IRREDUCIBLE SPECIALIZATION IN GENUS 0

BRIAN CONRAD, KEITH CONRAD, AND ROBERT GROSS

Abstract. For irreducible $f(T) \in \mathbf{Z}[T]$, a conjecture of Hardy–Littlewood predicts how often f has prime values. The asymptotic frequency of these prime values is believed to be controlled by local obstructions. We discuss an analogue of the Hardy–Littlewood conjecture for irreducible $f(T) \in \kappa[u][T]$, with κ a finite field. Here, local obstructions are not sufficient. When f is inseparable over $\kappa(u)$, there is a new obstruction that is global, and it is quantified and effectively computable through the average of the Möbius function on specializations of $f(T)$.

We build on some old results of Swan to prove the surprising fact that the "Möbius average" for $f(g)$'s with $g \in \kappa[u]$ of large degree n has periodic behavior in n when f is inseparable over $\kappa(u)$. The periodicity enables us to prove in specific examples that the naive version of the Hardy–Littlewood conjecture over $\kappa[u]$ is false. We use periodic Möbius average behavior to formulate a modified conjecture that agrees well with numerical data.

1. Introduction

A well-known conjecture going back to Bouniakowsky [5] says that a nonconstant irreducible polynomial in $\mathbf{Z}[T]$ has infinitely many prime values in $\mathbf{Z}$ unless there is a divisibility obstruction, meaning that all values of the polynomial on $\mathbf{Z}$ are divisible by a nontrivial common factor. For example, $3T^2 - T + 2$ is irreducible in $\mathbf{Z}[T]$ but $3n^2 - n + 2$ is always even (and thus hardly ever prime) for $n \in \mathbf{Z}$.

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

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

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

The relation (1.1) is a special case of the Hardy–Littlewood conjecture (also called the Bateman–Horn conjecture). The only proved case of (1.1) is in degree 1: the prime number theorem is the case $f(T) = T$ and Dirichlet's theorem is the case $f(T) = aT + b$ with a and b nonzero and relatively prime. The Hardy–Littlewood conjecture allows several polynomials, such as twin-prime pairs. No version of the conjecture for several polynomials has been proved, even qualitatively.

Date: 2004-05-05 Wed 15:20:15 EDT.

1991 *Mathematics Subject Classification.* 11N32.

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

In this paper, we discuss an analogue of the Hardy–Littlewood conjecture in $\kappa[u][T]$ with κ a finite field. An extension of this work, with $\kappa[u]$ replaced by the coordinate ring of any smooth affine curve over κ with one geometric point at infinity, will be the subject of [9]. The proofs in [9] do not supersede the material here, but rather will depend upon it.

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

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

as $n \to \infty$, where the constant $C(f) \neq 0$ is similar to the classical paradigm over $\mathbf{Z}$. (For the definition of $C(f)$, see (3.7) with $r = 1$.) Note that sampling in (1.2) is over all polynomials in $\kappa[u]$ of degree n, not just monics; this is why $q-1$ occurs in (1.2). Although it is traditional to believe that problems over $\mathbf{Z}$ become more accessible when they are reformulated over $\kappa[u]$, the only proved instance of (1.2) is $\deg_T f = 1$, just as in the classical situation. Counting g in (1.2) with $\deg g \leq n$ (or $(\deg g)|n$) instead of $\deg g = n$ does not simplify matters, and in fact we shall see that counting by separate degrees is essential for a proper understanding of the situation.

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

Example 1.1. In Table 1.1, we count prime values of $f(g)$, where $f(T) = T^{12}+(u+1)T^6+u^4$ and g runs over polynomials of degree n in $\mathbf{F}_3[u]$, with $9 \leq n \leq 17$. (Here and in later examples, checking the Bouniakowsky conditions for f is left to the reader. All computations in this paper were carried out using PARI, NTL, and MAGMA, with deterministic primality testing.) An estimate for $C(f)$ is 3.52138375. Our range of degrees in Table 1.1 is small, but the sampling sets are substantial; *e.g.*, there are 9,565,938 polynomials of degree 14 in $\mathbf{F}_3[u]$. After each count of prime values in the table, we give the naive estimate for that count according to (1.2) and we give the ratio of these quantities. These data suggest the ratio tends to a number ≈ 1.33 rather than to 1.

Remark 1.2. To keep the presentation of data in our tables clean and informative, we round naive estimates (that is, the right side of (1.2)) to one digit after the decimal point – as a simple reminder that they are only estimates – and we round ratios between the two sides of (1.2) to three digits after the decimal point. In some tables, this has the effect of making the ratio of the two sides of (1.2) appear (for small n) to be more accurate than is justified by our rounded estimate presented on the right side of (1.2). Our policy has been to compute $C(f)$ (as described in the appendix) to high enough accuracy to convince ourselves that we have correctly rounded all estimates presented in the tables; we have not worried about giving rigorous proofs of the correctness of the rounding in these tables, since the data in the tables merely serve to illustrate and motivate theorems and conjectures. *We will not comment on this issue again.*

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

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

Example 1.3. Consider $T^3 + u$ over $\mathbf{F}_3[u]$. Here $C(f) = 1/\log 3$. In Table 1.2, the ratio between the count and the naive estimate in (1.2) seems to fall into four interlaced statistics with limiting values 1, 2, 1, 0. In particular, it appears that $g^3 + u$ is reducible when $n = \deg g$ is a positive multiple of 4. (This includes $n = 4$, which is not in the table.)

n	Count	Naive Est.	Ratio
5	36	32.4	1.111
6	144	81.0	1.778
7	216	208.3	1.037
8	0	546.8	0
9	1404	1458.0	0.963
10	7776	3936.6	1.975
11	10746	10736.2	1.001
12	0	29524.5	0
13	82140	81760.2	1.005
14	455256	227760.4	1.999
15	637440	637729.2	1.000
16	0	1793613.4	0

TABLE 1.2. $T^3 + u$ over $\mathbf{F}_3[u]$

Example 1.4. Let us extend the constant field in Example 1.3: consider T^3+u over $\mathbf{F}_9[u]$. In Table 1.3, there appear to be two interlaced statistics, rather than four. These look as expected in odd degree (the ratio is near 1) but no prime values are arising in positive even degree. (The expected behavior over any $\mathbf{F}_{3^r}[u]$ is described in Example 9.8, using notation that will be introduced later in this Introduction.)

Example 1.5. Consider $T^8 + u^3$ over $\mathbf{F}_2[u]$. Although (1.2) predicts an exponentially growing number of prime values in each degree, $T^8 + u^3$ has no prime values on $\mathbf{F}_2[u]$! This

n	Count	Naive Est.	Ratio
1	24	24.0	1.000
2	0	108.0	0
3	648	648.0	1.000
4	0	4374.0	0
5	31104	31492.8	0.988
6	0	236196.0	0
7	1815696	1822083.4	0.996
8	0	14348907.0	0

TABLE 1.3. $T^3 + u$ over $\mathbf{F}_9[u]$

was proved by Swan [31] in 1962, but the context of his work was sufficiently different from the Hardy–Littlewood conjecture that a link between the two has not been identified before. We recall Swan's proof in Example 4.3, and we establish an analogous result in $\kappa[u]$ for any finite field κ in Example 4.15.

Example 1.5 is surprising from a classical point of view, but we regard Examples 1.3 and 1.4 as more instructive because they suggest that the ratio of the two sides of (1.2) can have interlaced limiting values as a periodic function of n. These examples also suggest that the limiting behavior of the ratio is sensitive to extension of the constant field κ.

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

- $f(T)$ is a polynomial in T^p, where p is the characteristic of κ.
- The ratio of the two sides in (1.2) appears to have 1, 2, or 4 limiting values as a function of $n \bmod 4$ when $n \to \infty$.
- The numbers $\mu(f(g))$, where μ is the Möbius function on $\kappa[u]$ (see Definition 4.1) and g runs over $\kappa[u]$, exhibit unusual statistics. Essentially, this means the nonzero values of $\mu(f(g))$ are not equally often 1 and -1. We call this idea the *Möbius bias.* One of the basic results in this paper is a theorem that lets us rigorously prove such a bias can occur for polynomials in T^p when $p \neq 2$ and in T^4 when $p = 2$.

For an algebraist, it is comforting to find apparent counterexamples to (1.2) only among polynomials in T^p, since irreducible polynomials in T^p are already well-known to exhibit peculiar algebraic properties in characteristic p. These are the irreducible $f \in \kappa[u][T]$ that have positive degree in T and are inseparable over the field $\kappa(u)$. While inseparable irreducibles have no classical analogue, there is no reason to dismiss them from consideration in (1.2). For instance, the nonvanishing of $C(f)$ in (1.2) is unrelated to whether or not $f(T)$ is inseparable in T. Moreover, (1.2) does look good for many inseparable irreducibles. A simple example is $T^p + u^2$ (see Example 4.14).

By studying apparent counterexamples to (1.2) in the context of our three observations above, we were led to a new heuristic idea: statistics for irreducible values of $f(g)$ as g varies are influenced by an appropriate average value of $\mu(f(g))$ as g varies. Averaging the Möbius bias in the right way enables us to predict the 1, 2, or 4 apparent limits suggested in all numerical examples that we have examined, and moreover these predicted values are effectively computable rational numbers. Whereas Bouniakowsky's divisibility obstruction is of a *local* character (a divisibility obstruction, by its definition, comes from divisibility by a common prime), the consideration of Möbius averages is fundamentally *global.* We are not aware of an explanation of the above phenomena in $\kappa[u]$ by a heuristic use of the circle method in characteristic p.

Let us illustrate our Möbius-bias heuristic for Example 1.1. For

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

over $\mathbf{F}_3[u]$, we will show in Example 5.2 that

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

for all g in $\mathbf{F}_3[u]$, where $(\frac{\cdot}{3})$ is the Legendre symbol. (The term $g(0)^2$ should not be omitted from the Legendre symbol, since it could be 0.) As g runs over polynomials of a given degree ≥ 2 in $\mathbf{F}_3[u]$, (1.3) shows that $\mu(f(g))$ is -1 twice as often as it is 1. The average nonzero value of $\mu(f(g))$ in each degree ≥ 2 is therefore $(-1-1+1)/3 = -1/3$ (not just asymptotically, but exactly). Note that $1-(-1/3) = 4/3 = 1.33\dots$ seems to fit the deviation from (1.2) in Table 1.1. Such agreement is purely numerical; we have no proof linking $\mu(f(g))$ to the primality statistics of $f(g)$.

Remark 1.6. Since the Möbius bias is a global parity condition on squarefree factorizations (with the squarefreeness of $f(g)$ considered to be a preliminary local condition), it is natural to ask if there are higher-order heuristic global obstructions to primality, such as a mod-3 condition on squarefree factorizations. We have studied many examples over small finite fields (of characteristics 2, 3, 5, and 7) and have found that the Möbius bias leads to a correction factor that gives an excellent numerical fit to all observed deviations from (1.2). Without guidance provided by examples giving evidence to the contrary, it seems to us that the Möbius-bias heuristic provides a satisfactory theory to account for all deviations from (1.2).

To convert our heuristic into a correction term in (1.2), we now describe some new theorems about the Möbius function on $\kappa[u]$. More accurately, our results concern the behavior of $\mu(f(g))$, where $f(T)$ is fixed in $\kappa[u][T^p]$ and g runs through $\kappa[u]$.

The key result, to be made precise in Theorem 1.8 below, is that $\mu(f(g))$ is essentially a periodic function of g and we can provide a formula for a modulus of periodicity. When $f(T)$ is monic in T, for instance, a modulus of periodicity is the radical of the $\kappa[u]$-resultant of $f(T)$ and the u-partial derivative $(\partial_u f)(T)$ (this means the resultant of polynomials in T with coefficients in $\kappa[u]$). As an example, let $f(T)$ be the polynomial in Example 1.1. The $\mathbf{F}_3[u]$-resultant of f and $\partial_u f$ is $u^{18}(u-1)^{18}$, whose radical is $u(u-1)$. This is consistent with (1.3), where we see $\mu(f(g))$ depends on g modulo $u(u-1)$.

Our results on $\mu(f(g))$, which are inspired by our study of (1.2), do not require that $f(T)$ be irreducible in $\kappa[u][T]$. We only need $f(T)$ to be squarefree, which unlike irreducibility is a stable property under finite extension of the constant field ($f(T)$ in Example 1.1 is reducible in $\mathbf{F}_9[u][T]$, but still squarefree). Therefore we now fix $f(T) \in \kappa[u][T^p]$ that is squarefree in $\kappa[u][T]$ and, to avoid trivialities, $f \notin \kappa$.

To generalize (to nonmonic polynomials) the resultant construction of a modulus of periodicity for $\mu(f(g))$, we use geometric language. Let $Z_f = \{f(u,T) = 0\}$ be the affine plane curve corresponding to $f \in \kappa[u,T]$. The projection from Z_f onto the T-axis is flat and generically étale, so this projection is non-étale at a finite set of points on Z_f, say at the set B. Projecting B onto the u-axis gives a finite set of points. Define M_f^{geom} to be the monic polynomial in $\kappa[u]$ whose roots are this finite set on the u-axis, each root having multiplicity 1 (that is, $M_f^{\rm geom}$ is squarefree). We label this polynomial $M_f^{\rm geom}$ since it is unaffected by replacing κ with a finite extension.

Remark 1.7. Concretely, an element u_0 in an algebraic closure of κ is a root of $M_f^{\rm geom}$ precisely when the specializations $f(u_0, T)$ and $(\partial_u f)(u_0, T)$ have a common T-root. This condition is the same as u_0 being a root of the $\kappa[u]$-resultant of f and $\partial_u f$ *only* when the u_0-specialization of either f or $\partial_u f$ has the same respective T-degree as f or $\partial_u f$. An equivalent description of this latter condition is: u_0 is not a double root of the leading coefficient of f as a polynomial in T.

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

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

Theorem 1.8. *Let κ have odd characteristic p and $f(T) \in \kappa[u][T^p]$ be squarefree in $\kappa[u][T]$ and not lie in κ.*

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

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

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

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

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

Examples 1.3 and 1.4 showed replacing the constant field with a finite extension is intriguing in the context of (1.2). (Example 5.9 gives an example where $M_{f,\kappa}^{\rm min} \neq M_f^{\rm geom}$.) Motivated by our examples and the technical needs of proofs, throughout the paper we will keep track of the behavior of bounds and other parameters with respect to replacing κ with an arbitrary finite extension κ' while using the same f.

In the proof of Theorem 1.8, the importance of $f(T)$ being a polynomial in T^p is that its T-partial derivative is 0. That implies, for any $g \in \kappa[u]$, the u-derivative of $f(g(u)) \in \kappa[u]$ is $(\partial_u f)(g(u))$. In other words, $\partial_u(f(u, g(u))) = (\partial_u f)(u, g(u))$ if we consider f as a function of two variables u and T. Therefore the u-derivative of $f(g)$ is a polynomial in g with no dependence on $g'(u)$ in such cases.

Remark 1.9. From the geometric point of view, it is surprising to have an implication as in (1.4) that can relate polynomials g_j with different degrees. Since the quadratic nature of -1 in $\kappa^\times$ influences whether or not (1.4) depends on $\deg g \bmod 4$ or on $\deg g \bmod 2$, it seems unlikely that there can be a purely geometric proof of (1.4), although geometric ideas do play a prominent role in our proof.

Example 1.10. Let $f(T) = T^{12}+(u+1)T^6+u^4$ in $\mathbf{F}_3[u][T]$, as in Example 1.1. Remark 1.7 and an earlier calculation imply that $M_f^{\text{geom}} = u(u-1)$, so Theorem 1.8 says that $\mu(f(g))$ depends on $g \bmod u(u-1)$ for $\deg g \gg 0$. This is consistent with the formula in (1.3), which can be proved in an elementary way without Theorem 1.8, and moreover the mod-4 and quadratic-character conditions in (1.4) drop out and the condition that $\deg g \gg 0$ can be made explicit: $\deg g \geq 2$.

Remark 1.11. For both theoretical and numerical purposes, it would be useful to establish a lower bound on n_1 and n_2 beyond which (1.4) holds even if κ is replaced by any finite extension. In every example that we have checked for odd p, the integer $1 + \deg(M_f^{\text{geom}})$ has been such a lower bound. We do not have any theorems in this direction.

Example 1.12. Let

$$f(T) = T^9 + (2u^2 + u)T^6 + (2u + 2)T^3 + u^2 + 2u + 1 \in \kappa[u][T]$$

with finite κ of characteristic 3. As a preparation for the proof of Theorem 1.8, in Example 5.3 we will show

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

when $g = cu^n + \cdots$, with $n \geq 1$. Therefore $\mu(f(g))$ depends on $g \bmod (u-1)(u-2)$, $\deg g \bmod 4$, and the quadratic character of $c = \operatorname{lead} g$. (One checks that f and $\partial_u f$ have $\kappa[u]$-resultant $-(u-1)^6(u-2)^9$ whose monic radical is $(u-1)(u-2)$, so (1.5) and Remark 1.7 recover Theorem 1.8 in this case.) Formula (1.5) shows that Möbius behavior can change upon extension of the ground field: when -1 is a square in κ, the term $\chi(-1)^{n(n-1)/2}$ drops out, so dependence of $\mu(f(g))$ on $\deg g \bmod 4$ drops to dependence on $\deg g \bmod 2$.

The case of characteristic 2 lies deeper than the case of odd characteristic. Our treatment of characteristic 2 uses liftings to characteristic 0, via Witt vectors. (Readers not interested in characteristic 2 can skip ahead to the paragraph after Remark 1.14.) Here is an analogue of Theorem 1.8 in characteristic 2 for the case of polynomials in T^4; in §8 we will state and prove a more technical theorem that applies to polynomials in T^2.

Theorem 1.13. *Let κ be a finite field with characteristic 2. Fix a nonzero $f(T) \in \kappa[u][T^4]$ that is squarefree in $\kappa[u][T]$ and assume $f \notin \kappa$. There is a nonzero $M_{f,\kappa}$ in $\kappa[u]$ such that for $g \in \kappa[u]$ with sufficiently large degree, $\mu(f(g))$ is determined by $g \bmod M_{f,\kappa}$ and $\deg g \bmod 2$. If $[\kappa : \mathbf{F}_2]$ is even or $\deg_T f \equiv 0 \bmod 8$, then there is no dependence on $\deg g \bmod 2$.*

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

An interesting example of Theorem 1.13 is $f(T) = T^8 + (u^3 + u)T^4 + u$ over any finite field κ with characteristic 2. For $g \in \kappa[u]$, the proof of Theorem 1.13 implies $M_{f,\kappa} = 1$ and $\mu(f(g)) = 1$ for $\deg g \gg 0$. Thus (1.2) fails in this example. See Example 8.14 for further information.

It seems likely that the modulus $M_{f,\kappa}$ in Theorem 1.13 need not be squarefree, which is a contrast with Theorem 1.8. For example, when κ has characteristic 2 and

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

then the proof of Theorem 1.13 yields

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

when $\deg g_j \geq 2$ (see Example 8.15), and numerical evidence suggests (but we cannot prove) that the modulus in (1.6) cannot be replaced with its radical, even if we restrict attention to $\deg g_j \gg 0$ instead of to $\deg g_j \geq 2$. Over some fields it seems probable that $u^9(u+1)^4$ in (1.6) can be replaced with a proper factor; for example, when $\kappa = \mathbf{F}_2$ the data suggest (but we cannot prove) that $u^3(u+1)$ may be used as a modulus in (1.6) when taking $\deg g_j \geq 0$.

Remark 1.14. It appears from numerics that there is not always a periodicity property for $\mu(f(g))$ when $f \in \kappa[u][T^2]$. For this reason, generalizing Theorem 1.13 beyond the case $f \in \kappa[u][T^4]$ will require further work.

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

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

where $M \in \kappa[u]$ is any modulus $M_{f,\kappa}$ from Theorem 1.8 or Theorem 1.13; both sums run over g, and the denominator is nonzero for large n by Lemma 9.3. Note $\Lambda_{\kappa,M}(f;n)$ is a rational number in $[0,2]$.

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

(1) for any two choices of modulus M, we will prove in Corollary 9.11 that the corresponding sequences $\Lambda_{\kappa,M}(f;n)$ agree for large n,
(2) in many (but not all!) examples, $\Lambda_{\kappa,M}(f;n) = \Lambda_{\kappa,1}(f;n)$ for large n (that is, the constraint $(f(g), M) = 1$ in (1.7) can be dropped), even when 1 is not a genuine modulus for $g \mapsto \mu(f(g))$.

In Remark 9.14, we will give a general criterion for (2) to hold, which in particular applies to Example 1.1. (We will also explain in that remark why we use the condition $\gcd(f(g), M) = 1$ in the definition of $\Lambda_{\kappa,M}(f;n)$.) Because of (1), we may abbreviate $\Lambda_{\kappa,M}(f;n)$ to $\Lambda_\kappa(f;n)$, provided the properties that we care about are limited to large n, as they usually are.

The marvelous surprise (Theorem 9.10) is that $\Lambda_\kappa(f;n)$ is *periodic* in n with period 1, 2, or 4 for sufficiently large n; intuitively, this is a consequence of Theorems 1.8 and 1.13, and consequently $\Lambda_\kappa(f;n)$ is far simpler than it at first appears to be. The deviation of this periodic function from the constant function 1, or equivalently the deviation of $\sum \mu(f(g))/\sum |\mu(f(g))|$ from 0, measures the tendency of $\mu(f(g))$ to be biased toward -1 or 1 when $f(g)$ is squarefree and $\deg g \bmod 4$ is fixed (and $\deg g$ is large). This makes the following proposed correction to (1.2) simple to appreciate: we conjecture that when $f \in \kappa[u][T^p]$ satisfies the Bouniakowsky conditions, with the extra restriction that $f \in \kappa[u][T^4]$

when $p = 2$,

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

as $n \to \infty$. If Theorem 1.13 can be generalized to allow $f \in \kappa[u][T^2]$ (see Remark 1.14), it should be possible to use it to formulate a version of (1.8) in characteristic 2 for any $f \in \kappa[u][T^2]$ that satisfies the Bouniakowsky conditions.

Example 1.15. Let $f_1(T) = T^{12} + (u+1)T^6 + u^4$ and $f_2(T) = T^3 + u$. We will later establish (in Examples 5.2 and 9.8) the periodic behavior $\Lambda_{\mathbf{F}_3}(f_1;n) = 4/3$ for $n \geq 2$, $\Lambda_{\mathbf{F}_3}(f_2;n) = 1,2,1,0,1,2,1,0,\dots$ for $n \geq 1$, and $\Lambda_{\mathbf{F}_9}(f_2;n) = 1,0,1,0,1,0,\dots$ for $n \geq 1$. Thus, (1.8) appears to fix the discrepancies in the ratio columns for Tables 1.1, 1.2, and 1.3 when we stay away from (periodic!) $n \gg 0$ such that $\Lambda_\kappa(f;n) = 0$.

As a consistency check between (1.2) and (1.8), we have always been able to prove $\Lambda_\kappa(f;n) = 1$ for large n in examples where data suggest that (1.2) holds. Our experience suggests that the naive estimate (1.2) is correct for many T-inseparable f as in Theorem 1.8, and likewise our computer work suggests that $\Lambda_\kappa(f;n) = 1$ for $n \gg 0$ for the same f. Nevertheless, other possibilities do occur, as we have seen.

The possibility that 0 lies in the period of $\Lambda_\kappa(f;n)$ requires a clarification on the meaning of (1.8) as an asymptotic relation. When 0 is in the period of $\Lambda_\kappa(f;n)$, what does (1.8) mean for the n's where $\Lambda_\kappa(f;n)$ is periodically 0? The vanishing of $\Lambda_\kappa(f;n)$ implies that for all g of degree n, either $\mu(f(g)) = 1$ or $(f(g), M_{f,\kappa}) \neq 1$. When n is large, both cases imply that $f(g)$ is reducible. Therefore, the appearance of 0 in the period for $\Lambda_\kappa(f;n)$ implies that both sides of (1.8) vanish for such n, which proves there is a periodic lack of irreducible specializations $f(g)$. For instance, the mod-4 and mod-2 patterns of 0's in Tables 1.2 and 1.3 provably continue for all larger n. For other large degrees n, where $\Lambda_\kappa(f;n) \neq 0$, we only conjecture that (1.8) is a genuine asymptotic relation.

