a definition that is sensitive to the constant field κ, and in characteristic 2 we have to stay away from the inseparable cases that are not polynomials in T^4.

Definition 7.1. Let κ be a finite field. Pick $f(T)$ in $\kappa[u][T^p]$ with $p \neq 2$ (resp. in $\kappa[u][T^4]$ with $p = 2$) such that $f \notin \kappa$ and f is squarefree in $\kappa[u][T]$. Define $M_{f,\kappa}^{\mathrm{min}}$ to be the unique

monic polynomial M in $\kappa[u]$ of minimal degree that satisfies the property of M_f^{geom} in (4.15) (resp. the property of $\overline{M}_H^{\mathrm{geom}}$ in (6.24), with $f(T) = h(T^4)$).

By the Chinese remainder theorem, all nonzero $M \in \kappa[u]$ satisfying (4.15) (resp. (6.24)) are divisible by $M_{f,\kappa}^{\min}$. If κ'/κ is a finite extension then it seems to be a rather subtle problem to relate $M_{f,\kappa}^{\min}$ and $M_{f,\kappa'}^{\min}$. In odd characteristic we always have $M_{f,\kappa}^{\min} | M_f^{\mathrm{geom}}$, so $M_{f,\kappa}^{\min}$ is squarefree. In characteristic 2 we have $M_{f,\kappa}^{\min} | R_{\kappa[u]}(h, \partial_u h)$ with $R_{\kappa[u]}(h, \partial_u h) \neq 0$ (by Lemma 4.4(2)), so again the polynomials $M_{f,\kappa'}^{\min}$ have only finitely many possibilities as κ' varies over finite extensions of κ. For characteristic 2 the situation is more subtle, so we will return to the nature of $M_{f,\kappa}^{\min}$ in characteristic 2 later.

The definition of the correction factor for the naive conjecture over $\kappa[u]$ requires a lemma:

Lemma 7.2. *Let κ be a finite field of characteristic p and let $f \in \kappa[u][T^p]$ be squarefree in $\kappa[u, T]$, and assume that f has no local obstructions (so in particular, f has no irreducible factors in $\kappa[u]$). For any nonzero $M \in \kappa[u]$, there exist elements $g \in \kappa[u]$ with any sufficiently large degree (depending on M and f) such that $f(g)$ is squarefree in $\kappa[u]$ and $\gcd(f(g), M) = 1$.*

Proof. The case $f \in \kappa^\times$ is trivial, so we may assume $f \notin \kappa$. We must find g in large degree n with $f(g)$ relatively prime to $M \cdot f(g)' = M \cdot (\partial_u f)(g)$. Obviously $\partial_u f \neq 0$ since $f \notin \kappa$. By Lemma 4.4(1), f and $\partial_u f$ have no common irreducible factor in $\kappa[u][T]$. For any irreducible monic $\pi \in \kappa[u]$, define

$$c_\pi = \#\{t \in \kappa[u]/(\pi) : f(t) \equiv M \cdot (\partial_u f)(t) \equiv 0 \bmod \pi\}.$$

The absence of local obstructions ensures $1 - c_\pi/\mathrm{N}\pi > 0$ for each π.

Poonen [19] proved that the statistics for squarefree specializations of a squarefree polynomial over $\kappa[u]$ do agree with local-probability heuristics. More specifically, since $1 - c_\pi/\mathrm{N}\pi > 0$ for each π, [19, Thm. 3.1] yields

$$\lim_{n\to\infty} \frac{\#\{g \in \kappa[u] \mid \deg g \leq n, f(g) \text{ squarefree}, \ \gcd(f(g), M) = 1\}}{(q-1)q^n} = \prod_\pi \left(1 - \frac{c_\pi}{\mathrm{N}\pi}\right),$$

where the product is absolutely convergent (and in particular, nonzero). Letting $P > 0$ denote the value of the infinite product, we obtain

$$\lim_{n\to\infty} \frac{\#\{g \in \kappa[u] \mid \deg g = n, f(g) \text{ squarefree}, \gcd(f(g), M) = 1\}}{(q-1)q^n} = \left(1 - \frac{1}{q}\right) P > 0.$$

