TRANSACTIONS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 360, Number 6, June 2008, Pages 2867–2908
S 0002-9947(08)04283-9
Article electronically published on January 30, 2008

PRIME SPECIALIZATION IN GENUS 0

BRIAN CONRAD, KEITH CONRAD, AND ROBERT GROSS

ABSTRACT. For a prime polynomial $f(T) \in \mathbf{Z}[T]$, a classical conjecture predicts how often f has prime values. For a finite field κ and a prime polynomial $f(T) \in \kappa[u][T]$, the natural analogue of this conjecture (a prediction for how often f takes prime values on $\kappa[u]$) is not generally true when $f(T)$ is a polynomial in T^p (p the characteristic of κ). The explanation rests on a new global obstruction which can be measured by an appropriate average of the nonzero Möbius values $\mu(f(g))$ as g varies. We prove the surprising fact that this "Möbius average," which can be defined without reference to any conjectures, has a periodic behavior governed by the geometry of the plane curve $f = 0$.

The periodic Möbius average behavior implies in specific examples that a polynomial in $\kappa[u][T]$ does not take prime values as often as analogies with $\mathbf{Z}[T]$ suggest, and it leads to a modified conjecture for how often prime values occur.

1. INTRODUCTION

A conjecture of Bouniakowsky [5] says that a nonconstant prime polynomial $f(T)$ in $\mathbf{Z}[T]$ has infinitely many prime values in $\mathbf{Z}$ unless there is a local obstruction: all values of $f(T)$ on $\mathbf{Z}$ are divisible by a nontrivial common factor. For example, $3T^2 - T + 2$ is prime in $\mathbf{Z}[T]$ but has an obstruction at 2 to taking prime values: $3n^2 - n + 2$ is even for every $n \in \mathbf{Z}$. (We allow negative primes, so we don't need to assume $f(T)$ has a positive leading coefficient.) When no local obstruction occurs, there is an asymptotic conjecture (as $x \to \infty$) for how many $1 \leq m \leq x$ (or $|m| \leq x$) give prime values $f(m)$; this is due to Hardy and Littlewood [12] in special cases and Bateman and Horn [1] more generally. The only proved case 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.

Let κ be a finite field and pick a prime polynomial $f(T)$ in $\kappa[u][T]$ which is nonconstant in T such that the values of $f(T)$ on $\kappa[u]$ do not all share a nontrivial common factor (we say $f(T)$ has no local obstruction). Qualitatively, it is natural to expect under these conditions that $f(g)$ is prime for infinitely many g in $\kappa[u]$. Quantitatively, a conjectural asymptotic estimate for how often $f(g)$ is prime as g varies is easy to write down using analogies between $\mathbf{Z}$ (which we will call the classical case) and $\kappa[u]$. The only proved instance of this asymptotic conjecture is the case $\deg_T f = 1$, just as in the classical situation. What if $\deg_T f > 1$?

Received by the editors June 19, 2005 and, in revised form, February 11, 2006.
2000 *Mathematics Subject Classification.* Primary 11N32.
Key words and phrases. Bateman–Horn conjecture, Hardy–Littlewood conjecture, Möbius function.

When κ has characteristic 2 it is shown in [15] that the values of T^8+u^3 on $\kappa[u]$ are never prime when $[\kappa : \mathbf{F}_2]$ is odd and are prime only finitely many times (at noncubes in κ) when $[\kappa : \mathbf{F}_2]$ is even. Here are some interesting examples in odd characteristic.

Example 1.1. Let $f(T) = T^{12} + (u+1)T^6 + u^4$ over $\mathbf{F}_3[u]$. Numerical evidence in [6, Table 3] suggests that the count of prime values of $f(g)$ as g runs over $\mathbf{F}_3[u]$ exceeds the amount predicted from analogies with the classical case by a ratio ≈ 1.33.

Example 1.2. Let $f(T) = T^9 + (2u^2+u)T^6 + (2u+2)T^3 + u^2 + 2u + 1$ over $\mathbf{F}_3[u]$. Numerical data in [6, Table 5] suggest that $f(g)$ is prime as often as expected (by analogies with the classical case) when $\deg g$ is even, it is never prime when $\deg g \equiv 1 \bmod 4$, and it is prime about twice as often as expected when $\deg g \equiv 3 \bmod 4$. In particular, the statistical behavior seems to depend on the mod 4 value of the degree in which we are sampling f. The absence of a prime $f(g)$ for $\deg g \equiv 1 \bmod 4$ will be proved in Example 6.10.

We observed three common features of prime polynomials in $\kappa[u][T]$ whose primality statistics do not seem to match expectations based on analogies to the classical case:

- The polynomial $f(T)$ lies in $\kappa[u][T^p]$, where p is the characteristic of κ; equivalently (since $f(T)$ is irreducible over $\kappa(u)$) $f(T)$ is inseparable over $\kappa(u)$.
- The ratio between the actual number of prime values $f(g)$, as g runs over polynomials with a common degree n, and the conjectural asymptotic estimate for that number based on analogies to the classical case appears to have a limit as $n \to \infty$ if we fix $n \bmod 4$. However, the apparent limit (which is not necessarily 1) may be different for different classes mod 4, and there are 1, 2, or 4 apparent limits (never 3).
- The Möbius function for $\kappa[u]$ (see Definition 2.1) has unusual statistics on the values $f(g)$. Roughly speaking, the nonzero values of $\mu(f(g))$ may fail to be 1 and -1 equally often.

The main impact of the third observation is that statistics for prime values of $f(g)$ as g varies can be linked to *appropriate* averages of the nonzero values of $\mu(f(g))$ as g varies (and the averages we define are effectively computable in any example). Subtracting these averages from 1 enables us to predict the 1, 2, or 4 apparent limits in the second observation. A simple example is the polynomial $f(T)$ in Example 1.1. As g runs over the polynomials in $\mathbf{F}_3[u][T]$ having a common degree ≥ 2 the value $\mu(f(g))$ is -1 twice as often as it is 1 (by Example 3.2), so the average nonzero value of $\mu(f(g))$ in each degree (at least 2) is exactly $-1/3$. Subtracting this from 1 gives $4/3 = 1.33\ldots$, which matches the deviation we found in Example 1.1.

The theme of this paper is the study of $\mu(f(g))$ when $f(T)$ is a polynomial in $\kappa[u][T^p]$ which is squarefree in $\kappa[u][T]$. The intended application of this work, which we will give in the final section, is the formulation of a conjecture for the frequency of prime values of $f(T)$ when $f(T)$ is a prime polynomial. However, our work on $\mu(f(g))$ does not require that $f(T)$ be prime. Letting $f(T)$ be squarefree instead provides greater technical flexibility. For instance, squarefreeness is preserved under finite extension of the constant field κ.

Here is an outline of the paper. In §2 and §3 we discuss some general features of the Möbius function, discriminants, and resultants relative to $\kappa[u]$. This builds on ideas of Swan [15] and is applied in §4 where we prove that $\mu(f(g))$, under suitable hypotheses, is periodic in g when κ has odd characteristic (see Theorem 4.8). In §5 we treat characteristic 2, which is more difficult. For instance, we need to use 2-adic lifts and residues of differential forms (Theorem 5.10). Finally, §6 uses these ideas to define a notion of Möbius average that seems to work well as a new correction factor for asymptotically estimating how often a prime polynomial in $\kappa[u][T]$ takes prime values on $\kappa[u]$ (Conjecture 6.2).

While the classical conjectures on prime values of prime polynomials in $\mathbf{Z}[T]$ involve only local obstructions, the Möbius average in the $\kappa[u]$-setting is a fundamentally *global* obstruction. The Möbius function could be considered as a "parity" obstruction to prime values once we know $f(g)$ is squarefree (crudely put, if $\mu(f(g)) = 1$ then $f(g)$ is definitely not prime). We have not found a role for a similar type of obstruction to prime values based on the number of irreducible factors of the squarefree values of $f(g)$ counted modulo 3 or higher, and our correction based on Möbius averages gives an excellent fit with all observed numerics.

Notation and terminology. Throughout the paper, κ denotes a finite field of characteristic p and size q. We often write χ instead of χ_κ to denote the quadratic character on the multiplicative group $\kappa^\times$ when $p \neq 2$.

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 $\operatorname{lead}_T f \in R[u]$. We write $\deg_{u,T} f$ to denote the total degree of such an f. An element in $R[u]$ is *primitive* when its coefficients generate the unit ideal in R. When R is a domain, the discriminant of a nonzero one-variable polynomial h with coefficients in R is denoted $\operatorname{disc} h$, or $\operatorname{disc}_R h$ for emphasis. Our definition of polynomial discriminants does not match the usual definition when the polynomial is not monic; see (2.3) and (2.4).

Let R be 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. The length of an R-module M is denoted $\ell(M)$, and if Z is an Artinian scheme then $\ell(Z)$ denotes the length of the ring $\Gamma(Z, \mathcal{O}_Z)$.

2. The Möbius function over finite fields

Definition 2.1. Let R be a Dedekind domain. The *Möbius function* on nonzero ideals of R is $\mu_R(\mathfrak{p}_1 \cdots \mathfrak{p}_m) = (-1)^m$ 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 the context, we write μ rather than μ_R.

The first step in the analysis of $\mu_{\kappa[u]}(f(g))$ when $f \in \kappa[u][T^p]$ is fixed and $g \in \kappa[u]$ varies is the derivation of a formula for $\mu_{\kappa[u]}(h)$ (for any nonzero $h \in \kappa[u]$) in terms of the discriminant of h. (This has no known analogue for the Möbius function on $\mathbf{Z}$.) The Möbius function of h (for $h \neq 0$) can be viewed as the Möbius function of the finite κ-algebra $\kappa[u]/(h)$ via the following definition.

Definition 2.2. 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.

Theorem 2.3. *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,*

$$\mu(A) = (-1)^{\dim_\kappa A}\chi_\kappa(\mathrm{disc}_\kappa A). \tag{2.1}$$

Proof. We immediately reduce to the case when A is a finite extension field of κ, and this is settled as in the proof of [6, Lemma 4.1] by computing the sign of a Galois-theoretic Frobenius as a permutation. □

We need an analogue of Theorem 2.3 in characteristic 2 that still 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, extending work of Swan [15]. (Swan essentially obtained Theorem 2.4 below and applied it to compute $\mu_{\kappa[u]}(h)$ in a few examples, although he expressed his answer directly in terms of the parity of the number of irreducible factors instead of in terms of the Möbius function.)

Let κ be a finite field of characteristic p and $W = W(\kappa)$ be the ring of Witt vectors of κ. Consider 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 [4, Example 3.2(4)]).

Suppose κ has characteristic 2 and A is étale over κ, so $\mathrm{disc}_W \widetilde{A} \in W^\times/(W^\times)^2$. Embed $\kappa^\times$ into $W^\times$ by the Teichmüller lifting, so $W^\times = \kappa^\times \times (1+2W)$. 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 $(1 + 2w)^2 \in 1 + 4W$ for all $w \in W$.)

Theorem 2.4. *For any finite κ-algebra A that admits a finite flat lifting $\widetilde{A}$ of A over W,*

$$\mu(A) = (-1)^{\dim_\kappa A}\widetilde{\chi}(\mathrm{disc}_W \widetilde{A}), \tag{2.2}$$

where $\widetilde{\chi}$ is the unique quadratic character on $W^\times$ when κ has odd characteristic and is the unique quadratic character on $\kappa^\times \times (1 + 4W)$ killing $(W^\times)^2$ when κ has characteristic 2. (An explicit formula for $\widetilde{\chi}$ in the characteristic 2 case is given in (5.24), where the formula is first needed.) In both cases, $\widetilde{\chi}$ is extended by 0 to pW.

The proof, which we omit, goes exactly as for Theorem 2.3, except that unramified extensions of the fraction field of W are used instead of finite extensions of κ.

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

$$\mathrm{disc}\, h := \prod_{i<j}(\gamma_i - \gamma_j)^2 \in F, \tag{2.3}$$

whether or not h is monic. (The discriminant of a nonzero constant polynomial is understood to be 1.) The usual definition of $\mathrm{disc}\, h$ has an additional square factor $(\mathrm{lead}\, h)^{2d-2}$, which makes the discriminant a homogeneous function of the coefficients of h. We prefer to use our definition since it agrees with the universally accepted definition of the discriminant of the finite F-algebra $F[u]/(h)$ relative to

the ordered basis $\{1, u, \dots, u^{d-1}\}$, whether or not h is monic. In terms of the derivative of h, (2.3) is the same as

$$\operatorname{disc} h = \frac{(-1)^{d(d-1)/2}}{(\operatorname{lead} h)^d} \prod_{i=1}^{d} h'(\gamma_i) = \frac{(-1)^{d(d-1)/2}}{(\operatorname{lead} h)^d} \operatorname{N}_{(F[u]/(h))/F}(h'). \tag{2.4}$$

Theorems 2.3 and 2.4 give

$$\mu_{\kappa[u]}(h) = \begin{cases} (-1)^{\deg h}\chi(\operatorname{disc}_\kappa h), & \text{if } \kappa \text{ has odd characteristic,} \\ (-1)^{\deg h}\widetilde{\chi}(\operatorname{disc}_W H), & \text{if } \kappa \text{ has any characteristic,} \end{cases} \tag{2.5}$$

where H is a lifting of h into $W[u]$ with $\deg H = \deg h$. (Since our polynomial discriminants differ from the usual definition by a square factor, the usual definition can also be used in (2.5).) The formula in (2.5) for the case of characteristic 2 uses a discriminant in characteristic 0. There is an intrinsic variant of the discriminant in characteristic 2 (see [3]), but we have not found this to be useful for our purposes.

Example 2.5. Let κ be a finite field with characteristic $p \neq 2$. For nonconstant $g = cu^n + \cdots \in \kappa[u]$ we see via (2.4) and (2.5) that

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

When n is odd, this equals 1 and -1 equally often as g varies in degree n. When n is even, $\mu(g^p + u)$ equals $\chi(-1)^{n/2}$ for all g of degree n. For instance, when $n \equiv 0 \bmod 4$, $\mu(g^p + u) = 1$ for all g of degree n. Thus $g^p + u$ is not prime when $4 | \deg g$ and $\deg g > 0$.

3. Discriminants and resultants

We wish to understand the behavior of $\mu(f(g))$ when $f \in \kappa[u][T^p]$ is fixed with $\deg_T f > 0$ and g varies in $\kappa[u]$ with large degree. Formula (2.5) suggests, at least for $p \neq 2$, that we should study $\operatorname{disc}(f(g))$ as an algebraic function of g with a specified degree n, where n is large. (We need n at least large enough that $\deg(f(g))$ depends on g only through its degree n.) Following Swan [15], we will find it useful to work with resultants rather than discriminants.

For an integral domain C, the *resultant* of two nonzero polynomials h_1 and h_2 in $C[u]$, denoted $R_C(h_1, h_2) = R(h_1, h_2)$, is defined to be

$$R(h_1, h_2) = (\operatorname{lead} h_1)^{\deg h_2} \prod_{h_1(\alpha)=0} h_2(\alpha) \tag{3.1}$$

with the product running over the roots of h_1 (counted with multiplicity) in a splitting field over the fraction field of C. There is a classical expression for $R(h_1, h_2)$ given as the determinant of a universal matrix in the coefficients of h_1 and h_2, and the size of the matrix depends on the degrees of h_1 and h_2. 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. When some h_j vanishes we define $R(h_1, h_2) = 0$, which is compatible with universal determinants that define resultants (letting the zero polynomial be assigned whatever positive degree we please).

If nonzero h_1 and h_2 have actual degrees d_1 and d_2, then for any $d_3 \geq d_2$,

$$R_{d_1,d_3}(h_1,h_2) = (\operatorname{lead} h_1)^{d_3-d_2} R_{d_1,d_2}(h_1,h_2). \tag{3.2}$$

Though (3.1) 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.

Warning. Failure to remember that resultants are sensitive to degrees can lead to errors when universal formulas over $\mathbf{Z}$ (such as (3.3) below) are used in characteristic $p > 0$.

The relation between discriminants and resultants is given by the formula

$$\operatorname{disc} h = \frac{(-1)^{d(d-1)/2} R_{d,d-1}(h,h')}{(\operatorname{lead} h)^{2d-1}}, \tag{3.3}$$

where $d = \deg h \geq 1$. If $h' \neq 0$ then this formula simplifies by (3.2) to

$$\operatorname{disc} h = \frac{(-1)^{d(d-1)/2} R(h,h')}{(\operatorname{lead} h)^{d+\deg h'}}, \tag{3.4}$$

where $R(h,h')$ 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 $d-1$ in positive characteristic).

Example 3.1. 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. (Example 1.2 is $\kappa = \mathbf{F}_3$.) For nonconstant $g = cu^n + \dots$ in $\kappa[u]$ with $c \neq 0$, $f(g)$ has degree $9n$ with leading coefficient c^9 and $f(g)' = (\partial_u f)(g)$ has degree $6n+1 < 9n-1$. By (3.4),

$$\operatorname{disc} f(g) = \frac{(-1)^{n(n-1)/2} R(f(g),(\partial_u f)(g))}{(c^9)^{15n+1}}. \tag{3.5}$$

Here $R(f(g),(\partial_u f)(g)) = R_{9n,6n+1}(f(g),(\partial_u f)(g))$. If instead we use $(c^9)^{2d-1} = (c^9)^{18n-1}$ in the denominator (see (3.3)), then an erroneous factor of $(c^9)^{3n-2}$ is introduced in (3.5), and this extra power of c affects the quadratic character of the right side of (3.5). In view of (2.5), such an error would lead to incorrect calculations of $\mu(f(g))$.

Resultants have several useful algebraic properties. We state five of them without proof, as in [15]. In this list, polynomials are nonzero and have coefficients in a domain C.