When κ has characteristic $p \neq 2$ and $f(T) \in \kappa[u][T]$ is irreducible with positive T-degree, we believe the correct $\kappa[u]$-variant on Bouniakowsky's (qualitative) conjecture is the following: $f(g)$ is irreducible for infinitely many $g \in \kappa[u]$ except in the following two cases: $f(T)$ has a divisibility obstruction or $f(T)$ is a polynomial in T^p with $\Lambda_\kappa(f;n) = 0$ for $n \gg 0$. Both types of obstructions can be checked with a finite amount of computation. An example which fits the second case but not the first is $f(T) = T^{4p} + u$; for any nonconstant g in $\kappa[u]$, $f(g)$ is reducible. For details, see Example 4.13. (We do not make any analogous conjecture in characteristic 2 because the case of characteristic 2 is still not adequately understood when f is a polynomial in T^2 but not a polynomial in T^4.)

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

Here is an outline of the paper. In §2, we recall the usual Hardy–Littlewood conjecture in $\mathbf{Z}[T]$, in a form slightly different from (1.1), and §3 gives additional apparent counterexamples to (1.2). In §4 and §5 we build on work of Swan to develop an understanding of $\mu(f(g))$ as in Theorem 1.8, which we restate as Theorem 5.7. Our proof of Theorem 5.7 is given in §6 and §7, and it uses a mixture of polynomial algebra and algebraic geometry. In particular, we only begin to prove some non-trivial results in §6*ff*; the development in §1–§5 is largely a discussion of examples and some classical facts. Since the phenomena we are coming to terms with is unrelated to any classical ideas about prime values of polynomials, we feel that this preliminary discussion will help the reader to understand the nature of the theorems that we prove in §6*ff.*

In §8, we treat characteristic 2. Theorem 1.13 appears in a more precise form as Theorem 8.11 and Corollary 8.12. Its proof uses ideas from our treatment of odd characteristic and some considerations with residues of differential forms on the projective lines in characteristics 2 and 0. Finally, §9 returns to conjectures, discussing the new factor $\Lambda_\kappa(f;n)$ in (1.8). This leads to our modified Hardy–Littlewood conjecture, given as Conjecture 9.12. In an appendix we explain how to compute constants such as $C(f)$ in (1.1) and (1.2); this is important for numerical work, without which most of the nontrivial phenomena in this paper would not have been discovered.

Let us conclude this Introduction by using our work in characteristic p to suggest a link between some classical conjectures in analytic number theory. The classical Hardy–Littlewood conjecture (1.1) is not expected to have any counterexamples. Since counterexamples to (1.2) in characteristic p appear to be explained by non-vanishing Möbius averages, it seems reasonable to conjecture that if $f(T)$ is irreducible in $\mathbf{Z}[T]$ and has no divisibility obstructions then it has a vanishing Möbius average:

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

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

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

For linear f, (1.10) is true [30]. Numerical evidence for (1.10) in other cases is encouraging. After being led to (1.10) by analogy with our work in characteristic p, we learned that it is a folklore conjecture for all non-constant $f(T)$. The case $f(T)=T^2+1$ is posed in [12, p. 417].

The interesting aspect of the above considerations in the classical case is that since some (inseparable) polynomials in characteristic p have nonzero Möbius average, we arrive at a new perspective on (1.10): the way that we were just led to (1.10) suggests that any counterexample to (1.10) is probably a counterexample to the classical Hardy–Littlewood conjecture. That is, the classical Hardy–Littlewood conjecture should imply (1.9) and (1.10). Can such an implication be proved, perhaps assuming some other standard conjectures?

Notation and Terminology. Throughout the paper, κ denotes a finite field of size q. For nonzero $g\in\kappa[u]$, we set $\mathrm{N}g=q^{\deg g}$. We often let μ, rather than $\mu_{\kappa[u]}$, denote the Möbius function on $\kappa[u]$, relying on the context to make clear the ring in which we are computing the Möbius function; see Definition 4.1. We likewise often write χ instead of χ_κ to denote the quadratic character on the multiplicative group $\kappa^\times$ of a finite field with odd characteristic.

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

For a nonzero polynomial h in one variable, we write the leading coefficient as $\operatorname{lead} h$. For a nonzero polynomial f in two variables u and T over a ring R, the T-degree of f and the leading coefficient of f as a polynomial in T are indicated with a subscript: $\deg_T f\geq 0$ and $\operatorname{lead}_T f\in R[u]$. An element in $R[u]$ is *primitive* when its coefficients generate the unit ideal in R. For a domain K, the discriminant of a one-variable polynomial with coefficients in K is denoted $\operatorname{disc} h$, or $\operatorname{disc}_K h$ for emphasis. Our definition of discriminants does not match the usual definition when the polynomial is not monic; see (4.1) and (4.2). Our notation for resultants is introduced in §5.

All algebras in this paper are assumed to be commutative.

When R is a local ring with residue field k, a *lift* of a polynomial $h \in k[u,T]$ is a polynomial $H \in R[u,T]$ whose reduction to $k[u,T]$ is h. We call the lift *unitary* when $\deg_T H$ equals $\deg_T h$ and $\mathrm{lead}_T H \in R[u]$ is a lift of $\mathrm{lead}_T h \in k[u]$ with the same u-degree. In particular, $\mathrm{lead}_T H \in R[u]$ has unit leading coefficient, hence the terminology.

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

2. The Classical Case

This section is intended for readers who are unfamiliar with the Hardy–Littlewood conjecture, and it also serves to fix some terminology. Experts can start in §3.

For irreducible $f_1(T), \ldots, f_r(T)$ in $\mathbf{Z}[T]$, where none is a unit multiple of the others, let

$$\pi_{f_1,\ldots,f_r}(x) = \#\{1 \le n \le x : f_1(n), \ldots, f_r(n) \text{ are all prime}\}. \tag{2.1}$$

This counts how often the f_j's all take prime values over positive integers up to x.

For $g(T) \in \mathbf{Z}[T]$ and a prime p, set

$$\omega_g(p) := \#\{\overline{n} \in \mathbf{Z}/(p) : g(n) \equiv 0 \bmod p\}. \tag{2.2}$$

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

Conjecture 2.1 (Hardy–Littlewood). *Pick $f_1(T), \ldots, f_r(T) \in \mathbf{Z}[T]$ with no f_i a unit multiple of f_j for $i \neq j$, and let $f(T)$ be their product. Assume the following two conditions:*

1) *The $f_j(T)$'s are irreducible and pairwise coprime in $\mathbf{Q}[T]$.*

2) *The product $f = \prod_{i=1}^r f_i$ has no local obstructions, i.e., $\omega_f(p) < p$ for all p.*

Then

$$\pi_{f_1,\ldots,f_r}(x) \overset{?}{\sim} C(f) {\sum_{n \le x}}' \frac{1}{\log|f_1(n)| \cdots \log|f_r(n)|}, \tag{2.3}$$

where

$$C(f) = \prod_p \frac{1 - \omega_f(p)/p}{(1 - 1/p)^r} \tag{2.4}$$

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

The second hypothesis in the Hardy–Littlewood conjecture is equivalent to f having a pair of relatively prime values, which is how the second hypothesis is checked in practice. For example, $T(T^2+2)$ has a local obstruction at 3 because $n(n^2+2)$ is divisible by 3 for any n. This implies that n and n^2+2 are not simultaneously prime infinitely often. The infinite product $C(f)$, taken in order of increasing p, is usually only conditionally convergent.

Replacing each $f_j(n)$ with its leading term as a polynomial in n simplifies (2.3) to

$$\pi_{f_1,\ldots,f_r}(x) \overset{?}{\sim} \frac{C(f)}{(\deg f_1)\cdots(\deg f_r)}\frac{x}{(\log x)^r}. \tag{2.5}$$

While (2.3) is a more complicated estimate than (2.5), it is the estimate that follows more directly from heuristics based on probability or the circle method, and it is numerically more accurate than (2.5). As Hardy and Littlewood [16, p. 38] write,

> For the *asymptotic* formula, naturally, it is indifferent which [formula] we adopt. But, for purposes of *verification within the limits of calculation*, it is by no means indifferent [and] it will be found that it makes a vital difference in the plausibility of the results.

We conclude this section with some remarks on the literature. The second condition in Conjecture 2.1 was first recognized by Bouniakowsky [5]. He stated Conjecture 2.1 for $r = 1$, with only the qualitative conclusion $\pi_f(x) \to \infty$; *i.e.*, $f(n)$ is prime infinitely often. Extensions of this qualitative conjecture are due to Dickson [11] for any number of polynomials in $\mathbf{Z}[T]$ of degree 1, and to Schinzel [24] for any number of polynomials in $\mathbf{Z}[T]$ of any degree. Schinzel's conjecture is usually called Hypothesis H. A precise qualitative number-field extension to S-integers is due to Serre [7, §4]. Applications of these qualitative conjectures are in [7], [21], [23], [24], [32], and [34].

3. A conjecture in $\kappa[u]$ and counterexamples

In this section we consider a $\kappa[u]$-analogue of (2.3). Pick $f \in \kappa[u]$ with $\deg_T f > 0$. Write

$$f(T) = \alpha_d(u)T^d + \alpha_{d-1}(u)T^{d-1} + \cdots + \alpha_0(u), \tag{3.1}$$

with $\alpha_d(u) \neq 0$ and $d > 0$. When $g \neq 0$ in $\kappa[u]$ and

$$\deg(\alpha_d g^d) > \deg(\alpha_i g^i) \tag{3.2}$$

for each $i < d$ such that $\alpha_i \neq 0$ ((3.2) is vacuous when $\alpha_i = 0$), then $f(g) \neq 0$ and the degree and leading coefficient of $f(g)$ in $\kappa[u]$ are the same as those for $\alpha_d g^d$:

$$\deg(f(g)) = d \cdot \deg g + \deg \alpha_d = (\deg_T f)n + \deg(\operatorname{lead}_T f), \tag{3.3}$$

$$\operatorname{lead}(f(g)) = (\operatorname{lead} \alpha_d)(\operatorname{lead} g)^d, \tag{3.4}$$

where $n = \deg g$. The lower bounds (3.2) hold provided $\deg g > \nu(f)$, where

$$\nu(f) = \max_{0 \le i \le d-1} \frac{\deg \alpha_i - \deg \alpha_d}{d - i}. \tag{3.5}$$

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

Definition 3.1. The polynomial $f(T)$ has a *local obstruction* at an irreducible $\pi \in \kappa[u]$ when $f(g) \equiv 0 \bmod \pi$ for all $g \in \kappa[u]$.

In practice, one checks that $f(T)$ has no local obstructions by finding two specializations of $f(T)$ on $\kappa[u]$ that are relatively prime.

Suppose $f_1(T), \ldots, f_r(T) \in \kappa[u][T]$ are each irreducible over $\kappa(u)$ with no f_i a unit multiple of f_j in $\kappa[u][T]$ for $i \neq j$, and assume that their product $f(T)$ has no local obstructions. Define

$$\pi_{f_1,\ldots,f_r}(n) = \#\{g \in \kappa[u] : \deg g = n, \text{all } f_j(g) \text{ prime}\}.$$

A conjecture analogous to (2.3) is

$$\pi_{f_1,\dots,f_r}(n) \overset{?}{\sim} C(f) \sum_{\deg g=n} \frac{1}{\log \mathrm{N}(f_1(g))\cdots \log \mathrm{N}(f_r(g))}, \tag{3.6}$$

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

$$C(f) = (\log q)^r \prod_{(\pi)} \frac{1-\omega_f(\pi)/\,\mathrm{N}\pi}{(1-1/\,\mathrm{N}\pi)^r}, \quad \omega_f(\pi) = \#\{g \bmod \pi : f(g) \equiv 0 \bmod \pi\}, \tag{3.7}$$

the product running over nonzero prime ideals in $\kappa[u]$. We sometimes write $C(f)$ as $C_{\kappa[u]}(f)$ to emphasize the base ring $\kappa[u]$ (especially the choice of κ). By (3.3), $\deg(f_j(g))$ depends on g only through $\deg g$ when $\deg g \gg 0$, so all terms in the sum in (3.6) are equal for large n. When $r = 1$, (3.6) is essentially the same as (1.2).

For computational purposes, it is expedient to remove the implicit factors of $\log q$ that arise in (3.6), both in $C(f)$ and in each $\log \mathrm{N}(f_j(g)) = (\log q)\deg(f_j(g))$. These factors of $\log q$ cancel, and we record the corresponding alternative form of (3.6) in the case of interest to us here, $r = 1$ (and writing f_1 as f):

$$\#\{g \in \kappa[u] : \deg g = n, f(g) \text{ prime}\} \overset{?}{\sim} \prod_{(\pi)} \frac{1-\omega_f(\pi)/\,\mathrm{N}\pi}{1-1/\,\mathrm{N}\pi} \cdot \frac{(q-1)q^n}{(\deg_T f)n + \deg(\mathrm{lead}_T f)}, \tag{3.8}$$

provided $f(T)$ satisfies the Bouniakowsky conditions. The term $\deg(\mathrm{lead}_T f)$ could of course be dropped off for asymptotic purposes, although to avoid confusion we will not drop this term when we provide numerical data related to (3.8).

To avoid any conceptual misunderstandings, it is worth stressing that the product over (π) on the right side of (3.8) is *not* the correct $\kappa[u]$-analogue of the classical product in (2.4), even though they formally look the same. The correct analogue is $C(f)$ in (3.7), with the factor $(\log q)^r$. This logarithmic power is crucial for good base-change properties of constants like $C(f)$, and the importance of $(\log q)^r$ is not evident from (2.4) because $\mathrm{Res}_{s=1}\,\zeta_{\mathbf{Z}}(s) = 1$, whereas $\mathrm{Res}_{s=1}\,\zeta_{\kappa[u]}(s) = 1/\log q \neq 1$. (We have no use for base-change properties of such constants in this paper, however.) Since, when $r = 1$, (3.8) is computationally preferable to (3.6), in numerical work (with $r = 1$) we compute $C(f)/(\log q)$ rather than $C(f)$.

We call (3.8), or the more general (3.6), the *naive Hardy–Littlewood conjecture* over $\kappa[u]$. It is an obvious conjecture to make, but in the Introduction we saw apparent (and proved) counterexamples. For future reference, we now look at additional apparent (and proved) counterexamples to (3.8). We omit the verification of the Bouniakowsky conditions.

Example 3.2. Let $f(T) = T^{p^b} + u$ over $\mathbf{F}_p[u]$ with $b \geq 1$. Here $\omega_f(\pi) = 1$ for all π, so (3.8) says

$$\#\{g \in \mathbf{F}_p[u] : \deg g = n, g^{p^b} + u \text{ prime}\} \overset{?}{\sim} \frac{(p-1)p^{n-b}}{n}. \tag{3.9}$$

In Table 3.1, we see (3.9) over $\mathbf{F}_2[u]$ looks correct for $T^2 + u$, but there is a deviation for $T^4 + u$: there appears to be a discrepancy factor that has periodic limits 2,0. In Table 1.2, we already saw that the discrepancy factor for $T^3 + u$ seems to have periodic limits 1,2,1,0. In Table 3.2, the prediction for $T^5 + u$ looks good in high odd degree, but fails spectacularly in even degree.

The periodic absence of prime specializations for $T^3 + u$, $T^4 + u$, and $T^5 + u$ is proved in Examples 4.5 and 8.13, showing (3.9) is wrong in these cases.

	T^2+u			T^4+u		
n	Count	Naive Est.	Ratio	Count	Naive Est.	Ratio
9	32	28.4	1.127	24	14.2	1.690
10	48	51.2	0.938	0	25.6	0
11	96	93.1	1.031	92	46.5	1.978
12	136	170.7	0.797	0	85.3	0
13	336	315.1	1.066	336	157.5	2.133
14	568	585.1	0.971	0	292.6	0
15	1136	1092.3	1.040	1076	546.1	1.970
16	1904	2048.0	0.930	0	1024.0	0
17	3824	3855.1	0.992	3904	1927.5	2.025

Table 3.1. T^2+u and T^4+u in $\mathbf{F}_2[u]$

n	Count	Naive Est.	Ratio
1	4	4.0	1.000
2	0	10.0	0
3	40	33.3	1.200
4	0	125.0	0
5	680	500.0	1.360
6	0	2083.3	0
7	9080	8928.6	1.017
8	0	39062.5	0
9	173340	173611.1	0.998
10	0	781250.0	0
11	3546020	3551136.4	0.999
12	0	16276041.7	0
13	75117240	75120192.3	1.000

Table 3.2. T^5+u over $\mathbf{F}_5[u]$

Example 3.3. Let $f(T) = uT^8 + 1$ over $\mathbf{F}_2[u]$, Since $\omega_f(\pi) = 1$ for $\pi \neq u$ and $\omega_f(u) = 0$, (3.8) suggests $\#\{g \in \mathbf{F}_2[u] : \deg g = n, f(g) \text{ prime}\} \overset{?}{\sim} 2^{n+1}/(8n+1)$, but Table 3.3 suggests an asymptotic discrepancy factor of 2. Example 4.4 and the Möbius bias (as discussed in the Introduction) give a good heuristic explanation for this doubling.

Example 3.4. Let $f(T) = T^9 + (2u^2+u)T^6 + (2u+2)T^3 + u^2 + 2u + 1$ over $\mathbf{F}_3[u]$. According to (3.8), $\#\{g \in \mathbf{F}_3[u] : \deg g = n, f(g) \text{ prime }\} \overset{?}{\sim} C \cdot 2 \cdot 3^{n-2}/n$, where $C = C_{\mathbf{F}_3[u]}(f)/\log 3 \approx 1.01570541$. Table 3.4 suggests this asymptotic relation is wrong for odd n. The absence of prime $f(g)$ for $\deg g \equiv 1 \bmod 4$ is proved in Example 9.9.

Example 3.5. Let $f(T) = T^{12} + (2u^4 + 2u^3 + 2u^2 + u + 1)T^6 + 2u^3 + 2u^2 + u$ over $\mathbf{F}_3[u]$. According to (3.8), $\#\{g \in \mathbf{F}_3[u] : \deg g = n, f(g) \text{ prime}\} \overset{?}{\sim} C \cdot 3^{n-1}/2n$, where $C = C_{\mathbf{F}_3[u]}(f)/\log 3 \approx 2.13579992$. In Table 3.5, the count systematically falls below the estimate from (3.8) by a factor that seems to be converging to .66 . . . . In §9 (see Table 9.1) we will propose general correction factors, and in the present case this correction factor is 2/3.

n	Count	Naive Est.	Ratio
13	310	156.0	1.987
14	542	290.0	1.869
15	1111	541.6	2.051
16	2000	1016.1	1.968
17	3855	1913.5	2.015
18	7202	3615.8	1.992
19	13657	6853.4	1.993
20	26296	13025.8	2.019

Table 3.3. $uT^8 + 1$ over $\mathbf{F}_2[u]$

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

Table 3.4. $T^9 + (2u^2 + u)T^6 + (2u + 2)T^3 + u^2 + 2u + 1$ over $\mathbf{F}_3[u]$

n	Count	Naive Est.	Ratio
9	526	778.5	0.676
10	1346	2101.9	0.640
11	3910	5732.6	0.682
12	10456	15764.6	0.663
13	28956	43655.8	0.663
14	80720	121612.7	0.664
15	227434	340515.5	0.668

Table 3.5. $T^{12} + (2u^4 + 2u^3 + 2u^2 + u + 1)T^6 + 2u^3 + 2u^2 + u$ over $\mathbf{F}_3[u]$

Example 3.6. For an example in characteristic 5, with a nonmonic polynomial, let $f(T) = (2u^2 + u + 3)T^{15} + (4u^2 + u + 3)T^5 + 4u^2 + u + 3$ over $\mathbf{F}_5[u]$. The prediction from (3.8) is $\#\{g \in \mathbf{F}_5[u] : \deg g = n, f(g) \text{ prime }\} \overset{?}{\sim} C \cdot 4 \cdot 5^n/(15n + 2)$, where $C \approx 1.82326856$. Table 3.6 looks good in odd degree, but not in even degree. We will return to this in Example 9.6.

n	Count	Naive Est.	Ratio
5	297	296.0	1.003
6	1563	1238.6	1.262
7	5264	5325.0	0.989
8	30436	23351.3	1.303
9	103702	103972.9	0.997
10	609531	468562.0	1.301

Table 3.6. $(2u^2+u+3)T^{15}+(4u^2+u+3)T^5+4u^2+u+3$ over $\mathbf{F}_5[u]$

4. The Möbius function over finite fields

In the Introduction we gave a heuristic explanation of the data in Example 1.1 as an effect of a Möbius bias. We speak of a Möbius bias when, roughly speaking, $\mu_{\kappa[u]}(f(g))$ does not take its nonzero values 1 and -1 equally often on average as g varies. In this section, we begin the systematic investigation of Möbius fluctuations in characteristic p, with the ultimate goal of using this work to correct the faulty (1.2). The first step in the analysis of $\mu_{\kappa[u]}(f(g))$ as g varies is the description of a formula for $\mu_{\kappa[u]}(h)$ $(h \in \kappa[u])$ other than its definition; the existence of an alternative Möbius formula on $\kappa[u]$ has no parallel in $\mathbf{Z}$. We will then apply the formula to compute $\mu_{\kappa[u]}(f(g))$ for varying g in some simple cases.

Definition 4.1. Let R be a Dedekind domain. The *Möbius function* on nonzero ideals of R is given by $\mu_R(\mathfrak{p}_1\cdots\mathfrak{p}_m)=(-1)^m$ for distinct nonzero prime ideals $\mathfrak{p}_j$, $\mu_R((1))=1$, and $\mu_R(\mathfrak{b})=0$ for any nonzero ideal $\mathfrak{b}\subseteq R$ divisible by the square of a prime. For nonzero $r\in R$, we define $\mu_R(r)=\mu_R(rR)$. If R is understood from context, we write μ rather than μ_R.

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

$$\operatorname{disc} h := \prod_{i<j}(\gamma_i-\gamma_j)^2 \in F, \tag{4.1}$$

whether or not h is monic. (For nonzero constant h, this empty product is understood to be 1.) In terms of the derivative of h, (4.1) is the same as

$$\operatorname{disc} h = \frac{(-1)^{d(d-1)/2}}{(\operatorname{lead} h)^d}\prod_{i=1}^{d} h'(\gamma_i). \tag{4.2}$$

The factor $(\operatorname{lead} h)^d$ in (4.2) reflects our convention on discriminants of nonmonic polynomials in (4.1). When h is not monic, a variant on (4.1) is often used in the literature to define $\operatorname{disc} h$ (*e.g.*, [18, p. 204]). This variant equals (4.1) multiplied by $(\operatorname{lead} h)^{2d-2}$. In particular, the two competing definitions of the discriminant of a polynomial differ by a nonzero square factor in $F^\times$. We prefer (4.1) for nonmonic h since it agrees with the universally accepted definition of $\operatorname{disc}_F(F[u]/(h))\in F$ relative to the ordered basis $\{1,u,\dots,u^{d-1}\}$.