■

Definition 7.3. Let κ be a finite field. Let $f \in \kappa[u][T]$ be squarefree and assume it is is a polynomial in T^p when $p \neq 2$ (resp. in T^4 for $p = 2$), and that $f \notin \kappa$. Assume also that f has no local obstructions (so f has no irreducible factors in $\kappa[u]$). Define

$$\Lambda_\kappa(f; n) := 1 - \frac{\sum_{\deg g = n, (f(g), M_{f,\kappa}^{\min}) = 1} \mu(f(g))}{\sum_{\deg g = n, (f(g), M_{f,\kappa}^{\min}) = 1} |\mu(f(g))|}. \tag{7.1}$$

By Lemma 7.2, the denominator in (7.1) is positive for large n. Clearly $\Lambda_\kappa(f; n)$ lies in the interval $[0, 2]$ (when its denominator is nonzero) and it differs from 1 by a restricted average on the nonzero Möbius value of $f(g)$ in degree n. Loosely, the closer $\Lambda_\kappa(f; n)$ is to 1 (resp. to 0, to 2), the more equally distributed (resp. skewed towards -1, skewed towards 1) the nonzero Möbius values of $f(g)$ are for g in degree n. We only care about $\Lambda_\kappa(f; n)$ for

large n. Note that $\Lambda_\kappa(f;n) = 0$ if and only if, for all g of degree n, $f(g)$ has a nontrivial factor in common with $M_{f,\kappa}^{\min}$ or $\mu(f(g)) \in \{0,1\}$. Therefore, the vanishing of $\Lambda_\kappa(f;n)$ for a sufficiently large n (uniform in finite extensions of κ) implies that for all g of degree n in $\kappa[u]$ the polynomial $f(g)$ is reducible in $\kappa[u]$.

We should address a uniformity for the nonvanishing of the denominator in (7.1) for large n as we vary the constant field. There exists nonzero $M \in \kappa[u]$ such that $M_{f,\kappa'}^{\min} | M$ in $\kappa'[u]$ for all finite extensions κ' of κ: take $M = M_f^{\mathrm{geom}}$ in odd characteristic and $M = \overline{M}_H^{\mathrm{geom}}$ in characteristic 2 (where H is a lift of h as in Theorem 6.10, with $f = h(T^4)$). Since f has no local obstructions, by applying Lemma 7.2 to f and M we see that for large n there do exist (many) $g \in \kappa[u]$ of degree n such that $f(g) \in \kappa[u]$ is squarefree and relatively prime to M. Since the inclusion $\kappa[u] \hookrightarrow \kappa'[u]$ for any finite extension κ'/κ preserves separability and relative primality, it follows that the denominator in the definition of $\Lambda_{\kappa'}(f;n)$ is nonzero for large n uniformly with respect to κ'/κ.

Clearly (7.1) is not affected by replacing $M_{f,\kappa}^{\min}$ with its radical. For large n depending on f but uniform with respect to finite extensions of κ, Definition 7.3 is not affected by replacing $M_{f,\kappa}^{\min}$ with *any* fixed nonzero multiple of its radical (see Theorem 7.5). This makes the computation of $\Lambda_\kappa(f;n)$ easier both in theory and in practice, since in odd characteristic we can replace $M_{f,\kappa}^{\min}$ with the radical polynomial M_f^{geom} and in characteristic 2 with $f = h(T^4)$ we can likewise replace $M_{f,\kappa}^{\min}$ with the radical of $R_{\kappa[u]}(h, \partial_u h)$.

Remark 7.4. In fact, in characteristic 2 we may be able to do much better than work with $R_{\kappa[u]}(h, \partial_u h)$, in the following sense. As we have noted earlier, by Lemma 4.4(2) in characteristic 2 we have $M_{f,\kappa}^{\min} | R_{\kappa[u]}(h, \partial_u h)$ with $R_{\kappa[u]}(h, \partial_u h) \neq 0$. The polynomial $R_{\kappa[u]}(h, \partial_u h)$ generally has factors with rather high multiplicities. It would therefore be desirable to find better upper bounds on the multiplicities in $M_{f,\kappa}^{\min}$ and to find an *a priori* construction of the least common multiple of all $M_{f,\kappa'}^{\min}$'s (or at least its radical) as the extension κ'/κ varies in characteristic 2. A nice "upper bound" on the radical of $M_{f,\kappa}^{\min}$ in characteristic 2, akin to the upper bound provided by M_f^{geom} in odd characteristic, is suggested by the following question: in characteristic 2 with $f = h(T^4)$, is M_h^{geom} the radical of the least common multiple of the $\overline{M}_H^{\mathrm{geom}}$'s over all lifts H of h as in Theorem 6.10? By Lemma 6.9 we know that M_h^{geom} divides this radical, and that this divisibility is an equality in the "generic" case when $\mathrm{lead}_T h \in \kappa[u]$ is separable. If this question has an affirmative answer for h then we can replace $M_{f,\kappa}^{\min}$ with the radical polynomial M_h^{geom} in the definition of $\Lambda_\kappa(f;n)$ for large n (depending on f, but uniform with respect to finite extensions of κ).