(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:

$$R(h_1h_3,h_2) = R(h_1,h_2)R(h_3,h_2), R(h_1,h_2h_3) = R(h_1,h_2)R(h_1,h_3).$$

(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 $c \in C$.
(4) $R(c,h) = R(h,c) = c^{\deg h}$ for $c \in C$, $h \neq 0$. Thus, $R(c_1,c_2) = 1$ for $c_1, c_2 \neq 0$ in C.
(5) For nonzero M, h_1, h_2 in $C[u]$,

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

We call property (5) the *quasi-periodicity* of the resultant (in its second argument). For monic M in $C[u]$ and any $f(T) \in C[u][T]$, $R(M, f(h))$ is genuinely periodic in h. The notation $R(h_1, h_2)$ in [15] differs by a sign factor from the definition of resultants generally used today; it denotes our $R(h_2, h_1)$, so comparisons with [15] must keep this distinction in mind.

Example 3.2. Let $f(T) = T^{12} + (u+1)T^6 + u^4 \in \kappa[u][T]$ with $\mathrm{char}(\kappa) = 3$, as in Example 1.1. Let $q = \#\kappa$. We shall compute $\mu(f(g))$ when $n = \deg g \geq 1$.

Let $h = f(g)$, so $h' = g^6 + u^3 = (g^2+u)^3$. Since $\deg h = 12n$, $\deg h' = 6n$, and $\mathrm{lead}\, h$ is a square, (2.5) and (3.4) give $\mu(f(g)) = \chi(R(f(g), (g^2+u)^3)) = \chi(R(g^2+u, f(g)))$. Since $f(g) \equiv u^6 - u^3 \bmod g^2 + u$, quasi-periodicity of the resultant gives

$$R(g^2+u, f(g)) = (\mathrm{lead}\, g)^{2(12n-6)} R(g^2+u, u^6-u^3).$$

But $R(g^2+u, u^6-u^3) = R(g^2+u, u)^3 R(g^2+u, u-1)^3 = g(0)^6(g(1)^2+1)^3$, so

$$\mu(f(g)) = \chi(g(0))^2 \chi(g(1)^2+1). \tag{3.6}$$

(Further calculations show that $\mathrm{disc}\, f(g) = (\mathrm{lead}\, g)^{-36(4n+1)} g(0)^{18}(g(1)^2+1)^9$.) 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 3.3. Let $f(T) = T^9 + (2u^2+u)T^6 + (2u+2)T^3 + u^2 + 2u + 1 \in \kappa[u][T]$, with $\mathrm{char}(\kappa) = 3$, as in Example 1.2. For nonconstant $g(u) = cu^n + \cdots$ with degree $n \geq 1$, $\deg(f(g)) = 9n$ and $\deg(f(g)') = 6n+1$, so $\mu(f(g)) = (-1)^n \chi(\mathrm{disc}\, f(g))$ by (2.5). By (3.5),

$$\mu(f(g)) = (-1)^n (\chi(-1))^{n(n-1)/2} \chi(c)^{n+1} \chi(R(f(g), (\partial_u f)(g))). \tag{3.7}$$

Since the algebraic properties of resultants are analogous to the algebraic properties of intersection numbers of plane curves, a recursive algebraic procedure that imitates the computation of such intersection numbers (as is given in detail in the proof of Theorem 4.1) yields

$$R(f(g), (\partial_u f)(g)) = c^{54n-6}(g(1)^2 + g(1) + 2)^3 g(2)^9. \tag{3.8}$$

Inserting (3.8) into (3.7), we find our Möbius formula:

$$\mu(f(g)) = (-1)^n \chi(-1)^{n(n-1)/2} \chi(c)^{n+1} \chi(g(1)^2 + g(1) + 2)) \chi(g(2)) \tag{3.9}$$

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. This formula 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 $\deg g \bmod 2$.

Definition 3.4. 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, define $M^{\mathrm{geom}}_{f_1,f_2}$ to be the monic separable polynomial in $F[u]$ whose zero locus is the image of $Z_{f_1} \cap Z_{f_2}$ in 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 nonzero, 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).

For applications to prime values of polynomials in $\kappa[u][T]$ we will use the last part of Definition 3.4, so let us describe M_f^{geom} more geometrically. The projection from the zero locus 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. The image of B in the u-axis is a finite set of geometric points. Then M_f^{geom} is the monic polynomial in $\kappa[u]$ that vanishes precisely at this finite set on the u-axis, with each root having multiplicity 1 (so M_f^{geom} is squarefree). If the leading coefficient $c_d(u)$ of $f(T) \in \kappa[u][T]$ does not have double roots, then M_f^{geom} is the radical of the $\kappa[u]$-resultant of f and $\partial_u f$.

For $f \in F[u][T]$ with $f \notin F$, the following lemma gives sufficient conditions for M_f^{geom} to be defined when F has positive characteristic.

Lemma 3.5. *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) *If $f \in F[u][T]$ is nonzero and $f(T^p)$ is squarefree in $F[u][T]$ (so f is squarefree in $F[u][T]$), then f and $\partial_u f$ have no nonconstant common factor 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 3.5 will be used only when $p = 2$.

Proof. It suffices to check the second part, since if f is in $F[u][T^p]$ and is squarefree in $F[u][T]$ then $f(T^p)$ is squarefree in $F[u][T]$.

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 ($Z_{\partial_u f}$ 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 (2) implies that $f(T^p)$ is a squarefree element in $F(T^p)[u]$ with nonzero u-derivative, so the image of $Z_f \cap Z_{\partial_u f}$ in 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 $f(T^p)$ is divisible by $T^p - c = (T - c^{1/p})^p$, contrary to the squarefreeness hypothesis. □

4. A resultant formula

Our goal in this section is to show that for $p \neq 2$, polynomials $f(T)$ in $\kappa[u][T^p]$ which are squarefree in $\kappa[u][T]$ have "periodic" Möbius values: for $g \in \kappa[u]$ with sufficiently large degree, $\mu(f(g))$ is determined by (i) the reduction of g modulo some nonzero $M \in \kappa[u]$, (ii) the mod-4 congruence class of the degree of g, and (iii) the quadratic character of the leading coefficient of g (*cf.* Example 3.3). This will be proved as Theorem 4.8 via 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 $0 \leq i < \deg M$. 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

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

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

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

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 non-constant $M \in F[u]$ and any $n \geq \deg M$, the formation of remainders under long division by M defines an algebraic morphism

$$\rho_{n,M} : \mathrm{Poly}_{n/F} \to F[u]/(M) \simeq \mathrm{Poly}_{\leq (\deg M - 1)/F} \tag{4.1}$$

that is a trivial $\mathrm{Poly}_{d/F}$-bundle with $d = n - \deg M$, by the division algorithm. We interpret $\mathrm{Poly}_{\leq -1/F}$ to be a 1-point space, so when $M \in F^\times$ the map

$$\rho_{n,M} : \mathrm{Poly}_{n/F} \to \mathrm{Spec}\, F \tag{4.2}$$

is the structure map to a point.

Let $f \in F[u][T]$ be a nonzero element with T-degree d. It will be useful later to record the simple formulas for the degree and leading coefficient of $f(g) \in F[u]$ when $d > 0$ and $g \in F[u]$ has $\deg g \gg 0$, and to make the condition $\deg g \gg 0$ effective. Write

$$f(T) = \alpha_d(u) T^d + \alpha_{d-1}(u) T^{d-1} + \cdots + \alpha_0(u), \tag{4.3}$$

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

$$\deg(f(g)) = d \cdot \deg g + \deg \alpha_d = (\deg_T f) n + \deg(\mathrm{lead}_T f), \tag{4.4}$$

$$\mathrm{lead}(f(g)) = (\mathrm{lead}\, \alpha_d)(\mathrm{lead}\, g)^d, \tag{4.5}$$

where $n = \deg g$. In particular, $\deg(f(g))$ is a linear polynomial in $\deg g$ when $\deg g \gg 0$. An explicit lower bound on $\deg g$, in terms of f, such that (4.4) and (4.5) apply to $f(g)$ is

$$\deg g > \nu(f) := \max_{0 \leq i \leq d-1} \frac{\deg \alpha_i - \deg \alpha_d}{d - i}, \tag{4.6}$$

where terms with $\alpha_i = 0$ are omitted or use the convention that $\deg 0 = -\infty$. For completeness, set $\nu(f) = 0$ in the vacuous case that $f(T) = \alpha(u) T^d$ is a T-monomial (no i's to consider), with the case $d = 0$ permitted.

Since $\deg(f(g))$ is determined by $n = \deg g$ for g of large degree (depending on f, as in (4.4) and (4.6)), there is a well-posed algebraic discriminant function

$$\mathrm{disc} \circ f : \mathrm{Poly}_{n/F} \to \mathbf{A}_F^1 \tag{4.7}$$

defined by $g \mapsto \mathrm{disc}(f(g))$ when n is sufficiently large; note that (4.7) does *not* extend to an algebraic function on $\mathrm{Poly}_{\leq n/F}$. Our aim is to understand the structure of the algebraic function (4.7) for f as in Lemma 3.5, and in particular the extent to which this function factors through the remainder morphism $\rho_{n,M}$ for some nonzero $M \in F[u]$.

To exploit inductive arguments, it is convenient to reinterpret the study of $\mathrm{disc}(f(g))$ as the study of the resultant $R(f(g), (\partial_u f)(g))$. 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 3.5 when $f \notin F$). The merit of this generality is that we may separately vary f_1 and f_2. Restricting attention to F of positive characteristic is not adequate: our later work in characteristic 2 will use the present considerations with a 2-adic field F.

Theorem 4.1. *Let F be a perfect field with any characteristic and let $f_1, f_2 \in F[u][T]$ be nonzero (possibly constant) elements with zero loci Z_{f_1} and Z_{f_2} in $\mathbf{A}^2_F$ that have finite intersection. For each $x = (u_x, t_x) \in Z_{f_1} \cap Z_{f_2}$, let $i_x(Z_{f_1}, Z_{f_2}) = \ell(\mathcal{O}_{Z_{f_1} \cap Z_{f_2}, x})$ be the local intersection number.*

There exist $c_0, c_1 \in F^\times$ and $m_0, m_1 \in \mathbf{Z}$ 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

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

where $\mathrm{lead}\, g \in F^\times$ is the leading coefficient of g (and the product is 1 if $Z_{f_1} \cap Z_{f_2} = \emptyset$); in fact, $m_1 = (\deg_T f_1) \cdot (\deg_T f_2)$. The "sufficiently large" condition on n only depends on the total degrees $\deg_{u,T} f_1$ and $\deg_{u,T} f_2$.

Theorem 4.1 will be proved near the end of this section after a lot of preparatory work. The identity in Theorem 4.1 is really a universal algebraic identity for g with coefficients in any F-algebra domain. The formulation of this identity is not well-suited to our method of proof, so we will first prove a less precise version after setting up some notation. Fixing f_1 and f_2 as in Theorem 4.1, for $g \in F[u]$ of degree n the degree of $f_j(g) \in F[u]$ is provided by (4.4) when $n > \nu(f_j)$ (see (4.6)): it is equal to

$$d_{j,n} := (\deg_T f_j)n + \deg(\mathrm{lead}_T f_j). \tag{4.8}$$

For $n > \max(\nu(f_1), \nu(f_2))$ 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 $\mathrm{Poly}_{\le n/F} = \mathrm{Spec}\, F[a_0, \ldots, a_n]$ of polynomials of degree $\le n$ over F-algebras; we are not requiring a_n to be a unit. Consider the following universal polynomial depending on f_1 and f_2:

$$R_n(G) := R_{F[a_0,\ldots,a_n]}(f_1(G), f_2(G)) \in F[a_0, \ldots, a_n], \tag{4.9}$$

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

Lemma 4.2. *For any $n > \max(\nu(f_1), \nu(f_2))$, $R_n(G)$ in $F[a_0, \ldots, a_n]$ is nonzero.*

Proof. Specializing G to $u^n + a_0$ commutes with formation of the resultant and carries $f_j(G)$ to the polynomial $f_j(u, u^n + a_0)$ that is obtained from $f_j(u, T)$ via the automorphism of the plane $(u, a_0) \mapsto (u, u^n + a_0)$. Hence, the zero loci of $f_1(u, u^n + a_0)$ and $f_2(u, u^n + a_0)$ in the (u, a_0)-plane have finite intersection, so the specialized $F[a_0]$-resultant is nonzero. □

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

$$\mathrm{N}_{F(x)[a_0,\ldots,a_n]/F[a_0,\ldots,a_n]}(a_0 + a_1 u_x + \cdots + a_n u_x^n - t_x) \in F[a_0, \ldots, a_n].$$

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

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

Lemma 4.4. *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.*

Proof. The extension $F(x)/F$ is finite separable and $P_{x,n}$ is the norm of a polynomial in $F(x)[a_0, \ldots, a_n]$ whose coefficients generate $F(x)$ over F (since $n \geq 1$), so 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 4.4 applies to all $x \in Z_{f_1} \cap Z_{f_2}$. Each $P_{x,n}$ is a non-unit in $F(a_1, \ldots, a_n)[a_0]$ and so is not an $F^\times$-multiple of a_n. Thus, the conclusions of Lemma 4.4 also hold in the coordinate ring of $\mathrm{Poly}_{n/F}$. By Definition 3.4 we have

$$M^{\mathrm{geom}}_{f_1,f_2}(u) = \prod_{u_x} \mathrm{N}_{F(u_x)/F}(u - u_x) \in F[u] - \{0\}, \tag{4.10}$$

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 (4.10) and the definition $P_{x,n}(g) = \mathrm{N}_{F(x)/F}(g(u_x) - t_x)$ for $n \geq \deg g$ we see that

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

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 (4.1) and (4.2). For each $x \in Z_{f_1} \cap Z_{f_2}$ we have $M(u_x) = 0$, so $P_{x,n} = P_{x,\deg M - 1} \circ \rho_{n,M}$; note that $\deg M > 0$ if such an x exists. The following theorem is a weak version of Theorem 4.1 in the sense that the dependence on n for the parameters b_n and e_n in (4.11) is not made explicit; the proof of Theorem 4.1 will rest on this weaker result.

Theorem 4.5. *With the hypotheses as in Theorem* 4.1*, assume that $n > 2\deg_{u,T}(f_1)\deg_{u,T}(f_2)$ if $f_1, f_2 \notin F$ and assume $n > \max(\nu(f_1), \nu(f_2))$ otherwise (with $\nu(f)$ as in* (4.6)*).*

There exists a unique $b_n \in F^\times$ and integer $e_n \geq 0$ such that

$$R_n(G) = b_n a_n^{e_n} \cdot \prod_{x \in Z_{f_1} \cap Z_{f_2}} P_{x,n}^{i_x(Z_{f_1}, Z_{f_2})} = b_n a_n^{e_n} \cdot \prod_{x \in Z_{f_1} \cap Z_{f_2}} P_{x,\deg M - 1}^{i_x(Z_{f_1}, Z_{f_2})} \circ \rho_{n,M} \tag{4.11}$$

as algebraic functions on $\mathrm{Poly}_{n/F}$*, where* $M = M^{\mathrm{geom}}_{f_1,f_2}$ *as in Definition* 3.4 *and it is understood that the products over* x *are taken to be* 1 *if* $Z_{f_1} \cap Z_{f_2}$ *is empty.*

Beware that b_n is generally sign-dependent on the ordering of the pair f_1 and f_2.

Proof. We shall first establish a weaker form of (4.11) in which the intersection number at each x is replaced with an unknown positive exponent $e_{x,n}$. Since the $P_{x,n}$'s are irreducible and not scalar multiples of each other in the coordinate ring of $\mathrm{Poly}_{n/F}$, to establish this weaker form of (4.11) it suffices (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 $\mathbf{A}^2_F$ 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}.

It remains to prove that $e_{x,n} = i_x(Z_{f_1}, Z_{f_2})$ for all $x \in Z_{f_1} \cap Z_{f_2}$ and for n as large as in the theorem. By perfectness of F it is harmless to assume that F is algebraically closed, so we now assume this to be the case. If $\deg_T f_1 = \deg_T f_2 = 0$ or if some f_j lies in $F^\times$, then there are no x's and $R_n(G)$ lies in $F^\times$ and has no dependence on n. Hence, we may suppose at least one of the f_j's has positive T-degree, say f_2, and $f_1 \notin F^\times$. First assume $\deg_T f_1 = 0$, so $f_1 \in F[u]$ with positive degree. By multiplicativity of local intersection numbers and resultants, we can assume $f_1 = u - u_0$ for some $u_0 \in F$. Thus, by quasi-periodicity of resultants and the analogous property for local intersection numbers we can replace f_2 with the nonzero $f_2(u_0, T) \in F[T]$ that we may assume has positive degree. We likewise reduce to the case $f_2 = T - t_0$ with $t_0 \in F$. Hence, $Z_{f_1} \cap Z_{f_2} = \{(u_0, t_0)\}$ with intersection number 1 and clearly $R_n(G) = (-1)^n(G(u_0) - t_0)$, so $e_{(u_0,t_0),n} = 1$ as desired.

We may now assume that $\deg_T f_1$ and $\deg_T f_2$ are positive. Let us first argue by deformation theory that $e_{x,n} \geq i_x(Z_{f_1}, Z_{f_2})$ for all x. (We are grateful to de Jong for suggesting this approach, which generalizes to the case of higher genus [8].) Consider any deformation $\widetilde{f}_j$ of f_j over $F[\![\tau]\!]$ with the same T-degree such that $\mathrm{lead}_T(\widetilde{f}_j) \in F[\![\tau]\!][u]$ has leading $F[\![\tau]\!]$-coefficient that is a unit, so its u-degree is the same as that of its reduction $\mathrm{lead}_T(f_j) \in F[u]$. Require also that $\deg_{u,T}(\widetilde{f}_j) = \deg_{u,T}(f_j)$. For example, we may define $\widetilde{f}_j = f_j + \tau f_{j,0}$ for any $f_{j,0} \in F[u][T]$ such that $\deg_T f_{j,0} = \deg_T f_j$, $\deg_u \mathrm{lead}_T(f_{j,0}) = \deg_u \mathrm{lead}_T(f_j)$, and $\deg_{u,T}(f_{j,0}) = \deg_{u,T}(f_j)$; a good choice of such $f_{j,0}$ will be made later. The T-degree and u-degree conditions on $\widetilde{f}_1$ and $\widetilde{f}_2$ ensure that $\widetilde{R}_n(G) := R_{F[\![\tau]\!]}(\widetilde{f}_1(u, G(u)), \widetilde{f}_2(u, G(u)))$ for universal G with large degree n specializes to $R_n(G)$ modulo τ.