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

When $R = \mathbf{F}_p[u]$, there is a classical formula for the Möbius function on R in terms of discriminants and the Legendre symbol $(\frac{\cdot}{p})$ (recall that $(\frac{a}{2}) = (-1)^{(a^2-1)/8}$ for odd a and $(\frac{a}{2}) = 0$ for even a):

Theorem 4.2 (Pellet [19])**.** *For a nonzero polynomial $h \in \mathbf{F}_p[u]$,*

$$\mu(h) = (-1)^{\deg h}\left(\frac{\operatorname{disc} h}{p}\right) \tag{4.3}$$

with the convention that when $p = 2$ we replace $\operatorname{disc} h$ *with* $\operatorname{disc} H$ *for any monic $H \in \mathbf{Z}[u]$ lifting h.*

Note that (4.3) is trivial when h has a multiple factor, because then $\mu(h) = \operatorname{disc} h = 0$. When h is squarefree in $\mathbf{F}_p[u]$ with m_h irreducible factors, (4.3) can be rewritten as

$$\left(\frac{\operatorname{disc} h}{p}\right) = (-1)^{\deg h - m_h}. \tag{4.4}$$

In this form, (4.4) is essentially a famous formula of Stickelberger's from algebraic number theory, on the quadratic character of the discriminant of a number field [6, Prop. 4.8.10]. What is crucial for us is not simply the formula itself but its interpretation: classically, one uses (4.3) – or rather the special case, (4.4) – with fixed $h \in \mathbf{Z}[u]$ and varying p. Instead we will use (4.3) with fixed p and varying $h \in \mathbf{F}_p[u]$. This is an idea that goes back to Swan [31], although he only considered separable h and did not bring out the Möbius aspect of the formula.

We will prove a generalization of Theorem 4.2 soon (see Theorems 4.7 and 4.9), but now we want to illustrate the utility of Theorem 4.2 in some simple examples. These examples show that $\mu(f(g))$ is sometimes an easily computable function of g.

Example 4.3 (Swan)**.** For g in $\mathbf{F}_2[u]$ let $h(u) = g(u)^8 + u^3$ in $\mathbf{F}_2[u]$, as in Example 1.5. If $g(0) = 0$ then $\mu(h) = 0$ and $h(u)$ is reducible. We shall now prove that if $g(0) = 1$ then $\mu(h) = 1$, so $h(u)$ is again reducible. (This explains Example 1.5.) The case $g = 1$ is trivial, so assume $\deg g > 0$. By (4.3), $\mu(h) = \left(\frac{\operatorname{disc} H}{2}\right)$ where H is any monic lift of h to $\mathbf{Z}[u]$. Choose $H = G^8 + u^3$, where G is a monic lift of g to $\mathbf{Z}[u]$. Note $8|\deg H$. Let $E/\mathbf{Q}$ be a splitting field of H. By (4.2), with h there equal to the polynomial H here,

$$\operatorname{disc} H = \prod_{H(\gamma)=0} H'(\gamma) \equiv \prod_{H(\gamma)=0} 3\gamma^2 = 3^{\deg H} H(0)^2 \bmod 8\mathscr{O}_E.$$

Since $H(0)$ is odd, we get $\operatorname{disc} H \equiv 1 \bmod 8$. Thus $\mu(h) = 1$.

Example 4.4. Let $h(u) = ug(u)^8 + 1$ in $\mathbf{F}_2[u]$ with $g(u) \neq 0$, as in Example 3.3. We show $\mu(h) = -1$. By Theorem 4.2, $\mu(h) = -\left(\frac{\operatorname{disc} H}{2}\right)$, with $H = uG^8 + 1$ and G a monic lift of g to $\mathbf{Z}[u]$. Let $E/\mathbf{Q}$ be a splitting field of H. As in the previous example, (4.2) implies

$$\operatorname{disc} H = \prod_{H(\gamma)=0} H'(\gamma) \equiv \prod_{H(\gamma)=0} G(\gamma)^8 \bmod 8\mathscr{O}_E.$$

Since $G(\gamma)^8 = -1/\gamma$, $\operatorname{disc} H \equiv H(0)^{-1} \equiv 1 \bmod 8$. Thus $\mu(h) = -1$.

Example 4.5. Let $p \neq 2$. We treat some examples related to Example 3.2 with $b = 1$. Pick g in $\mathbf{F}_p[u]$ with $\deg g \geq 1$. We will compute a formula for $\mu(g^p + u)$ that exhibits dependence only on $\deg g \bmod 4$ and on the quadratic character of the leading coefficient of g. (There is also dependence on $p \bmod 4$, but p is fixed.) Set $h = g^p + u$, so h is separable over $\mathbf{F}_p$, and let $n = \deg g$. Since $n \geq 1$, the degree of h is $d = pn$. Let g have leading coefficient c, so h has leading coefficient $c^p = c$. Since $h'(u) = 1$, (4.2) shows

$$\operatorname{disc} h = \frac{(-1)^{pn(pn-1)/2}}{c^{pn}} = \frac{(-1)^{n(pn-1)/2}}{c^n}. \tag{4.5}$$

Since $d \equiv n \bmod 2$, (4.3) says

$$\mu(g^p + u) = (-1)^n \left(\frac{\operatorname{disc} h}{p}\right) = (-1)^n \left(\frac{c}{p}\right)^n \left(\frac{-1}{p}\right)^{n(n+1)/2}. \tag{4.6}$$

When $n = \deg g$ is odd, $\mu(g^p + u) = -(\frac{c}{p})(\frac{-1}{p})^{(n+1)/2}$. As g runs over polynomials of odd degree n, this formula shows $\mu(g^p + u)$ is 1 and -1 equally often, with the Möbius value determined by the leading coefficient of g.

When $n = \deg g$ is even, $\mu(g^p + u) = (\frac{-1}{p})^{n/2}$. In particular, $\mu(g^p + u) = 1$ when $n \equiv 0 \bmod 4$. This explains the 0's in Table 1.2 and some of the 0's in Table 3.2. When $n \equiv 2 \bmod 4$, $\mu(g^p+u) = 1$ for all g of degree n when $(\frac{-1}{p}) = 1$. This explains the remaining 0's in Table 3.2. On the other hand, if $(\frac{-1}{p}) = -1$ then $\mu(g^p+u) = -1$ for all g with degree $\equiv 2 \bmod 4$. This is a total bias towards Möbius value -1, and numerics suggest that for varying g with fixed large degree n such that $n \equiv 2 \bmod 4$ (and $p \equiv 3 \bmod 4$), $g^p + u$ is irreducible approximately twice as often as predicted by (3.8). We saw such doubling in Table 1.2 when $n \equiv 2 \bmod 4$.

Definition 4.6. Let κ be a finite field. For a finite κ-algebra A, let $\mu(A) = (-1)^{\# \operatorname{Spec} A}$ if A is étale over κ (*i.e.*, reduced) and let $\mu(A) = 0$ otherwise.

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

Theorem 4.7. *Suppose κ is finite with odd characteristic, and let χ_κ be the quadratic character on $\kappa^\times$, with $\chi_\kappa(0) = 0$. For any finite κ-algebra A,*

$$\mu(A) = (-1)^{\dim_\kappa A} \chi_\kappa(\operatorname{disc}_\kappa A). \tag{4.7}$$

Proof. Both sides of (4.7) vanish when A is not étale over κ, so we may assume A is étale over κ. Both sides are multiplicative with respect to finite products in A. The case $A = 0$ is trivial, so we reduce to the case when $A = \kappa'$ is a finite extension of κ, and we want to prove

$$\chi_\kappa(\operatorname{disc}_\kappa \kappa') = (-1)^{d-1} \tag{4.8}$$

in $\mathbf{Z}$, where $d = [\kappa' : \kappa]$. Let γ be a field generator for κ' over κ. Since κ does not have characteristic 2, $\operatorname{disc}_\kappa \kappa'$ is a square in κ precisely when a generator for $\operatorname{Gal}(\kappa'/\kappa)$ acts as an even permutation on the κ-conjugates of γ. Since this permutation of the roots is a d-cycle, its sign is $(-1)^{d-1}$. ■

Remark 4.8. Theorem 4.7 and its proof carry over *verbatim* to finite algebras over any perfect field k with characteristic not 2 and with only cyclic Galois extensions; *e.g.*, we could take $k = \mathbf{C}((X))$. See [25, XIII, Exercise 3] for artificial examples in positive characteristic.

The proof of Theorem 4.7 works for étale algebras A in characteristic 2 if we formulate the result in terms of signs of certain permutations rather than in terms of quadratic characters of certain discriminants. (See [13, p. 237] for an application of this idea.) For our purposes, the role of discriminants is critical and therefore we need an analogue of Theorem 4.7 in characteristic 2 that involves discriminants. This analogue will use a lifting of A into characteristic 0. We shall now formulate a setup for finite κ with arbitrary characteristic (which for odd characteristic will recover a reformulation of Theorem 4.7).

Let κ be any finite field (of characteristic p, say), F the unramified extension of $\mathbf{Q}_p$ with residue field κ, and $W = W(\kappa)$ the valuation ring of F. (In other words, W is the ring of Witt vectors of κ.) We extend Theorem 4.7 to all characteristics by using finite flat liftings of A over W; *i.e.*, finite flat W-algebras $\widetilde{A}$ such that $\widetilde{A}/p\widetilde{A}$ is isomorphic to A as κ-algebras. For instance, a finite flat lifting of $\kappa[u]/(h(u))$ over W is $W[u]/(H(u))$, where $H \in W[u]$ satisfies $H \bmod p = h$ and $\deg H = \deg h$. By Hensel's lemma, if A is étale over κ then $\widetilde{A}$ exists (and is finite étale over W) and is unique up to unique W-isomorphism. If A is not étale over κ, a finite flat lifting of A over W may not exist (see [2, Example 3.2(4)]).

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

Here is a Möbius formula using liftings to characteristic 0.

Theorem 4.9. *For any finite κ-algebra A that admits a finite flat lifting $\widetilde{A}$ of A over W,*

$$\mu(A) = (-1)^{\dim_\kappa A}\widetilde{\chi}(\operatorname{disc}_W \widetilde{A}), \tag{4.9}$$

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

$$(\kappa^\times \times (1+4W))/((\kappa^\times \times (1+4W)) \cap (W^\times)^2) \simeq (1+4W)/((1+4W) \cap (W^\times)^2) \tag{4.10}$$

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

Before we prove Theorem 4.9, we make some remarks on the case $\operatorname{char}(\kappa) = 2$.

Remark 4.10. When κ has characteristic 2, we do not need to extend $\widetilde{\chi}$ to $1 + 2W$ or to all of $W^\times$, and there is no canonical extension anyway. Note that $(1 + 4W) \cap (W^\times)^2$ is the index-2 kernel of

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

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

Proof. (of Theorem 4.9) The case $A = 0$ is trivial. Since the reduction of $\operatorname{disc}_W \widetilde{A}$ modulo pW is $\operatorname{disc}_\kappa A$, (4.9) is trivial when A is non-étale over κ. (All we need to know about $\widetilde{\chi}$ here is that, by definition, it vanishes on pW.)

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

$$\widetilde{\chi}(\operatorname{disc}_W(W_d)) = (-1)^{d-1}. \tag{4.11}$$

By the definition of $\widetilde{\chi}$, this formula says that $\mathrm{disc}_W W_d$ is a square in $W^\times$ if and only if d is odd. This criterion for being a square is proved via the argument used to prove (4.8). ∎

Remark 4.11. Theorem 4.9 and its proof apply with κ replaced with any perfect field k of positive characteristic such that all finite Galois extensions of k are cyclic. When k has characteristic 2, Artin-Schreier theory ensures that the subgroup $\{x^2+x \mid x \in k\}$ has index ≤ 2 in k. However, there is no description of this subgroup akin to Remark 4.10 when k is infinite.

Taking $A = \kappa[u]/(h)$ for nonzero $h \in \kappa[u]$, Theorems 4.7 and 4.9 specialize to say

$$\mu(h) = \begin{cases} (-1)^{\deg h}\chi(\mathrm{disc}_\kappa h), & \text{if } \kappa \text{ has odd characteristic,} \\ (-1)^{\deg h}\widetilde{\chi}(\mathrm{disc}_W H), & \text{if } \kappa \text{ has any characteristic,} \end{cases} \tag{4.12}$$

where χ and $\widetilde{\chi}$ are described in Theorems 4.7 and 4.9, and H is a lifting of h into $W[u]$ with $\deg H = \deg h$.

Remark 4.12. Our formula in (4.12) for the case of characteristic 2 uses a discriminant in characteristic 0. There is an intrinsic characteristic 2 variant of the discriminant, due to Berlekamp [1] (and developed by later authors, such as Wadsworth [33]), but we have not found this to be useful for our purposes.

Example 4.13. Let κ be a finite field with characteristic p. For nonconstant g in $\kappa[u]$, $\mu(g^{4p}+u) = 1$. Indeed, for $p \neq 2$, this follows from (4.12) because $\mathrm{disc}(g^{4p}+u)$ is a square in κ by (4.2). For $p = 2$, let $W = W(\kappa)$. By (4.2),

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

Therefore, $\widetilde{\chi}(\mathrm{disc}_W(G^8+u)) = 1$ when G is a polynomial in $W[u]$ with positive degree and unit leading coefficient. Thus, by (4.12), $\mu(g^{4p}+u) = 1$ for nonconstant $g \in \kappa[u]$ when $p = 2$.

Example 4.14. Let κ be a finite field with characteristic $p \neq 2$. Generalizing (4.6), for nonconstant $g = cu^n + \cdots \in \kappa[u]$ we see via (4.12) that

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

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

We also find

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

In particular, for fixed $n \geq 1$, $\mu(g^p+u^2)$ is equal to 1 as often as it is equal to -1. Therefore there is no Möbius bias, in contrast with $\mu(g^p+u)$ when $\deg g$ is even and $-1 \in \kappa^\times$ is a square. Numerical tests over $\mathbf{F}_3$, $\mathbf{F}_5$, $\mathbf{F}_7$, and $\mathbf{F}_9$ suggest that (3.8) is correct for $f(T) = T^p + u^2$.

Example 4.15. Let κ have size q and characteristic p. Choose an integer b such that $1 < b < 4q$ and $(b, p(q-1)) = 1$ (*e.g.*, $b = 2q-1$). Then the polynomial $f(T) = T^{4q} + u^b$ is irreducible in $\kappa(u)[T]$ by [18, p. 297] and has no local obstructions, but $f(g)$ is reducible in $\kappa[u]$ for every $g \in \kappa[u]$. Indeed, this holds when $g = c$ is constant since $u^b + c$ is non-linear and has a root. When g is nonconstant, then $f(g)$ has u as a multiple factor if $g(0) = 0$ and (4.12) implies $\mu(f(g)) = 1$ if $g(0) \neq 0$.

When $q = 2$ and $b = 3$, we recover Example 4.3.

Example 4.16. Consider $f(T) = T^8 + u^3$ on $\kappa[u]$ with κ of size $q = 2^m$. Extending the work in Example 4.3, the reader can check that as g runs over $\kappa[u]$, $\mu(f(g))$ only has values 0 or 1, unless m is even and g is a (constant) noncube in $\kappa^\times$. In particular, the specializations of $T^8 + u^3$ on $\kappa[u]$ are irreducible only finitely many times (with fixed κ).

This completes our discussion on generalities about the Möbius function over finite fields. Theorems 4.7 and 4.9 are also important in a version of the Hardy–Littlewood conjecture for function fields of higher genus curves, as we will explain in [9]. In the present paper, we will prove a refinement of (4.12) when $h = f(g)$ with $f \in \kappa[u][T^p]$ nonzero and fixed and $g \in \kappa[u]$ varying. Our main results in this direction are Theorems 5.7, 6.4, 7.1, and 8.11 (and Corollary 8.12).

5. Discriminants and Resultants

For nonconstant $f \in \kappa[u][T^p]$, we wish to understand the behavior of $\mu(f(g))$ as g varies in $\kappa[u]$ with large degree. The formulas (3.3) and (4.12) suggest that we should study $\operatorname{disc}(f(g))$ as an algebraic function of varying g with large but fixed degree. Following Swan [31], we will find it useful to work with resultants and not discriminants. The relation between resultants and discriminants is given by the formula

$$\operatorname{disc} h = \frac{(-1)^{d(d-1)/2} R(h, h')}{(\operatorname{lead} h)^{2d-1}}. \tag{5.1}$$

Here $d = \deg h \geq 1$ and $R(h, h')$ is the resultant of h and h'. We now review some of the basic formalism of resultants.

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

$$R(h_1, h_2) = (\operatorname{lead} h_1)^{\deg h_2} \prod_{h_1(\alpha)=0} h_2(\alpha) \tag{5.2}$$

with the product running over the roots of h_1 (counted with multiplicity) in a splitting field over the fraction field of A. In [18, p. 200], an expression for $R(h_1, h_2)$ is given as a *universal* determinant in the coefficients of h_1 and h_2. An essential aspect of this universal formula is that the size of the determinant defining the resultant depends on the degrees of h_1 and h_2. We may write $R_{d_1,d_2}(h_1, h_2)$ to indicate that h_j is being treated as a polynomial of degree d_j for the resultant calculation via a universal determinant; in some later considerations it will be very natural to use $R_{d_1,d_2}(h_1, h_2)$ when $d_j > \deg h_j$ for some j, and so we make the *convention* that when a resultant $R(h_1, h_2)$ appears without degree subscripts then it is defined in terms of the actual degrees of its arguments if h_1 and h_2 are nonzero. We also agree to define $R(h_1, h_2) = 0$ when at least one h_j vanishes. This latter definition is compatible with universal determinants that define resultants (when the zero polynomial is assigned whatever nonnegative degree we please).

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

$$R_{d_1,d_3}(h_1, h_2) = (\operatorname{lead} h_1)^{d_3 - d_2} R_{d_1,d_2}(h_1, h_2). \tag{5.3}$$

Thus, giving the second polynomial h_2 a fake higher degree d_3 may change the resultant, although it does not change the property of vanishing or nonvanishing for the resultant (we only work with resultants over domains, not over arbitrary commutative rings). The two resultants in (5.3) agree when h_1 is monic, no matter what d_3 is. Beware that (5.2) is valid as written when h_2 is given a fake higher degree (still denoted $\deg h_2$), but it is generally not valid when h_1 is given a fake higher degree; also keep in mind that in general $R(h_1, h_2)$ and $R(h_2, h_1)$ are related by a sign (the precise sign-factor will be recorded shortly in a list of standard algebraic properties of resultants).

Warning. Failure to remember that the construction of resultants is sensitive to degrees can lead to errors when standard formulas are used carelessly, especially in specialization arguments. For instance, (5.1) is a specialization of a universal polynomial identity over $\mathbf{Z}$, where h has degree $d > 0$ and h' has degree $d-1$. The resultant in (5.1) is a $(d+(d-1))$-dimensional determinant, even under specialization to characteristic p where h' may have degree less than $d-1$ (perhaps even $h' = 0$). We should write the resultant in (5.1) as $R_{d,d-1}(h, h')$ to remind us of the dimensions of the determinant. If $\deg h' < d-1$, then h' must be given fake degree $d-1$, using initial coefficients that are equal to 0. This is consistent with (5.1) when $h' = 0$ (and $h \neq 0$ with $\deg h > 0$).

Example 5.1. Let $f(T) = T^9 + (2u^2+u)T^6 + (2u+2)T^3 + u^2 + 2u + 1$ in $\kappa[u][T]$, where κ has characteristic 3 ($\kappa = \mathbf{F}_3$ is Example 3.4). For nonconstant $g = cu^n + \dots$ in $\kappa[u]$ with $c \neq 0$, $f(g)$ has degree $9n$ and $f(g)' = (\partial_u f)(g)$ has degree $6n+1 < 9n-1$. The "true" resultant of $f(g)$ and $(\partial_u f)(g)$ is $R_{9n,6n+1}(f(g), f(g)')$, but the resultant needed to compute $\operatorname{disc} f(g)$ in (5.1) is

$$R_{9n,9n-1}(f(g), f(g)') = (c^9)^{3n-2} R_{9n,6n+1}(f(g), (\partial_u f)(g)). \tag{5.4}$$

Thus

$$\operatorname{disc} f(g) = \frac{(-1)^{n(n-1)/2}(c^9)^{3n-2} R_{9n,6n+1}(f(g), (\partial_u f)(g))}{c^{9(18n-1)}}. \tag{5.5}$$

If we forget that (5.1) is a universal identity over $\mathbf{Z}$ then we overlook the factor of $(c^9)^{3n-2}$. This power of c affects the quadratic nature of the right side of (5.5), so in view of (4.12) such an error would be serious.

Resultants have several useful algebraic properties. We summarize six of them without proof, as in [31], and we include (5.3) in the list. In this list, polynomials are nonzero and have coefficients in a domain A.

(1) $R(h_1, h_2) = (-1)^{(\deg h_1)(\deg h_2)} R(h_2, h_1)$.
(2) $R(h_1, h_2)$ is bimultiplicative: $R(h_1h_3, h_2) = R(h_1, h_2)R(h_3, h_2)$ and $R(h_1, h_2h_3) = R(h_1, h_2)R(h_1, h_3)$.
(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 $c \in A$.
(4) $R(c, h) = R(h, c) = c^{\deg h}$ for $c \in A$, $h \neq 0$. Thus, $R(c_1, c_2) = 1$ for $c_1, c_2 \neq 0$ in A.
(5) When h_1 has degree d_1, h_2 has degree d_2, and $d_3 \geq d_2$,

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

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

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

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

We call property (6) the *quasi-periodicity* of the resultant (in its second argument). When M is monic, $R(M, h)$ is genuinely periodic in h, with period (M). More generally (and of greater relevance to our work), for monic M in $A[u]$ and any $b(T) \in A[u][T]$, $R(M, b(h))$ is genuinely periodic in h. Swan's definition of $R(h_1, h_2)$ in [31] is what we call $R(h_2, h_1)$, so property (1) warns us that any comparison with [31] must keep this distinction in mind.

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

Example 5.2. Let $f(T) = T^{12} + (u+1)T^6 + u^4 \in \kappa[u][T]$ with finite κ of characteristic 3. (Example 1.1 is $\kappa = \mathbf{F}_3$.) Let q denote the size of κ and let χ be the quadratic character on $\kappa^\times$, with $\chi(0) = 0$. We shall compute $\mu(f(g))$ when $n = \deg g \geq 1$.

Since $4|\deg(f(g))$ and $\operatorname{lead}(f(g))$ is a square, (4.12) and (5.1) with $h = f(g)$ give

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

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

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

so

$$\mu(f(g)) = \chi(\operatorname{disc} f(g)) = \chi(g(0))^2 \chi(g(1)^2+1). \tag{5.6}$$

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

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

in $\kappa[u][T]$. This polynomial was already met over $\mathbf{F}_3[u]$ in Examples 3.4 and 5.1. We will compute a formula for $\mu(f(g))$ as g runs over nonconstant polynomials in $\kappa[u]$. The argument is long compared to Example 5.2, but at the same time it is more indicative of the general case, and thus is more instructive.

For nonconstant $g(u) = cu^n + \cdots$ with degree $n \geq 1$, we have $\deg(f(g)) = 9n$, so $\mu(f(g)) = (-1)^{9n}\chi(\operatorname{disc} f(g))$ by (4.12). By (5.1), (5.3), and (5.4),

$$\begin{aligned}\operatorname{disc} f(g) &= \frac{(-1)^{9n(9n-1)/2} R_{9n,9n-1}(f(g), f(g)')}{(c^9)^{18n-1}} \\ &= \frac{(-1)^{n(n-1)/2}(c^9)^{3n-2} R(f(g), (\partial_u f)(g))}{(c^9)^{18n-1}},\end{aligned}$$

so

$$\mu(f(g)) = (-1)^n (\chi(-1))^{n(n-1)/2} \chi(c)^{n+1} \chi(R(f(g), (\partial_u f)(g))). \tag{5.7}$$

We now compute a universal formula for $R(f(g), (\partial_u f)(g))$ in five steps, working over any field (or even domain) of characteristic 3. The formula is given in (5.12) as an algebraic identity, so for the purposes of the following calculation we may take g to be the universal polynomial of degree n over a field of characteristic 3 (so g has coefficients in a rational function field of transcendence degree $n+1$ over $\mathbf{F}_3$). In particular, the operation of division by $g(2)$ in Step 1 is not problematic.