Our work in §3–§6 leads to the following important periodicity.

Theorem 7.5. *Let κ be finite, and $f(T)$ be as in Definition* 7.3*. For any finite extension κ'/κ, the sequence $\Lambda_{\kappa'}(f;n)$ is periodic with period dividing* 4 *for $n \gg 0$, and the largeness is uniform with respect to κ'.*

With f and κ fixed, for any large $n \gg 0$ that may be chosen uniformly with respect to κ'/κ and the degree of a nonzero multiple M of $M_{f,\kappa'}^{\min}$ in $\kappa'[u]$ we may define $\Lambda_{\kappa'}(f;n)$ by using M in place of $M_{f,\kappa'}^{\min}$ in Definition 7.3*.*

Proof. See [10, Thm. 8.1], where the argument is carried out in a more general setting, with $\kappa[u]$ replaced by the coordinate ring of any smooth affine κ-curve with one geometric point at infinity. ∎

We refer the reader to Remark 1.16 for a discussion of the asymptotic properties of $\Lambda_{\kappa'}(f;n)$ as $[\kappa':\kappa] \to \infty$ with f fixed.

Example 7.6. Let κ be finite with odd characteristic p, and $f(T) = T^p + u \in \kappa[u][T]$. Using Example 3.12,

$$\Lambda_\kappa(f;n) = \begin{cases} 1, & \text{if } n \text{ is odd,} \\ 0, & \text{if } n \equiv 0 \bmod 4, \\ 1-\chi(-1), & \text{if } n \equiv 2 \bmod 4, \end{cases}$$

for $n \geq 1$, where χ is the quadratic character on $\kappa^\times$. In particular, $\Lambda_{\mathbf{F}_3}(T^3+u;n)$ is $1,2,1,0,1,2,1,0,\ldots$ and $\Lambda_{\mathbf{F}_9}(T^3+u;n)$ is $1,0,1,0,1,0,1,0,\ldots$ over $\mathbf{F}_9[u]$. This is consistent with numerical data, *e.g.*, g^3+u appears to fit (2.3) as g runs through polynomials with odd degree in both $\mathbf{F}_3[u]$ and $\mathbf{F}_9[u]$, while g^3+u seems to be irreducible about twice as often as (2.3) predicts for $g \in \mathbf{F}_3[u]$ when $\deg g \equiv 2 \bmod 4$. See [8, Table 2] for data on irreducible values of T^3+u over $\mathbf{F}_3[u]$.

Example 7.7. Let $f(T) = T^{12} + (u+1)T^6 + u^4$ be the polynomial from Example 1.1, but considered over any finite field κ of characteristic 3, not just $\mathbf{F}_3$ as in Example 1.1. Using (4.7), for $n \geq 2$ we have

$$\begin{aligned} \Lambda_\kappa(f;n) &= 1 - \frac{\sum_{a,b\in\kappa} \chi(a)^2\chi(b^2+1)}{\sum_{a,b\in\kappa} |\chi(a)^2\chi(b^2+1)|} \\ &= \begin{cases} 1+1/q, & \text{if } \chi(-1)=1, \\ 1+1/(q-2), & \text{if } \chi(-1)=-1. \end{cases} \end{aligned}$$

In particular, $\Lambda_{\mathbf{F}_3}(f;n) = 4/3$ for all $n \geq 2$. This agrees well with the data in Table 1.1.

Example 7.8. Let $f(T)$ be the polynomial from Example 1.3, but viewed in $\kappa[u][T]$ for any κ of characteristic 3. This example will illustrate the *importance* of the condition $(f(g), M_{f,\kappa}^{\min}) = 1$ in the definition of $\Lambda_\kappa(f;n)$.