The affine $F[\![\tau]\!]$-scheme $Z_{\widetilde{f}_1} \cap Z_{\widetilde{f}_2}$ is necessarily quasi-finite. Indeed, $R_{F[\![\tau]\!][u]}(\widetilde{f}_1, \widetilde{f}_2)$ specializes to $R_{F[u]}(f_1, f_2) \neq 0$, so if the $F(\!(\tau)\!)$-fibers of the $Z_{\widetilde{f}_j}$'s have a common irreducible component, then the $\widetilde{f}_j$'s (which are τ-primitive) must have a common τ-primitive factor in $F[\![\tau]\!][u]$ whose reduction modulo τ is necessarily some $c \in F^\times$ (as f_1 and f_2 do not have a common factor in $F[u]$). Such a

common factor lies in $c + \tau F[\![\tau]\!][u]$ and must have positive u-degree (as otherwise it is a unit), so each $\mathrm{lead}_T(\widetilde{f}_j) \in F[\![\tau]\!][u]$ is forced to have its top-degree coefficient divisible by τ. This contradicts how the $\widetilde{f}_j$'s were chosen, so the desired quasi-finiteness holds. Also, the $\deg_{u,T}$-condition on the $\widetilde{f}_j$'s ensures that n is "sufficiently large" for comparing generic and specialized resultants for $\widetilde{f}_1$ and $\widetilde{f}_2$ over $F[\![\tau]\!]$ in degree n.

Since $\{f_1, f_2\}$ is a regular sequence in $F[u][T]$, by [13, Cor. to 22.5] the $F[\![\tau]\!]$-scheme $Z_{\widetilde{f}_1} \cap Z_{\widetilde{f}_2}$ is flat in a neighborhood of its closed fiber and so is flat. Hence, the "finite part" of this quasi-finite flat $F[\![\tau]\!]$-scheme (in the sense of the structure theorem [11, 18.5.11] is a finite flat $F[\![\tau]\!]$-scheme whose points are naturally labelled by the points x of $Z_{f_1} \cap Z_{f_2}$, and the x-component has rank $i_x(Z_{f_1}, Z_{f_2})$ over $F[\![\tau]\!]$ because F is algebraically closed.

We next claim that the deformations $\widetilde{f}_j$ can be chosen so that the $F(\!(\tau)\!)$-fibers of the $Z_{\widetilde{f}_j}$'s are smooth and their finite intersection is $F(\!(\tau)\!)$-étale. Fix $j \in \{1, 2\}$ and consider the F-vector space spanned by u, T, and the monomials that appear in f_j. This vector space of functions contains f_j and (since $\deg_T f_j > 0$) its generic member has T-degree equal to $\deg_T f_j$, leading T-coefficient with u-degree $\deg_u(\mathrm{lead}_T(f_j))$, and total (u,T)-degree $\deg_{u,T}(f_j)$. This defines a closed immersion of the affine (u,T)-plane into an affine space $\mathbf{A}^{N_j}$. By Bertini's theorem there is a Zariski-dense open locus W_j of affine hyperplanes in $\mathbf{A}^{N_j}$ whose members have smooth intersection with the (u,T)-plane. In particular, if $f_{j,0}$ is generically chosen in W_j, then the pencil $f_j + \tau f_{j,0}$ has all but finitely many members with smooth zero locus in the (u,T)-plane and it satisfies the degree requirements (over $F[\tau]$) that we have demanded for $\widetilde{f}_j$ (over $F[\![\tau]\!]$). Taking this pencil to define $\widetilde{f}_j$ over $F[\![\tau]\!]$ therefore makes the $Z_{\widetilde{f}_j}$'s have smooth $F(\!(\tau)\!)$-fibers, so it remains to arrange the choices of $f_{1,0}$ and $f_{2,0}$ to ensure that $Z_{\widetilde{f}_1} \cap Z_{\widetilde{f}_2}$ is $F(\!(\tau)\!)$-étale.

We fix any $f_{2,0} \in W_2$ as above and will choose $f_{1,0} \in W_1$ appropriately. The zero-scheme of $f_{2,0}$ is smooth in the (u,T)-plane, so a Zariski-dense open locus of affine hyperplanes in $\mathbf{A}^{N_1}$ therefore meets this zero-scheme with étale overlap and does not meet the part of its closure in $\mathbf{P}^{N_1}$ that lies in the hyperplane at infinity. Thus, we can choose the pencil of affine hyperplanes $\widetilde{f}_1 = f_1 + \tau f_{1,0}$ so that $Z_{f_{1,0}} \cap Z_{f_{2,0}}$ is étale and coincides with the overlap of the closures of the $Z_{f_{j,0}}$'s in $\mathbf{P}^{N_1}$. The overlap of the associated pencils of projective hyperplanes in $\mathbf{P}^{N_1}$ (with parameter $\tau \in \mathbf{P}^1$ for the two pencils) is therefore finite and flat over a neighborhood of $\tau = \infty$ in $\mathbf{P}^1$, so it is (finite and) étale over such a neighborhood as well. With this choice of $f_{1,0}$, the intersection $Z_{\widetilde{f}_1} \cap Z_{\widetilde{f}_2}$ over $F(\!(\tau)\!)$ is étale.

Let K be an algebraic closure of $F(\!(\tau)\!)$. With the $\widetilde{f}_j$'s as just chosen, the quasi-finite flat overlap $Z_{\widetilde{f}_1} \cap Z_{\widetilde{f}_2}$ over $F[\![\tau]\!]$ with étale generic fiber contains exactly $i_x(Z_{f_1}, Z_{f_2})$ points on the geometric generic fiber (over K) that specialize to x. The universal resultant $\widetilde{R}_n(G)$ associated to $\widetilde{f}_1$ and $\widetilde{f}_2$ over K is a determinant that lies in $F[\![\tau]\!][a_0, \ldots, a_n]$ and (as we noted earlier) it specializes to $R_n(G)$ under reduction modulo τ. Thus, by applying the proved weak form of (4.11) (with exponents $e_{x,n}$ positive but not yet known) to the $\widetilde{f}_j$'s over K, each K-point of $Z_{\widetilde{f}_1} \cap Z_{\widetilde{f}_2}$ that specializes to x contributes a linear factor to $\widetilde{R}_n(G)$ over $K[a_0, \ldots, a_n]$ with coefficients in the integral closure $\mathcal{O}$ of $F[\![\tau]\!]$ in K, and this element of $\mathcal{O}[a_0, \ldots, a_n]$ has reduction modulo the maximal ideal of $\mathcal{O}$ equal to the irreducible $P_{x,n} =$

$\sum a_i u_x^i - t_x$ (that involves a_0) over F. By Gauss' lemma over a sufficiently large finite extension of $F(\!(\tau)\!)$, each such K-point of $Z_{\widetilde{f}_1} \cap Z_{\widetilde{f}_2}$ therefore contributes a factor to $\widetilde{R}_n(G)$ in $\mathcal{O}[a_0, \ldots, a_n]$ with reduction $P_{x,n}$, and these factors are pairwise relatively prime over K. There are $i_x(Z_{f_1}, Z_{f_2})$ such K-points, so overall we see that $P_{x,n}$ divides $R_n(G)$ with multiplicity at least as large as this intersection number. This gives the desired inequality $e_{x,n} \geq i_x(Z_{f_1}, Z_{f_2})$.

To deduce equality for all x, by adding these inequalities for all x it is enough to prove that $\sum_x e_{x,n}$ is equal to the F-length of $Z_{f_1} \cap Z_{f_2}$. Each $P_{x,n}$ is a linear polynomial in the a_j's with exact degree 1 in a_0. Hence, $\sum_x e_{x,n}$ is the a_0-degree of $R_n(G)$. Consider the specialization map $F[a_0, \ldots, a_n][1/a_n] \to F[T]$ that carries G to $h_n(u) + T \in F[u][T]$ with $h_n(u) \in F[u]$ defined to be

$$h_n(u) := u^{n - \sum_x (i_x(Z_{f_1}, Z_{f_2}) + 1)} \cdot \prod_x (u - u_x)^{i_x(Z_{f_1}, Z_{f_2}) + 1}. \tag{4.12}$$

The factor given by the power of u in this definition has a nonnegative exponent due to Bézout's theorem and the largeness hypothesis on n. This specialization mapping carries a_0-degree to T-degree and $R_n(G)$ to $R_{F[T]}(f_1(u, h_n(u)+T), f_2(u, h_n(u)+T))$ (due to the largeness of n), so

$$\sum_x e_{x,n} = \deg_T R_{F[T]}(f_1(u, h_n(u) + T), f_2(u, h_n(u) + T)).$$

Our problem is therefore to prove

$$\deg_T R_{F[T]}(f_1(u, h_n(u) + T), f_2(u, h_n(u) + T)) \stackrel{?}{=} \ell(Z_{f_1} \cap Z_{f_2})$$

for n as large as in the theorem. Due to the largeness of n, each $g_j(u, T) = f_j(u, h_n(u)+T)$ has leading coefficient in $F^\times$ when considered as a polynomial in u (with coefficients in $F[T]$). Hence, by the following lemma, the above T-degree of the resultant is equal to the length of the (necessarily F-finite) intersection scheme of the zero loci of the polynomials $g_j(u, T)$.

Lemma 4.6 (Zeuthen's rule)**.** *Let $g_1(u, T) = 0$ and $g_2(u, T) = 0$ be (possibly empty) plane curves over an algebraically closed field K, and assume that these zero loci do not share a common irreducible component and that the leading u-coefficients $\mathrm{lead}_u(g_1), \mathrm{lead}_u(g_2) \in K[T]$ do not have a common zero at $t_0 \in K$. The resultant $R_{K[T]}(g_1, g_2)$ vanishes at t_0 to order*

$$\mathrm{ord}_{t_0} R_{K[T]}(g_1, g_2) = \sum_{c \in K} i_{(c,t_0)}(g_1, g_2),$$

where $i_x(g_1, g_2) = \dim_K \mathcal{O}_{\mathbf{A}^2,x}/(g_1, g_2)$ is the intersection number at x.

In characteristic 0 this identity can proved by using the Puiseux series description of finite extensions of $K(\!(T - t_0)\!)$ [10, 1.2.5(f)]. In positive characteristic there is no simple description like this, so a different method is required.

Proof. We may assume $t_0 = 0$. Both sides of the desired identity are finite, and each side is unaffected by switching the roles of g_1 and g_2. In $K[\![T]\!][u]$ at least one of g_1 or g_2 has a unit leading coefficient, so we may assume g_1 has this property and we wish to prove that the nonzero element $R_{K[\![T]\!]}(g_1, g_2) \in K[\![T]\!]$ has order $\sum_c \ell(K[\![u - c, T]\!]/(g_1, g_2))$ as c ranges through the finitely many common roots of $g_1(u, 0)$ and $g_2(u, 0)$ (with $g_1(u, 0) \neq 0$). The g_j's shall now only matter through

their associated ideals in $K[\![T]\!][u]$, so in particular it is permissible to replace each of them with a $K[\![T]\!]^\times$-multiple (that may not lie in $K[T][u]$).

Clearly $K[\![T]\!][u]/(g_1)$ is a finite flat $K[\![T]\!]$-algebra and replacing g_1 with a $K[\![T]\!]^\times$-multiple allows us to assume that g_1 is monic in u, so by universal identities,

$$R_{K[\![T]\!]}(g_1, g_2) = \mathrm{N}_{(K[\![T]\!][u]/(g_1))/K[\![T]\!]}(g_2 \bmod (g_1))$$

in $K[\![T]\!]$. If $\deg_u g_1 = 0$, then $g_1 = 1$, so the desired result is trivial in this case. Now assume $\deg_u g_1 > 0$. The decomposition of $K[\![T]\!][u]/(g_1)$ into a finite product of finite local $K[\![T]\!]$-algebras matches the decomposition of $K[u]/(g_1(u,0))$ into a finite product of finite local K-algebras. For each root $c \in K$ of the nonconstant $g_1(u,0)$, the image of g_2 in the corresponding factor is a unit if $g_2(c,0) \neq 0$ and is not a unit otherwise, so the $K[\![T]\!]$-norm of $g_2 \bmod (g_1)$ is a unit multiple of the product of the norms of g_2 with respect to the finite local $K[\![T]\!]$-algebras $K[\![u-c, T]\!]/(g_1)$ as c ranges through the common roots of $g_1(u,0)$ and $g_2(u,0)$. It therefore suffices to prove that for each such c,

$$\mathrm{ord}(\mathrm{N}_{(K[\![u-c,T]\!]/(g_1))/K[\![T]\!]}(g_2)) \stackrel{?}{=} \ell(K[\![u-c, T]\!]/(g_1, g_2)).$$

We may assume $c = 0$. The $K[\![T]\!]$-algebra $K[\![u, T]\!]/(g_1)$ is finite and flat, with g_2 a regular element that is not a unit, so the right side is the Herbrand quotient for multiplication by g_2 on the $K[\![T]\!]$-module $K[\![u, T]\!]/(g_1)$ and the left side is the Herbrand quotient for its determinant acting on the $K[\![T]\!]$-module $K[\![T]\!]$. Hence, the desired equality of order and length is a special case of the general behavior of Herbrand quotients with respect to determinants [10, Lemma A.2.6]. □

Returning to the proof of Theorem 4.5, since $Z_{f_1} \cap Z_{f_2} = Z_{g_1} \cap Z_{g_2}$ as sets (by the definition of $h_n(u)$ in (4.12)) it remains to improve this to an equality of schemes. That is, we want $i_x(Z_{f_1}, Z_{f_2}) = i_x(Z_{g_1}, Z_{g_2})$ for all x in this common overlap. By construction, $h_n(u)$ vanishes at u_x to order exceeding $i_x(Z_{f_1}, Z_{f_2})$. Hence, we just have to prove that if two curves meet properly with intersection number $r > 0$ at a point x on a smooth surface S, then the intersection number at x is invariant under any local deformation of the defining equations of the curves if the deformation is the identity to order r. That is, if k is an algebraically closed field and $\{f_1, f_2\}$ is a system of parameters in $k[\![u, T]\!]$ with $r = \ell(k[\![u, T]\!]/(f_1, f_2))$, then we claim that for any $\varepsilon_1, \varepsilon_2 \in (u, T)^{r+1}$ the k-finite quotient $k[\![u, T]\!]/(f_1 + \varepsilon_1, f_2 + \varepsilon_2)$ has length r as well. In fact, even the ideals $(f_1 + \varepsilon_1, f_2 + \varepsilon_2)$ and (f_1, f_2) in $k[\![u, T]\!]$ coincide: since $(u, T)^r \subseteq I := (f_1, f_2)$ we have $\varepsilon_1, \varepsilon_2 \in (u, T)I$, so the elements $f_1 + \varepsilon_1, f_2 + \varepsilon_2 \in I$ are a pair of generators of I by Nakayama's Lemma. □

Corollary 4.7. *Let F be a perfect field with positive characteristic p and $f(T) \in F[u][T]$ a nonconstant squarefree element. Let M_f^{geom} be as in Definition* 3.4.

1) *If f lies in $F[u][T^p]$, then for $\deg g > 2 \deg_{u,T}(f) \deg_{u,T}(\partial_u f)$ 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 $\deg g > 2p^2 \deg_{u,T}(f) \deg_{u,T}(\partial_u f)$ the property of $f(g^p)$ being separable in $F[u]$ is determined by $g \bmod M_f^{\mathrm{geom}}$.*

Since $f \notin F$, Lemma 3.5 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). For the study of $p = 2$ we will need the second case in this corollary.

Proof. If f as in the first case is written in the form $f = h(T^p)$, then $M_h^{\mathrm{geom}} = M_f^{\mathrm{geom}}$ because on geometric points the map $(u,t) \mapsto (u,t^p)$ sets up a bijection between $Z_f \cap Z_{\partial_u f}$ and $Z_h \cap Z_{\partial_u h}$. Replacing T^p with T decreases the total degree at most by a proportion $1/p$, so as in the proof of Lemma 3.5 it suffices to prove the second case.

We may apply Theorem 4.5 with $f_1 = f(T^p)$ and $f_2 = (\partial_u f)(T^p)$ to conclude that for g with large degree as in (2), the vanishing of the resultant of $f(g^p)$ and $(\partial_u f)(g^p)$ only depends on $g \bmod M_f^{\mathrm{geom}}$. Also, $f(g^p)$ is inseparable in $F[u]$ precisely when it has a common geometric root with its derivative $f(g^p)' = (\partial_u f)(g^p)$. Hence, the separability of $f(g^p)$ only depends on $g \bmod M_f^{\mathrm{geom}}$ for g with large degree as in (2). □

Before we prove Theorem 4.1, we apply it to prove a periodicity property for $\mu(f(g))$:

Theorem 4.8. *Let κ have odd characteristic p, and let χ be the quadratic character on $\kappa^\times$; define $\chi(0) = 0$. Let $f(T) \in \kappa[u][T^p]$ be squarefree in $\kappa[u][T]$, and assume $f \notin \kappa$.*

There is a nonzero polynomial $M = M_{f,\kappa}$ in $\kappa[u]$ such that for $g_1 = c_1 u^{n_1} + \cdots$ and $g_2 = c_2 u^{n_2} + \cdots$ in $\kappa[u]$ with sufficiently large degrees n_1 and n_2,

$$g_1 \equiv g_2 \bmod M, \quad n_1 \equiv n_2 \bmod 4, \quad \chi(c_1) = \chi(c_2) \Longrightarrow \mu(f(g_1)) = \mu(f(g_2)). \tag{4.13}$$

If -1 is a square in κ or $\deg_T f$ is even, the second congruence in (4.13) *may be relaxed to $n_1 \equiv n_2 \bmod 2$.*

A choice for the modulus M is M_f^{geom} as in Definition 3.4. *Moreover, the "sufficiently large" condition on n_1 and n_2 only depends on the total degree $\deg_{u,T} f$ and $\deg M$ (not on $\#\kappa$).*

The monic modulus $M_{f,\kappa}^{\min}$ of minimal degree in $\kappa[u]$ for which (4.13) *is true for all large n_1 and n_2 is a factor of any other M. 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'.*

Proof. In this proof, the importance of $f(T)$ being a polynomial in T^p is that 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, $\partial_u(f(u, g(u))) = (\partial_u f)(u, g(u))$ is a polynomial in g with no dependence on $g'(u)$.