Step 1. Explicitly,

$$f(g) = g^9 + (2u^2+u)g^6 + (2u+2)g^3 + u^2 + 2u + 1, \quad (\partial_u f)(g) = (u+1)g^6 + 2g^3 + 2u + 2. \tag{5.8}$$

Using (5.8), write $R(f(g), (\partial_u f)(g)) = (-1)^n R((\partial_u f)(g), f(g))$ to make the lower-degree term $(\partial_u f)(g)$ appear as the first argument. We want to simplify the resultant by quasi-periodicity, but the leading terms in (5.8) suggest it is easier to reduce $(u+1)f(g)$, rather than $f(g)$, modulo $(\partial_u f)(g)$. Therefore we apply bimultiplicativity to introduce a factor of $u+1$:

$$R(f(g), (\partial_u f)(g)) = \frac{(-1)^n R((\partial_u f)(g), (u+1)f(g))}{R((\partial_u f)(g), u+1)} = \frac{(-1)^n R((\partial_u f)(g), (u+1)f(g))}{g(2)^3}. \tag{5.9}$$

Treating g as if it is generic (so $g(2)$ is a unit) ensures that (5.9) is a meaningful (and correct) algebraic formula. Our derivation of (5.9) used bimultiplicativity to create convenient leading terms for quasi-periodicity. This computational idea will be used again in Step 3.

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

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

The nonzero constant in front will disappear when we apply χ as part of (5.7).

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

$$\frac{R((u^2+2u+2)g^3 + u + 1, (u^2+2u+2)(g^6 + u^2 g^3 + u + 1))}{R((u^2+2u+2)g^3 + u + 1, u^2 + 2u + 2)}. \tag{5.10}$$

Step 4. The denominator in (5.10) is 1 by quasi-periodicity (switch the two terms, which introduces no sign, and then reduce mod $u^2 + 2u + 2$). As for the numerator,

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

so the numerator of (5.10) is $(c^3)^{3n+1} R((u^2+2u+2)g^3 + u + 1, (2u+2)g^3 + 2u^2 + u + 2)$. The resultant factor equals

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

Putting everything together into (5.9), we have a cancellation of $g(2)^3$ and obtain

$$R(f(g), (\partial_u f)(g)) = c^{45n-3}(g(1)^2 + g(1) + 2)^3 R(g^3 + u + 1, (u^2+2u+2)g^3 + u + 1). \tag{5.11}$$

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

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

Feeding this into (5.11) gives the resultant formula

$$R(f(g),(\partial_u f)(g)) = c^{54n-6}(g(1)^2+g(1)+2)^3 g(2)^9. \tag{5.12}$$

Inserting (5.12) into (5.7), we find our Möbius formula:

$$\mu(f(g)) = (-1)^n \chi(-1)^{n(n-1)/2}\chi(c)^{n+1}\chi(g(1)^2+g(1)+2))\chi(g(2)) \tag{5.13}$$

for nonconstant g in $\kappa[u]$. This depends on $g \bmod (u-1)(u-2)$, $\deg g \bmod 4$, and the quadratic character of the leading coefficient of g. Taking $\kappa = \mathbf{F}_3$, we will show in Example 9.9 that (5.13) is numerically compatible with the statistics in Table 3.4.

Motivated by the goal of making patterns in $\mu(f(g))$ provable when $f(T)$ is irreducible and inseparable, as in Examples 5.2 and 5.3, we discovered that the function $g \mapsto \mu(f(g))$ admits a periodicity in g when f is squarefree with irreducible factors that are inseparable (in T). Before stating our periodicity theorem, we need a lemma.

Lemma 5.4. *Let F be perfect of characteristic $p > 0$.*

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

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

Note that if $f \notin F$ then $f(T)$ cannot lie in $F[u^p,T]$ under either hypothesis in the lemma, so $\partial_u f \neq 0$ in such cases. It may happen that $\partial_u f$ is constant; *e.g.*, $f = u^pT^p+u+1$.

The second case in Lemma 5.4 will be used only when $p=2$.

Proof. The case $f \in F^\times$ is trivial, so we may assume $f \notin F$.

1) Since f is squarefree, the irreducible factors ϕ of f in $F[u,T]$ must all lie in $F[u,T^p]$, so by perfectness of F none can lie in $F[u^p,T]$; thus, such ϕ are separable over $F(T)$. Hence, if $Z_f \subseteq \mathbf{A}^2_F$ is the (reduced) zero scheme of f then the projection pr_T from Z_f onto the T-axis is quasi-finite and flat, as well as generically étale. Thus, the non-étale locus of pr_T is finite. This locus is exactly where Z_f meets the zero locus of $\partial_u f$.

2) Now suppose $f \in F[u,T]$ is nonconstant and $f(T^p)$ is squarefree. We first check that f and $\partial_u f$ have no common factor in $F[u]$. Write $f = b(u)h$ where $b \in F[u]$ and $h \in F[u,T]$ has no irreducible factors in $F[u]$. Since b must be squarefree, $\gcd(b, \partial_u b) = 1$. Since $\partial_u f \equiv (\partial_u b)\cdot h \bmod b$, no irreducible factor of b can divide $\partial_u f$. This rules out the possibility of irreducible common factors of f and $\partial_u f$ in $F[u]$.

Since f and $\partial_u f$ are nonzero, to show that they have no common factor in $F[u,T]$ with positive T-degree it is equivalent to proving $R_{F[u]}(f,\partial_u f) \neq 0$. The case $\deg_T f = 0$ is clear, so we assume f has positive T-degree. We induct on $\deg_T f$ and the number of irreducible factors of f. If f is irreducible and $R_{F[u]}(f,\partial_u f) = 0$, then f and $\partial_u f$ have a common root in an extension of $F(u)$, so $\partial_u f = \beta(u) f$ for some $\beta(u) \in F(u)^\times$ since $\partial_u f \neq 0$. Since f is irreducible in $F[u,T]$ with positive T-degree, $\beta \in F[u]$. A comparison of u-degrees on both sides of the equation $\partial_u f = \beta(u) f$ gives a contradiction.

More generally, suppose $f = f_1 f_2$, where the f_j's are non-constant, so each $f_j(T^p)$ is squarefree in $F[u, T]$ and $\gcd(f_1, f_2) = 1$ in $F(u)[T]$. Bimultiplicativity and periodicity for resultants gives

$$R_{F[u]}(f, \partial_u f) = (*) R_{F[u]}(f_1, \partial_u f_1) R_{F[u]}(f_2, \partial_u f_2) R_{F[u]}(f_1, f_2)^2,$$

where $(*)$ is a nonzero factor in F built up from signs and leading coefficients. The nonvanishing follows by induction. ■

Definition 5.5. If $f_1, f_2 \in F[u, T]$ are two nonzero polynomials over a perfect field F such that their zero loci Z_{f_1} and Z_{f_2} in $\mathbf{A}^2_F$ have finite intersection, $M^{\rm geom}_{f_1,f_2} \in F[u]$ is the monic separable polynomial whose zero locus is the projection of $Z_{f_1} \cap Z_{f_2}$ onto the u-axis (so $M^{\rm geom}_{f_1,f_2} = 1$ if $Z_{f_1} \cap Z_{f_2}$ is empty, such as when some f_j lies in $F^\times$). When $f \in F[u, T]$ is nonconstant, define $M^{\rm geom}_f = M^{\rm geom}_{f,\partial_u f}$ when this makes sense (*i.e.*, when $\partial_u f \neq 0$ and $Z_f \cap Z_{\partial_u f}$ is finite).

Note that the formation of $M^{\rm geom}_{f_1,f_2}$ commutes with extension of the perfect ground field. When $\mathrm{lead}_T\, f$ is separable, $M^{\rm geom}_{f_1,f_2}$ is the radical of the resultant $R_{F[u]}(f_1, f_2)$. (We saw in Remark 1.7 that this need not hold when $\mathrm{lead}_T\, f$ is not separable).

For $f \in F[u, T]$ with $f \notin F$, Lemma 5.4 gives some sufficient conditions for $M^{\rm geom}_f$ to be defined when F has positive characteristic. The next lemma gives a general geometric criterion in any characteristic.

Lemma 5.6. *If F is perfect with arbitrary characteristic and $f \in F[u, T]$ is not in F, then the zero loci of f and $\partial_u f$ in $\mathbf{A}^2_F$ have finite intersection if and only if f is squarefree in $F[u, T]$ with no irreducible factors in $F[T]$ and the projection*

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

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

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

Proof. Necessity of the conditions that f be squarefree and have no irreducible factors in $F[T]$ is clear. Granting these conditions, the plane curve Z_f is reduced (hence geometrically reduced since F is perfect) and its projection to the T-axis is quasi-finite and hence flat. Thus, the property of pr_T being étale at a point of Z_f may be checked on the geometric fibers of pr_T. Extending scalars to an algebraic closure of F, we thereby see that the non-étale locus for pr_T is where $f = 0$ meets $\partial_u f = 0$ in $\mathbf{A}^2_F$. This completes the proof of the desired equivalence, and also yields the asserted relationship between $M^{\rm geom}_f$ and the non-étale locus of pr_T. ■

Here is our main result in odd characteristic. The proof will be given in §7, using Theorem 6.4 and Theorem 7.1.

Theorem 5.7. *Let κ be a finite field with odd characteristic p, and let χ be the quadratic character of $\kappa^\times$. Fix a nonzero $f(T) \in \kappa[u][T^p]$ that is squarefree in $\kappa[u][T]$. Assume $f \notin \kappa$.*

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

$$g_1 \equiv g_2 \bmod M^{\rm geom}_f, \quad n_1 \equiv n_2 \bmod 4, \quad \chi(c_1) = \chi(c_2) \Longrightarrow \mu(f(g_1)) = \mu(f(g_2)). \tag{5.14}$$

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

If -1 is a square in κ or $\deg_T f$ is even, the second congruence in (5.14) *may be relaxed to $n_1 \equiv n_2 \bmod 2$.*

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

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

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

Remark 5.8. The finiteness of κ in Theorem 5.7 may be relaxed in odd characteristic exactly as in Remark 4.8 without changing the proofs, though we do not know any interesting examples of this generalized theorem with infinite κ.

Although Theorem 5.7 does not say that $g \mapsto \mu(f(g))$ is genuinely periodic in g, we will refer to any nonzero M satisfying the role of M^{geom}_f in (5.14) as a *modulus* for $\mu(f(g))$. Since any congruence class in $\kappa[u]/(M)$ may be represented by a polynomial of any large degree with any desired leading coefficient, it is a trivial exercise with the Chinese remainder theorem to check that for any two moduli M_1 and M_2 for $\mu(f(g))$, $\gcd(M_1, M_2)$ is also a modulus. It therefore follows trivially from (5.14) that $M^{\min}_{f,\kappa}$ divides all other moduli for $\mu(f(g))$. The fact that $M^{\min}_{f,\kappa''} = M^{\mathrm{geom}}_f$ for all finite extensions κ''/κ containing some finite extension κ'/κ will require an understanding of how we prove (5.14).

Examples 5.2 and 5.3 illustrated some techniques that will be used in the proof of Theorem 5.7. The following examples focus only on explicit Möbius formulas, illustrating the conclusions of Theorem 5.7.

Example 5.9. The variation of $M^{\min}_{f,\kappa}$ as κ grows is interesting. Since $M^{\min}_{f,\kappa} | M^{\mathrm{geom}}_f$, there are only finitely many possibilities for $M^{\min}_{f,\kappa'}$ as κ' varies over finite extensions of κ. We now give an example where $M^{\min}_{f,\kappa} \neq M^{\mathrm{geom}}_f$

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 5.7 shows

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

Note that $\chi(g(0)^2 + 1)^2$ is not always 1 because it may vanish. This Möbius formula, like (5.6), has no dependence on $\deg g \bmod 4$ or on the quadratic character of the leading coefficient of g. Since $R(u^2+1, f(g))$ only depends on g modulo u^2+1 (by quasi-periodicity of resultants), we see that $\mu(f(g))$ only depends on g modulo $u(u-1)(u^2+1)$. (Since $R_{\kappa[u]}(f, \partial_u f) = u^{12}(u-1)^{18}(u^2+1)^{12}$, we have $M^{\mathrm{geom}}_f = u(u-1)(u^2+1)$.) When $[\kappa : \mathbf{F}_3]$ is odd, $g(0)^2 + 1$ is nonzero, so $\mu(f(g))$ only depends on g modulo $(u-1)(u^2+1)$ for such κ. This illustrates that the minimal modulus in Theorem 5.7 can be sensitive to a change in the base field κ.

Example 5.10. Let $f(T) = T^{12} + (2u^4 + 2u^3 + 2u^2 + u + 1)T^6 + 2u^3 + 2u^2 + u$ in $\kappa[u][T]$, where κ has characteristic 3. (See Example 3.5 for $\kappa = \mathbf{F}_3$.) For nonconstant g in $\kappa[u]$, the proof of Theorem 5.7 shows

$$\mu(f(g)) = \chi(g(0)^2+1)\chi(R(u^2+u+2, (2u+2)g^6+u+1))\chi(R(u^2+2u+2, 2ug^6+u+1))^2.$$

Thus $\mu(f(g))$ only depends on $g \bmod u(u^2+u+2)(u^2+2u+2)$. In this case, $M^{\mathrm{geom}}_f = u(u^2+u+2)(u^2+2u+2)$.

Example 5.11. Let $f(T) = (2u^2+u+3)T^{15} + (4u^2+u+3)T^5 + 4u^2+u+3$ in $\kappa[u][T]$, where κ has characteristic 5. (See Example 3.6 for $\kappa = \mathbf{F}_5$.) For $g(u) = cu^n + \cdots$ in $\kappa[u]$ with degree $n \geq 1$, the proof of Theorem 5.7 shows

$$\mu(f(g)) = (-1)^n \chi(3)^{n+1} \chi(c)^n \chi(g(0)^3 + g(0) + 1)\chi(g(-1)^3 - g(-1) - 1). \tag{5.15}$$

In this formula, we see dependence on n mod 2, on $\chi(c)$, and on a congruence condition on g modulo $u(u+1)$. In this case, $M_f^{\mathrm{geom}} = u(u+1)$.

6. A resultant formula via geometry

We will obtain Theorem 5.7 from a periodicity property for resultants over arbitrary perfect fields. We indulge in the following notational device: for a field F and a nonzero $M \in F[u]$, we write $F[u]/(M)$ to denote the vector-scheme of remainders upon long division by M over F-algebras A. That is, $F[u]/(M)$ is viewed as an affine space of dimension $\deg M$, whose coordinates arise from coefficients of u^i for $0 \leq i < \deg M$. (This space is $\operatorname{Spec} F$ when $\deg M = 0$.) Such abuse of notation is standard for vector-schemes in the theory of algebraic groups. The context will indicate whether $F[u]/(M)$ denotes the affine space over $\operatorname{Spec} F$ or its F-valued points, the "usual" F-vector space $F[u]/(M)$.

We will also work with the scheme

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

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

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

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

$$\rho_{n,M} : \mathrm{Poly}_{n/F} \to F[u]/(M) \simeq \mathrm{Poly}_{\leq (\deg M - 1)/F} \tag{6.1}$$

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

$$\rho_{n,M} : \mathrm{Poly}_{n/F} \to \operatorname{Spec} F \tag{6.2}$$

is the structure map to a point.

Since $\deg(f(g))$ is determined by $n = \deg g$ for g of large degree (depending on f, as in (3.3)), there is a well-posed algebraic function

$$\mathrm{disc} \circ f : \mathrm{Poly}_{n/F} \to \mathbf{A}_F^1 \tag{6.3}$$

defined by $g \mapsto \mathrm{disc}(f(g))$ when n is sufficiently large; note that (6.3) does *not* extend to an algebraic function on $\mathrm{Poly}_{\leq n/F}$ (*cf.* Remark 1.9). Our aim is to understand the structure of the algebraic function (6.3) for f as in Lemma 5.6, and in particular the extent to which it factors through some remainder morphism $\rho_{n,M}$ for some nonzero $M \in F[u]$.

To exploit inductive arguments, it is convenient to re-interpret our discriminant problem as the study of the resultant $R(f(g), (\partial_u f)(g))$ for varying g of large (fixed) degree; the utility of this point of view is that it allows us to consider the more general algebraic function $\mathrm{Poly}_{n/F} \to \mathbf{A}_F^1$ 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 5.4 when $f \not\in F$). The merit of this generality is that we may separately vary f_1 and f_2. Restricting attention to finite or perfect F of positive characteristic is not adequate: our later work in characteristic 2 will use the present considerations with a 2-adic field F.

Let us now fix a pair of nonzero relatively prime elements $f_1, f_2 \in F[u,T]$, so the zero loci $Z_{f_1} = \{f_1 = 0\}$ and $Z_{f_2} = \{f_2 = 0\}$ are (possibly empty) curves in $\mathbf{A}^2_F$ with no common irreducible components. For $g \in F[u]$ of degree n, (3.3) gives the degree of $f_j(g) \in F[u]$ when $n \gg 0$. We give this formula the label $d_{j,n}$. That is,

$$d_{j,n} := (\deg_T f_j)n + \deg(\mathrm{lead}_T\, f_j). \tag{6.4}$$

The largeness of $n = \deg g$ that makes (3.3) hold for both f_1 and f_2 depends only on $\deg_T f_1$, $\deg_T f_2$, and the u-degrees of the coefficients of f_1 and f_2 when the f_j's are viewed as polynomials in T. See (3.5) for an explicit universal lower bound on n that makes (3.3) valid when g is a point of $\mathrm{Poly}_{n/F}$ with values in any F-algebra domain.

Fixing such large n, let

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

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

$$R_n(G) := R_{F[a_0,\dots,a_n]}(f_1(G), f_2(G)) \in F[a_0, \dots, a_n], \tag{6.5}$$

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

Lemma 6.1. *$R_n(G) \neq 0$.*

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

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

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

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

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

Definition 6.2. 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

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

Lemma 6.3. *Assume $n \geq 1$. For each $x \in Z_{f_1} \cap Z_{f_2}$ such that $F(x)/F$ is separable, $P_{x,n}$ is irreducible in the coordinate ring of* $\mathrm{Poly}_{\leq n/F}$. *If x and x' are two such distinct points, then $P_{x,n}$ and $P_{x',n}$ are not unit-multiples of each other in this coordinate ring.*

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

Proof. Since the extension $F(x)/F$ is finite separable and $P_{x,n}$ is a norm-form of a polynomial in $F(x)[a_0, \ldots, a_n]$ whose coefficients generate $F(x)$ over F (since $n \geq 1$), the irreducibility is obvious. If L/F is a finite Galois extension into which $F(x)$ admits an F-embedding, then over L we see that $P_{x,n}$ factors as a product of linear forms $P_{x_i,n}$ defined by the L-points x_i of $\mathbf{A}_F^2$ 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 6.3 applies to all $x \in Z_{f_1} \cap Z_{f_2}$. Recall that in Definition 5.5 we defined

$$M_{f_1,f_2}^{\mathrm{geom}}(u) := \prod_{u_x} \mathrm{N}_{F(u_x)/F}(u - u_x) \in F[u] - \{0\}, \tag{6.6}$$

where u_x runs over the distinct images of the x's on the u-axis. In particular, $M_{f_1,f_2}^{\mathrm{geom}} = 1$ 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 (6.6) and the definition $P_{x,n}(g) = \mathrm{N}_{F(x)/F}(g(u_x) - t_x)$ for $n \geq \deg g$ we see

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

where $n_j = \deg g_j$.

For $M := M_{f_1,f_2}^{\mathrm{geom}} \neq 0$, consider the division-algorithm morphism $\rho_{n,M}$ as in (6.1) and (6.2). Assume $Z_{f_1} \cap Z_{f_2}$ is nonempty, so $M \notin F$. Choose $x \in Z_{f_1} \cap Z_{f_2}$, so $M(u_x) = 0$. Clearly $P_{x,n} = P_{x,M} \circ \rho_{n,M}$ for the algebraic function $P_{x,M}$ on $\mathrm{Poly}_{\leq(\deg M - 1)/F}$ given by the norm construction $g \mapsto \mathrm{N}_{(F(x)\otimes_F F')/F'}(g(u_x) - t_x)$ for F-algebras F' and $g \in F'[u]$ with degree $\leq \deg M - 1$.

Theorem 6.4. *Let $f_1, f_2 \in F[u, T]$ be nonzero and relatively prime such that the zero-loci Z_{f_1} and Z_{f_2} of f_1 and f_2 in $\mathbf{A}_F^2$ have finite intersection. Assume that F is perfect. For $x \in Z_{f_1} \cap Z_{f_2}$ and n sufficiently large, there exist unique $b_n \in F^\times$ and integers $e_n \geq 0$ and $e_x > 0$ such that*

$$R_n(G) = b_n a_n^{e_n} \cdot \prod_x P_{x,n}^{e_x} = b_n a_n^{e_n} \cdot \prod_x P_{x,M}^{e_x} \circ \rho_{n,M} \tag{6.7}$$

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

The functorial construction of $R_n(G)$ as a universal resultant for large n (an alternative to the explicit definition (6.5)) only makes sense over $\mathrm{Poly}_{n/F}$ and not over $\mathrm{Poly}_{\leq n/F}$, so it does not seem possible to use geometric methods alone to determine how the discrete parameters e_n and b_n depend on n (though clearly b_n is generally sign-dependent on the ordering of the pair f_1 and f_2). In §7 we shall prove via algebraic methods that for large n, e_n is a linear polynomial in n and $b_n = \beta_0\beta_1^n$ for some $\beta_0, \beta_1 \in F^\times$. The products over $x \in Z_{f_1} \cap Z_{f_2}$ in (6.7) are understood to be 1 if $Z_{f_1} \cap Z_{f_2}$ is empty.

Proof. We will first establish a weaker identity

$$R_n(G) = b_n a_n^{e_n} \cdot \prod_x P_{x,n}^{e_{x,n}} \tag{6.8}$$

on $\mathrm{Poly}_{\leq n/F}$ for large n, with $b_n \in F^\times$, exponents $e_{x,n} > 0$ that *a priori* might depend on n, and an exponent $e_n \geq 0$ that is positive if and only if both $\deg_T f_j$'s are positive.

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

Since the irreducible $P_{x,n}$'s are not scalar multiples of a_n, to establish (6.8) it remains (by the Nullstellensatz) to show that the restriction of $R_n(G)$ to $\mathrm{Poly}_{n/F}$ has geometric zero locus equal to the union of the geometric zero loci of the $P_{x,n}$'s. If $\overline{F}/F$ is an algebraic closure, then since $F(x)/F$ is separable (as F is perfect) the irreducible factorization of $P_{x,n}$ in $\overline{F}[a_0, \ldots, a_n]$ is as the product of the $P_{x_i,n}$'s for the $\overline{F}$-points x_i of $\mathbf{A}_F^2$ over the physical point x. Thus, we may assume F is algebraically closed and we wish to prove that if $g \in F[u]$ has large exact degree n then the resultant of $f_1(u, g(u))$ and $f_2(u, g(u))$ vanishes if and only if $g(u_x) = t_x$ for some x in the intersection of the zero-loci Z_{f_j}. But this is obvious since the vanishing of the resultant says that $f_1(u, g(u))$ and $f_2(u, g(u))$ have a common root $u_0 \in F$, and then $x = (u_0, g(u_0))$ lies on both zero-loci Z_{f_j}. This completes the proof of (6.8).

It remains to prove that $e_{x,n}$ in (6.8) is independent of n for large n. Due to the simple behavior of $P_{x,n}$ under extension F'/F (since $F(x)/F$ is separable), we have $e_{x,n} = e_{x',n}$ for any F'-point x' of $\mathbf{A}_F^2$ lying over x. Thus, we may (and do) now assume that F is algebraically closed.