In Example 1.3, where $\kappa = \mathbf{F}_3$, we observed that $f(g)$ seems to be reducible when $n = \deg g$ satisfies $n \equiv 1 \bmod 4$, and $f(g)$ has approximately twice as many irreducible values as the naive conjecture (2.3) predicts when $n \equiv 3 \bmod 4$. We now compute $\Lambda_\kappa(f;n)$ for any κ of characteristic 3, and we will find consistency with Table 1.2 for $\kappa = \mathbf{F}_3$.

We recall (4.14) from Example 4.3: when $g = cu^n + \cdots \in \kappa[u]$ with $n = \deg g \geq 1$,

$$\mu(f(g)) = (-1)^n(\chi(-1))^{n(n-1)/2}\chi(c)^{n+1}\chi(g(1)^2+g(1)+2)\chi(g(2)). \tag{7.2}$$

This formula implies $M_{f,\kappa}^{\min} = (u-1)(u-2)$. Call this M for simplicity.

To compute $\Lambda_\kappa(f;n)$, we only count g of degree n such that $(f(g),M) = 1$, a condition we want to make explicit in terms of g. Clearly $(f(g),M) = 1$ if and only if $f(g)|_{u=1} \neq 0$ and $f(g)|_{u=2} \neq 0$. Since

$$f(g)|_{u=1} = (g(1)-1)^3(g(1)^2+g(1)-1)^3, \quad f(g)|_{u=2} = (g(2))^6(g(2)+1)^3, \tag{7.3}$$

the condition $(f(g),M) = 1$ is equivalent to the combined conditions that $g(1)$ is not 1 or $1 \pm \sqrt{-1}$ (the term $1 \pm \sqrt{-1}$ appears only if $[\kappa:\mathbf{F}_3]$ is even) and $g(2)$ is not 0 or -1.

If κ has size $q = 3^m$, then by separately treating the cases when m is even or odd and when n is even or odd, elementary arguments resting on the preceding formulas (7.2) and

(7.3) show that

$$\Lambda_\kappa(f;n) = \begin{cases} 1, & \text{if } n > 0 \text{ is even,} \\ 1 + 2 \cdot (-1)^{(n+1)/2}/((q-1)(q-2)), & \text{if } n \text{ is odd,} \end{cases}$$

for odd m and

$$\Lambda_\kappa(f;n) = \begin{cases} 1, & \text{if } n > 0 \text{ is even,} \\ 1 + 2/((q-2)(q-3)), & \text{if } n \text{ is odd,} \end{cases}$$

for even m. As a special case, for $n \geq 1$ the periodic sequence of values $\Lambda_{\mathbf{F}_3}(f;n)$ is

$$0, 1, 2, 1, 0, 1, 2, 1, \cdots,$$

which is an excellent fit with the discrepancies between Table 1.2 and the naive conjecture for $f(T)$ on $\mathbf{F}_3[u]$. Here, if $n \equiv 1 \bmod 4$, then $\mu(f(g)) = -1$ only when $(f(g), M) \neq 1$. If $n \equiv 3 \bmod 4$ then $\mu(f(g)) = 1$ only when $(f(g), M) \neq 1$. In particular, since $M = (u-1)(u-2)$ and $\deg f(g) > 1$ when $\deg g \geq 1$, it follows that $f(g)$ is reducible in $\mathbf{F}_3[u]$ whenever $g \in \mathbf{F}_3[u]$ satisfies $\deg g \equiv 1 \bmod 4$.

If we did not include the condition $(f(g), M) = 1$ in the definition of $\Lambda_\kappa(f;n)$ then this sequence would be constant: $\{1, 1, \dots\}$ for $n \geq 1$. In other words, while the nonzero values of $\mu(f(g))$ for g of a fixed degree $n \geq 1$ are equally often 1 and -1, what matters for the link to irreducibility statistics appears to be the nonzero values of $\mu(f(g))$ constrained by the additional local condition $(f(g), M) = 1$.

Here, finally, is our correction to the naive conjecture.

Conjecture 7.9. *Let κ be a finite field and let $f \in \kappa[u, T^p]$ be irreducible in $\kappa[u][T]$ with no local obstructions. When $f \notin \kappa[u, T^p]$, then*

$$\#\{g \in \kappa[u] : \deg g = n, f(g) \text{ prime}\} \overset{?}{\sim} \frac{C(f)}{\deg_T f} \frac{(q-1)q^n}{\log(q^n)}$$

$n \to \infty$. If $p \neq 2$ and $f \in \kappa[u, T^p]$, or $p = 2$ and $f \in \kappa[u, T^4]$, then

$$\#\{g \in \kappa[u] : \deg g = n, f(g) \text{ prime}\} \overset{?}{\sim} \Lambda_\kappa(f;n) \frac{C(f)}{\deg_T f} \frac{(q-1)q^n}{\log(q^n)}$$

as $n \to \infty$. Here $\Lambda_\kappa(f;n)$ is defined in Definition 7.3 *and is periodic in large n by Theorem* 7.5.