For g in $\kappa[u]$ of sufficiently large degree, $f(g)$ is nonzero and (2.5) and (3.4) yield

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

with $d = \deg f(g)$. Note that $f(g)' = (\partial_u f)(g)$ since $f \in \kappa[u][T^p]$. Since f is squarefree and $f \notin \kappa$, so $\partial_u f \neq 0$ by Lemma 3.5, we have $(\partial_u f)(g) \neq 0$ when $\deg g > \nu(\partial_u f)$.

When $\deg g \geq \deg_{u,T} f$, both $d = \deg f(g)$ and $\deg((\partial_u f)(g))$ are linear in $\deg g$ by (4.4) and (4.6). Using (4.4) and (4.5) to compute $\deg f(g)$ and $\operatorname{lead} f(g)$ in terms of $\deg g$ and $\operatorname{lead} g$ for such g, we have by Theorem 4.1 that there exist $\varepsilon_0, \varepsilon_1 \in \{\pm 1\}$ and integers m_0 and m_1 such that for $\deg g$ sufficiently large (depending only on $\deg_{u,T} f$),

$$\mu(f(g)) = \varepsilon_0 \varepsilon_1^{\deg g} (\chi(-1))^{(\deg f(g))(\deg f(g)-1)/2} \chi(\operatorname{lead} g)^{m_0 + m_1 \deg g} \chi(L(g)) \tag{4.14}$$

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 (4.14) depends on $\deg g$ modulo 2.

Since any congruence class in $\kappa[u]/(M)$ with $M \neq 0$ may be represented by a polynomial of any degree $\geq \deg M$ 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 $\mu(f(g))$, the greatest common divisor (M_1, M_2) is also a modulus. Hence, $M_{f,\kappa}^{\min}$ (defined to be the monic modulus of least degree) divides all other moduli for $\mu(f(g))$.

Now let us establish the final part of Theorem 4.8 concerning the behavior of $M_{f,\kappa'}^{\min}$ for sufficiently large finite extensions κ' of κ. Let κ'/κ be a finite extension such that all points in the finite set $Z_f \cap Z_{\partial_u f} \subseteq \mathbf{A}_\kappa^2$ are κ'-rational, and so in particular M_f^{geom} splits into linear factors in $\kappa'[T]$. This rationality property is inherited by all finite extensions of κ'. We claim that $M_{f,\kappa''}^{\min} = M_f^{\mathrm{geom}}$ when κ'' is any finite extension of κ'. To prove this, we can assume M_f^{geom} is nonconstant since $M_{f,\kappa''}^{\min} | M_f^{\mathrm{geom}}$.

Choose a monic linear factor of M_f^{geom} in $\kappa'[u]$, say $u - u_x$ 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 in $\kappa'[u]$ 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 $x' \neq x$. By (4.11) and the positivity of the intersection numbers in the exponents, for sufficiently large n the resultant of $f(g_1)$ and $(\partial_u f)(g_1) = f(g_1)'$ vanishes and the resultant of $f(g_2)$ and $f(g_2)'$ is nonzero; 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 $M_f^{\mathrm{geom}}/(u - u_x)$, we conclude that this divisor of M_f^{geom} cannot be a modulus for $\mu_{\kappa''[u]}(f(g))$ and so cannot be divisible by $M_{f,\kappa''}^{\min}$. Thus, $M_{f,\kappa''}^{\min} = M_f^{\mathrm{geom}}$. □

Example 4.9. 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 4.8 gives

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

Note that $\chi(g(0)^2+1)^2$ is not always 1 because it may vanish. Since $R(u^2+1, f(g))$ only depends on $g \bmod u^2 + 1$ (by quasi-periodicity of resultants), $\mu(f(g))$ only depends on $g \bmod u(u-1)(u^2+1)$. (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, $M_{f,\kappa}^{\min} = (u-1)(u^2+1) \neq M_f^{\mathrm{geom}}$.

As preparation for the proof of Theorem 4.1, which uses an algebraic method that rests on variation in the ordered pair (f_1, f_2), we need to establish an alternative formulation of the result. Fix a nonzero $M \in F[u]$. Choose

$$n_0 > 2 \deg_{u,T} f_1 \cdot \deg_{u,T} f_2$$

such that $n_0 \geq \deg(\mathrm{lcm}(M, M^{\mathrm{geom}}_{f_1,f_2}))$ and (in case some f_j is in $F^\times$) $n_0 > \max(\nu(f_1), \nu(f_2))$ (see (4.6)). Consider the claim that there exist $c \in F^\times$, integers m_0 and m_1, and an algebraic function $L_{f_1,f_2} : F[u]/(M) \to \mathbf{A}^1_F$ such that for $n \geq n_0$ there is an equality of algebraic functions

$$R_n(G) = c^n a_n^{m_0+m_1 n} \cdot (L_{f_1,f_2} \circ \rho_{n,M}) \tag{4.15}$$

on $\mathrm{Poly}_{n/F}$, with $\rho_{n,M}$ as in (4.1) or (4.2). The key point is that a formula as in (4.15) holds for some such M and all $n \geq n_0$ if and only if e_n is a $\mathbf{Z}$-polynomial of degree ≤ 1 in such n and $b_n = \beta_0 \beta_1^n$ for some $\beta_0, \beta_1 \in F^\times$ for such n (with e_n and b_n as in Theorem 4.5). Sufficiency is obvious by Theorem 4.5, and for necessity we may replace M with $\mathrm{lcm}(M, M^{\mathrm{geom}}_{f_1,f_2})$ to get to the case where $M^{\mathrm{geom}}_{f_1,f_2} | M$ and $n_0 \geq \deg M$, so we have formulas

$$R_n(G) = b_n a_n^{e_n} \prod_x P^{i_x(Z_{f_1},Z_{f_2})}_{x,\deg M-1} \circ \rho_{n,M}$$

and

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

as rational functions on $\mathrm{Poly}_{n/F}$ for all $n \geq n_0$. Thus, 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 all such n, so

$$b_n c^{-n} \prod_x P^{i_x(Z_{f_1},Z_{f_2})}_{x,\deg M-1} \circ \rho_{n,M} = L_{f_1,f_2} \circ \rho_{n,M}$$

for all such n. Hence,

$$b_n c^{-n} \prod_x P^{i_x(Z_{f_1},Z_{f_2})}_{x,\deg M-1} = L_{f_1,f_2}$$

on $\mathrm{Poly}_{\leq(\deg M-1)/F}$. Since L_{f_1,f_2} and the $P_{x,\deg M-1}$'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 our large n. Thus, $b_n = c' c^n$ for $c, c' \in F^\times$ and all $n \geq n_0$; this is the desired result.

We shall now aim to prove an identity of the form (4.15) for large n (without making the lower bound on n explicit) 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 $M^{\mathrm{geom}}_{f_1,f_2}$ we must have

$$L_{f_1,f_2} = c_0 \prod_x P^{i_x(Z_{f_1},Z_{f_2})}_{x,\deg M-1}$$

for some $c_0 \in F^\times$. In what follows we will (for expository simplicity) 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.

Remark 4.10. Since $\deg M^{\mathrm{geom}}_{f_1,f_2} \leq \sum_x i_x(Z_{f_1}, Z_{f_2}) \leq (\deg_{u,T} f_1)(\deg_{u,T} f_2)$, the interested reader may easily check that the inductive argument below determines an $n_0 = n_0(\deg_{u,T} f_1, \deg_{u,T} f_2)$ so that (4.15) with $M = M^{\mathrm{geom}}_{f_1,f_2}$ holds for all $n \geq n_0$. The main point is that the number of steps in the recursive procedure is bounded

above in terms of $\deg_{u,T} f_1$ and $\deg_{u,T} f_2$. We have not attempted to make such an n_0 explicit.

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 (4.15) does not depend on this ordering. Indeed, for generic g of any sufficiently large degree (such as $\deg g > \max(\nu(f_1), \nu(f_2))$ with notation as in (4.6)),

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

where $e_0 = (\deg \alpha_{1,d_1})(\deg \alpha_{2,d_2})$ and $e_1 = d_1 \deg \alpha_{2,d_2} + d_2 \deg \alpha_{1,d_1} + d_1 d_2$, 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 and f_2 are interchanged. We will use this repeatedly.

Our proof of Theorem 4.1 will roughly be a series of identities

$$R(f_1(g), f_2(g)) = c_0 c_1^{\deg g} (\operatorname{lead} g)^{\mu_0 + \mu_1 \deg g} R(f_3(g), f_4(g))$$

for generic g of large positive degree (or universal g with a unit leading coefficient and large degree), where the elements $c_0, c_1 \in F^\times$ and $\mu_0, \mu_1 \in \mathbf{Z}$ depend on the ordered pair (f_1, f_2), 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 of "smaller" that we use, depending on the stage of our argument.) In this way, induction and the reformulation via (4.15) will establish Theorem 4.1, provided that along the way we also prove that m_1 in (4.15) is equal to $(\deg_T f_1) \cdot (\deg_T f_2)$.

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,

$$R(a(u), f_2(g)) = R(c, f_2(g)) R(a_1(u), f_2(g)) = c^{\deg f_2(g)} R(a_1(u), f_2(g)). \tag{4.16}$$

For generic g with degree exceeding $\nu(f_2)$, $c^{\deg f_2(g)} = c_0 c_1^{\deg g}$ 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 (4.15) in the present case for large n (with $m_1 = 0$), and so proves Theorem 4.1 when some $\deg_T f_j$ vanishes.

To prove Theorem 4.1 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

$$R(f_1(g), f_2(g)) = R(a(u), f_2(g)) R(h(g), f_2(g)), \tag{4.17}$$

with the first factor on the right side satisfying the induction hypothesis by the preceding discussion. Removing a common factor from the coefficients of f_2 as a polynomial in T goes the same way.

We will prove Theorem 4.1 (in the guise of (4.15)) 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 4.11. *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 $\alpha, \beta \notin F$). Then 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

$$\min(\deg\widetilde{\alpha},\deg\widetilde{\beta}) < \min(\deg\alpha,\deg\beta) \tag{4.18}$$

such that for all extensions F'/F and all g in $F'[u]$ with sufficiently large degree (depending only on d and the u-degrees of the monomials appearing in the h_j's) we have

$$R(h_1(g),h_2(g)) = c\varepsilon^{\deg g}(\operatorname{lead} g)^m R(\widetilde{h}_1(g),\widetilde{h}_2(g)). \tag{4.19}$$

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 $\deg\alpha \le \deg\beta$. (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 $\deg g$ is at least as large as $\nu(h_1)$, $\nu(h_2)$, and $\nu(k)$ (notation as in (4.6)), 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 4.2, the identity (4.19) 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 that the leading T-coefficients of $h_1(T)$ and $h_2(T)$ do not divide each other, we assume $h_1(T)$ and $h_2(T)$ are relatively prime.

Lemma 4.12. *Let $h_1(T)$ and $h_2(T)$ in $F[u][T]$ have common T-degree $d \ge 1$:*

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

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 (depending only on d and the u-degrees of the monomials appearing in the h_j's), we have $R(h_1(g),h_2(g)) = c\varepsilon^{\deg g}(\operatorname{lead} g)^{m+m_1\deg(g)} R(\widetilde{h}_1(g),\widetilde{h}_2(g))$ with $m_1 = d(d-\deg_T\widetilde{h}_1)$.

Proof. If neither α nor β divides the other in $F[u]$, apply Lemma 4.11 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 such as $c\varepsilon^{\deg g}(\operatorname{lead} g)^m$ behave well under multiplication: the c's and ε's are multiplicative and the m's are additive.) The condition (4.18) 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 $(h_1,h_2) = 1$. Proceed as in the proof of Lemma 4.11, taking $\widetilde{h}_1 = k$ and $\widetilde{h}_2 = h_1$. □

We are now ready to prove Theorem 4.1:

Proof of Theorem 4.1. We argue via (4.15) by induction on $\max(\deg_T f_1, \deg_T f_2)$, and we refer the reader to Remark 4.10. 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 (4.16). Remove any nontrivial common factor from the $F[u]$-coefficients of $f_1(T)$ as a polynomial in T, using (4.17), so $f_1(T)$ is primitive over $F[u]$. Similarly make f_2 primitive. By Lemma 4.12 and induction, we may assume $d_1 \neq d_2$, and without loss of generality $0 < d_1 < d_2$. Writing

$$f_1(T) = \alpha(u)T^{d_1} + \dots, \quad f_2(T) = \beta(u)T^{d_2} + \dots, \tag{4.20}$$

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 $\operatorname{lead} f_1(g)$ has 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 $m_0, m_1 \in \mathbf{Z}$ that do not depend on g, with $m_1 = d_1(d_2 - \deg_T k)$, so we are now reduced to proving (4.15) 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 we have $\max(\deg_T f_1, \deg_T k) < d_2$, so (4.15) (and hence Theorem 4.1) with f_1 and k has already been proved by the induction hypothesis. Thus, it remains to treat the case (4.20) 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 4.11. To circumvent this problem, we shall use a trick that puts us in the case in which $\alpha | \beta$: consider the universal identity

$$R(f_1(g), \alpha(u))R(f_1(g), f_2(g)) = R(f_1(g), \alpha(u)f_2(g)) \tag{4.21}$$

with g the universal polynomial of large degree n with a unit leading coefficient. The first term in (4.21) is nonzero, since primitivity of f_1 forces $(f_1(g), \alpha(u)) = 1$. Since all three resultants admit expressions as in Theorem 4.5 for a common modulus M, if (4.15) is proved for two of the three pairs (f_1, α), (f_1, f_2), and $(f_1, \alpha f_2)$ in (4.21) then it follows for the third. Since the case of a polynomial of T-degree zero has already been settled, it suffices to treat the ordered pair $(f_1, \alpha(u)f_2)$.

The right side of (4.21) 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 $(f_1, f_2) = 1$ 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 $(f_1, h) = 1$. 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 (4.21) 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 $m_0, m_1 \in \mathbf{Z}$. Explicitly, $m_1 = d_1(d_2 - \deg_T h)$. 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. We may now infer the desired result for the pair $(f_1, \alpha f_2)$. □

Example 4.13. For f as in Lemma 3.5 with $f \notin F$, the pair $f_1 = f$ and $f_2 = \partial_u f$ satisfies the hypotheses in Theorem 4.1. A local calculation shows that in this case $B = Z_{f_1} \cap Z_{f_2}$ is the non-étale locus for projection from $\{f = 0\}$ to the T-axis, and $i_x(Z_{f_1}, Z_{f_2})$ is the length of B at x. As an illustration (via Example 3.2), for f as in Example 1.1 the projection from the plane curve $\{f = 0\}$ to 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. Theorems 4.1 and 4.5 thereby explain why $\mu(f(g))$ has the form given in (3.6); note the appearance of $g(0)^{18}(g(1)^2 + 1)^9$ in the discriminant formula immediately following (3.6).

Corollary 4.14. *Let F be a perfect field of characteristic $p > 0$ and let f_1 and f_2 be nonzero and relatively prime in $F[u][T]$. Assume, for some $m \geq 0$, $f_j = h_j(u, T^{p^m})$ for $j = 1$ and 2. For each $x \in Z_{f_1} \cap Z_{f_2}$, the multiplicity 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 as in Theorem 4.5 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. We just have to prove $i_x(Z_{f_1}, Z_{f_2}) = p^m \cdot i_{(1 \times \phi^m)(x)}(Z_{h_1}, Z_{h_2})$. As a radicial self-map of the affine plane, $1 \times \phi^m$ is finite flat with degree p^m and it identifies $Z_{f_1} \cap Z_{f_2}$ with the fiber over $Z_{h_1} \cap Z_{h_2}$. Thus, it multiplies lengths by p^m since F is perfect. □

Corollary 4.14 is useful in the study of variation of Möbius periodicity as we vary f:

Example 4.15. Consider the algebraic family of polynomials $\sum_{i=0}^n \alpha_i(u) T^{e_i}$ with a fixed strictly increasing sequence of nonnegative integers $\{e_i\}$ and varying $\alpha_i(u)$'s with specified degrees $d_i \geq 0$ such that $d_i > 0$ for some i. We are interested in the case when $p \nmid e_{i_0}$ for some i_0, but we do not impose such a condition yet. Fix $m \geq 0$. Denoting a polynomial in our algebraic family as h, is the branch scheme nonempty for projection from $h(T^{p^m}) = 0$ to the T-axis for generic h in the family when $p \neq 2$? The generic existence of branch points amounts to the generic resultant $R(h, \partial_u h)$ of h and $\partial_u h$ as polynomials in T having positive u-degree. In the special case that $d_i \leq 1$ for all i, the generic resultant is nonzero with u-degree 0 and hence there are no branch points for generic h. In all other cases there are branch points for generic h, as we now explain.