We will use deformation-theoretic reasoning to prove that the sequence $\{e_{x,n}\}$ for fixed x is monotone decreasing for large n, so this sequence eventually becomes constant. Our original proof proceeded by constructive methods. We are grateful to de Jong for suggesting that we work out a (non-effective) deformation-theoretic approach, since it turns out to adapt to higher genus (see [9]) while the constructive method does not. The reader who prefers constructive methods can rediscover our original proof by developing an expanded version of the proof of Theorem 7.1. This will lead to a constructive algebraic proof that $e_{x,n}$ is independent of large n. Such a proof appears to give a poor bound on how large n must be for the sequence $\{e_{x,n}\}$ to become constant.

The first step in the study of $e_{x,n}$ for fixed x is a description of resultants in terms of norms. For large n, the polynomial $f_j(G)$ has leading coefficient that is an $F^\times$-multiple of a power of a_n (possibly the power $a_n^0 = 1$ when $f_j \in F[u]$), so over the coordinate ring of $\mathrm{Poly}_{n/F}$ we see that $f_j(G)$ has degree $d_{j,n}$ (see (6.4)) and a unit leading coefficient. In particular,

$$F[a_0, \ldots, a_n, 1/a_n][u]/(f_j(G))$$

is a finite free module of rank $d_{j,n}$ over the coordinate ring $F[a_0, \ldots, a_n, 1/a_n]$ of $\mathrm{Poly}_{n/F}$.

Lemma 6.5. *For sufficiently large n, there exist $c_n \in F^\times$ and $\widetilde{e}_n \in \mathbf{Z}$ such that*

$$R_n(G) = c_n a_n^{\widetilde{e}_n} \mathrm{N}_{(\mathscr{O}_n[u]/(f_1(G)))/\mathscr{O}_n}(f_2(G))$$

in $\mathscr{O}_n = F[a_0, \ldots, a_n, 1/a_n]$.

Proof. Let $d_j = \deg_T f_j \geq 0$. By reduction to the universal case of unitary polynomials, and then factoring out unit leading coefficients, it suffices to prove that for universal monic polynomials

$$h_1 = u^{d_1} + a_{d_1-1}u^{d_1-1} + \cdots + a_0, \;\; h_2 = u^{d_2} + b_{d_2-1}u^{d_2-1} + \cdots + b_0$$

of degrees $d_1 \geq 0$ and $d_2 \geq 0$ over $\mathscr{O} = \mathbf{Z}[a_i, b_j]$ (so $h_k = 1$ if $d_k = 0$), the resultant $R_{\mathscr{O}}(h_1, h_2) \in \mathscr{O}$ is equal to the norm $\mathrm{N}_{(\mathscr{O}[u]/(h_1))/\mathscr{O}}(h_2)$.

Let K be the fraction field of the domain $\mathscr{O}$, so clearly h_1 and h_2 are separable over K. Let K'/K be a splitting field for the h_j's, so if $\{\alpha\}$ is the set of roots of h_1 in K' then by definition

$$R_{\mathscr{O}}(h_1, h_2) = R_K(h_1, h_2) = \prod_{\alpha} h_2(\alpha).$$

Since $K'[u]/(h_1) \simeq \prod_\alpha K'$ via $u \mapsto (\alpha)$ we also have

$$\mathrm{N}_{(\mathscr{O}[u]/(h_1))/\mathscr{O}}(h_2) = \mathrm{N}_{(K[u]/(h_1))/K}(h_2) = \mathrm{N}_{(K'[u]/(h_1))/K'}(h_2) = \prod_{\alpha} h_2(\alpha).$$

■

Consider the algebraic function $\mathrm{Poly}_{n/F} \to \mathbf{A}^1_F$ defined by

$$N_n : g \mapsto \mathrm{N}_{(F[u]/(f_1(g)))/F}(f_2(g)) \tag{6.9}$$

for $g \in \mathrm{Poly}_{n/F}(F')$ for F-algebras F'. Let $\eta_{x,n}$ be the codimension-1 generic point of the hypersurface $\{P^0_{x,n} = 0\}$ in the F-smooth variety $\mathrm{Poly}_{n/F}$, where $P^0_{x,n} = P_{x,n}|_{\mathrm{Poly}_{n/F}}$. By Lemmas 6.1 and 6.5, $N_n \neq 0$ and for large n we may identify $e_{x,n}$ with the order of N_n at $\eta_{x,n}$. We will prove that this order is a monotonically decreasing function of large n for a fixed $x = (u_x, t_x) \in Z_{f_1} \cap Z_{f_2}$. Since F is algebraically closed, so $x \in \mathbf{A}^2_F$ is F-rational, we may make an additive translation on T so that $t_x = 0$. Hence, the locus $\{P^0_{x,n} = 0\}$ is the space of g's of exact degree n such that $g(u_x) = 0$. For generic such g, clearly g has a simple zero at u_x and $g(u_{x'}) \neq t_{x'}$ for $x' \neq x = (u_x, 0)$. Obviously

$$\mathrm{ord}_{u_x}(f_1(g)) \geq r_x := \min_{j \leq \deg_T f_1} (\mathrm{ord}_{u_x}(\alpha_j) + j) > 0$$

where $f_1 = \sum \alpha_j(u)T^j \in F[u,T]$. Note that r_x is independent of n. By replacing g with a generic $F^\times$-multiple we can eliminate any cancellation of contributions from parts of order r_x in the sum $\sum \alpha_j g^j$ (work in the ring $F[u]/(u^{r_x+1})$). Thus, $\mathrm{ord}_{u_x}(f_1(g)) = r_x$ for a generic choice of closed point g in $\{P^0_{x,n} = 0\}$. Likewise, if we write $f_2 = \sum \beta_j(u)T^j \in F[u,T]$ and define

$$s_x = \min_{j \leq \deg_T f_2} (\mathrm{ord}_{u_x}(\beta_j) + j),$$

then $\mathrm{ord}_{u_x}(f_2(g)) = s_x$ for a generic choice of closed point $g \in \{P^0_{x,n} = 0\}$.

Fix large n_0, with $n_0 > \max(r_x, s_x)$, and choose a closed point $g_{n_0} \in \{P^0_{x,n_0} = 0\}$ such that $\mathrm{ord}_{u_x}(f_1(g_{n_0})) = r_x$ and $\mathrm{ord}_{u_x}(f_2(g_{n_0})) = s_x$. Taking $\varepsilon_j = c_j(u - u_x)^j$ for generic $c_j \in F^\times$, we may arrange that for all $n > n_0$ we have that

$$g_n := g_{n_0} + \varepsilon_{n_0+1} + \cdots + \varepsilon_n \in \{P^0_{x,n} = 0\} \tag{6.10}$$

satisfies $\mathrm{ord}_{u_x}(f_1(g_n)) = r_x$ and $\mathrm{ord}_{u_x}(f_2(g_n)) = s_x$ and that $g_n \notin \{P_{x',n} = 0\}$ for all $x' \neq x$, so in particular $f_2(g_n)$ is non-vanishing at all roots of $f_1(g_n)$ away from u_x. To be precise about the genericity of the c_j's, for any fixed n we may suppose that $(c_1, \ldots, c_n) \in F^n$ is generic (*i.e.*, it avoids any desired proper Zariski-closed condition on $\mathbf{A}^n_F$).

Let $V_n = \mathrm{Poly}_{\leq n/F}$ and $V^0_n = \mathrm{Poly}_{n/F}$, so a system of linear coordinates on V_n is given by the basis $\{y_0, \ldots, y_n\}$ dual to the basis $\{\varepsilon_0, \ldots, \varepsilon_n\}$ of $V_n(F)$, where $\varepsilon_0 \in F^\times$ and necessarily y_0 is an $F^\times$-multiple of the linear functional $P_{x,n} : g \mapsto g(u_x)$, and the element

$$g^{\mathrm{univ}}_n = 1 \otimes g_n + \sum_{0 \leq j \leq n} (y_j - y_j(g_n)) \otimes \varepsilon_j \in \mathscr{O}_{V^0_n, g_n} \otimes_F F[u]$$

is an algebraized universal deformation of g_n. Obviously $y_0(g_n) = 0$ and $y_1, \ldots, y_n$ restrict to a system of linear coordinates on the hyperplane $\{P_{x,n} = 0\} = \{y_0 = 0\}$ in V_n. In particular, the residue field $F(\eta_{x,n})$ at the generic point $\eta_{x,n}$ of $\{P^0_{x,n} = 0\}$ is canonically identified with $F(y_1, \ldots, y_n)$.

The quotient ring

$$(\mathscr{O}_{V^0_n, g_n} \otimes_F F[u])/(f_1(g^{\mathrm{univ}}_n)) \tag{6.11}$$

is finite and free as an $\mathscr{O}_{V^0_n, g_n}$-module since the leading coefficient of $f_1(g^{\mathrm{univ}}_n)$ is

$$1 + (y_n - y_n(g_n)) = y_n \in \mathscr{O}^\times_{V^0_n, g_n}$$

(as $y_n(g_n) = 1$, due to (6.10)). Thus, the maximal ideals of (6.11) are in bijective correspondence with the maximal ideals of the quotient $F[u]/(f_1(g_n))$ of (6.11) by the maximal ideal of $\mathscr{O}_{V^0_n, g_n}$. As we noted above, $f_2(g_n)$ is non-vanishing at all roots of $f_1(g_n)$ away from $u = u_x$, so the image of $f_2(g_n)$ in $F[u]/(f_1(g_n))$ has a unit component in all local factor-rings away from $u = u_x$. In particular, upon localizing (6.11) at $\eta_{x,n}$ and extending scalars to the completion $\widehat{\mathscr{O}}_{V^0_n, \eta_{x,n}} \simeq F(\eta_{x,n})[\![y_0]\!]$ (an isomorphism of F-algebras), we get a product-decomposition of rings

$$(\widehat{\mathscr{O}}_{V^0_n, \eta_{x,n}} \otimes_F F[u])/(f_1(g^{\mathrm{univ}}_n)) \simeq (\widehat{\mathscr{O}}_{V^0_n, \eta_{x,n}}[\![u - u_x]\!]/(f_1(g^{\mathrm{univ}}_n))) \times \Delta_n \tag{6.12}$$

such that the first factor-ring is local and finite free over the discrete valuation ring $\widehat{\mathscr{O}}_{V^0_n, \eta_{x,n}}$ and (by computing modulo the uniformizer y_0) has rank $\rho_{x,n}$ equal to the $F(\eta_{x,n})$-dimension of the u_x-factor of the finite $F(\eta_{x,n})$-algebra

$$(F(\eta_{x,n}) \otimes_F F[u])/(f_1(1 \otimes g_n + \sum_{0 < j \leq n} (y_j - y_j(g_n)) \otimes \varepsilon_j)). \tag{6.13}$$

Lemma 6.6. *The image of $f_2(g_n^{\rm univ})$ in Δ_n is a unit and $\rho_{x,n} = r_x$. In particular, the image of $f_1(g_n^{\rm univ})$ in $\widehat{\mathscr{O}}_{V_n^0,\eta_{x,n}} \widehat{\otimes}_F \widehat{\mathscr{O}}_{\mathbf{A}^1_F,u_x} = (F(\eta_{x,n})[\![y_0]\!])[\![u-u_x]\!]$ is a unit multiple of a unique Weierstrass polynomial of degree r_x in $(F(\eta_{x,n})[\![y_0]\!])[u-u_x]$.*

Proof. Consider the unique splitting of (6.13) into a product of two $F(\eta_{x,n})$-algebras with one factor having support equal to the section induced by $u = u_x$. This splitting is the reduction of (6.12), and by denominator-chasing in $F(\eta_{x,n})$ it may be extended to a splitting over some dense open $W_{x,n}$ of $\overline{\{\eta_{x,n}\}} = \{P^0_{x,n} = 0\}$ with factor rings that are finite and free over $\mathscr{O}_{W_{x,n}}$, so the rank $\rho_{x,n}$ can be computed upon generically specializing $y_1, \dots, y_n$ into F.

We have seen that at a generically-chosen closed point $g \in \{P^0_{x,n} = 0\}$ the F-finite quotient $F[u]/(f_1(g))$ has u_x-factor with F-dimension equal to r_x and $f_2(g)$ has a unit component in all other local factors. Thus, $\rho_{x,n} = r_x$ and the element $f_2(g_n^{\rm univ})$ in (6.12) has component in Δ_n with reduction modulo $\mathfrak{m}_{\eta_{x,n}}$ that has unit specialization at generically-chosen closed points of the irreducible $\overline{\{\eta_{x,n}\}}$. Hence, $f_2(g_n^{\rm univ})$ has unit image in Δ_n. ■

For N_n defined as in (6.9), Lemma 6.6 ensures that the image of N_n in $\widehat{\mathscr{O}}_{V_n^0,\eta_{x,n}}$ is a unit multiple of the norm

(6.14)
$$N_n(g_n^{\rm univ}) = {\rm N}_{((\widehat{\mathscr{O}}_{V_n^0,\eta_{x,n}} \widehat{\otimes}_F \widehat{\mathscr{O}}_{\mathbf{A}^1_F,u_x})/(f_1(\widehat{g}_n^{\rm univ})))/\widehat{\mathscr{O}}_{V_n^0,\eta_{x,n}}}(f_2(\widehat{g}_n^{\rm univ})) \in \widehat{\mathscr{O}}_{V_n^0,\eta_{x,n}} \simeq F(\eta_{x,n})[\![y_0]\!],$$

where
$$\widehat{g}_n^{\rm univ} = 1 \widehat{\otimes} g_{n_0} + y_0 \widehat{\otimes} \varepsilon_0 + \sum_{0<j\le n} (y_j - y_j(g_n)) \widehat{\otimes} \varepsilon_j$$

(recall $y_0(g_n) = 0$ and $\varepsilon_0 \in F^\times$) and the norm is taken with respect to the local ring extension
$$\begin{aligned} F(\eta_{x,n})[\![y_0]\!] &= \widehat{\mathscr{O}}_{V_n^0,\eta_{x,n}} \\ &\to (\widehat{\mathscr{O}}_{V_n^0,\eta_{x,n}} \widehat{\otimes}_F \widehat{\mathscr{O}}_{\mathbf{A}^1_F,u_x})/(f_1(\widehat{g}_n^{\rm univ})) \\ &= (F(\eta_{x,n})[\![y_0]\!])[\![u-u_x]\!]/(f_1(\widehat{g}_n^{\rm univ})) \end{aligned}$$

that is finite free of rank equal to r_x. Thus, the integer $e_{x,n} > 0$ is the y_0-adic order of (6.14), and this order is what we will now prove is monotonically decreasing for large n.

To compute (6.14) up to unit multiple, we may replace both $f_1(g_n^{\rm univ})$ and $f_2(g_n^{\rm univ})$ with their Weierstrass-polynomial parts $w_{1,n}(u-u_x)$ and $w_{2,n}(u-u_x)$ with respective degrees necessarily equal to r_x and s_x. To compute the Weierstrass-polynomial part $w_{1,n}(u-u_x)$ of

$$f_1(g_{n_0} + y_0\varepsilon_0 + \sum_{0<j\le n} (y_j - y_j(g_n))\varepsilon_j) \in (F(y_1, \dots, y_n)[\![y_0]\!])[\![u-u_x]\!] \tag{6.15}$$

observe that since ${\rm ord}_{u_x}(\varepsilon_j) = j$, for $j > r_x$ the ε_j-term makes no contribution to $(u-u_x)$-monomials in (6.15) in degree $\le r_x$. Thus, when computing $w_{1,n}(u-u_x)$ by recursive substitution in $(u-u_x)$-degrees $> r_x$ we may work in the subring
$$(K[y_{r_x+1}, \dots, y_n][\![y_0]\!])[\![u-u_x]\!] \subseteq (F(y_1, \dots, y_n)[\![y_0]\!])[\![u-u_x]\!]$$
with $K = F(y_1, \dots, y_{r_x})$. In particular,
$$w_{1,n} \in (K[y_{r_x+1}, \dots, y_n][\![y_0]\!])[u-u_x]$$

and specializing y_n to 1 yields $w_{1,n-1}$ since $y_j(g_n) = y_j(g_{n-1})$ for all $j \le n-1$. Similarly, such specialization carries $w_{2,n}$ to $w_{2,n-1}$.

Thus, if $m_x = \max(r_x, s_x)$ and $K' = F(y_1, \ldots, y_{m_x})$ then the norm

$$N_n(g_n^{\mathrm{univ}}) \in K'[y_{m_x+1}, \ldots, y_n][\![y_0]\!] \subseteq F(y_1, \ldots, y_n)[\![y_0]\!]$$

specializes to the norm

$$N_{n-1}(g_{n-1}^{\mathrm{univ}}) \in K'[y_{m_x+1}, \ldots, y_{n-1}][\![y_0]\!] \subseteq F(y_1, \ldots, y_{n-1})[\![y_0]\!]$$

when y_n is specialized to 1. Since ord_{y_0}'s cannot increase under specialization of coefficients,

$$e_{x,n} = \mathrm{ord}_{y_0}(N_n(g_n^{\mathrm{univ}})) \le \mathrm{ord}_{y_0}(N_{n-1}(g_{n-1}^{\mathrm{univ}})) = e_{x,n-1},$$

proving the desired monotonic decay for large n. This completes the proof of Theorem 6.4. ■

Corollary 6.7. *Let F be a perfect field with positive characteristic p and $f(T) \in F[u,T]$ a nonzero squarefree element.*

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

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

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

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

Proof. The case $f \in F^\times$ is trivial, so we may assume $f \notin F$. Thus, in either case, Lemma 5.4 assures us that $\partial_u f \neq 0$ and that f and $\partial_u f$ have no nonconstant common factor in $F[u,T]$ (so M_f^{geom} makes sense). Hence, we may apply Theorem 6.4 with $f_1 = f$ and $f_2 = \partial_u f$ to conclude that for g with large degree, the vanishing of the resultant of $f(g)$ and $(\partial_u f)(g)$ only depends on $g \bmod M_f^{\mathrm{geom}}$. Also, $f(g)$ is inseparable in $F[u]$ precisely when it has a common geometric root with its derivative $f(g)'$.

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

7. A refined resultant formula via algebra

A defect in Theorem 6.4 is that it does not provide a description of how b_n and e_n depend on n. These deficiencies are settled by:

Theorem 7.1. *Let F be a perfect field and $f_1, f_2 \in F[u,T]$ nonzero and relatively prime. Let b_n and e_n be as in Theorem 6.4 for the ordered pair (f_1, f_2). There exist unique $\beta_0, \beta_1 \in F^\times$ such that $b_n = \beta_0 \beta_1^n$ for large n, and e_n is a linear polynomial in n for large n.*

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

$$R_n(G) = c^n a_n^{m_0 + m_1 n} \cdot (L_{f_1,f_2} \circ \rho_{n,M}) \tag{7.1}$$

on $\mathrm{Poly}_{n/F}$, with $\rho_{n,M}$ as in (6.1) or (6.2).

Before we prove Theorem 7.1, we use it to prove Theorem 5.7.

Proof. (of Theorem 5.7) For g in $\kappa[u]$ of sufficiently large degree, $f(g)$ is nonzero and (4.12) and (5.1) yield

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

with $d = \deg f(g)$. Since f is squarefree and $f \notin \kappa$, so $(\partial_u f)(T) \neq 0$ by Lemma 5.4, we have $(\partial_u f)(g) \neq 0$ when $\deg g \gg 0$.

Taking into account that $f(g)' = (\partial_u f)(g)$ may have smaller degree than $d-1$, (5.3) gives

$$\chi(R_{d,d-1}(f(g),f(g)')) = \chi((\operatorname{lead} f(g))^{k_g}R(f(g),(\partial_u f)(g))),$$

where $k_g = \deg f(g) - 1 - \deg(\partial_u f)(g)$. When $\deg g \gg 0$, k_g is linear in $\deg g$. Combining this with Theorem 6.4 and Theorem 7.1, there exists $\varepsilon_1 \in \{\pm 1\}$ and integers m_0 and m_1 such that for $\deg g \gg 0$,

$$\mu(f(g)) = \varepsilon_1^{\deg g}(\chi(-1))^{(\deg f(g))(\deg f(g)-1)/2}\chi(\operatorname{lead} g)^{m_0+m_1\deg g}\chi(L(g)) \tag{7.2}$$

where L is an algebraic function on the affine space $\kappa[u]/M_f^{\text{geom}}$ 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 (7.2) depends on $\deg g$ modulo 2.

Now let us establish the final part of Theorem 5.7 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 $Z_f \cap Z_{\partial_u f} \subseteq \mathbf{A}^2_\kappa$ 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 no proper factor of M_f^{geom} can serve as a modulus for $\mu_{\kappa''[u]}(f(g))$ with κ'' any finite extension of κ'. Since $M_{f,\kappa''}^{\min}|M_f^{\text{geom}}$, we can assume M_f^{geom} is nonconstant.

Choose a monic linear factor of M_f^{geom} in $\kappa'[u]$, so it has the form $h = u - u_x$ for some (κ'-rational) point $x = (u_x, t_x) \in Z_f \cap Z_{\partial_u f}$. We can find polynomials g_1 and g_2 with any large degree n and a common leading coefficient such that $g_1(u_x) = t_x \neq g_2(u_x)$ and $g_1(u_{x'}) = g_2(u_{x'}) \neq t_{x'}$ for all $x' \in Z_f \cap Z_{\partial_u f}$ with $x' \neq x$, in which case (6.7) and the positivity of the exponents $e_{x'}$ ensure that for sufficiently large n we have the vanishing of the resultant of $f(g_1)$ and $(\partial_u f)(g_1) = f(g_1)'$ and the non-vanishing of the resultant of $f(g_2)$ and $f(g_2)'$; that is, $\mu_{\kappa'[u]}(f(g_1)) = 0$ and $\mu_{\kappa'[u]}(f(g_2)) \neq 0$. The same properties persist after replacing κ' with any finite extension κ''. Since g_1 and g_2 are clearly congruent modulo $M_f^{\text{geom}}/(u-u_x)$, 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, the monic factor $M_{f,\kappa''}^{\min}$ of the monic M_f^{geom} must equal M_f^{geom}. ■

Let us now prepare for the proof of Theorem 7.1. We first establish a key point: the existence of a formula of the shape (7.1) for some M is equivalent to the claim that e_n is linear in n for large n and that $b_n = \beta_0\beta_1^n$ for some $\beta_0, \beta_1 \in F^\times$ for large n. Necessity is obvious by Theorem 6.4, and for sufficiency we may replace M with $MM_{f_1,f_2}^{\text{geom}}$ to get to the case where $M_{f_1,f_2}^{\text{geom}}|M$, so we have formulas

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

and

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

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

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

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

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

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

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

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

We shall now aim to prove an identity of the form (7.1) by means of induction on the ordered pair (f_1, f_2), and the flexibility in the choice of M will be essential for the success of the induction. In what follows we will work with a generic field-valued point g of the geometrically integral F-variety $\mathrm{Poly}_{n/F}$ for a large $n = \deg g$, though one can instead work throughout with the universal unitary case in large degree n.

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 (7.1) does not depend on this ordering. Indeed, for $\deg g > \nu(f_1), \nu(f_2)$ (see (3.5)),

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

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

Our proof of Theorem 7.1 will roughly be a series of algebraic identities

$$R(f_1(g), f_2(g)) = c_0 c_1^{\deg g} (\mathrm{lead}\, g)^{m_0+m_1 \deg g} R(f_3(g), f_4(g)) \tag{7.3}$$

for generic (or universal unitary) g of large degree, where $c_0, c_1 \in F^\times$ and $m_0, m_1 \in \mathbf{Z}$, and the ordered pair (f_3, f_4) of nonzero relatively prime polynomials in $F[u, T]$ is in some sense smaller than (f_1, f_2). (There is more than one sense that we use, depending on the stage of our argument.) In this way, induction will establish (7.1).