Remark 7.10. In characteristic 2 our conjecture is incomplete because it does not make a prediction for $f = h(T^2)$ with $h \in \kappa[u][T]$ when h is not a polynomial in T^2. Due to (6.23), our lack of understanding of the (generally nonzero) function $g \mapsto s_2(\omega_{h,g})$ is the obstruction to formulating a conjecture that covers such cases at the present time.

When 0 occurs in the period for $\Lambda_\kappa(f;n)$, we interpret the asymptotic in Conjecture 7.9 to mean the easily proved consequence (for such large n) that there is no $g \in \kappa[u]$ in those degrees such that $f(g)$ is irreducible.

We collect sample periodic parts of $\Lambda(f;n)$ in Table 7.1. When the period is not 1, we write the period so that the first term occurs when $n \gg 0$ and $n \equiv 1 \bmod 4$.

Each of the polynomials in Table 7.1 has been tested against (2.3) over the indicated finite field. The last polynomial was tested over $\mathbf{F}_3$, $\mathbf{F}_5$, $\mathbf{F}_7$, and $\mathbf{F}_9$. The values of $\Lambda(f;n)$ in each example are in excellent numerical agreement with the data.

$f(T)$	$\Lambda(f;n)$
$T^3+u/\mathbf{F}_3[u]$	$1,2,1,0$
$T^5+u/\mathbf{F}_5[u]$	$1,0$
$T^{12}+\cdots/\mathbf{F}_3[u]$ (Examples 1.1, 4.2)	$\frac{4}{3}$
$T^9+\cdots/\mathbf{F}_3[u]$ (Examples 1.3, 4.1, 4.3, 7.8)	$0,1,2,1$
$T^{12}+(2u^4+2u^3+2u^2+u+1)T^6+2u^3+2u^2+u/\mathbf{F}_3[u]$	$\frac{2}{3}$
$(2u^2+u+3)T^{15}+(4u^2+u+3)T^5+4u^2+u+3/\mathbf{F}_5[u]$	$1,\frac{13}{10}$
$T^p+u^2/\mathbf{F}_q[u]$, $p\neq 2$ (Example 3.12)	1

TABLE 7.1. Examples of $\Lambda(f;n)$ for $n\gg 0$

Remark 7.11. For f as in Conjecture 7.9, the definition of $\Lambda_\kappa(f;n)$ involves the constraint $(f(g),M_{f,\kappa}^{\min})=1$. We do not have a conceptually satisfying explanation for this relative primality condition, so let us explain how it was found.

Initial deviations from (1.2) were discovered in situations like $f(T)=T^3+u$ over $\mathbf{F}_3[u]$, which seem to require correction factors 0 or 2. Factorizations of $f(g)$ in such cases revealed extreme parity behavior: in certain degrees, the number of irreducible factors of $f(g)$ had the same parity for all g with a fixed positive degree, and (trivially) $f(g)$ was always squarefree. This suggested a link to Möbius fluctuations, and our first guess at a correction factor was an expression, say $\widetilde{\Lambda}_\kappa(f;n)$, defined like $\Lambda_\kappa(f;n)$ but *without* the condition $\gcd(f(g),M_{f,\kappa}^{\min})=1$. Periodicity of $\widetilde{\Lambda}_\kappa(f;n)$ follows by the same arguments as for $\Lambda_\kappa(f;n)$ in the proof of Theorem 7.5.

When we found numerically, for the polynomial in Example 1.3, that $\widetilde{\Lambda}_\kappa(f;n)$ was not the correct correction factor in (1.2), the reason that it failed (as seen in Example 7.8) led to the consideration of the gcd constraint. Table 7.2 gives several examples over $\mathbf{F}_3[u]$ where $\widetilde{\Lambda}_{\mathbf{F}_3}(f;n)\neq\Lambda_{\mathbf{F}_3}(f;n)$. The first is a polynomial we have already met. The last example is particularly interesting, since $\widetilde{\Lambda}_{\mathbf{F}_3}(f;n)$ and $\Lambda_{\mathbf{F}_3}(f;n)$ lie on opposite sides of 1.