Assume $d_{i_1} \geq 2$ for some i_1. As we have noted, the generic existence of branch points amounts to the resultant $R(h, \partial_u h)$ of h and $\partial_u h$ (as polynomials in T) having positive u-degree for generic α_i with degree d_i. If $\deg \alpha_i = 0$ for all $i > 0$ (so $\deg \alpha_0 \geq 2$), then the resultant is $\alpha_0'(u)^{\deg_T h}$, and this has positive degree in the generic case since $p > 2$. Suppose instead that $\deg \alpha_i > 0$ for some $i > 0$, so $d := \deg_T \partial_u h$ is positive. Choose any i_1 with $d_{i_1} \geq 2$, so α_{i_1}' has positive degree

in the generic case (as $p \neq 2$). Let $c = \mathrm{lead}_u(\alpha'_{i_1})$ in the generic case. Consider the expansion of the universal determinant that defines the u-polynomial $R(h, \partial_u h)$ (over the ring of generic coefficients of the α_i's). Since the generic constant term of each α_j does not appear in $\partial_u f$, this expansion involves exactly one appearance of

$$c^{\deg_T h}\alpha_n(0)^r\alpha_0(0)^s u^{\deg_T h \cdot \deg_u(\alpha'_{i_1})}$$

with $r, s \geq 0$ and r maximal ($r = 0$ in case $i_1 = n$). Hence, this term with positive u-degree does not cancel out.

If some e_{i_0} is not divisible by p and $d_{i_1} \geq 2$ for some i_1, then when $p \neq 2$ the projection from Z_h to the T-axis has an étale point on its branch scheme for generic h; this is proved via deformation theory and Bertini theorems in [9] (with the u-line replaced by an arbitrary smooth affine curve over κ with one geometric point at infinity). Thus, except for the cases when $d_i \leq 1$ for all i, if $p \neq 2$ then the branch scheme of projection to the T-axis from the zero-scheme of a generic member of any algebraic family $\{\sum \alpha_i T^{\mu_i} \mid \deg \alpha_i = d_i\}$ as above has a point with *odd* (and even p-power) length when $p|\mu_i$ for all i, due to Corollary 4.14.

5. Characteristic 2

The analogue of Theorem 4.8 in characteristic 2 is subtle because (2.5) in characteristic 2 requires liftings into characteristic 0. When κ has characteristic 2, ideally we want a result about the periodicity of $\mu(f(g))$ for squarefree $f \in \kappa[u][T^2]$. We will be able to prove something about $\mu(f(g))$ when f is a polynomial in T^2 (Theorem 5.10), but not periodicity: Möbius formulas in particular examples (even the polynomial $T^2 + u$, as in [7]) do not seem to satisfy simple periodicity properties. However, we shall prove (Theorem 5.12) that squarefree polynomials in T^4 have periodic Möbius values.

In odd characteristic, a modulus of Möbius periodicity for a squarefree polynomial $f(T)$ is given by a geometrically constructed polynomial M_f^{geom} as in Definition 3.4. In characteristic 2 we have to use a slightly different procedure, as follows. Writing $f(T) = h(T^2)$, $h(T)$ is squarefree since $f(T)$ is, and $f(T)$ has no local obstructions if and only if $h(T)$ has none. The relevant "modulus" for $g \mapsto \mu(f(g))$ will turn out to be not M_f^{geom}, but a polynomial closely related to M_h^{geom} (and in some examples it does not seem that there is a squarefree modulus, as can always be found in odd characteristic). The work we carry out in characteristic 2 will involve an interplay between resultants in characteristic 2 and in characteristic 0. For the application to finite fields κ, we will need to work in the Witt vectors of κ. Since finiteness of the field won't matter until we reach the application to Möbius values, our work on resultants will be carried out over any perfect field k of characteristic 2 and over its Witt vector ring $W = W(k)$.

Hypothesis. Our running convention throughout §5 is that k is a perfect field with characteristic 2 and h is a polynomial in $k[u][T]$ such that $h \notin k$ and $h(T^2)$ is squarefree in $k[u][T]$. If k is finite then we also assume that $h(T)$ has no local obstructions; that is, $h(T)$ is nonzero as a function on the finite fields $k[u]/(\pi)$ for all π.

Our hypotheses force h to be squarefree in $k[u][T]$ and not to have any prime factors in $k[T]$ (otherwise $h(T^2)$ has a factor in $k[T^2] = k[T]^2$ since k is perfect), and also force $h(g^2) \neq 0$ for all $g \in k[u]$. Since $h \notin k$, Lemma 3.5(2) ensures that

$\partial_u h \neq 0$ and that h and $\partial_u h$ are relatively prime in $k[u][T]$. Thus, $R_{k[u]}(h, \partial_u h) \neq 0$ and we may define M_h^{geom} as in Definition 3.4. Corollary 4.7(2) ensures that whether or not $h(g^2)$ is separable in $k[u]$ only depends on $g \bmod M_h^{\mathrm{geom}}$, provided $\deg g$ is sufficiently large. This largeness only depends on the total degree of h.

Let us check that our hypotheses on h ensure that for any sufficiently large d there exists $g \in k[u]$ with degree d such that $h(g^2)$ is nonconstant and separable in $k[u]$. Since $h(T^2)$ is squarefree in $k[u][T]$ and $h \notin k$, for infinite k the existence of such a g in any degree exceeding $\max(\nu(h), \nu(\partial_u h))$ (see (4.6)) follows from Lemma 4.2 and the Zariski-denseness of the locus of k-rational points in any affine space over k. In the case of finite k, we use the additional hypothesis (for such k) that h has no local obstructions, as then the existence of such a g in any large degree is ensured by [14, Thm. 3.4] (where the hypothesis of no local obstructions is replaced by the weaker assumption that $h(T^2)$ is nonzero as a function on $k[u]/(\pi^2)$ for all prime $\pi \in k[u]$). The condition that $h(g^2)$ is separable is a congruence condition on $g \bmod M_h^{\mathrm{geom}}$. Since $\deg M_h^{\mathrm{geom}}$ is bounded in terms of $\deg_{u,T} h$, it follows that we may find such a g with any desired degree exceeding a lower bound determined by $\deg_{u,T} h$ (and not depending on k). Taking this universal bound large enough forces $\deg h(g^2) > 0$, so $(\partial_u h)(g^2) = h(g^2)'$ is nonzero and $R_k(h(g^2), h(g^2)')$ is nonzero.

Let H be a lift of h to $W[u][T] = W(k)[u][T]$ such that $\deg_T H = \deg_T h$ and $\mathrm{lead}_T(H) \in W[u]$ has the same u-degree as $\mathrm{lead}_T(h) \in k[u]$, so $\mathrm{lead}_T(H) \in W[u]$ has unit leading coefficient and reduces to $\mathrm{lead}_T(h) \in k[u]$. We also require $\deg_{u,T} H = \deg_{u,T} h$; this condition on the total degrees can certainly be satisfied, and its only purpose is to ensure that various largeness conditions below only depend on $\deg_{u,T} h$ and not on H.

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 a generic formula as in (4.4), and likewise for the degree of $H(G^2)$; this condition on $\deg g$ only depends on $\deg_{u,T} h$ (since $\deg_{u,T} H = \deg_{u,T} h$). 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 $k[u]/(h(g^2))$. Keeping in mind (2.5) for applications to finite k, we want to understand how the unit discriminant $\mathrm{disc}_W(H(G^2)) \bmod 8W$ depends on G.

Let F denote the fraction field of $W = W(k)$. Since $\mathrm{lead}_T H \in W[u]$ has leading coefficient in $W^\times$ and $h = H \bmod 2 \in k[u][T]$ is not in k and has no prime factors in $k[T]$ (as $h(T^2)$ is squarefree), we conclude that H is not in W and that H has no prime 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 $\partial_u H \neq 0$. The zero loci $Z_H = \{H = 0\}$ and $Z_{\partial_u H} = \{\partial_u H = 0\}$ in $\mathbf{A}_F^2$ have finite intersection by:

Lemma 5.1. *If K is perfect with arbitrary characteristic and $f \in K[u][T]$ is not in K, then the zero loci of f and $\partial_u f$ in $\mathbf{A}_K^2$ have finite intersection if and only if the following three conditions hold: f is squarefree in $K[u][T]$, f has no irreducible factors in $K[T]$, and the projection*

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

to the T-axis is generically étale on Z_f. When this happens, the non-étale locus of pr_T is finite and its image in the u-axis is the zero locus of M_f^{geom} in $\mathbf{A}_K^1$, with M_f^{geom} as in Definition 3.4.

The generically-étale property is always satisfied for squarefree nonzero $f \in K[u][T]$ in characteristic 0 since pr_T is *a priori* quasi-finite and flat. The case of 2-adic fields is of most interest for our present purposes.

Proof. The necessity of the conditions that f be squarefree and have no irreducible factors in $K[T]$ is clear. Granting these conditions, the plane curve Z_f is reduced (hence geometrically reduced since K 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 K, we thereby see that the non-étale locus for pr_T is where Z_f meets $Z_{\partial_u f}$ in $\mathbf{A}^2_K$. 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. □

We conclude that $R_{W[u]}(H, \partial_u H) \in W[u]$ is nonzero and that the monic squarefree polynomial $M_H^{\mathrm{geom}} \in F[u]$ can be formed as in Definition 3.4. The geometric roots of M_H^{geom} are the u-coordinates of the intersection points of Z_H and $Z_{\partial_u H}$ in $\mathbf{A}^2_F$. Since $\deg_{u,T} H = \deg_{u,T} h$, the degree of M_H^{geom} is bounded in terms of $\deg_{u,T} h$ (by Bézout's theorem over F).

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

$$R_F(H(G^2), H(G^2)'), \;\; R_F(H(G^2), (\partial_u H)(G^2)) \tag{5.1}$$

lie in W and have reductions in k that are both zero or both nonzero (see (3.2) and the Warning above (3.3)). Both reductions therefore lie in $k^\times$ since $h(g^2)$ is separable, so both terms in (5.1) lie in $W^\times$. When k is finite, the quadratic character of the first resultant in (5.1) is related to $\mathrm{disc}_W(H(G^2))$ and intervenes in the study of $\mu(h(g^2))$ (see (2.5)). The second resultant in (5.1), for finite k, is one to which Theorem 4.1 may be applied over the field F of characteristic zero since $Z_H \cap Z_{\partial_u H} \subseteq \mathbf{A}^2_F$ is finite. We are going to show that the ratio of the resultants in (5.1), which is in $W^\times$, can be made explicit in $(W/8W)^\times$ modulo unit-square factors, so we will be able to use Theorem 4.1 to study the quadratic character of $\mathrm{disc}_W(H(G^2))$. (A lot of algebraic calculations are coming up; a special case where the main ideas can be seen without complications is in [7].)

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 $\overline{F}$ and let $\overline{\alpha}$ denote the reduction of each such root 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 such α.

Since $H(G^2)' = (\partial_u H)(G^2) + 2(\partial_T H)(G^2)GG'$, the classical formula (3.1) for resultants in terms of products over geometric roots gives

$$\frac{R_F(H(G^2), H(G^2)')}{R_F(H(G^2), (\partial_u H)(G^2))} = \mathrm{lead}(H(G^2))^{d_G} \prod_\alpha \left(1 + 2 \cdot \frac{(\partial_T H)(G^2)GG'}{(\partial_u H)(G^2)}\bigg|_\alpha\right), \tag{5.2}$$

where $d_G := \deg(H(G^2)') - \deg((\partial_u H)(G^2))$.

Remark 5.2. 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

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

In either case, the largeness condition on $\deg g$ is determined by $\deg_{u,T} h$.

We want to understand the product in (5.2) modulo $8W$. The remarkable surprise is that there is a very simple formula for this product mod $8W$ (see (5.5)), and the formula only depends on g and h (not on G or H). We need to make two definitions before we can state the formula of interest.

Definition 5.3. For any perfect field K and any rational differential form ω on $\mathbf{P}^1_K$, set

$$s_2(\omega) := \sum_{\{y_1,y_2\}} \mathrm{Res}_{y_1}\omega \cdot \mathrm{Res}_{y_2}\omega \in K, \tag{5.3}$$

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

Our interest in $s_2(\omega)$ will be restricted largely to cases when ω has at worst simple poles. For ω varying with only simple poles, $s_2(\omega)$ is not algebraic in ω if we do not fix the number of simple geometric poles of ω. For example, if $\omega = b \cdot \mathrm{d}u/u + \mathrm{d}u/(u-a)$ with $b \neq 0, -1$, then $s_2(\omega) = -b(b+1) - 1$ if $a \neq 0$ and $s_2(\omega) = -(b+1)^2$ if $a = 0$.

Definition 5.4. For $g \in k[u]$, define $\omega_{h,g} = ((\partial_T h)(g^2)g/h(g^2))\mathrm{d}g$; the initial hypotheses on $h \in k[u][T]$ in this section ensure that $h(g^2) \neq 0$.

When g is a square in $k[u]$ or h is a polynomial in T^2, $\omega_{h,g}$ vanishes. For $g \in k[u]$ with large degree such that $h(g^2)$ is separable, the equation

$$\omega_{h,g} = \frac{(\partial_T h)(g^2)g^2}{h(g^2)} \cdot \frac{\mathrm{d}g}{g} \tag{5.4}$$

shows that this rational differential form on $\mathbf{P}^1_k$ has simple poles. We will see in Theorem 5.10 that $s_2(\omega_{h,g})$ intervenes in the behavior of $\mu(h(g^2))$ when k is finite. For finite k, the vanishing of $s_2(\omega_{h,g^2})$ will therefore make the behavior of $\mu(h(g^4))$ very accessible; this is $\mu(f(g))$ when f is a polynomial in T^4.

Theorem 5.5. *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 and $\deg_{u,T} H = \deg_{u,T} h$. For $g \in k[u]$ of large degree with $h(g^2)$ separable and $G \in W[u]$ lifting g with $\mathrm{lead}(G) \in W^\times$, the product in* (5.2) *satisfies*

$$\prod_\alpha \left(1 + 2 \cdot \frac{(\partial_T H)(G^2)GG'}{(\partial_u H)(G^2)}\bigg|_{u=\alpha}\right) \equiv 1 + 2\deg g \deg_T h + 4s_2(\omega_{h,g}) \bmod 8W, \tag{5.5}$$

where α runs over the geometric roots of $H(G^2)$. The largeness of $\deg g$ only depends on $\deg_{u,T} h$ and not on H or 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

$$\frac{(\partial_T H)(G^2)GG'}{(\partial_u H)(G^2)}\bigg|_{u=\alpha} = \mathrm{Res}_\alpha\left(\frac{(\partial_T H)(G^2)GG'}{(\partial_u H)(G^2)} \cdot \frac{\mathrm{d}P}{P}\right). \tag{5.6}$$

We will first show that

$$2\left.\frac{(\partial_T H)(G^2)GG'}{(\partial_u H)(G^2)}\right|_{u=\alpha} \equiv 2\mathrm{Res}_\alpha \omega_{H,G} + 4(\mathrm{Res}_\alpha \omega_{H,G})^2 \bmod 8\overline{W}, \tag{5.7}$$

where $\overline{W}$ is the integral closure of W in an algebraic closure $\overline{F}$ of F and we define $\omega_{H,G}$ by the formula (5.4) with H and G replacing h and g respectively. Note that we can replace the second residue in (5.7) 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, using $P = H(G^2)$,

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

so by (5.6) we conclude that in $\overline{W}/8\overline{W}$

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

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 (5.7).

Using (5.7), expanding the product on the left side of (5.5) modulo 8 gives the element

$$1 + 2\sum_\alpha \mathrm{Res}_\alpha \omega_{H,G} + 4\sum_{\alpha_1\neq\alpha_2}\mathrm{Res}_{\overline{\alpha}_1}\omega_{h,g}\mathrm{Res}_{\overline{\alpha}_2}\omega_{h,g} + 4\sum_\alpha \mathrm{Res}_{\overline{\alpha}}(\omega_{h,g})^2 \in W/8W, \tag{5.8}$$

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 (5.8) 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 $\deg_T H$.

The final sum in (5.8) lies in k, where it equals

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

by the residue theorem in characteristic 2. The second and third sums in (5.8) therefore combine to give $4s_2(\omega_{h,g})$ in (5.5). □

By (3.4), (5.2), and Theorem 5.5, if $\deg g$ is sufficiently large (in a manner determined by $\deg_{u,T} h$) and $h(g^2)$ is *separable* then the discriminant $\mathrm{disc}_W(H(G^2))$ is congruent modulo $8W$ to

(5.9)
$$\frac{(-1)^{\delta_g(\delta_g-1)/2}}{(\mathrm{lead}\, H(G^2))^{2\delta_g-1-d_G}} R_W(H(G^2),(\partial_u H)(G^2))(1+2\deg g \deg_T h + 4s_2(\omega_{h,g})),$$

where

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

and d_G is given by Remark 5.2; the exponent $2\delta_g - 1 - d_G$ of $\mathrm{lead}\, H(G^2)$ in (5.9) is linear in $\deg g = \deg G$ when $\deg g$ is large (depending only on $\deg_{u,T} h$). Since $-4 \equiv 4 \bmod 8$, $\mathrm{disc}_W(H(G^2)) \bmod 8W$ is equal to

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

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

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

and (by checking cases for ab modulo 4)

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

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

(5.10)
$$\frac{R_W(H(G^2),(\partial_u H)(G^2))}{(\mathrm{lead}\, H(G^2))^{2\delta_g-1-d_G}}(-1)^{\deg(\mathrm{lead}_T h)(\deg(\mathrm{lead}_T h)-1)/2}(1+4(m_g+s_2(\omega_{h,g}))),$$

where $m_g = \lfloor(1+(\deg g)(\deg_T h))/2\rfloor$ and $\deg g \gg 0$ (depending only on $\deg_{u,T} h$).