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 $a(u) = ca_1(u)$ with $c \in F^\times$ and $a_1(u)$ monic,

$$R(a(u), f_2(g)) = R(c, f_2(g)) R(a_1(u), f_2(g)) = c^{\deg f_2(g)} R(a_1(u), f_2(g)). \tag{7.4}$$

For $\deg g > \nu(f_2)$, $c^{\deg f_2(g)} = c_0 c_1^{\deg g}$ for suitable c_0 and c_1 in $F^\times$ that are independent of g. The factor $R(a_1(u), f_2(g))$ is an algebraic function of g modulo $a_1(u)$, since $a_1(u)$ is monic.

To prove Theorem 7.1 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 $F[u]$, then

$$R(f_1(g), f_2(g)) = R(a(u), f_2(g)) R(h(g), f_2(g)), \tag{7.5}$$

with the first factor satisfying (7.1), by (7.4). Removing a common factor from the coefficients of f_2 as a polynomial in T is also compatible with Theorem 7.1.

We will prove Theorem 7.1 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 7.2. *Let $h_1(T)$ and $h_2(T)$ in $F[u][T]$ have common T-degree $d \geq 1$:*

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

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

$$\min(\deg \widetilde{\alpha}, \deg \widetilde{\beta}) < \min(\deg \alpha, \deg \beta) \tag{7.6}$$

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

$$R(h_1(g), h_2(g)) = c\varepsilon^{\deg g}(\operatorname{lead} g)^m R(\widetilde{h}_1(g), \widetilde{h}_2(g)). \tag{7.7}$$

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

Proof. We will prove the lemma when $\deg \alpha \leq \deg \beta$. (When $\deg \alpha > \deg \beta$, we can reduce to the other case by interchanging h_1 and h_2, at the cost of changing c and ε in the conclusion.) In $F[u]$, write $\beta(u) = \alpha(u)q(u) + r(u)$, where $r \neq 0$ and $\deg r < \deg \alpha$. Since $r \neq 0$, $k(T) := h_2(T) - q(u)h_1(T)$ has leading term $r(u)T^d$ as a polynomial in T with coefficients in $F[u]$. For all g, clearly $h_2(g) \equiv k(g) \bmod h_1(g)$. When $\deg g$ exceeds $\nu(h_1)$, $\nu(h_2)$, and $\nu(k)$ (see (3.5)), 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 6.1, the identity (7.7) forces relative primality of the $\widetilde{h}_j$'s when the h_j's are relatively prime. ■

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

Lemma 7.3. *Let $h_1(T)$ and $h_2(T)$ in $F[u][T]$ have common T-degree $d \geq 1$:*

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

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

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

Proof. If neither α nor β divides the other in $F[u]$, apply Lemma 7.2 to get a second pair of polynomials in $F[u][T]$ with T-degree d. Repeat this process if again neither leading coefficient as a polynomial in T divides the other. (Note that terms like $c\varepsilon^{\deg g}(\operatorname{lead} g)^m$ behave well under multiplication: the c's and ε's are multiplicative, while the m's are additive.) The condition (7.6) ensures that we eventually reach the case where $\alpha(u)|\beta(u)$ or $\beta(u)|\alpha(u)$. Thus, we may interchange h_1 and h_2 if necessary to suppose $\alpha(u)|\beta(u)$. Write $\beta(u) = \alpha(u)q(u)$. The polynomial $k(T) := h_2(T) - q(u)h_1(T)$ has T-degree less than d. This polynomial is nonzero and is relatively prime to h_1 since $\gcd(h_1, h_2) = 1$. Proceed as in the proof of Lemma 7.2, taking $\widetilde{h}_1 = k$ and $\widetilde{h}_2 = h_1$. ■

We are finally ready to prove Theorem 7.1:

Proof. (of Theorem 7.1). We argue by induction on $\max(\deg_T f_1, \deg_T f_2)$. Set $d_1 = \deg_T f_1$ and $d_2 = \deg_T f_2$. We can assume both d_1 and d_2 are positive, by (7.4). Remove any common factors from the coefficients of $f_1(T)$ as a polynomial in T, using (7.5), so $f_1(T)$ is primitive over $F[u]$. Similarly make f_2 primitive. By Lemma 7.3 we may assume $d_1 \neq d_2$, and without loss of generality $0 < d_1 < d_2$. Writing

$$f_1(T) = \alpha(u)T^{d_1} + \dots, \quad f_2(T) = \beta(u)T^{d_2} + \dots, \tag{7.8}$$

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

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

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

The power of $\operatorname{lead} f_1(g)$ can be written in the form $c_0 c_1^{\deg g}(\operatorname{lead} g)^{m_0 + m_1 \deg g}$ for suitable c_0, c_1 in $F^\times$ and integers m_0 and m_1 that do not depend on g. (The number m_1 is nonzero when $\deg_T k < d_2$.) We are now reduced to proving Theorem 7.1 with f_2 replaced by k.

Either $\deg_T k = d_2$ and the leading coefficient of k as a polynomial in T has smaller degree than $\deg\alpha$, or $\deg_T k < d_2$. In the latter case, $\max(\deg_T f_1, \deg_T k) < d_2$, so Theorem 7.1 with f_1 and k has already been proved by the inductive hypothesis. Thus, it remains to treat the case (7.8) 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 7.2.

We now apply a generalization of the trick with $u + 1$ in (5.9). Consider the universal identity

$$R(f_1(g), \alpha(u))R(f_1(g), f_2(g)) = R(f_1(g), \alpha(u)f_2(g)) \tag{7.9}$$

with universal unitary g of large degree n. The first term in (7.9) is nonzero, since primitivity of f_1 forces $\gcd(f_1(g), \alpha(u)) = 1$. Since all three resultants admit expressions as in Theorem 6.4 for a common modulus M, and since we know that an identity as in (7.1) is equivalent to linearity of e_n in n and an identity of the form $b_n = \beta_0\beta_1^n$ for large n, it is obvious that (7.1) for two of the resultants in (7.9) implies (7.1) for the third resultant in (7.9). Since (7.1) with a polynomial of T-degree zero has already been settled, it suffices to prove (7.1) for the ordered pair $(f_1, \alpha(u)f_2)$.

The right side of (7.9) has the form $R(\alpha(u)g^{d_1} + \cdots, \alpha(u)\beta(u)g^{d_2} + \cdots)$. Let $h(T) = \alpha(u)f_2(T) - \beta(u)f_1(T)T^{d_2-d_1}$. Since $\gcd(f_1, f_2) = 1$ and f_1 is primitive over $F[u]$, and we may assume $\deg_T f_1 > 0$, it follows that h is nonzero and satisfies $\deg_T h < d_2$ and $\gcd(f_1, h) = 1$. Since $h(g) \equiv \alpha(u)f_2(g) \bmod f_1(g)$ for all g, when $\deg g \gg 0$ the right side of (7.9) is

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

for suitable c_0, c_1 in $F^\times$ and integers m_0 and m_1. (For instance, $m_1 = d_2 - \deg_T h$.) Since $\deg_T f_1$ and $\deg_T h$ are both less than d_2, there is a formula for $R(f_1(g), h(g))$ as in (7.1), by induction on the maximum T-degree. ■

8. Characteristic 2

The analogue of Theorem 5.7 in characteristic 2 is subtle because (4.9) in characteristic 2 requires liftings into characteristic 0. Fix a perfect field k of characteristic 2, and let $W = W(k)$ (the Witt vectors of k) and $F = \operatorname{Frac}(W)$.

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

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

Since $h \notin k$, Lemma 5.4(2) ensures $\partial_u h \neq 0$ and that there is no common irreducible factor of h and $\partial_u h$ in $k[u, T]$. Thus, $R_{k[u]}(h, \partial_u h) \neq 0$ and we may define M_h^{geom} as in Definition 5.5. We emphasize that M_h^{geom} is not to be confused with $M_{h(T^2)}^{\text{geom}}$; in our study of Möbius bias for specializations of $f(T) = h(T^2)$ in characteristic 2, it is M_h^{geom} that will turn out to be of more interest than M_f^{geom}. Corollary 6.7(2) ensures that the separability property of $h(g^2)$ in $k[u]$ only depends on $g \bmod M_h^{\text{geom}}$ provided that $\deg g$ is sufficiently large, with largeness that depends on h and is uniform with respect to all perfect extensions of k.

Since $h(T^2)$ is squarefree in $k[u, T]$ and $h \notin k$, we can find $g \in k[u]$ of any sufficiently large degree such that $h(g^2)$ is nonconstant and separable in $k[u]$: use [20, Theorem 3.1] if k is finite, and use Lemma 6.1 and the denseness of the locus of k-rational points in an affine space over k if k is infinite. In particular, $(\partial_u h)(g^2) = h(g^2)'$ is nonzero and $R_k(h(g^2), h(g^2)')$ is nonzero. Fix such a choice of g; concretely, g is a representative of some (nonempty) collection of residue classes modulo M_h^{geom}.

Definition 8.1. A lift $H \in W[u, T]$ of $h \in k[u, T]$ is called *unitary* if H has the same T-degree as h and $\operatorname{lead}_T H \in W[u]$ is a lift of $\operatorname{lead}_T h \in k[u]$ with the same u-degree. In particular, $\operatorname{lead}_T H \in W[u]$ has unit leading coefficient.

Let H be a unitary lift of h and 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 (3.3), and likewise for the degree of $H(G^2)$. Note that $H(G^2) \in W[u]$ has unit leading coefficient (and hence the same degree as $h(g^2)$). Hence, $W[u]/(H(G^2))$ is a finite flat W-algebra that lifts the finite étale k-algebra $k[u]/(h(g^2))$; by (4.12), we need to understand how the unit discriminant $\mathrm{disc}_W(H(G^2)) \bmod 8W$ depends on G.

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

$$R_F(H(G^2), H(G^2)'),\ \ R_F(H(G^2), (\partial_u H)(G^2)) \tag{8.1}$$

lie in W and have reductions in k that are $k^\times$-multiples of each other (see (5.3) and the Warning above Example 5.1). Both reductions therefore lie in $k^\times$ since $h(g^2)$ is separable, so both terms in (8.1) lie in $W^\times$. The quadratic nature of the first resultant in (8.1) intervenes in the study of $\mathrm{disc}_W(H(G^2))$, and the second resultant in (8.1) is a form to which Theorem 6.4 and Theorem 7.1 may be applied (over the field F of characteristic zero). We are going to show that the unit ratio of the resultants in (8.1) can be made explicit in $(W/8W)^\times$ modulo unit-square factors, so we will be able to use Theorems 6.4 and 7.1 to study the quadratic nature of $\mathrm{disc}_W(H(G^2))$.

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

Since $H(G^2)' = (\partial_u H)(G^2) + 2(\partial_T H)(G^2)GG'$, the classical formula (5.2) for resultants in terms of products over geometric roots gives

$$\frac{R_F(H(G^2), H(G^2)')}{R_F(H(G^2), (\partial_u H)(G^2))} = \mathrm{lead}(H(G^2))^{d_G} \prod_{\alpha} \left(1 + 2 \cdot \left. \frac{(\partial_T H)(G^2)GG'}{(\partial_u H)(G^2)} \right|_{\alpha} \right), \tag{8.2}$$

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

Remark 8.2. 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

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

We need to understand the product in (8.2) modulo $8W$. The remarkable surprise is that there is a very simple formula for this product mod $8W$ (see (8.4)), and the formula only depends on g and h (not on G or H). This formula uses residues of a certain differential form. We need to make two definitions before we can state the formula of interest.

Definition 8.3. For any perfect field K and any rational differential form ω on $\mathbf{P}^1_K$, set

$$s_2(\omega) := \sum_{\{y_1, y_2\}} \mathrm{Res}_{y_1}\omega \cdot \mathrm{Res}_{y_2}\omega \in K \tag{8.3}$$

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

In words, $s_2(\omega)$ is the second symmetric function of the geometric residues of ω. Our interest in $s_2(\omega)$ will be restricted largely to cases when ω has simple poles. We are grateful to Gabber for pointing out to us that, for ω varying with only simple geometric poles, $s_2(\omega)$ is not algebraic in ω if we do not fix the number of simple geometric poles of ω. For example, let

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

with $b, b+1 \neq 0$. This has three simple poles when $a \neq 0$ and two simple poles when $a = 0$. When $a \neq 0$, $s_2(\omega) = -b(b+1) - 1$, but when $a = 0$, $s_2(\omega)$ changes to $-(b+1)^2$. This non-algebraicity is analogous to the fact that (6.3) does not extend to an algebraic function on $\mathrm{Poly}_{\leq n/F}$.

Definition 8.4. For $\gamma \in k[u]$, define

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

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

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

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

so this rational differential form on $\mathbf{P}^1_k$ has simple poles. We will see in Theorem 8.11 that $s_2(\omega_{h,\gamma})$ intervenes in the behavior of $\mu(h(\gamma^2))$ when k is finite. The vanishing of $s_2(\omega_{h,\gamma^2})$ will therefore make the behavior of $\mu(h(\gamma^4))$ quite tractable for finite k.

Theorem 8.5. *Let H be a unitary lift of h, in the sense of Definition 8.1. For $g \in k[u]$ of large degree with $h(g^2)$ separable and $G \in W[u]$ lifting g with* $\mathrm{lead}(G) \in W^\times$*,*

$$\prod_\alpha \left(1 + 2 \cdot \left.\frac{(\partial_T H)(G^2)GG'}{(\partial_u H)(G^2)}\right|_{u=\alpha}\right) \equiv 1 + 2\deg g \deg_T h + 4s_2(\omega_{h,g}) \bmod 8W, \tag{8.4}$$

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

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

$$\left.\frac{(\partial_T H)(G^2)GG'}{(\partial_u H)(G^2)}\right|_{u=\alpha} = \mathrm{Res}_\alpha\left(\frac{(\partial_T H)(G^2)GG'}{(\partial_u H)(G^2)} \cdot \frac{\mathrm{d}P}{P}\right). \tag{8.5}$$

We will first show that

$$2\left.\frac{(\partial_T H)(G^2)GG'}{(\partial_u H)(G^2)}\right|_{u=\alpha} \equiv 2\mathrm{Res}_\alpha\omega_{H,G} + 4(\mathrm{Res}_\alpha\omega_{H,G})^2 \bmod 8\overline{W}, \tag{8.6}$$

where $\overline{W}$ is the integral closure of W in an algebraic closure $\overline{F}$ of F. Note that we can replace the residue in the final term in the mod-8 equation (8.6) with a residue in characteristic 2, namely $\mathrm{Res}_{\overline{\alpha}}(\omega_{h,g})$ with $\overline{\alpha}$ the reduction of α.

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

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

However,

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

and $P = H(G^2)$, so by (8.5) we conclude that in $\overline{W}/8\overline{W}$

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

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

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

since $\mathrm{Res}_x(s^p\mathrm{d}r/r) = \mathrm{Res}_x(s\mathrm{d}r/r)^p$ in characteristic $p > 0$. This establishes (8.6).

Using (8.6), expanding the product on the left side of (8.4) modulo 8 gives

$$1 + 2\sum_\alpha \mathrm{Res}_\alpha\omega_{H,G} + 4\sum_{\alpha_1\neq\alpha_2}\mathrm{Res}_{\overline{\alpha}_1}\omega_{h,g}\mathrm{Res}_{\overline{\alpha}_2}\omega_{h,g} + 4\sum_\alpha \mathrm{Res}_{\overline{\alpha}}(\omega_{h,g})^2, \tag{8.7}$$

where α_1 and α_2 in the second sum run over unordered pairs of distinct $\overline{F}$-roots of $H(G^2)$. By the residue theorem in characteristic 0, the first sum over the zeros α of $H(G^2)$ in (8.7) is equal to

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

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

Since (8.7) is being considered in $W/8W$, the final sum in (8.7) only matters in $W/2W$, where it can be computed to be

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

by the residue theorem in characteristic 2. The second and third sums in (8.7) therefore combine to give $4s_2(\omega_{h,g})$ in (8.4). ■

By (5.1), (8.2), and Theorem 8.5, since $h(g^2)$ is *separable* the discriminant $\mathrm{disc}_W(H(G^2))$ is congruent modulo $8W$ to

$$\frac{(-1)^{\delta_g(\delta_g-1)/2}}{(\mathrm{lead}\,H(G^2))^{2\delta_g-1-d_G}}R_W(H(G^2),(\partial_u H)(G^2))(1+2\deg g\deg_T h+4s_2(\omega_{h,g})), \tag{8.8}$$

where

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

and d_G is given by Remark 8.2; the exponent $2\delta_g - 1 - d_G$ of $\mathrm{lead}\,H(G^2)$ in (8.8) is linear in $\deg g = \deg G$ when $\deg g$ is large. Since $-4 \equiv 4 \bmod 8$, $\mathrm{disc}_W(H(G^2)) \bmod 8W$ is therefore equal to

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

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

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

and (by checking cases for ab modulo 4)

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

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

$$\frac{R_W(H(G^2),(\partial_u H)(G^2))}{(\operatorname{lead} H(G^2))^{2\delta_g-1-d_G}}(-1)^{\deg(\operatorname{lead}_T h)(\deg(\operatorname{lead}_T h)-1)/2}(1+4(m_g+s_2(\omega_{h,g}))), \tag{8.9}$$

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

If we had instead chosen g of large degree such that $h(g^2)$ is *not* separable and $G\in W[u]$ is a lift of g with $\operatorname{lead}(G)\in W^\times$, then since $H(G^2)$ has the same degree as its reduction $h(g^2)$ we see via (5.3) that $R_W(H(G^2),(\partial_u H)(G^2))$ has reduction that is a $k^\times$-multiple (depending on G) of

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

Thus, $R_W(H(G^2),(\partial_u H)(G^2))\in 2W$ in such cases, so although $\operatorname{disc}_W(H(G^2))$ may not be congruent modulo 8 to (8.9) when $h(g^2)$ is not separable, the expression (8.9) *always makes sense* in W and is a non-unit precisely when $\operatorname{disc}_W(H(G^2))$ is a nonunit. Thus, we can use the resultant $R_W(H(G^2),(\partial_u H)(G^2))$ from characteristic 0 to study $\operatorname{disc}_W(H(G^2)) \bmod 8W$ even though usually $(\partial_u H)(G^2)\neq H(G^2)'$ in characteristic 0.

Since $\operatorname{lead}_T H\in W[u]$ has leading coefficient in $W^\times$ and $h=H \bmod 2\in k[u,T]$ is not in k and has no irreducible factors in $k[T]$ (as $h(T^2)$ is squarefree), we conclude that H is not in W and H has no irreducible factors in $W[T]$. Moreover, since h is squarefree in $k[u,T]$ we see that its unitary lifting H is squarefree in $W[u,T]$. The same therefore holds using F-coefficients, so $\partial_u H\neq 0$ and the zero loci $Z_H=\{H=0\}$ and $Z_{\partial_u H}=\{\partial_u H=0\}$ in $\mathbf{A}^2_F$ have finite intersection by Lemma 5.6. In particular,

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

is *nonzero* and we may form the monic squarefree polynomial $M_H^{\mathrm{geom}}\in F[u]$ as in Definition 5.5, where the geometric roots of M_H^{geom} are the u-coordinates of intersection points of Z_H and $Z_{\partial_u H}$ in $\mathbf{A}^2_F$.

We may use Theorems 6.4 and 7.1 to obtain the identity of algebraic functions

$$R_F(H(G),(\partial_u H)(G))=\beta_0\beta_1^n\cdot\operatorname{lead}(G)^{m_0+m_1 n}\cdot\prod_x P_{x,n}(G)^{e_x} \tag{8.10}$$

on $\operatorname{Poly}_{n/F}$ for large n, where the integers $m_0,m_1\in\mathbf{Z}$ and the scalars $\beta_0,\beta_1\in F^\times$ are independent of n, the indexing set $\{x\}$ is the set of intersection points of Z_H and $Z_{\partial_u H}$ in $\mathbf{A}^2_F$, the e_x's are positive integers, and $P_{x,n}(G)=\mathrm{N}_{F(x)/F}(G(u_x)-t_x)$ where (u_x,t_x) are the coordinates of $x\in\mathbf{A}^2_F$. Of course, all of the parameters in (8.10) may depend on the (fixed) choice of unitary H lifting h. When $G\in W[u]$, the left side of (8.10) is a resultant over W. We now show that the identity (8.10) over F can be factored in a manner that is well-behaved with respect to W.

Lemma 8.6. *For large n (uniform with respect to perfect extensions of k), the algebraic maps*

$$\beta_0 \cdot \prod_{|u_x|\leq 1, |t_x|>1} P_{x,n}^{e_x}(\cdot), \;\; \beta_1^n \cdot \prod_{|u_x|>1} P_{x,n}^{e_x}(\cdot) : \mathrm{Poly}_{\leq n/F} \to \mathbf{A}^1_F \tag{8.11}$$

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

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

$$P_{x,n}(G) = \mathrm{N}_{F(x)/F}(G(u_x) - t_x) = \mathrm{N}_{F(x)/F}(t_x) \cdot \mathrm{N}_{F(x)/F}(t_x^{-1}G(u_x) - 1) \tag{8.12}$$

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

$$P_{x,n}(G) = \mathrm{N}_{F(x)/F}(G(u_x) - t_x) = \mathrm{N}_{F(x)/F}(u_x)^n \cdot \mathrm{N}_{F(x)/F}(G^*(1/u_x) - u_x^{-n}t_x) \tag{8.13}$$

with $|u_x^{-n}t_x| \ll 1$ for large n. Hence, to see that (8.11) extends over W, it is enough to show that the elements

$$b_0 := \beta_0 \cdot \prod_{|u_x|\leq 1, |t_x|>1} \mathrm{N}_{F(x)/F}(t_x)^{e_x}, \;\; b_1 := \beta_1 \cdot \prod_{|u_x|>1} \mathrm{N}_{F(x)/F}(u_x)^{e_x} \tag{8.14}$$

in F are integral. We shall prove these are in fact units in W. It then follows trivially that the first map in (8.11) extends over W and has constant reduction $\overline{b}_0 \in k^\times$. Likewise, the second map in (8.11) then extends over W and has reduction

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

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

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

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

For $G_n \in W[u]$ lifting any such g_n with $\mathrm{lead}(G_n) \in W^\times$, clearly the W-resultant of $H(G_n)$ and $(\partial_u H)(G_n)$ is a unit in W. Thus, the left side of (8.10) is a unit in W when evaluated at G_n. Now consider the right side of (8.10) when evaluated at G_n. The contribution of $\mathrm{lead}(G_n)$ is an integral unit, so we conclude

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

By the norm-scaling calculations (8.12) and (8.13), we thereby obtain

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

or equivalently

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

Obviously a $\overline{W}$-point $x = (u_x, t_x)$ in the zero loci of H and $\partial_u H$ reduces to a geometric point in the zero loci of h and $\partial_u h$. Thus, for such x we conclude via Theorem 6.4 that the reduction of $P_{x,n}(G_n) \in W$ must be nonzero, since the resultant of $h(g_n)$ and $(\partial_u h)(g_n)$ is nonzero. Hence, $P_{x,n}(G_n) \in W^\times$ for such x. Thus, $b_0 b_1^n \in W^\times$ for all large n. Hence, $b_0, b_1 \in W^\times$. ■

In the study of (8.10) on G^2 for $G \in W[u]$ with unit leading coefficient, we will be able to ignore x's with $|u_x| > 1$ due to:

Theorem 8.7. *For $G \in W[u]$ with unit leading coefficient and large degree n (uniform with respect to perfect extensions of k),*

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

Proof. By Lemma 8.6, 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 leaves us with

$$\prod_{|u_x|>1} \mathrm{N}_{F(x)/F}(G^*(1/u_x)^2 - u_x^{-2n} t_x)^{e_x}, \tag{8.15}$$

where G^* is the polynomial of (possibly fake) degree n obtained by reversing the order of the coefficients of G. Note that the square $G^*(1/u_x)^2$ is a unit when $|u_x| > 1$, as its reduction is $\mathrm{lead}(g)^2 \neq 0$. Since $u_x^{-2n} t_x \to 0$ as $n \to \infty$, for large n we see that $G^*(1/u_x)^2 - u_x^{-2n} t_x$ is very close to a unit square in the valuation ring $W(x)$ of $F(x)$. Hence, depending *just* on the amount of ramification in $F(x)$ (bounded by $[F(x) : F]$), we can make n large enough, uniformly with respect to perfect extensions of k, such that $G^*(1/u_x)^2 - u_x^{-2n} t_x$ is a square in $W(x)^\times$. Passing to n so uniformly large for all finitely many x's such that $|u_x| > 1$, the norm-product (8.15) is a unit square in W. ■

To emphasize that $b_0 \in W^\times$ in (8.14) depends on H, we now rename it: define

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

so η_H depends on H since the algebraic factorization on the right side of (8.10) depends on H. Using Lemma 8.6 and Theorem 8.7, 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 (8.10) yields an identity

$$R_H(G) \in \eta_H \cdot \prod_{|u_x| \le 1, |t_x| > 1} \mathrm{N}_{W(x)/W}(t_x^{-1} G(u_x)^2 - 1)^{e_x} \cdot \prod_{|u_x|, |t_x| \le 1} P_{x,2n}(G^2)^{e_x} \cdot (W^\times)^2 \tag{8.16}$$

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

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

Since $\eta_H \in W^\times$ and all terms in the products in (8.16) are visibly integral, the resultant $R_H(G)$ is a unit in W if and only if each of the terms in the products in (8.16) is a unit, in which case the image of $R_H(G)$ in $W^\times/(W^\times)^2$ is represented by the expression in (8.16).