$f(T)$	$\widetilde{\Lambda}_{\mathbf{F}_3}(f;n)$	$\Lambda_{\mathbf{F}_3}(f;n)$
$T^9+\cdots$ (Example 1.3)	1 $(n\geq 2)$	$0,1,2,1,\ldots$ $(n\geq 1)$
$T^{12}+(2u^4+2u^3+2u^2+u+1)T^6+2u^3+2u^2+u$	$\frac{20}{21}$ $(n\geq 3)$	$\frac{2}{3}$ $(n\geq 3)$
$(u^2+2u+1)T^6+(u^2+2u)T^3+2u^2$	1 $(n\geq 2)$	$0,2,0,2,\ldots$ $(n\geq 1)$
$(u+2)T^{12}+u^2T^6+u^3+2$	$\frac{6}{7}$ $(n\geq 3)$	$\frac{6}{5}$ $(n\geq 4)$

TABLE 7.2. Examples where $\widetilde{\Lambda}_{\mathbf{F}_3}(f;n)\neq\Lambda_{\mathbf{F}_3}(f;n)$ for $n\gg 0$

Whenever $\widetilde{\Lambda}_\kappa(f;n)=\Lambda_\kappa(f;n)$ for $n\gg 0$ in examples, we have found a common explanation: $\mu(f(g))=0$ when $(f(g),M_{f,\kappa}^{\min})\neq 1$. This implies $\widetilde{\Lambda}_\kappa(f;n)=\Lambda_\kappa(f;n)$ for $n\gg 0$ since it tells us that for any irreducible π dividing $M_{f,\kappa}^{\min}$, any root of $f(T)$ in $\kappa[u]/(\pi)$ is a multiple root. (This includes the vacuous case f has no values on $\kappa[u]$ divisible by π.) Is this always an explanation when $\widetilde{\Lambda}_\kappa(f;n)=\Lambda_\kappa(f;n)$ for $n\gg 0$?

Remark 7.12. If we search for irreducible values of $f(g)$ not over all g in each degree, but just monic g in each degree (say), then we need a monic version of $\Lambda_\kappa(f;n)$. This is a possibly new periodic sequence (with mod 4 periodicity, *etc.*, by the same arguments).

Numerical data support the use of this new sequence as correction factors in a "monic g" version of Conjecture 7.9.

The most delicate part of numerical testing of Conjecture 7.9 is estimating the constant $C(f) = C_{\kappa[u]}(f)$. We therefore conclude this paper with a consequence of our conjecture that does not involve $C(f)$ and so is much easier to check in practice.

Suppose $p \neq 2$ and $f(T) \in \kappa[u][T^p]$ satisfies the Bouniakowsky conditions (we can also take $p = 2$ if $f(T) \in \kappa[u][T^4]$). We have $f(T) = F(T^{p^m})$ for a maximal $m \geq 1$, and this *p-free part* $F(T)$ of $f(T)$ satisfies the Bouniakowsky conditions and is separable in T. We expect that $F(T)$ satisfies the first asymptotic formula in Conjecture 7.9 and that $f(T)$ satisfies the second one. Easily $C(f) = C(F)$, so dividing the asymptotic estimates for f and F in Conjecture 7.9 leads to the prediction

$$\frac{\#\{g \in \kappa[u] : \deg g = n, f(g) \text{ prime}\}}{\#\{g \in \kappa[u] : \deg g = n, F(g) \text{ prime}\}} \stackrel{?}{\to} \frac{\Lambda_\kappa(f;n)}{p^m} \tag{7.4}$$

as $n \to \infty$, where the contribution of the constant $C(f) = C(F)$ has cancelled out. The right side of (7.4) is periodic in n mod 4 for $n \gg 0$, so this limit is understood to be taken for (large) n running through a fixed congruence class modulo 4 (and it is implicitly part of the prediction that the denominator on the left side is nonzero for large n). The two sides of (7.4) can be computed independently for increasing n. If one accepts Conjecture 1.2 for F, then (7.4) is equivalent to Conjecture 7.9 for f.