If we had instead chosen g of large degree as above such that $h(g^2)$ is not separable and $G \in W[u]$ is a lift of g with $\mathrm{lead}(G) \in W^\times$, then since $H(G^2)$ has the same degree as its reduction $h(g^2)$ we see via (3.2) 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 $\mathrm{disc}_W(H(G^2))$ may not be congruent modulo 8 to (5.10) when $h(g^2)$ is not separable, the expression (5.10) *always makes sense* in W and is a non-unit precisely when $\mathrm{disc}_W(H(G^2))$ is a non-unit. We can therefore use the resultant $R_W(H(G^2),(\partial_u H)(G^2))$ from characteristic 0 to study $\mathrm{disc}_W(H(G^2)) \bmod 8W$ even though usually $(\partial_u H)(G^2) \neq H(G^2)'$ in $W[u]$.

Since $R_{F[u]}(H,\partial_u H) \in F[u]$ is nonzero, by Theorem 4.1 for large n we obtain an identity of algebraic functions on $\mathrm{Poly}_{n/F}$,

$$R_F(H(G),(\partial_u H)(G)) = \beta_0\beta_1^n \cdot \mathrm{lead}(G)^{m_0+m_1 n} \cdot \prod_x P_{x,n}(G)^{e_x}, \tag{5.11}$$

with $m_0, m_1 \in \mathbf{Z}$ and $\beta_0, \beta_1 \in F^\times$ independent of n, the product taken over the set of $x = (u_x, t_x) \in Z_H \cap Z_{\partial_u H} \subseteq \mathbf{A}^2_F$, $P_{x,n}(G) = \mathrm{N}_{F(x)/F}(G(u_x) - t_x)$, and $e_x = i_x(Z_H, Z_{\partial_u H})$. The parameters in (5.11) may depend on the fixed choice of H lifting h (subject to the conditions $\deg_T H = \deg_T h$, $\deg_u(\mathrm{lead}_T(H)) = \deg_u(\mathrm{lead}_T(h))$, and $\deg_{u,T} H = \deg_{u,T} h$). When $G \in W[u]$, the left side of (5.11) is a resultant over W. We now show that the identity (5.11) over F can be factored in a manner that is well-behaved with respect to W.

Lemma 5.6. *For large n (depending only on $\deg_{u,T} h$ and not k or H), the algebraic maps*

$$\beta_0 \cdot \prod_{|u_x| \le 1, |t_x| > 1} P_{x,n}^{e_x}, \quad \beta_1^n \cdot \prod_{|u_x| > 1} P_{x,n}^{e_x} : \mathrm{Poly}_{\le n/F} \to \mathbf{A}^1_F \tag{5.12}$$

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

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

$$P_{x,n}(G) = \mathrm{N}_{F(x)/F}(G(u_x) - t_x) = \mathrm{N}_{F(x)/F}(t_x) \cdot \mathrm{N}_{F(x)/F}(t_x^{-1} G(u_x) - 1) \tag{5.13}$$

as algebraic functions of $G \in \mathrm{Poly}_{\le n/F}$. 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

$$P_{x,n}(G) = \mathrm{N}_{F(x)/F}(G(u_x) - t_x) = \mathrm{N}_{F(x)/F}(u_x)^n \cdot \mathrm{N}_{F(x)/F}(G^*(1/u_x) - u_x^{-n} t_x) \tag{5.14}$$

and $|u_x^{-n} t_x| < 1$ for large n. Note the nth power of $\mathrm{N}_{F(x)/F}(u_x)$ in (5.14). Since $H(u_x, t_x) = 0$ and $\mathrm{lead}_T H \in W[u]$ has unit leading coefficient, for $|u_x| > 1$ the largeness on n that is required to force $|u_x^{-n} t_x| < 1$ is determined by $\deg_{u,T} H = \deg_{u,T} h$ (use an integrality argument).

To see that (5.12) extends over W it suffices to show that the elements

$$b_0 := \beta_0 \cdot \prod_{|u_x| \le 1, |t_x| > 1} \mathrm{N}_{F(x)/F}(t_x)^{e_x}, \quad b_1 := \beta_1 \cdot \prod_{|u_x| > 1} \mathrm{N}_{F(x)/F}(u_x)^{e_x} \tag{5.15}$$

in F are integral. We shall prove these are in fact in $W^\times$ (this claim has nothing to do with n), so for large n determined by $\deg_{u,T} h$ the first map in (5.12) extends over W and has constant reduction $\overline{b}_0 \in k^\times$. Likewise, for the same large n the second map in (5.12) 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 $g = \sum_{i \le n} a_i(g) u^i$, since for $G \in \mathrm{Poly}_{\le n/F}(\overline{F}) = \overline{F}^{n+1}$ with coefficients in $\overline{W}$ the value $G^*(1/u_x)$ has the same reduction as $G^*(0) = a_n(G)$ when $|u_x| > 1$.

We may assume k is algebraically closed, so k is infinite. Thus, as we have seen in the beginning of this section, for $n > \max(\nu(h), \nu(\partial_u h))$ there exists $g_n \in k[u]$ of degree n such that $R_k(h(g_n), (\partial_u h)(g_n)) \ne 0$. For $G_n \in W[u]$ lifting any such g_n with $\mathrm{lead}(G_n) \in W^\times$, the W-resultant of $H(G_n)$ and $(\partial_u H)(G_n)$ is a unit in W. Thus, the left side of (5.11) is a unit in W when evaluated at G_n. Now consider the right side of (5.11) 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 (5.13) and (5.14), we thereby obtain that the product

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

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

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 Theorem 4.5 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 (5.11) with G replaced by G^2, where G has unit leading coefficient, we will be able to ignore the x's with $|u_x| > 1$ due to:

Theorem 5.7. *For $G \in W[u]$ with $\mathrm{lead}(G) \in W^\times$ and $n = \deg G$ large (determined by $\deg_{u,T} h$), $\beta_1^{2n} \cdot \prod_{|u_x|>1} P_{x,2n}(G^2)^{e_x} \in (W^\times)^2$.*

Proof. By Lemma 5.6, 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 reduces us to proving

$$\prod_{|u_x|>1} \mathrm{N}_{F(x)/F}(G^*(1/u_x)^2 - u_x^{-2n} t_x)^{e_x} \in (W^\times)^2, \tag{5.16}$$

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 $\mathrm{lead}(g)^2 \ne 0$. 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 of $F(x)$. Hence, depending only on the absolute ramification degree of $F(x)$ (bounded by $[F(x) : F]$) and not on k, we can make n large enough such that $G^*(1/u_x)^2 - u_x^{-2n} t_x$ is a square in the integral units of $F(x)$. Since

$$[F(x) : F] \le \deg_F(Z_H \cap Z_{\partial_u H}) \le \deg_{u,T} H \cdot \deg_{u,T} \partial_u H$$

and $\deg_{u,T} H = \deg_{u,T} h$, the largeness condition on n only depends on $\deg_{u,T} h$. Passing to an n that is uniformly large for all the finitely many x's such that $|u_x| > 1$, the norm-product (5.16) is a unit square in W. □

To emphasize that the unit b_0 in W from (5.15) depends on H, we now rename it: define

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

so b_H depends on H since the algebraic factorization on the right side of (5.11) depends on H. Also define

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

Using Lemma 5.6 and Theorem 5.7, 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 (5.11) implies

$$R_H(G) \in b_H \cdot \prod_{|u_x|\le 1, |t_x|>1} \mathrm{N}_{W(x)/W}(t_x^{-1} G(u_x)^2 - 1)^{e_x} \cdot \prod_{|u_x|,|t_x|\le 1} P_{x,2n}(G^2)^{e_x} \cdot (W^\times)^2 \tag{5.17}$$

when the reduction g of G has degree that is sufficiently large (depending on $\deg_{u,T} h$).

Since $b_H \in W^\times$ and all terms in the products in (5.17) are integral, the resultant $R_H(G)$ is a unit in W if and only if each of the terms in the products in (5.17) is a unit, in which case the image of $R_H(G)$ in $W^\times/(W^\times)^2$ is represented by the expression in (5.17).

Define

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

where $e_H = 1$ if $\mathrm{lead}_T H \notin W^\times$ and $e_H = \deg(\mathrm{lead}_T \partial_u H)$ if $\mathrm{lead}_T H \in W^\times$; $\widetilde{b}_H$ absorbs both the constant sign-factor and (by Remark 5.2) the odd-exponent power of the unit $\mathrm{lead}(H(G^2))$ in (5.10) modulo $(W^\times)^2$. Choose $g \in k[u]$ with large degree as required in (5.10) and (5.17), and choose $G \in W[u]$ lifting g with $\deg G = \deg g$. When $h(g^2)$ is *separable* it follows from (5.10) that $\mathrm{disc}_W(H(G^2)) \in W^\times$ is a unit-square multiple of the integral product

$$\widetilde{b}_H \cdot (1 + 4(m_g + \widetilde{s}_2(\omega_{h,g}))) \cdot \prod_{|u_x|\le 1, |t_x|>1} \mathrm{N}_{W(x)/W}(t_x^{-1} G(u_x)^2 - 1)^{e_x} \cdot \prod_{|u_x|,|t_x|\le 1} P_{x,2n}(G^2)^{e_x}, \tag{5.18}$$

with $m_g = \lfloor (1 + \deg g \deg_T h)/2 \rfloor$ and $\widetilde{s}_2(\omega_{h,g})$ denoting any lift of $s_2(\omega_{h,g})$ from k to W (see (5.3)). On the other hand, if $h(g^2)$ is not separable, then (5.17) implies that one of the terms $P_{x,2n}(G^2)$ with $|u_x|, |t_x| \le 1$ is in the maximal ideal of W, so (5.18) is also in the maximal ideal of W in such cases.

Motivated by (5.18), consider the W-scheme map $L_{H,n} : \mathrm{Poly}_{\le n/W} \to \mathbf{A}^1_W$ defined by

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

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

$$\widetilde{\rho}_{n,H} := \rho_{n,(M_H^{\mathrm{geom}})^{\le 1}} : \mathrm{Poly}_{\le n/W} \to W[u]/((M_H^{\mathrm{geom}})^{\le 1}) \tag{5.19}$$

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

$$(M_H^{\mathrm{geom}})^{\le 1} := \prod_{|u_x|\le 1} (u - u_x) \in W[u], \tag{5.20}$$

which is separable over F. Here we are viewing $W[u]/((M_H^{\mathrm{geom}})^{\le 1})$ 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

$$L_{H,n} = L_H \circ \widetilde{\rho}_{n,H} \tag{5.21}$$

for a unique W-scheme map $L_H : W[u]/((M_H^{\mathrm{geom}})^{\le 1}) \to \mathbf{A}^1_W$ that is independent of n.

Summarizing the conclusions of the above efforts, for any $g \in k[u]$ with large degree (determined by $\deg_{u,T} h$) and any $G \in W[u]$ lifting g with $\deg G = \deg g$, we have

$$\begin{aligned}\operatorname{disc}_W(H(G^2)) \equiv (1 + 4(\lfloor (1 + \deg g \deg_T h)/2 \rfloor + s_2(\omega_{h,g}))) \\ \cdot L_H(\widetilde{\rho}_{n,H}(G)) \cdot (W^\times)^2 \bmod 8W\end{aligned} \tag{5.22}$$

when $h(g^2)$ is separable, and otherwise both sides lie in $2W/8W$.

When k is finite, we will use the quadratic character of (5.22) to investigate $\mu(h(g^2))$. Before passing to the case of finite k we study the relationship between $(M_H^{\mathrm{geom}})^{\leq 1}$ and M_h^{geom}. We may factor M_H^{geom} in $F[u]$ as a product of 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 $\overline{W}$ (see (5.20)) and $(M_H^{\mathrm{geom}})^{>1}$ contains the other roots. Each root of the squarefree polynomial $(M_H^{\mathrm{geom}})^{\leq 1}$ is integral over W and is a root of the resultant $R_{W[u]}(H, \partial_u H)$, so this resultant is divisible by $(M_H^{\mathrm{geom}})^{\leq 1}$ in $W[u]$.

Definition 5.8. The reduction of $(M_H^{\mathrm{geom}})^{\leq 1}$ is denoted by $\overline{M}_H^{\mathrm{geom}} \in k[u]$.

To compute $\overline{M}_H^{\mathrm{geom}}$, choose $c \in F^\times$ such that cM_H^{geom} is primitive in $W[u]$ and has its unit coefficient in highest degree equal to 1. The reduction of cM_H^{geom} modulo 2 is $\overline{M}_H^{\mathrm{geom}}$. By reduction of divisibility over W, $\overline{M}_H^{\mathrm{geom}}$ divides $R_{k[u]}(h, \partial_u h)$ in $k[u]$; the polynomial $\overline{M}_H^{\mathrm{geom}}$ need not be squarefree (see Example 5.15).

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

Lemma 5.9. *For all lifts $H \in W[u][T]$ of $h \in k[u][T]$ such that $\deg_T H = \deg_T h$, $\deg_{u,T} H = \deg_{u,T} h$, and $\mathrm{lead}_T(H) \in W[u]$ has the same u-degree as $\mathrm{lead}_T(h) \in k[u]$, we have $M_h^{\mathrm{geom}} | \overline{M}_H^{\mathrm{geom}}$ in $k[u]$. In particular, for $\deg g$ sufficiently large (depending only on $\deg_{u,T} h$ and not on H or k), 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 4.7(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 3.5(2)). Thus the k-algebra $k[u][T]/(h, \partial_u h)$ is finite, $\partial_u H$ is nowhere a zero divisor on $\operatorname{Spec} W[u][T]/(H)$ at points over the closed point of $\operatorname{Spec} W$, and the W-scheme $\operatorname{Spec} W[u][T]/(H, \partial_u H)$ is flat at points over the closed point of $\operatorname{Spec} W$. (We used the local flatness criterion for the second and third assertions.) On the generic fiber over $\operatorname{Spec} F$ (F is the fraction field of W), $F[u][T]/(H, \partial_u H)$ is a finite (flat) F-algebra since $\{H = 0\}$ meets $\{\partial_u H = 0\}$ at only finitely many points in $\mathbf{A}^2_F$ (Lemma 5.1). 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 [11, 18.5.11], $W[u][T]/(H, \partial_u H)$ equals $R^{\mathrm{f}} \times R'$, 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

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

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 $\mathbf{A}^1_W$ over $\operatorname{Spec} W$ are the roots of M_h^{geom} and M_H^{geom} respectively. Thus, each root of M_h^{geom} is the reduction of an integral root of M_H^{geom} because each geometric closed point of a finite flat W-scheme (specifically, $\operatorname{Spec} R^{\mathrm{f}}$) is the specialization of an integral generic-fiber geometric point.

To prove that M_h^{geom} is the radical of $\overline{M}_H^{\mathrm{geom}}$ when $\operatorname{lead}_T h$ is separable, we check that if (c,t) is a geometric point in the common zero locus of H and $\partial_u H$, and 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 $(\operatorname{lead}_T h)(c) = 0$ then we want $(\operatorname{lead}_T h)'(c) \neq 0$. Since $\operatorname{lead}_T h$ is separable, we are done. □

Now let $g \in k[u]$ be arbitrary with large degree as in Lemma 5.9. The property of $h(g^2)$ being 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 polynomial $\overline{M}_H^{\mathrm{geom}}$ constructed by reduction from characteristic 0 controls the separability of $h(g^2)$ in characteristic 2 when $\deg g$ is as large as required in (5.10) and (5.17).