Define

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

where $e_H = 1$ if $\mathrm{lead}_T H \notin W^\times$ and $e_H = \deg(\mathrm{lead}_T \partial_u H)$ if $\mathrm{lead}_T H \in W^\times$; $\widetilde{\eta}_H$ absorbs both the constant sign-factor and (by Remark 8.2) the odd-exponent power of the unit

$\mathrm{lead}(H(G^2))$ in (8.9) modulo $(W^\times)^2$. Choose $g \in k[u]$ with large degree and choose $G \in W[u]$ lifting g with $\deg G = \deg g$. When $h(g^2)$ is *separable* it follows from (8.9) that $\mathrm{disc}_W(H(G^2)) \in W^\times$ is a unit-square multiple of the visibly integral

(8.17)

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

with $m_g = \lfloor (1 + \deg g \deg_T h)/2 \rfloor$ and $\widetilde{s}_2(\omega_{h,g}) \in W$ any lift of $s_2(\omega_{h,g}) \in k$ (see (8.3)). On the other hand, if $h(g^2)$ is not separable, then (8.16) implies that one of the terms $P_{x,2n}(G^2)$ with $|u_x|, |t_x| \leq 1$ is in the maximal ideal of W, so (8.17) is also in the maximal ideal of W in such cases.

Motivated by (8.17), consider the W-scheme map $L_{H,n} : \mathrm{Poly}_{\leq n/W} \to \mathbf{A}^1_W$ defined by

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

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

$$\widetilde{\rho}_{n,H} := \rho_{n,(M_H^{\mathrm{geom}})^{\leq 1}} : \mathrm{Poly}_{\leq n/W} \to W[u]/(M_H^{\mathrm{geom}})^{\leq 1} \tag{8.18}$$

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

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

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

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

(8.19)

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

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

We will use the quadratic nature of (8.19) to investigate $\mu(h(g^2))$ in the case of finite k, but before passing to the finite case we need to study the relationship between $(M_H^{\mathrm{geom}})^{\leq 1}$ and M_h^{geom}. We may factor the separable monic M_H^{geom} in $F[u]$ into monic polynomials

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

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

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

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

Definition 8.8. The reduction of $(M_H^{\mathrm{geom}})^{\leq 1}$ is denoted $\overline{M}_H^{\mathrm{geom}}$.

Up to $k^\times$-multiple, $\overline{M}_H^{\mathrm{geom}}$ is the mod-2 reduction of a primitively-scaled multiple of M_H^{geom} in $W[u]$. By reduction of divisibility over W we conclude that $\overline{M}_H^{\mathrm{geom}}$ divides $R_{k[u]}(h, \partial_u h)$; note that $\overline{M}_H^{\mathrm{geom}}$ need not be squarefree (see Example 8.15).

Remark 8.9. Obviously M_h^{geom} divides the radical of $R_{k[u]}(h, \partial_u h)$. One can have proper divisibility here if the nonzero $\text{lead}_T h \in k[u]$ has a double root at some c, since the resultant $R_{k[u]}(h, \partial_u h)$ vanishes at such c for determinantal reasons but the specializations $h(c,T)$ and $(\partial_u h)(c,T)$ might not have a common geometric root; *cf.* Remark 1.7.

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

Lemma 8.10. *For all unitary lifts H of h, $M_h^{\text{geom}} | \overline{M}_H^{\text{geom}}$; in particular, the property of $h(g^2)$ being squarefree is determined by $g \bmod \overline{M}_H^{\text{geom}}$. If $\text{lead}_T h$ is separable (e.g., h is monic in T), then M_h^{geom} is the radical of $\overline{M}_H^{\text{geom}}$.*

Proof. Recall that by Corollary 6.7(2), $g \bmod M_h^{\text{geom}}$ determines whether or not $h(g^2)$ is squarefree. Since M_h^{geom} is squarefree, clearly $M_h^{\text{geom}} | \overline{M}_H^{\text{geom}}$ 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 5.4(2)). Thus, $k[u,T]/(h, \partial_u h)$ is a finite k-algebra. Moreover, since $W[u,T]$ is W-flat, it follows from the local flatness criterion that $\partial_u H$ is nowhere a zero divisor on $\operatorname{Spec} W[u,T]/(H)$ at points over the closed point of $\operatorname{Spec} W$ and that $\operatorname{Spec} W[u,T]/(H, \partial_u H)$ is W-flat at points over the closed point of $\operatorname{Spec} W$. On the generic fiber over $\operatorname{Spec} F$, $F[u,T]/(H, \partial_u H)$ is a finite (flat) F-algebra since $\{H = 0\}$ meets $\{\partial_u H = 0\}$ at only finitely many points in $\mathbf{A}_F^2$. To summarize, the finite-type separated morphism $\operatorname{Spec} W[u,T]/(H, \partial_u H) \to \operatorname{Spec} W$ is quasi-finite and flat.

By the structure theorem for quasi-finite separated schemes over a henselian local base [15, 18.5.11], it follows that $W[u,T]/(H, \partial_u H) = R^{\text{f}} \times R'$, 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

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

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

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

Now let $g \in k[u]$ be arbitrary with large degree. By Lemma 8.10, whether or not $h(g^2)$ is separable is determined by $g \bmod \overline{M}_H^{\text{geom}}$, and even by g modulo the radical of $\overline{M}_H^{\text{geom}}$. Thus, the monic $\overline{M}_H^{\text{geom}}$ constructed by reduction from characteristic 0 controls the separability of $h(g^2)$ in characteristic 2 when $\deg g$ is large.

Let us now specialize to the case of a *finite* field $k = \kappa$ of characteristic 2. We fix nonconstant $h \in \kappa[u,T]$ such that $h(T^2)$ is squarefree. Choose a unitary lift H of h. Pick $g \in \kappa[u]$ of large degree, and choose a lift $G \in W[u]$ of g with the same degree (*i.e.*, with unit leading coefficient). Hence, $H(G^2)$ is a lift of $h(g^2)$ with the same degree, and $\mathrm{disc}_W(H(G^2))$ is a unit precisely when $h(g^2)$ is separable. Recall also (as we explained above Theorem 4.9) that if $\mathrm{disc}_W(H(G^2)) \in W^\times$ then it lies in $\kappa^\times \times (1+4W)$; that is, its 1-unit part lies in $1+4W$, not merely in $1+2W$, when it is a unit in W.

By Theorem 4.9 and Remark 4.10,

$$\mu(h(g^2)) = (-1)^{\deg(\mathrm{lead}_T h)}\widetilde{\chi}(\mathrm{disc}_W(H(G^2))), \tag{8.20}$$

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

$$\widetilde{\chi}(c\cdot(1+4w)) = (-1)^{\mathrm{Tr}_{\kappa/\mathbf{F}_2}(w \bmod 2)}. \tag{8.21}$$

We can now prove an analogue of (7.2) in characteristic 2:

Theorem 8.11. *Let κ be finite of characteristic* 2*, and $h \in \kappa[u,T]$ be such that $h \notin \kappa$ and $h(T^2)$ is squarefree in $\kappa[u,T]$. Fix a unitary lift H of h.*

For g of sufficiently large degree n,

$$\mu(h(g^2)) = (-1)^{\deg \mathrm{lead}_T(h)+[\kappa:\mathbf{F}_2]\lfloor(1+n\deg_T h)/2\rfloor+\mathrm{Tr}_{\kappa/\mathbf{F}_2}(s_2(\omega_{h,g}))}\cdot\widetilde{\chi}(L_H(\widetilde{\rho}_{n,H}(G))), \tag{8.22}$$

where $G \in W[u]$ is any lift of g with degree n. Here, $s_2(\omega_{h,g})$ is defined by (8.3) *and $\widetilde{\rho}_{n,H}$ is defined by* (8.18)*. The "sufficient largeness" for* $\deg g$ *may be chosen uniformly with respect to finite extensions of κ.*

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

$$(-1)^{\mathrm{Tr}_{\kappa/\mathbf{F}_2}(s_2(\omega_{h,g_1}))}\mu(h(g_1^2)) = (-1)^{\mathrm{Tr}_{\kappa/\mathbf{F}_2}(s_2(\omega_{h,g_2}))}\mu(h(g_2^2)). \tag{8.23}$$

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

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

Proof. The preceding calculations ensure that $L_H(\widetilde{\rho}_{n,H}(G)) \in W$ lies in $\kappa^\times \times (1+4W)$ when it is a unit (because the same is true for both $\mathrm{disc}_W(H(G^2))$ and squares in $W^\times$). Thus, the asserted formula (8.22) for $\mu(h(g^2))$ makes sense and is immediate from (8.20), (8.21), and (8.19). Since any two elements $g_1, g_2 \in \kappa[u]$ that are congruent modulo the reduction $\overline{M}_H^{\mathrm{geom}}$ of the *monic* $(M_H^{\mathrm{geom}})^{\leq 1}$ 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 $n_j = \deg G_j = \deg g_j$), we conclude via (8.22) that the indicated congruence conditions on g_j's and $\deg g_j$'s are enough to imply (8.23). ■

An easy argument with the Chinese remainder theorem shows that Theorem 8.11 remains true with $\overline{M}_H^{\mathrm{geom}}$ replaced by the gcd of all $\overline{M}_H^{\mathrm{geom}}$'s as H runs over all unitary lifts of h to $W[u,T]$. This gcd is a multiple of M_h^{geom} (by Lemma 8.10) and is obviously a factor of $R_{\kappa[u]}(h, \partial_u h)$, but it probably can fail to be squarefree (see Example 8.15 below). We do not know if this gcd is the "minimal modulus" for $g \mapsto (-1)^{\mathrm{Tr}_{\kappa/\mathbf{F}_2}(s_2(\omega_{h,g}))}\mu(h(g^2))$ when specializing in $\kappa[u]$ (but see Question 9.2).

Corollary 8.12. *Let κ be finite of characteristic 2, and $h \in \kappa[u,T]$ be such that $h \notin \kappa$ and $h(T^2)$ is squarefree in $\kappa[u,T]$. Fix a unitary lift H of h.*

For g of sufficiently large degree n,

$$\mu(h(g^4)) = (-1)^{\deg \mathrm{lead}_T h + [\kappa:\mathbf{F}_2](\deg_T h)\cdot n} \cdot \widetilde{\chi}(L_H(\widetilde{\rho}_{n,H}(G))), \tag{8.24}$$

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

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

$$g_1 \equiv g_2 \bmod \overline{M}_H^{\mathrm{geom}},\ \ \deg g_1 \equiv \deg g_2 \bmod 2 \Rightarrow \mu(h(g_1^4)) = \mu(h(g_2^4)). \tag{8.25}$$

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

We now give some Möbius calculations in characteristic 2, using Corollary 8.12 (and omitting further tables of data). Our second and third example will justify what we said after Remark 1.14 in the Introduction about some characteristic 2 examples.

Example 8.13. 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 $\mu(f(g)) = (-1)^{[\kappa:\mathbf{F}_2]\deg g}$ for $\deg g$ sufficiently large. It is left to the reader to check that $\deg g \geq 1$ is "large enough". It follows that the conjecture in (3.8) fails in even degrees when $[\kappa : \mathbf{F}_2]$ is odd, and in all degrees when $[\kappa : \mathbf{F}_2]$ is even.

Example 8.14. 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 $M_H^{\mathrm{geom}} = 6u^5 + 2u^3 + 1$, so $\overline{M}_H^{\mathrm{geom}} = 1$ and $\deg_T H$ is even. Thus, $\mu(f(g)) = 1$ for $\deg g \gg 0$. A closer analysis, carried out in [8], shows that $\mu(f(g)) = 1$ for $\deg g \geq 3$ and $\mu(f(cu^2)) = -1$ for some $c \in \kappa^\times$, so the lower bound on $\deg g$ is sharp.

Example 8.15. In $\kappa[u][T]$, let $f(T) = T^{16} + (u^9 + u^4 + u^2 + u)T^8 + u^5 + u^3$. Using the proof of Theorem 8.11 to make sufficient largeness explicit, for g_1 and g_2 with degree at least 2 we have

$$g_1 \equiv g_2 \bmod u^9(u+1)^4 \Longrightarrow \mu(f(g_1)) = \mu(f(g_2)). \tag{8.26}$$

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

9. Conjectures over $\kappa[u]$

We return to the Hardy–Littlewood conjecture over $\kappa[u]$ for a finite field κ. Numerical testing supports the belief that (3.8) is correct when f is separable (in any characteristic). We have seen that (3.8) is not always true for inseparable f. To define a correction factor in the inseparable cases away from polynomials in T^2 that are not polynomials in T^4 in characteristic 2, we begin with a definition that is sensitive to the constant field κ.

Definition 9.1. Let κ be a finite field. Pick $f(T)$ in $\kappa[u][T^p]$ with $p \neq 2$ (resp. in $\kappa[u][T^4]$ with $p = 2$) such that $f \notin \kappa$ and f is squarefree in $\kappa[u][T]$. Define $M_{f,\kappa}^{\min}$ to be the unique monic polynomial M in $\kappa[u]$ of minimal degree that satisfies the property of M_f^{geom} in (5.14) (resp. the property of $\overline{M}_H^{\mathrm{geom}}$ in (8.25), with $f(T) = h(T^4)$).

By the Chinese remainder theorem, all nonzero $M \in \kappa[u]$ satisfying (5.14) (resp. (8.25)) are divisible by $M_{f,\kappa}^{\min}$. If κ'/κ is a finite extension, it seems to be a rather subtle problem to relate $M_{f,\kappa}^{\min}$ and $M_{f,\kappa'}^{\min}$. In odd characteristic, we always have $M_{f,\kappa}^{\min}|M_f^{\mathrm{geom}}$, so $M_{f,\kappa}^{\min}$ is squarefree. For characteristic 2, we have $M_{f,\kappa}^{\min}|R_{\kappa[u]}(h, \partial_u h)$ with $R_{k[u]}(h, \partial_u h) \neq 0$ (by Lemma 5.4(2)), so again the polynomials $M_{f,\kappa'}^{\min}$ have only finitely many possibilities as κ' varies over finite extensions of κ. However, the polynomial $R_{\kappa[u]}(h, \partial_u h)$ generally has factors with rather high multiplicities, so it would be desirable to find better upper bounds on the multiplicities in $M_{f,\kappa}^{\min}$ and to find an *a priori* construction of the least common multiple of all $M_{f,\kappa'}^{\min}$'s (or at least its radical) for characteristic 2 as the extension κ'/κ varies. The following suggests a nice "upper bound" on the radical of $M_{f,\kappa}^{\min}$ in characteristic 2, akin to the upper bound provided by M_f^{geom} in odd characteristic.

Question 9.2. In characteristic 2, is M_h^{geom} the radical of the least common multiple of the $\overline{M}_H^{\mathrm{geom}}$'s over all unitary lifts H of h? By Lemma 8.10 we know that M_h^{geom} divides this radical, and that this divisibility is an equality in the "generic" case when $\mathrm{lead}_T h \in \kappa[u]$ is separable.

We are almost ready to define our correction factor for the Hardy–Littlewood conjecture over $\kappa[u]$, but we first need a lemma.

Lemma 9.3. *Let κ be a finite field of characteristic p and let $f \in \kappa[u][T^p]$ be squarefree in $\kappa[u,T]$, and assume that f has no local obstructions* (*so in particular, f has no irreducible factors in $\kappa[u]$*). *For any nonzero $M \in \kappa[u]$, there exist elements $g \in \kappa[u]$ with any sufficiently large degree* (*depending on M and f*) *such that $f(g)$ is squarefree in $\kappa[u]$ and* $\gcd(f(g), M) = 1$.

Proof. The case $f \in \kappa^\times$ is trivial, so we may assume $f \notin \kappa$. We must find g in large degree n with $f(g)$ relatively prime to $M \cdot f(g)' = M \cdot (\partial_u f)(g)$. Obviously $\partial_u f \neq 0$ since $f \notin \kappa$. By Lemma 5.4(1), f and $\partial_u f$ have no common irreducible factor in $\kappa[u][T]$. For any irreducible monic $\pi \in \kappa[u]$, define

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

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

Poonen [20] proved that the statistics for squarefree specializations of a squarefree polynomial over $\kappa[u]$ do agree with local-probability heuristics. More specifically, since $1 - c_\pi/\mathrm{N}\pi > 0$ for each π, [20, Thm. 3.1] yields

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

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

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

■

With an eye toward future considerations with several polynomials, we now make a definition that is more general than we presently require.

Definition 9.4. Let κ be a finite field. Let $f_1, \dots, f_r$ be squarefree and pairwise relatively prime elements in $\kappa[u][T]$. Suppose that each f_j is a polynomial in T^p when $p \neq 2$ (resp. in T^4 for $p = 2$), and that $f_j \notin \kappa$ for all j. Assume that $\prod f_j$ has no local obstructions (so in fact each f_j has no irreducible factors in $\kappa[u]$). Let M_κ be the least common multiple of the polynomials $M_{f_j,\kappa}^{\min}$.

For $n \gg 0$, define

$$\Lambda_\kappa(f_1, \dots, f_r; n) := \frac{\sum_{\deg g = n, (f_j(g), M_\kappa)=1} \prod_j (|\mu(f_j(g))| - \mu(f_j(g)))}{\sum_{\deg g = n, (f_j(g), M_\kappa)=1} \prod_j |\mu(f_j(g))|} \in [0, 2^r] \cap \mathbf{Q}, \tag{9.1}$$

where the condition $\gcd(f_j(g), M_\kappa) = 1$ (a congruence condition on g modulo $\operatorname{rad}(M_\kappa)$) is imposed for all j; we take n large enough as in Lemma 9.3 so that the denominator in (9.1) is nonzero. When κ is understood from context, we write Λ rather than Λ_κ.

Example 9.5. The case of one polynomial is the one of most interest to us, and it illuminates the meaning of the ratio in (9.1):

$$\begin{aligned} \Lambda_\kappa(f; n) \quad &:= \quad \frac{\sum_{\deg g = n, (f(g), M_{f,\kappa}^{\min})=1} (|\mu(f(g))| - \mu(f(g)))}{\sum_{\deg g = n, (f(g), M_{f,\kappa}^{\min})=1} |\mu(f(g))|} \\ &= \quad 1 - \frac{\sum_{\deg g = n, (f(g), M_{f,\kappa}^{\min})=1} \mu(f(g))}{\sum_{\deg g = n, (f(g), M_{f,\kappa}^{\min})=1} |\mu(f(g))|}. \end{aligned}$$

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

We should address a uniformity for the nonvanishing of the denominator in (9.1) for large n as we vary the constant field. There exists nonzero $M \in \kappa[u]$ such that $M_{f_j,\kappa'}^{\min} | M$ in $\kappa'[u]$ for all finite extensions κ' of κ and all j: take $M = \prod M_{f_j}^{\mathrm{geom}}$ in odd characteristic and $M = \prod \overline{M}_{H_j}^{\mathrm{geom}}$ in characteristic 2 (where H_j is a unitary lift of h_j, with $f_j = h_j(T^4)$). Since $f = \prod f_j$ has no local obstructions, by applying Lemma 9.3 to f and M we see that for large n there do exist (many) $g \in \kappa[u]$ of degree n such that $\prod f_j(g) \in \kappa[u]$ is squarefree and relatively prime to M. Since the inclusion $\kappa[u] \hookrightarrow \kappa'[u]$ for any finite extension κ'/κ preserves separability and relative primality, it follows that the denominator in the definition of $\Lambda_{\kappa'}(f_1, \dots, f_r; n)$ is nonzero for large n uniform with respect to κ'/κ.

Clearly (9.1) is unaffected by replacing M_κ with its radical. In Corollary 9.11 we will see that Definition 9.4 is unaffected by replacing M_κ with *any* common (nonzero) multiple of the radicals of the $M_{f_j,\kappa}^{\min}$'s. This makes computation of Λ_κ easier both in theory and in practice, since in odd characteristic we can replace M_κ with the radical of the product of the $M_{f_j}^{\mathrm{geom}}$'s, and in characteristic 2 we can likewise replace M_κ with the radical of the product of the $R_{\kappa[u]}(h_j, \partial_u h_j)$'s (or even with the radical of the product of the $M_{h_j}^{\mathrm{geom}}$'s in characteristic 2 when Question 9.2 has an affirmative answer for each h_j).

We only care about $\Lambda_\kappa(f_1, \dots, f_r; n)$ for large n. Note that $\Lambda_\kappa(f_1, \dots, f_r; n) = 0$ if and only if, for all g of degree n, some $f_j(g)$ has a nontrivial factor in common with $M_{f,\kappa}^{\min}$ or $\mu(f_j(g)) \in \{0, 1\}$ for some j. Therefore the vanishing of $\Lambda_\kappa(f_1, \dots, f_r; n)$ implies that for all g of degree n in $\kappa[u]$, one of the the polynomials $f_1(g), \dots, f_r(g)$ is reducible in $\kappa[u]$.

Example 9.6. Let $f(T)$ be the polynomial from Example 3.6, viewed in $\mathbf{F}_5[u][T]$. We will compute $\Lambda(f;n)$. By (5.15), $M_{f,\mathbf{F}_5}^{\min} = u(u+1)$ and $\mu(f(g))$ is determined by $g \bmod u(u+1)$ when $\deg g$ is even.

As g runs over all quadratics in $\mathbf{F}_5[u]$, 28 times $\mu(f(g)) = 1$ and 52 times $\mu(f(g)) = -1$. In all of these cases, $(f(g), M_{f,\mathbf{F}_5}^{\min}) = 1$. (The Möbius formula and a derivative calculation show that $\mu(f(g)) = 0$ if and only if $(f(g), M_{f,\mathbf{F}_5}^{\min}) \neq 1$.) Thus $\Lambda(f;2) = 1 - (28-52)/(28+52) = 13/10$. Similarly, we find $\Lambda(f;3) = 1$. From the proof of Theorem 9.10 below, $\Lambda(f;n)$ has period 2 when $n \geq 2$, and we have just computed the two terms in the period. Compare the alternating sequence $\{1, 13/10, 1, 13/10, \dots\}$ with the data in Table 3.6.

Example 9.7. Let $f(T)$ be as in Example 5.2. Working over $\kappa[u]$ with κ of size $q = 3^m$, it is easy to check that $\Lambda_\kappa(f;n) = (q+1)/q$ for m odd and $n \geq 2$, and $\Lambda_\kappa(f;n) = (q-1)/(q-2)$ for m even and $n \geq 2$.