References

[1] P. T. . Bateman and R. A. . Horn, *A heuristic asymptotic formula concerning the distribution of prime numbers*, Math. Comp. **16** (1962), 363–367. MR 26 #6139

[2] ———, *Primes represented by irreducible polynomials in one variable*, Proc. Sympos. Pure Math., Vol. VIII, Amer. Math. Soc., Providence, R.I., 1965, pp. 119–132. MR 31 #1234

[3] A. O. . Bender and O. . Wittenberg, *A potential analogue of Schinzel's hypothesis for polynomials with coefficients in* $\mathbf{F}_q[t]$, `http://arxiv.org/abs/math.NT/0412303`.

[4] E. R. . Berlekamp, *An analog to the discriminant over fields of characteristic two*, J. Algebra **38** (1976), 315–317. MR 53 #8000

[5] P. Berthelot and A. Ogus, *Notes on crystalline cohomology*, Princeton University Press, Princeton, 1978. MR 58 #10908

[6] V. . Bouniakowsky, *Sur les diviseurs numériques invariables des fonctions rationnelles entières*, Mémoires sc. math. et phys. **6** (1854), 306–329.

[7] H. . Cohen, *A course in computational algebraic number theory*, Graduate Texts in Mathematics, vol. 138, Springer-Verlag, New York, 1993. MR 94i:11105

[8] K. . Conrad, *Irreducible values of polynomials: a non-analogy*, Number Fields and Function Fields – Two Parallel Worlds, Progress in Mathematics, vol. 239, Birkhäuser, Basel, 2005, to appear.

[9] B. Conrad and K. Conrad, *The Möbius function and the residue theorem*, Journal of Number Theory **110** (2005), 22–36.

[10] B. Conrad, K. Conrad, and R. Gross, *Irreducible specialization in higher genus*, in preparation.

[11] B. Conrad, K. Conrad, and H. Helfgott, *Root numbers and ranks in positive characteristic*, Adv. Math., to appear.

[12] P. D. T. A. . Elliott, *Arithmetic functions and integer products*, Grundlehren der Mathematischen Wissenschaften, vol. 272, Springer-Verlag, New York, 1985. MR 86j:11095

[13] R. J. . Evans, *The evaluation of Selberg character sums*, Enseign. Math. (2) **37** (1991), 235–248. MR 93c:11062

[14] A. Granville, *ABC allows us to count squarefrees*, Internat. Math. Res. Notices (1998), 991–1009. MR 99j:11104

[15] A. Grothendieck, *Éléments de géométrie algébrique* IV_4*. Étude locale des schémas et des morphismes de schémas*, Inst. Hautes Études Sci. Publ. Math. (1967), 361. MR 39 #220

[16] G. H. . Hardy and J. E. . Littlewood, *Some problems of Partitio Numerorum III: on the expression of a number as a sum of primes*, Acta Math. **44** (1923), 1–70.
[17] S. Lang, *Algebra*, 3rd ed., Graduate Texts in Mathematics, vol. 211, Springer-Verlag, New York, 2002. MR 2003e:00003
[18] A. E. . Pellet, *Sur la décomposition d'une fonction entière en facteurs irréductibles suivant un module premier p*, C. R. Acad. Sci. Paris **86** (1878), 1071–1072.
[19] B. Poonen, *Squarefree values of multivariable polynomials*, Duke Math. J. **118** (2003), 353–373. 1 980 998
[20] J -P. . Serre, *Local fields*, Graduate Texts in Mathematics, vol. 67, Springer-Verlag, New York, 1979. MR82e:12016
[21] H. N. . Shapiro, *Some assertions equivalent to the prime number theorem for arithmetic progressions*, Comm. Pure Appl. Math. **2** (1949), 293–308. MR 11,419d
[22] R. G. . Swan, *Factorization of polynomials over finite fields*, Pacific J. Math. **12** (1962), 1099–1106. MR 26 #2432
[23] A. R. . Wadsworth, *Discriminants in characteristic two*, Linear and Multilinear Algebra **17** (1985), 235–263. MR 86m:12004

Department of Mathematics, University of Michigan, Ann Arbor, MI 48109-1043, U.S.A.
E-mail address: bdconrad@umich.edu

Department of Mathematics, University of Connecticut, Storrs, CT 06269-3009, U.S.A.
E-mail address: kconrad@math.uconn.edu

Department of Mathematics, Boston College, Chestnut Hill, MA 02467-3806, U.S.A.
E-mail address: gross@bc.edu