For the rest of this section we specialize to the case of a *finite* field $k = \kappa$ of characteristic 2. We want to study Möbius behavior in $\kappa[u]$. Pick a nonconstant $f \in \kappa[u][T^2]$ which is squarefree in $\kappa[u][T]$ and has no local obstructions. Write $f(T) = h(T^2)$. Choose a lift H of h as in Lemma 5.9, and pick $g \in \kappa[u]$ of large degree (depending only on $\deg_{u,T} h$) as in Lemma 5.9. Finally, 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 $\operatorname{disc}_W(H(G^2))$ is a unit precisely when $h(g^2)$ is separable. Recall (as above Theorem 2.4) that if $\operatorname{disc}_W(H(G^2)) \in W^\times$ 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 2.4 we have

$$\mu(h(g^2)) = (-1)^{\deg(\operatorname{lead}_T h)} \widetilde{\chi}(\operatorname{disc}_W(H(G^2))), \tag{5.23}$$

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

$$\widetilde{\chi}(c \cdot (1+4w)) = (-1)^{\operatorname{Tr}_{\kappa/\mathbf{F}_2}(w \bmod 2)}. \tag{5.24}$$

We can now prove an analogue of (4.14) in characteristic 2:

Theorem 5.10. *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 $\kappa[u][T]$. Also assume that $h(T)$ has no local obstructions. Fix $H \in W[u][T]$ lifting h such that $\deg_T(H) = \deg_T(h)$, $\operatorname{lead}_T(H) \in W[u]$ has unit leading coefficient (so $\deg_u(\operatorname{lead}_T(H)) = \deg_u(\operatorname{lead}_T(h))$), and $\deg_{u,T} H = \deg_{u,T} h$.*

For g of sufficiently large degree n,

$$\mu(h(g^2)) = (-1)^{\deg \operatorname{lead}_T(h) + [\kappa:\mathbf{F}_2]\lfloor (1+n\deg_T h)/2 \rfloor + \operatorname{Tr}_{\kappa/\mathbf{F}_2}(s_2(\omega_{h,g}))} \cdot \widetilde{\chi}(L_H(\widetilde{\rho}_{n,H}(G))), \tag{5.25}$$

where $G \in W[u]$ is any lift of g with degree n. Here, $s_2(\omega_{h,g})$ is defined by (5.3)*, $\widetilde{\rho}_{n,H}$ is defined by* (5.19)*, and L_H is defined in* (5.21)*. The "sufficient largeness" for* $\deg g$ *only depends on* $\deg_{u,T} h$ *and not on κ or H.*

In particular, if g_1 and $g_2 \in \kappa[u]$ have sufficiently large degrees (depending only on $\deg_{u,T} h$*),* $\deg g_1 \equiv \deg g_2 \bmod 4$*, and $g_1 \equiv g_2 \bmod \overline{M}_H^{\mathrm{geom}}$, then*

$$(-1)^{\mathrm{Tr}_{\kappa/\mathbf{F}_2}(s_2(\omega_{h,g_1}))}\mu(h(g_1^2)) = (-1)^{\mathrm{Tr}_{\kappa/\mathbf{F}_2}(s_2(\omega_{h,g_2}))}\mu(h(g_2^2)). \tag{5.26}$$

If $\deg_T h$ *is even then the congruence on* $\deg g_j$*'s for* (5.26) *to hold 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 for* (5.26) *to hold.*

Proof. By the above calculations, $L_H(\widetilde{\rho}_{n,H}(G)) \in W$ 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 (5.25) for $\mu(h(g^2))$ makes sense and is immediate from (5.23), (5.24), and (5.22). 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 polynomial $(M_H^{\mathrm{geom}})^{\leq 1}$ may be respectively lifted to $G_1, G_2 \in W[u]$ with unit leading coefficients such that $G_1 \equiv G_2 \bmod (M_H^{\mathrm{geom}})^{\leq 1}$ (so $\widetilde{\rho}_{n_1,H}(G_1) = \widetilde{\rho}_{n_2,H}(G_2)$ with $n_j = \deg G_j = \deg g_j$), the identity (5.26) follows from the congruence conditions on the g_j's and $\deg g_j$'s using (5.25). □

Since $s_2(\omega_{h,g^2}) = 0$ we get

Corollary 5.11. *Let κ be finite of characteristic* 2*, and $h \in \kappa[u][T]$ be such that $h \notin \kappa$, h has no local obstructions, and $h(T^2)$ is squarefree in $\kappa[u][T]$. Fix $H \in W[u][T]$ lifting h as in Theorem* 5.10*.*

For g of sufficiently large degree n (depending only on $\deg_{u,T} h$*),*

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

where $G \in W[u]$ is any lift of g with degree n. In particular, for g_1 and g_2 in $\kappa[u]$ of such sufficiently large degrees in $\kappa[u]$,

$$g_1 \equiv g_2 \bmod \overline{M}_H^{\mathrm{geom}},\ \deg g_1 \equiv \deg g_2 \bmod 2 \Longrightarrow \mu(h(g_1^4)) = \mu(h(g_2^4)). \tag{5.27}$$

There is no dependence on $\deg g \bmod 2$ *if* $[\kappa : \mathbf{F}_2]$ *or* $\deg_T h$ *is even.*

For applications it is convenient to restate the final part of Corollary 5.11 as follows (a characteristic 2 analogue of Theorem 4.8):

Theorem 5.12. *Let κ be a finite field with characteristic* 2*. Fix $f(T) \in \kappa[u][T^4]$ that is squarefree in $\kappa[u][T]$ and assume $f \notin \kappa$. There is a nonzero polynomial $M = M_{f,\kappa}$ in $\kappa[u]$ such that for all g_1 and g_2 in $\kappa[u]$ with sufficiently large degrees n_1 and n_2,*

$$g_1 \equiv g_2 \bmod M,\quad n_1 \equiv n_2 \bmod 2 \Longrightarrow \mu(f(g_1)) = \mu(f(g_2)). \tag{5.28}$$

If $[\kappa : \mathbf{F}_2]$ is even or $\deg_T f \equiv 0 \bmod 8$*, then there is no dependence on* $\deg g \bmod 2$*.*

The modulus M may be chosen to be $\overline{M}_H^{\mathrm{geom}}$ where $f(T) = h(T^4)$ and H is any lift of h as in Theorem 5.10*.*

Example 5.15 below suggests that in characteristic 2 the modulus $M_{f,\kappa}^{\min}$ of minimal degree in Theorem 5.12 (which must divide any modulus in Theorem 5.12) is not always squarefree, which is in contrast with Theorem 4.8. However, it divides the reduction of a squarefree polynomial from characteristic 0.

Example 5.13. 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 Corollary 5.11 says that $\mu(f(g)) = \mu(h(g^4))$ only depends on $\deg g \bmod 2$ when $\deg g \gg 0$. In [7] there is a more precise study of $\mu(g^2 + u)$ which, replacing g with g^2, implies that $\mu(g^4 + u) = (-1)^{[\kappa:\mathbf{F}_2]\deg g}$ when $\deg g \geq 1$. It follows that $g^4 + u$ is never prime when $[\kappa : \mathbf{F}_2]$ is even and g is nonconstant or when $[\kappa : \mathbf{F}_2]$ is odd and $\deg g$ is even. See [6, Table 1] for data on prime values of $T^4 + u$ over $\mathbf{F}_2[u]$.

Example 5.14. 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 that $M_H^{\mathrm{geom}} = 6u^5 + 2u^3 + 1$, so $\overline{M}_H^{\mathrm{geom}} = 1$ and $\deg_T H$ is even. Thus Corollary 5.11 says that $\mu(f(g))$ is constant for $\deg g \gg 0$. A closer analysis, carried out in [7], shows that $\mu(f(g)) = 1$ for $\deg g \geq 3$ and this is sharp: $\mu(f(g)) = -1$ for some g of degree 2.

Example 5.15. In $\kappa[u][T]$, let $f(T) = T^{16} + (u^9 + u^4 + u^2 + u)T^8 + u^5 + u^3$ and define the lift $H = T^4 + (u^9 + u^4 + u^2 + u)T^2 + u^5 + u^3 \in W[u][T]$. Then $\overline{M}_H^{\mathrm{geom}} = u^{24}(u + 1)^8$, so for g_1 and g_2 with sufficiently large degree,

$$g_1 \equiv g_2 \bmod u^{24}(u + 1)^8 \Longrightarrow \mu(f(g_1)) = \mu(f(g_2)).$$

Further work shows that the lower-degree modulus $u^9(u+1)^4$ works in (5.15) when g_1 and g_2 have degree at least 2. Numerical evidence suggests that when $\kappa = \mathbf{F}_2$ we can even use $u^3(u+1)$ instead of $u^9(u+1)^4$ as a modulus (but not any proper factor of $u^3(u+1)$). The proof of Corollary 5.11 shows that in principle the determination of the minimal degree monic modulus in characteristic 2 is a finite calculation, but we have not carried it out explicitly in this example.

6. A revised characteristic-p conjecture

We want to describe applications of the preceding work in this paper to the formulation of a modified conjectural asymptotic estimate on primality statistics of prime polynomials in $\kappa[u][T]$. Our new correction factor makes sense on its own for squarefree polynomials, not just prime polynomials, so we define it in that setting:

Definition 6.1. Let κ be a finite field. Let $f \in \kappa[u][T]$ be squarefree. Assume $\deg_T f > 0$ and f has no local obstructions. For any nonzero $M \in \kappa[u]$ and $n \geq 0$, define

$$\Lambda_{\kappa,M}(f; n) := 1 - \frac{\sum_{\deg g = n, (f(g), M) = 1} \mu(f(g))}{\sum_{\deg g = n, (f(g), M) = 1} |\mu(f(g))|}. \tag{6.1}$$

By work of Poonen [14, Thm. 3.1] applied to $M \cdot f$, the denominator in (6.1) is positive for large n. Note that $\Lambda_{\kappa,M}(f; n)$ is a rational number in the interval $[0, 2]$. The closer $\Lambda_{\kappa,M}(f; n)$ 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 such that $(f(g), M) = 1$.

The particular case of most interest to us is when $f(T)$ is a polynomial in T^p for $p \neq 2$ or a polynomial in T^4 for $p = 2$. In this case $\mu(f(g))$ is periodic in g (Theorems 4.8 and 5.12) and there is a monic modulus of periodicity $M_{f,\kappa}^{\min}$ having least degree. For this particular modulus we set

$$\Lambda_\kappa(f; n) := \Lambda_{\kappa, M_{f,\kappa}^{\min}}(f; n). \tag{6.2}$$

(We will see in Example 6.10 that $\Lambda_\kappa(f;n)$ is not always the same as $\Lambda_{\kappa,1}(f;n)$, although it is in Examples 1.1 and 2.5.) For any two choices of modulus M_1 and M_2 as in Theorem 4.8 or Theorem 5.12, Theorem 6.5 below ensures that the corresponding sequences $\Lambda_{\kappa,M_1}(f;n)$ and $\Lambda_{\kappa,M_2}(f;n)$ agree for all $n \geq n_0$, where n_0 depends only on $\deg M_1$, $\deg M_2$, and $\deg_{u,T} f$. This provides a robustness that makes Definition 6.1 less sensitive to change in M than it may initially seem to be when $f(T)$ is a polynomial in T^p or T^4. (For example, if $p \neq 2$ then $\Lambda_\kappa(f;n)$ equals $\Lambda_{\kappa,M_f^{\mathrm{geom}}}(f;n)$ for large n depending only on $\deg_{u,T} f$.) The marvelous fact (Theorem 6.5) is that in this case $\Lambda_\kappa(f;n)$ is *periodic* in n with period 1, 2, or 4 for all sufficiently large n (determined by $\deg_{u,T} f$), so $\Lambda_\kappa(f;n)$ is far simpler than it at first appears to be. This periodicity makes the following conjecture on primality statistics simple to appreciate.

Conjecture 6.2. *Let κ be a finite field and let $f \in \kappa[u][T]$ be prime with positive T-degree and no local obstructions. Define $C(f) = \log q \cdot \prod_{v \neq \infty}(1 - \omega_f(v)/q_v)/(1-1/q_v)$, where $q = \#\kappa$, the product runs over all the places of $\kappa(u)$ except ∞, q_v is the size of the residue field at v, and $\omega_f(v)$ denotes the number of roots of f in the residue field at v. If $f \in \kappa[u][T^p]$ for $p \neq 2$ and if $f \in \kappa[u][T^4]$ for $p = 2$, then as $n \to \infty$,*

$$\#\{g \in \kappa[u] : \deg g = n, f(g) \text{ is prime}\} \overset{?}{\sim} \Lambda_\kappa(f;n)\frac{C(f)}{\deg_T f}\frac{(q-1)q^n}{\log(q^n)}. \tag{6.3}$$

The product $C(f)$ is defined according to increasing values of q_v, with all $v \neq \infty$ having a common q_v-value introduced into the product at the same time. (It is convergent because f has no local obstructions, and usually is not absolutely convergent.)

Example 6.10 will show that in (6.3) it is crucial to impose the relative primality condition $(f(g), M_{f,\kappa}^{\min}) = 1$ on the averaging process in the definition of $\Lambda_\kappa(f;n)$. If 0 is in the period of the sequence $\Lambda_\kappa(f;n)$ (*e.g.*, Example 6.6 below), then for each g of degree n where $\Lambda_\kappa(f;n) = 0$ either $\mu(f(g)) = 1$ or $(f(g), M_{f,\kappa}^{\min}) \neq 1$, so $f(g)$ is composite when n is large (in particular, we at least require $n > \deg M_{f,\kappa}^{\min}$). Thus, the appearance of 0 in the period for $\Lambda_\kappa(f;n)$ implies that the left side of (6.3) is 0 for such n. When n runs through a sequence in which $\Lambda_\kappa(f;n)$ does not vanish, we have not proved a connection between $\Lambda_\kappa(f;n)$ and primality counts for $f(g)$ with $\deg g = n$ as $n \to \infty$, but (6.3) agrees well with the extensive numerical testing that we have carried out.

Let us give a brief heuristic justification for including the relative primality condition $(f(g), M) = 1$ in the definition of $\Lambda_\kappa(f;n)$ for large n with M a nonzero multiple of $M_{f,\kappa}^{\min}$. When making asymptotic primality predictions based on a randomness model, it is reasonable to expect that imposing a local condition such as $(f(g), M) = 1$ on the sample space of the g's (with a fixed nonzero M) should not influence the statistics. But our theory and examples of Möbius periodicity show that for a nonzero $M \in \kappa[u]$ the formation of Möbius averages as in the definition of $\Lambda_{\kappa,M}(f;n)$ for large n can be sensitive to the particular choice of M, except when M is a multiple of $M_{f,\kappa}^{\min}$. Hence, it is natural to require M to be sufficiently divisible in this sense when we define $\Lambda_\kappa(f;n)$.

In Theorem 5.12 we gave a Möbius periodicity result for polynomials in T^4, but we did not address the possibility that there may also be periodicity for other polynomials in T^2 in characteristic 2. Numerical evidence suggests that Möbius

periodicity can arise in many such cases (including the possibility of periodicity depending on $\deg g \bmod 4$ rather than just on $\deg g \bmod 2$ as in Theorem 5.12), but the general picture appears to be complicated. For example, $T^2 + u$ over $\mathbf{F}_2[u]$ does not appear to exhibit non-trivial Möbius periodicities, though it also numerically fits the naive conjecture for predicting primality statistics (of $g^2 + u$ with $g \in \mathbf{F}_2[u]$ of large degree) and therefore a non-trivial correction factor in this case (as in Conjecture 6.2) is apparently not necessary.

Remark 6.3. Let us make some observations concerning omitted cases in Conjecture 6.2. If Theorem 5.12 can be generalized to allow $f \in \kappa[u][T^2]$ then it should be possible to formulate a version of (6.3) in characteristic 2 without the restriction that $f \in \kappa[u][T^4]$. Now suppose f is separable and irreducible over $\kappa(u)$ without local obstructions. In this case we expect $f(g)$ to be prime as often as analogies between $\mathbf{Z}$ and $\kappa[u]$ predict, and in the context of Conjecture 6.2 this corresponds to removing the factor $\Lambda_\kappa(f;n)$ from (6.3). Thus, let us formulate a conjecture for $p \neq 2$ that does not treat the separable and inseparable cases separately. Let $f(T)$ be any prime polynomial in $\kappa[u][T]$ with positive T-degree and no local obstructions. Consider the projection π from $Z_f = \{f = 0\}$ to the T-axis. The finite (possibly empty) set $B \subseteq Z_f$ of isolated points in the non-étale locus of π has finite image in the u-axis, so we may define M_f^{geom} in terms of B just as we can in the inseparable case. (This works even if Definition 3.4 does not apply to f.) Set $\Lambda_\kappa(f;n) = \Lambda_{\kappa,M_f^{\mathrm{geom}}}(f;n)$, which recovers (6.2) for large n when $f(T)$ is a polynomial in T^p (by Theorem 6.5). Equation (6.3) now makes sense as a conjecture in characteristic $p \neq 2$ even if f is not a polynomial in T^p, but nothing is proved about the behavior of $\Lambda_\kappa(f;n)$ when f is separable. (See [2] for some work in the case of separable f.) We expect that any reasonable averaging process for $\mu(f(g))$ should tend to 0 in the large-degree limit when f is separable over $\kappa(u)$, so we expect $\Lambda_\kappa(f;n) \to 1$ as $n \to \infty$ for separable f. This would make (6.3) for separable f equivalent to the conjecture based only on analogies between $\mathbf{Z}$ and $\kappa[u]$.

For inseparable f we have always been able to prove *a posteriori* that $\Lambda_\kappa(f;n) = 1$ for all large n in congruence classes $c \bmod 4$ for which the numerical data suggest that classical analogies are making correct asymptotic predictions when $n \equiv c \bmod 4$. In Examples 6.9 and 6.10 we will check that (6.3) appears to fix the discrepancies in Examples 1.1 and 1.2.

We turn now to justifying the properties of $\Lambda_{\kappa,M}(f;n)$ used above when f is inseparable in T. For $\Lambda_{\kappa,M}(f;n)$ to make sense, its denominator has to be positive: there needs to be a $g \in \kappa[u]$ of degree n such that $(f(g), M) = 1$ and $f(g)$ is squarefree. The following lemma addresses this problem and gives some information on the required largeness of n.

Lemma 6.4. *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 prime factors in $\kappa[u]$). For any nonzero $M \in \kappa[u]$, there exist polynomials $g \in \kappa[u]$ with any sufficiently large degree (depending only on $\deg M$ and $\deg_{u,T} f$) such that $(f(g), M) = 1$ and $f(g)$ is squarefree in $\kappa[u]$.*

Proof. By replacing M with its radical, we can assume M is squarefree. Hence, $f(g)$ is squarefree and relatively prime to M if and only if $Mf(g)$ is squarefree. Since $M \cdot f(T) \in \kappa[u][T]$ is squarefree (as f has no irreducible factors in $\kappa[u]$), the

absence of local obstructions allows us to apply [14, Thm. 3.4] to get the result except for determining how the required largeness of $\deg g$ depends on M and f.

By our work on Möbius periodicity (using extra care in characteristic 2), since f is a polynomial in T^p we know that if $\deg g$ is large enough (depending only on $\deg_{u,T} f$) then whether or not $f(g)$ is squarefree depends only on a congruence condition on g modulo a nonzero polynomial with degree bounded in terms of $\deg_{u,T} f$. The relative primality property $(f(g), M) = 1$ is also a congruence condition on $g \bmod M$, so overall the combined properties we are requiring for g are congruence conditions modulo a nonzero polynomial with degree bounded in terms of $\deg M$ and $\deg_{u,T} f$ when $\deg g$ is sufficiently large (depending only on $\deg_{u,T} f$). Since we have shown that this collection of congruence conditions in $\kappa[u]$ has a solution in some large degree, it has solutions in any degree exceeding a bound determined by $\deg M$ and $\deg_{u,T} f$. □

Our work in §2–§5 leads to the following important periodicity result.

Theorem 6.5. *Let κ be finite, and let $f(T)$ be as in Definition 6.1. Assume $f \in \kappa[u][T^p]$, and if $p = 2$ then assume $f \in \kappa[u][T^4]$. For any finite extension κ'/κ, the sequence $\Lambda_{\kappa'}(f; n)$ is periodic with period dividing 4 for $n \gg 0$, and this largeness for n only depends on $\deg_{u,T} f$ and not on κ'.*

For any nonzero M in $\kappa'[u]$ that is divisible by $M_{f,\kappa'}^{\min}$, $\Lambda_{\kappa'}(f; n) = \Lambda_{\kappa',M}(f; n)$ when n is sufficiently large. More precisely, for each $d \geq 1$ there is an n_0 depending only on $\deg_{u,T} f$ and d such that if $M \in \kappa'[u]$ is a nonzero multiple of $M_{f,\kappa'}^{\min}$ in $\kappa'[u]$ with $\deg M \leq d$, then $\Lambda_{\kappa'}(f; n) = \Lambda_{\kappa',M}(f; n)$ for all $n \geq n_0$.

Proof. This is a simple argument with quadratic character sums; in [9] the argument is given with $\kappa[u]$ replaced by the coordinate ring of any smooth affine κ-curve with one geometric point at infinity. The main requirement on the largeness of n is that it be large enough as in Theorem 4.8 (for $p \neq 2$), Theorem 5.12 (for $p = 2$), and Lemma 6.4. □

Periodicity of $\Lambda_{\kappa,1}(f; n)$ (omitting the relative primality condition on $f(g)$'s) follows by the same arguments as for $\Lambda_\kappa(f; n)$ in the proof of Theorem 6.5.

Example 6.6. Let $p \neq 2$ and $f(T) = T^p + u \in \kappa[u][T]$ (Example 2.5). Then $M_f^{\mathrm{geom}} = 1$, $\Lambda_\kappa(f; n) = 1$ for odd n, $\Lambda_\kappa(f; n) = 0$ for positive $n \equiv 0 \bmod 4$, and $\Lambda_\kappa(f; n) = 1 - \chi_\kappa(-1)$ if $n \equiv 2 \bmod 4$. In particular, $\Lambda_{\mathbf{F}_3}(T^3 + u; n)$ is $1, 2, 1, 0, 1, 2, 1, 0, \ldots$, $\Lambda_{\mathbf{F}_9}(T^3 + u; n)$ is $1, 0, 1, 0, 1, 0, 1, 0, \ldots$, and $\Lambda_{\mathbf{F}_5}(T^5 + u; n)$ is $1, 0, 1, 0, 1, 0, 1, 0, \ldots$ for $n \geq 1$. This is consistent with numerical data when (6.3) is tested for $T^p + u$ over $\mathbf{F}_3, \mathbf{F}_9$, and $\mathbf{F}_5$. In particular, $g^p + u$ is not prime for any g in degree n when $\Lambda_\kappa(T^p + u; n) = 0$.

Example 6.7. Let $p \neq 2$ and $f(T) = T^p + u^2 \in \kappa[u][T]$. Then $M_f^{\mathrm{geom}} = u$ and

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

where g has leading term cu^n and $n \geq 1$. In particular, $\Lambda_\kappa(T^p + u^2; n) = 1$ for $n \geq 1$; this is consistent with the numerical observation that if $\kappa = \mathbf{F}_3, \mathbf{F}_9, \mathbf{F}_5$, or $\mathbf{F}_7$, then asymptotically $g^p + u^2$ appears to be prime as often as classical analogies suggest it should be.

Example 6.8. When

$$f(T) = T^{12} + (2u^4 + 2u^3 + 2u^2 + u + 1)T^6 + 2u^3 + 2u^2 + u$$

in characteristic 3, $\Lambda_{\mathbf{F}_3}(f;n) = 2/3$ for $n \geq 3$. The numerical evidence for Conjecture 6.2 in this case looks good.

Example 6.9. Let $f(T) = T^{12}+(u+1)T^6+u^4$ be the polynomial from Example 1.1, considered over any finite field κ of characteristic 3; let $q = \#\kappa$. (The polynomial f is irreducible over $\mathbf{F}_3$ but it factors over $\mathbf{F}_9$.) We use the formula $\mu(f(g)) = \chi(g(0)^2(g(1)^2+1))$ from (3.6), and $M := M_f^{\mathrm{geom}} = u(u-1)$. We need to determine when $(f(g), u(u-1)) \neq 1$:

$$f(g)|_{u=0} = g(0)^6(g(0)^2+1)^3, \quad f(g)|_{u=1} = (g(1)^2+1)^6.$$

If $(u-1)|f(g)$, then $g(1)^2+1 = 0$, so $\mu(f(g)) = 0$. Therefore the condition that $(f(g), u-1) = 1$ in the sums in $\Lambda_\kappa(f;n)$ can be ignored. We have $u|f(g)$ if and only if $g(0) = 0$ or $g(0)^2+1 = 0$; if $g(0) = 0$ then $\mu(f(g)) = 0$. Thus, the condition $(f(g), u(u-1)) = 1$ in the sums of $\Lambda_\kappa(f;n)$ can be relaxed to $g(0)^2+1 \neq 0$. When $n \geq 2$, polynomials g of degree n in $\kappa[u]$ are equally spread out over the pairs $(g(0), g(1))$, so

$$\begin{aligned}\Lambda_\kappa(f;n) &= 1 - \frac{\sum_{a,b\in\kappa}\chi(a^2(b^2+1)) - \sum_{a^2+1\neq 0}\chi(a^2(b^2+1))}{\sum_{a,b\in\kappa}|\chi(a^2(b^2+1))| - \sum_{a^2+1\neq 0}|\chi(a^2(b^2+1))|} \\ &= \begin{cases} 1+1/(q-2), & \text{if } \chi(-1) = 1, \\ 1+1/q, & \text{if } \chi(-1) = -1, \end{cases}\end{aligned}$$

since $\sum_{b\in\kappa}\chi(b^2-d) = -1$ when $d \neq 0$. In particular, $\Lambda_{\mathbf{F}_3}(f;n) = 4/3$ for all $n \geq 2$. This agrees well with the data on primality statistics for $f(g)$ in [6, Table 3].

Example 6.10. Let $f(T)$ be the polynomial from Example 1.2, but viewed in $\kappa[u][T]$ for any κ of characteristic 3. We recall (3.9) from Example 3.3: when $g = cu^n + \cdots \in \kappa[u]$ with $n = \deg g \geq 1$,

$$\mu(f(g)) = (-1)^n(\chi(-1))^{n(n-1)/2}\chi(c)^{n+1}\chi(g(1)^2+g(1)+2)\chi(g(2)). \tag{6.4}$$

This formula implies that $M_{f,\kappa}^{\min} = (u-1)(u-2)$. 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

$$f(g)|_{u=1} = (g(1)-1)^3(g(1)^2+g(1)+2)^3, \quad f(g)|_{u=2} = (g(2))^6(g(2)+1)^3, \tag{6.5}$$

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 that $g(2)$ 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 (6.4) and (6.5) 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

$$0, 1, 2, 1, 0, 1, 2, 1, \ldots,$$

which is an excellent fit with the discrepancies between Example 1.2 and the conjectural asymptotic for primality statistics of $f(T)$ on $\mathbf{F}_3[u]$ based only on the analogy between $\mathbf{Z}$ and $\mathbf{F}_3[u]$. Taking $\kappa = \mathbf{F}_9$, $\Lambda_{\mathbf{F}_9}(f;n) = 1$ for even n and $\Lambda_{\mathbf{F}_9}(f;n) = 22/21$ for odd n; this fits the numerical evidence. Returning to $\kappa = \mathbf{F}_3$, 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 composite 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 $(f(g), M_{f,\kappa}^{\min}) = 1$ in the definition of $\{\Lambda_\kappa(f;n)\}_{n\geq 1}$, then this sequence would become constant: $\Lambda_{\kappa,1}(f;n) = 1$ for $n \geq 1$. In other words, the nonzero values of $\mu(f(g))$ for g of a fixed degree $n \geq 1$ are equally often 1 and -1, but what matters for the link to primality statistics appears to be the nonzero values of $\mu(f(g))$ *constrained by the additional local condition* $(f(g), M_{f,\kappa}^{\min}) = 1$. (We can replace $M_{f,\kappa}^{\min}$ with any nonzero multiple, such as M_f^{geom} when $p \neq 2$, by Theorem 6.5.)

A striking example of the distinction between $\Lambda_\kappa(f;n)$ and $\Lambda_{\kappa,1}(f;n)$ when used as correction factors in (6.3) is given by $f(T) = (u-1)T^{12} + u^2T^6 + u^3 - 1 \in \mathbf{F}_3[u][T]$: $\Lambda_{\mathbf{F}_3}(f;n) = 6/5 > 1$ for $n \geq 4$ whereas $\Lambda_{\mathbf{F}_3,1}(f;n) = 6/7 < 1$ for $n \geq 3$. Thus, for this f, if we replace $\Lambda_{\mathbf{F}_3}(f;n)$ with $\Lambda_{\mathbf{F}_3,1}(f;n)$ in (6.3), then we would be predicting a deficit of prime values for $f(T)$ in comparison to predictions based purely on analogies with the classical case (no Λ-factor), whereas numerical data and the actual prediction of Conjecture 6.2 suggest a surplus.

Remark 6.11. If we search for prime 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 periodic sequence (with mod 4 periodicity, *etc.*, by the same arguments), and it can differ from $\Lambda_\kappa(f;n)$. Numerical data support the use of this new sequence as correction factors in a "monic g sampling" version of Conjecture 6.2.

Fix κ with characteristic $p \neq 2$ and a polynomial $f \in \kappa[u][T^p]$ with $\deg_T f$ positive such that f is squarefree in $\kappa[u][T]$. Assume that f has no local obstructions on $\kappa[u]$. We now look at how $\Lambda_{\kappa,M}(f;n)$ varies when we let the constant field grow (initially keeping f fixed, but then we shall let f vary). Our constraints on f are inherited over any finite extension κ'/κ, and the absence of local obstructions automatically holds if we increase κ so that $\#\kappa > \deg_T f$. For $c \in \{0,1,2,3\}$, let $\lambda_{\kappa'}(f;c)$ be the common value of $\Lambda_{\kappa'}(f;n)$ for all large $n \equiv c \bmod 4$. (By Theorem 6.5, this largeness of n only depends on $\deg_{u,T} f$ and not on κ'/κ.) Theoretical considerations show that this function on $\mathbf{Z}/4\mathbf{Z}$ is usually very close to 1 when $[\kappa' : \kappa]$ is large enough:

Theorem 6.12. *For κ and f as above with $p \neq 2$ and a fixed $c \in \{0,1,2,3\}$, as $[\kappa' : \kappa] \to \infty$ the numbers $\lambda_{\kappa'}(f;c)$ either tend to 1 or else lie in the set $\{0,2\}$ for all κ'. In the latter case, $\lambda_{\kappa'}(f;c)$ only depends on the parity of $[\kappa' : \kappa]$. Moreover, for "generic" f the limiting value is 1 for each c.*

The meaning of genericity for f in Theorem 6.12 is that at least one of the local intersection numbers among the points in the intersection of the zero loci of

f and $\partial_u f$ is odd (and in particular, there *is* such an intersection point); we saw in Example 4.15 that if $p \neq 2$, then for most f (in a sense made precise with the Zariski topology) this property is satisfied.

Theorem 6.12 is the genus-0 case of a more general result that is proved in [9] (in the setting of arbitrary genus) by using the Lang–Weil estimate and a link between the Λ_κ's and zeta-functions. (In [9] we also prove an analogue of Theorem 6.12 for the case $p = 2$ and any genus.) To illustrate Theorem 6.12, let κ' run over finite fields of characteristic 3 and let $f(T)$ be the polynomial from Example 1.2. We saw in Example 6.10 that $\lambda_{\kappa'}(f; c)$ equals 1, $1 \pm 2/((q-1)(q-2))$, or $1+2/((q-2)(q-3))$, where q is the size of κ', so $\lambda_{\kappa'}(f; c) \to 1$ for each c as $[\kappa' : \mathbf{F}_3] \to \infty$. For any odd p and $f = \alpha T^p + (\beta u + \gamma)$ with $\alpha, \beta, \gamma \in \kappa$ and $\alpha, \beta \neq 0$, Example 2.5 gives $\lambda_{\kappa'}(f; c) = 1$ for odd c and $\lambda_{\kappa'}(f; c) = 1 - \chi_{\kappa'}(-1)^{c/2}$ for even c. Hence, values of 0 or 2 as in Theorem 6.12 do really occur.

What happens to $\lambda_{\kappa'}(f; c)$ if we vary f over large finite fields (with characteristic $p \neq 2$)? To be precise, fix a family $\{\sum \alpha_i T^{\mu_i} \mid \deg \alpha_i = d_i\}$ as at the end of Example 4.15 with common T-degree d and consider finite extensions $\kappa/\mathbf{F}_p$ with $\#\kappa > d$, so there are no local obstructions for κ-points f in the Zariski-dense open locus U satisfying the condition $\gcd_i(\alpha_i) = 1$. For $\varepsilon > 0$, does the proportion of $f \in U(\kappa)$ such that $|\lambda_\kappa(f; c) - 1| < \varepsilon$ for all $c \in \{0, 1, 2, 3\}$ (let us say that such an f is ε-*classical*) tend to 1 as $[\kappa : \mathbf{F}_p] \to \infty$? To rule out families such as $\{\alpha T^p + (\beta u + \gamma)\}$ that give a negative answer, Example 4.15 suggests requiring $d_i \geq 2$ for some i. The sufficiency of this condition is proved in [9] in the setting of odd characteristic and arbitrary genus g, with the lower bound of 2 on some d_i replaced with a lower bound of $2g + 2$ and with the locus of ε-classical f's even proved to contain all points in a Zariski-dense open subset $U^0 \subseteq U$ with values in any sufficiently large extension κ of $\mathbf{F}_p$ (where the largeness on $[\kappa : \mathbf{F}_p]$ depends on ε but U^0 does not). There we also prove that if moreover the T-degrees of f and $\partial_u f$ are odd for generic f in the family, then there exists a positive integer N and a Zariski-dense open subset $U' \subseteq U$ such that for $f \in U'(\kappa)$ the correction factor $\lambda_{\kappa'}(f; \cdot)$ is not identically 1 on $\mathbf{Z}/4\mathbf{Z}$ whenever $N|[\kappa' : \kappa]$, *except* for families of the form $\{\alpha_1 T^{p^r} + \alpha_0\}$ with $\deg \alpha_i \geq 2$ for some i (in which case the opposite extreme $\lambda_{\kappa'}(f; \cdot) = 1$ on $\mathbf{Z}/4\mathbf{Z}$ occurs whenever $N|[\kappa' : \kappa]$ and $f \in U(\kappa)$ lies in a suitable Zariski-dense open subset of U). In this sense, the theory of the correction factor $\lambda_\kappa(f; c)$ is often nontrivial.

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 version of this work. The first author was partially supported by the Alfred P. Sloan Foundation and NSF grant DMS-0093542 during work on this paper. The second author thanks the Clay Mathematics Institute and the Number Theory Foundation for support.

References

[1] Paul T. Bateman and Roger A. Horn, *Primes represented by irreducible polynomials in one variable*, Proc. Sympos. Pure Math., Vol. VIII, 119–132, Amer. Math. Soc., Providence, R.I., 1965, MR0176966 (31:1234).

[2] A. O. Bender and O. Wittenberg, *A potential analogue of Schinzel's hypothesis for polynomials with coefficients in* $\mathcal{F}_q[t]$, Int. Math. Res. Not., **36**, 2005, 2237–2248. MR2181456 (2006g:11230)
[3] E. R. Berlekamp, *An analog to the discriminant over fields of characteristic two*, J. Algebra, **38**, 1976, 2, 315–317, MR0404197 (53:8000).
[4] Pierre Berthelot and Arthur Ogus, *Notes on crystalline cohomology*, Princeton University Press, Princeton, 1978, vi+243, MR0491705 (58:10908).
[5] V. Bouniakowsky, *Sur les diviseurs numériques invariables des fonctions rationnelles entières*, Mémoires sc. math. et phys., **6**, 1854, 306–329.
[6] K. Conrad, *Irreducible values of polynomials: a non-analogy*, Number Fields and Function Fields – Two Parallel Worlds, 71–85, Progress in Mathematics, **239**, Birkhäuser, Basel, 2005. MR2176587 (2006i:11033)
[7] Brian Conrad and Keith Conrad, *The Möbius function and the residue theorem*, Journal of Number Theory, **110**, 2005, 22–36, issn=0022-314X, MR2114671 (2006b:11153).
[8] Brian Conrad and Keith Conrad, *Prime specialization in higher genus I*, In preparation.
[9] Brian Conrad, Keith Conrad and Robert Gross, *Prime specialization in higher genus II*, In preparation.
[10] W. Fulton, *Intersection theory*, Ergebnisse der Mathematik und ihrer Grenzgebiete, **2**, 2, Springer-Verlag, Berlin, 1997, xiv+470, MR732620 (85k:14004).
[11] Alexander Grothendieck *Éléments de géométrie algébrique* IV_4. *Étude locale des schémas et des morphismes de schémas*, Inst. Hautes Études Sci. Publ. Math., **32**, 1967, 361 pp., MR0238860 (39:220).
[12] G. H. Hardy and J. E. Littlewood, *Some problems of Partition Numerorum III: On the expression of a number as a sum of primes*, Acta Math., **44**, 1923, 1–70.
[13] Matsumura Hideyuki, *Commutative ring theory*, Cambridge Studies in Advanced Mathematics, **8**, 2, Cambridge Univ. Press, Cambridge, 1990, xiii+320, MR879273 (88h:13001).
[14] Bjorn Poonen, *Squarefree values of multivariable polynomials*, Duke Math. J., **118**, 2003, 2, 353–373, MR1980998 (2004d:11094).
[15] Richard G. Swan, *Factorization of polynomials over finite fields*, Pacific J. Math., **12**, 1962, 1099–1106, MR0144891 (26:2432).

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

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

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