There is no Hardy–Littlewood conjecture for even m because $T^4 + (u+1)T^2 + u^4$ is reducible over $\mathbf{F}_9(u)$ (one root is $\sqrt{u^2+2u+2} + i(u+2)$, with $i \in \mathbf{F}_9$ satisfying $i^2 = -1$). This serves to remind us that the polynomial has to be irreducible in order that a Hardy–Littlewood conjecture be meaningful.

Example 9.8. Let κ be finite with odd characteristic p, and $f(T) = T^p + u \in \kappa[u][T]$. Generalizing the Möbius calculation in Example 4.5, we find

$$\Lambda_\kappa(f;n) = \begin{cases} 1, & \text{if } n \text{ is odd,} \\ 0, & \text{if } n \equiv 0 \bmod 4, \\ 1 - \chi(-1), & \text{if } n \equiv 2 \bmod 4, \end{cases} \tag{9.2}$$

for $n \geq 1$, where χ is the quadratic character on $\kappa^\times$. In particular, $\Lambda_{\mathbf{F}_3}(T^3+u;n)$ is $1, 2, 1, 0, 1, 2, 1, 0, \dots$ and $\Lambda_{\mathbf{F}_9}(T^3+u;n)$ is $1, 0, 1, 0, 1, 0, 1, 0, \dots$ over $\mathbf{F}_9[u]$. The $\mathbf{F}_3[u]$-calculation is consistent with Table 1.2. The $\mathbf{F}_9[u]$-calculation tells us $g^3 + u$ will not be irreducible for g with (positive) even degree and suggests $g^3 + u$ will satisfy (3.8) as g runs through polynomials with odd degree. This is supported by Table 1.3.

Example 9.9. 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. This example will illustrate the importance of the condition $(f(g), M_{f,\kappa}^{\min}) = 1$ in the definition of $\Lambda_\kappa(f;n)$.

The case $\kappa = \mathbf{F}_3$ was considered numerically in Example 3.4. There we observed in Table 3.4 that $f(g)$ seems to be reducible when $n = \deg g$ satisfies $n \equiv 1 \bmod 4$, and $f(g)$ has approximately twice as many irreducible values as the naive Hardy–Littlewood conjecture (3.8) predicts when $n \equiv 3 \bmod 4$. We now compute $\Lambda_\kappa(f;n)$ for any κ of characteristic 3, and we will find consistency with the data from Table 3.4 for $\kappa = \mathbf{F}_3$.

We recall (5.13) from Example 5.3: when $g = cu^n + \cdots \in \kappa[u]$ with $n = \deg g \geq 1$,

$$\mu(f(g)) = (-1)^n(\chi(-1))^{n(n-1)/2}\chi(c)^{n+1}\chi(g(1)^2 + g(1) + 2)\chi(g(2)). \tag{9.3}$$

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

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

$$f(g)|_{u=1} = (g(1)-1)^3(g(1)^2 + g(1) - 1)^3, \quad f(g)|_{u=2} = (g(2))^6(g(2)+1)^3, \tag{9.4}$$

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

If κ has size $q = 3^m$, then by separately treating the cases when m is even or odd and when n is even or odd, elementary arguments resting on the preceding formulas (9.3) and (9.4) show that we have $\Lambda_\kappa(f;n) = 1$ for even $n > 0$ and

$$\Lambda_\kappa(f;n) = \begin{cases} 1 + 2\cdot(-1)^{(n+1)/2}/((q-1)(q-2)), & m \text{ odd}, \\ 1 + 2/((q-2)(q-3)), & m \text{ even}, \end{cases}$$

for odd n. In contrast, if we do not include the condition $(f(g), M) = 1$ in the definition of $\Lambda_\kappa(f;n)$ then we would instead get the constant sequence of values $\{1, 1, \dots\}$ for $n = 1, 2, \dots$. 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 these global values are constrained by the local condition $(f(g), M) = 1$.

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

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

which is an excellent fit with the discrepancies between Table 3.4 and the naive Hardy–Littlewood conjecture for $f(T)$ on $\mathbf{F}_3[u]$. Here, if $n \equiv 1 \bmod 4$, then $\mu(f(g)) = -1$ only when $(f(g), M) \neq 1$. If $n \equiv 3 \bmod 4$, then $\mu(f(g)) = 1$ only when $(f(g), M) \neq 1$.

Our work in §4–§8 leads to the following important periodicity:

Theorem 9.10. *Let κ be finite, and $f_1(T), \dots, f_r(T)$ be as in Definition 9.4. For any finite extension κ'/κ, the sequence $\Lambda_{\kappa'}(f_1, \dots, f_r; n)$ is periodic with period dividing 4 for $n \gg 0$, and the largeness is uniform with respect to κ'.*

Proof. In this proof, we will work with $\kappa' = \kappa$, and we leave it to the reader to check that all references to "sufficiently large" can be made uniformly with respect to finite extensions κ' of κ (keep in mind that, in odd characteristic, the monic polynomial $M^{\min}_{f_j,\kappa'}$ may be different from $M^{\min}_{f_j,\kappa}$, but for all j it must be a factor of a fixed nonzero "geometric" polynomial $M^{\text{geom}}_{f_j}$ constructed from f_j over κ, so $\deg M^{\min}_{f_j,\kappa'}$ is *bounded* as κ' varies; similar remarks apply in characteristic 2 using $M^{\text{geom}}_{h_j}$ with $f_j = h_j(T^4)$).

We first give the proof in odd characteristic. Fix $a \bmod 4$. By (7.2), there exist integers $m_{0,j}$ and $m_{1,j}$, signs $s_{a,j} \in \{\pm 1\}$, and an algebraic function L_j on the κ-scheme of remainders modulo $M^{\text{geom}}_{f_j}$ such that

$$\mu(f_j(g)) = s_{a,j}\chi(\operatorname{lead} g)^{m_{0,j}+m_{1,j}a}\chi(L_j(g)) \tag{9.5}$$

for $g \in \kappa[u]$ with $\deg g \gg 0$ and $\deg g \equiv a \bmod 4$. In particular,

$$|\mu(f_j(g))| = \begin{cases} 1, & \text{if } L_j(g) \neq 0, \\ 0, & \text{if } L_j(g) = 0. \end{cases}$$

We claim that $M^{\min}_{f_j,\kappa}$ is the least-degree monic divisor D of $M^{\text{geom}}_{f_j}$ such that the set-theoretic function

$$\chi_j = \chi \circ L_j : \kappa[u]/(M^{\text{geom}}_{f_j}) \to \{0, 1, -1\}$$

factors through projection to $\kappa[u]/(D)$; we emphasize that $\kappa[u]/(M^{\text{geom}}_{f_j})$ appears here as a finite-dimensional κ-vector space, not as an affine space over $\operatorname{Spec}\kappa$. Our problem is purely set-theoretic in nature: to show that $M^{\min}_{f_j,\kappa}$ is also the "modulus of definition" for the set-theoretic function χ_j on $\kappa[u]$. By (7.2), if a nonzero monic $D \in \kappa[u]$ is a modulus of definition for χ_j then D is a modulus of definition for $g \mapsto \mu(f_j(g))$ in the sense that

$$g_1 \equiv g_2 \bmod D,\ \ \chi(\text{lead } g_1) = \chi(\text{lead } g_2),\ \ \deg g_1 \equiv \deg g_2 \bmod 4 \Rightarrow \mu(f_j(g_1)) = \mu(f_j(g_2)).$$

The converse also holds because any congruence class in $\kappa[u]/(D)$ for any nonzero D may be represented by elements of arbitrary large degree with any desired leading coefficient. We conclude that $M^{\min}_{f_j,\kappa}$ is intrinsic to the set-theoretic function χ_j.

Let M be the least common multiple of the $M^{\min}_{f_j,\kappa}$'s. By definition,

$$\Lambda(f_1, \ldots, f_r; n) = \frac{\sum_{\deg g=n,(f_j(g),M)=1} \prod_j (|\mu(f_j(g))| - \mu(f_j(g)))}{\sum_{\deg g=n,(f_j(g),M)=1} \prod_j |\mu(f_j(g))|}. \tag{9.6}$$

To show that the ratio is periodic, we will first analyze the numerator and denominator separately with $n \gg 0$.

All χ_j's may be viewed as functions on $\kappa[u]/(M)$. When $n \geq \deg M$, the polynomials g with degree n are $g = MQ + R$ where Q has degree $n - \deg M$ and either $R = 0$ or $\deg R < \deg M$. Since M is monic, $\text{lead}\, g = \text{lead}\, Q$. Therefore, for $n \gg 0$ with $n \equiv a \bmod 4$ the numerator in (9.6) is

$$\sum_{\deg Q = n - \deg M} \ \sum_{\substack{(f_j(R),M)=1 \\ \text{all } \chi_j(R) \neq 0}} \prod_{j=1}^{r} (1 - s_{a,j}\chi(\text{lead}\, Q)^{m_{0,j}+m_{1,j}a}\chi_j(R)).$$

Expanding the product, this is

$$\sum_{Q} \ \sum_{\substack{(f_j(R),M)=1 \\ \text{all } \chi_j(R) \neq 0}} \left(\sum_{0 \leq k \leq r} (-1)^k \sum_{1 \leq j_1 < \cdots < j_k \leq r} s_{a,\underline{j}}\chi(\text{lead}\, Q)^{m_{0,\underline{j}}+m_{1,\underline{j}}a}\chi_{j_1}(R) \cdots \chi_{j_k}(R)) \right),$$

where $\underline{j}$ is the vector $(j_1, \ldots, j_k)$, $s_{a,\underline{j}}$ is the product of the s_{a,j_i} and $m_{0,\underline{j}}$ and $m_{1,\underline{j}}$ are the respective sums of the m_{0,j_i} and m_{1,j_i}.

Bringing the sum over Q to the inside, the numerator of (9.6) is

$$\sum_{0 \leq k \leq r} (-1)^k \sum_{1 \leq j_1 < \cdots < j_k \leq r} s_{a,\underline{j}} \left(\sum_{Q} \chi(\text{lead}\, Q)^{m_{0,\underline{j}}+m_{1,\underline{j}}a} \right) \left(\sum_{\substack{(f_j(R),M)=1 \\ \text{all } \chi_j(R) \neq 0}} \chi_{j_1}(R) \cdots \chi_{j_k}(R) \right).$$

(In the rightmost sum, the constraint $\chi_j(R) \neq 0$ runs over $j = 1, \ldots, r$, not just $j = j_1, \ldots, j_k$. Therefore we cannot eliminate this constraint unless $k = r$.)

The sum over Q is 0 if $m_{0,\underline{j}} + m_{1,\underline{j}}a$ is odd and is $(q-1)q^{n-\deg M}$ if $m_{0,\underline{j}} + m_{1,\underline{j}}a$ is even, where q is the size of κ. Thus, the numerator of (9.6) is

$$(q-1)q^{n-\deg M} \sum_{0 \leq k \leq r} (-1)^k \sum_{\substack{1 \leq j_1 < \cdots < j_k \leq r \\ m_{0,\underline{j}}+m_{1,\underline{j}}a \text{ even}}} s_{a,\underline{j}} \sum_{\substack{(f_j(R),M)=1 \\ \text{all } \chi_j(R) \neq 0}} \chi_{j_1}(R) \cdots \chi_{j_k}(R).$$

Aside from $q^{n-\deg M}$, the rest of the expression only depends on n through $a = n \bmod 4$, so as a function of n it has period dividing 4.

Now we turn to the denominator of (9.6). By Lemma 9.3, this denominator is nonzero for large n. Explicitly, this denominator is

$$(q-1)q^{n-\deg M}\#\{R \bmod M:\ \text{all } (f_j(R), M_j) = 1,\ \text{all } \chi_j(R) \neq 0\}.$$

Since (9.5) holds for g of sufficiently large degree, (9.6) with $n \gg 0$ and $n \equiv a \bmod 4$ is

$$\frac{\sum_{0\le k\le r}(-1)^k \sum_{\substack{1\le j_1<\cdots<j_k\le r\\ m_{0,\underline{j}}+m_{1,\underline{j}}a \text{ even}}} s_{a,\underline{j}} \sum_{\substack{(f_j(R),M)=1\\ \text{all } \chi_j(R)\neq 0}} \chi_{j_1}(R)\cdots\chi_{j_k}(R)}{\#\{R \bmod M:\ \text{all } (f_j(R), M) = 1,\ \text{all } \chi_j(R) \neq 0\},} \tag{9.7}$$

where the innermost sum runs over $R \bmod M$. This fraction is constant for fixed a, so the desired periodicity is proved for odd characteristic.

The same argument works in the case of characteristic 2, as we now explain. We have $f_j = h_j(T^4)$ where $h_j(T^2)$ is squarefree and $h_j \notin \kappa$ for all j; fix unitary lifts $H_1, \ldots, H_r$ of $h_1, \ldots, h_r$ respectively, and fix $a \bmod 2$. We will use the formula (8.24) as the replacement of (9.5). For our purposes, a more convenient way to write (8.24) is

$$\mu(f_j(g)) = s_{a,j} \cdot \widetilde{\chi}(L_j(G)) \tag{9.8}$$

for $\deg g \gg 0$ with $g \in \kappa[u]$ satisfying $\deg g \equiv a \bmod 2$, where

$$s_{a,j} = (-1)^{\deg \mathrm{lead}_T(h_j) + [\kappa:\mathbf{F}_2](\deg_T h_j)\cdot a}$$

is a sign, $G \in W[u]$ is a lift of g with unit leading coefficient, and L_j is an algebraic function on the W-scheme $W[u]/(M_{H_j}^{\mathrm{geom}})^{\le 1}$ of remainders modulo the monic $(M_{H_j}^{\mathrm{geom}})^{\le 1}$; L_j is denoted L_{H_j} in §8.

By (8.19), if L_j has unit value on some congruence class from $W[u]$ then this unit value lies in $\kappa^\times \times (1 + 4W)$. Thus, the composite map of sets

$$\widetilde{\chi} \circ L_j : W[u]/(M_{H_j}^{\mathrm{geom}})^{\le 1} \to \{0, 1, -1\}$$

makes sense as a set-theoretic function, where the source is viewed as a finite free W-module (not a W-scheme) and

$$\widetilde{\chi} : \kappa^\times \times (1 + 4W) \twoheadrightarrow \{\pm 1\}$$

is the unique quadratic character killing $(W^\times)^2$ (with $\widetilde{\chi}$ defined to be zero on $2W$).

Pick $\xi \in W[u]/(M_{H_j}^{\mathrm{geom}})^{\le 1}$. For representatives $G \in W[u]$ of ξ with unit leading coefficient, (9.8) says $(\widetilde{\chi} \circ L_j)(\xi) = s_{\deg g, j} \cdot \mu(f_j(g))$, where $g = G \bmod 2 \in \kappa[u]$ has the same degree as G. In particular, $(\widetilde{\chi} \circ L_j)(\xi)$ only depends on g, not G. Since $(M_{H_j}^{\mathrm{geom}})^{\le 1} \in W[u]$ is a *monic* polynomial, as $G \in W[u]$ varies over all large-degree representatives of ξ with unit leading coefficient we see that its reduction $g \in \kappa[u]$ varies over *all* large-degree representatives of the mod-2 residue class $\overline{\xi} \in \kappa[u]/\overline{M}_{H_j}^{\mathrm{geom}}$. Hence, we conclude that the set-theoretic function $\widetilde{\chi} \circ L_j$ factors through reduction mod 2 as a set-theoretic function

$$\chi_j : \kappa[u]/\overline{M}_H^{\mathrm{geom}} \to \{0, 1, -1\}.$$

In particular, for $g \in \kappa[u]$ of large degree we have a formula $\mu(f_j(g)) = s_{a,j}\chi_j(g)$ where $a = \deg g \bmod 2$.

Using these properties of χ_j, it is straightforward to adapt the argument from the odd-characteristic case to show that $M_{f_j,\kappa}^{\min}$ is the unique minimal "modulus of definition" for the set-theoretic function χ_j, and then the rest of the odd-characteristic argument carries over almost *verbatim* in the case of characteristic 2. For example, the absence of a lead-coefficient contribution in the formula $\mu(f_j(g)) = s_{a,j}\chi_j(g)$ in characteristic 2 corresponds to setting $m_0 = m_1 = 0$ in the odd-characteristic calculations. ■

Corollary 9.11. *Assume $f_1, \dots, f_r$ are as in Theorem* 9.10. *For κ'/κ any finite extension and large $n \gg 0$ (depending on the $M_{f_j,\kappa}$'s) that may be chosen uniformly with respect to κ', $\Lambda_{\kappa'}(f_1, \dots, f_r; n)$ may be defined by using any nonzero multiple $\widehat{M}$ of the $M^{\min}_{f_j,\kappa'}$'s in place of the least common multiple $M_{\kappa'}$ as in Definition* 9.4.

Proof. Once again, we present the argument for $\kappa' = \kappa$ and leave it to the reader to make the routine check that all "sufficiently large" statements may be made uniformly with respect to finite extensions of κ. We shall treat the case of odd characteristic, and we leave it to the reader to check that the techniques used in the proof of Theorem 9.10 for characteristic 2 allow us to adapt the argument to work nearly *verbatim* in the case of characteristic 2.

Let $M_j = M^{\min}_{f_j,\kappa}$, and fix a nonzero common multiple $\widehat{M}$ of the M_j's. Let M be the least common multiple. The results of Poonen cited in the proof of Lemma 9.3 ensure that the ratio $\widehat{\Lambda}_\kappa(f_1, \dots, f_r; n)$ defined using $\widehat{M}$ has nonvanishing denominator for $n \gg 0$. The proof of Theorem 9.10 carries over for $\widehat{\Lambda}_\kappa(f_1, \dots, f_r; n)$ except that (9.7) for $\widehat{\Lambda}_\kappa(f_1, \dots, f_r; n)$ has both the inner sum in the numerator and the count in the denominator running over $R \bmod \widehat{M}$ with the condition $(f_j(R), M) = 1$ replaced by $(f_j(R), \widehat{M}) = 1$. Therefore, it suffices to show that the fraction (9.7) using $\widehat{M}$ is equal to (9.7) in its original form using M.

Write $\widehat{M} = DM$ where $D = D_1D_2$, with D_1 having all of its prime factors dividing M and $\gcd(D_2, M) = 1$. Since the χ_j's admit the least common multiple M of the M_j's as a common modulus of definition, the Chinese remainder theorem yields the following comparison of inner sums in the numerator in (9.7) for Λ and $\widehat{\Lambda}$:

$$\sum_{\substack{(f_j(R),\widehat{M})=1 \\ \text{all } \chi_j(R)\neq 0}} \chi_{j_1}(R)\cdots\chi_{j_k}(R) = c \cdot \sum_{\substack{(f_j(R),M)=1 \\ \text{all } \chi_j(R)\neq 0}} \chi_{j_1}(R)\cdots\chi_{j_k}(R), \tag{9.9}$$

where $c = \#\{R \bmod D_2 \mid (f_j(R), D_2) = 1\}q^{\deg D_1}$ and the sum on the left side of (9.9) runs over $R \bmod \widehat{M}$ while the sum on the right side of (9.9) runs over $R \bmod M$.

Similarly,

$$\#\{R \bmod \widehat{M} : (f_j(R), \widehat{M}) = 1, \chi_j(R) \neq 0\} = c \cdot \#\{R \bmod M : (f_j(R), M) = 1, \chi_j(R) \neq 0\}$$

(the conditions imposed simultaneously over all $1 \le j \le r$). The nonvanishing for denominators ensures $c \neq 0$, so upon taking ratios we see that c cancels and hence

$$\widehat{\Lambda}_\kappa(f_1, \dots, f_r; n) = \Lambda_\kappa(f_1, \dots, f_r; n),$$

at least for n large. ∎

Conjecture 9.12. *Let κ be a finite field and let $f \in \kappa[u, T^p]$ be irreducible in $\kappa[u][T]$ with no local obstructions. If $p = 2$, assume $f \in \kappa[u, T^4]$. As $n \to \infty$,*

$$\#\{g \in \kappa[u] : \deg g = n, f_j(g) \text{ prime}\} \overset{?}{\sim} \Lambda_\kappa(f; n)C_{\kappa[u]}(f) \times {\sum_{\deg g = n}}' \frac{1}{\log(\mathrm{N}(f(g)))},$$

where $\Lambda_\kappa(f; n)$ is defined in Example 9.5 *and is provably periodic in large n by Theorem* 9.10.

Remark 9.13. In characteristic 2 our conjecture is incomplete because it does not make a prediction for $f = h(T^2)$ with $h \in \kappa[u][T]$ when h is not a polynomial in T^2. Due to (8.23), our lack of understanding of the properties of the (generally nonzero) function $g \mapsto s_2(\omega_{h,g})$ is the obstruction to formulating a conjecture that covers such cases at the present time.

When 0 occurs in the period for $\Lambda_\kappa(f;n)$, we interpret the asymptotic in Conjecture 9.12 to mean the easily proved consequence (for such large n) that there is no $g \in \kappa[u]$ in those degrees such that $f(g)$ is irreducible.

We collect sample periodic parts of $\Lambda(f;n)$ in Table 9.1 for $\kappa = \mathbf{F}_p$. When the period is not 1, we write the period so that the first term occurs when $n \gg 0$ and $n \equiv 1 \bmod 4$.

$f(T)$	$\Lambda(f;n)$
$T^3 + u$ (Examples 3.2, 4.5)	$1, 2, 1, 0$
$T^5 + u$ (Examples 3.2, 4.5)	$1, 0$
$T^{12} + \cdots$ (Examples 1.1, 5.2)	$4/3$
$T^9 + \cdots$ (Examples 3.4, 5.1, 5.3, 9.9)	$0, 1, 2, 1$
$T^{12} + \cdots$(Examples 3.5, 5.10)	$2/3$
$(2u^2 + u + 3)T^{15} + \cdots$ (Examples 3.6, 5.11)	$1, 13/10$
$T^3 + u^2$ (Example 4.14)	1

Table 9.1. Examples of $\Lambda(f;n)$ for $n \gg 0$

Each of the polynomials in Table 9.1, aside from the last one, appeared in §3 as a plausible counterexample to (3.8). (When $\Lambda(f;n)$ has 0 in its period, the counterexample is certain.) The reader can easily check that the values of $\Lambda(f;n)$ in each example are in excellent numerical agreement with the ratio column in the tables in Examples 3.2, 1.1, 3.4, 3.5, and 3.6.

Remark 9.14. For f as in Conjecture 9.12, the definition of $\Lambda_\kappa(f;n)$ involves the constraint $(f(g), M_{f,\kappa}^{\min}) = 1$. We do not have a conceptually satisfying explanation for this relative primality condition, so let us explain how it was found.

Initial deviations from (1.2) were discovered with Examples 1.3, 1.4, and 3.2, and seemed to require correction factors 0 or 2. Factorizations of $f(g)$ in these cases revealed extreme parity behavior: the number of irreducible factors of $f(g)$ had the same parity for all g (when $\deg g \geq 1$) and (trivially) $f(g)$ was always squarefree. This suggested a link to Möbius fluctuations, and our first guess at a correction factor was an expression, say $\widetilde{\Lambda}_\kappa(f;n)$, defined like $\Lambda_\kappa(f;n)$ but lacking the condition $\gcd(f(g), M_{f,\kappa}^{\min}) = 1$ in the sums. Periodicity of $\widetilde{\Lambda}_\kappa(f;n)$ follows by the same arguments as for $\Lambda_\kappa(f;n)$ in Theorem 9.10; in fact, that proof was first developed for $\widetilde{\Lambda}_\kappa(f;n)$.

When we found numerically, for the polynomial in Example 3.4, that $\widetilde{\Lambda}_\kappa(f;n)$ was not always the correct correction factor in (1.2), the reason that it failed (as seen in Example 9.9) led to the consideration of the gcd constraint. Table 9.2 gives several examples over $\mathbf{F}_3[u]$ where $\widetilde{\Lambda}_{\mathbf{F}_3}(f;n) \neq \Lambda_{\mathbf{F}_3}(f;n)$. The first two are polynomials we have already met and the remaining two are new nonmonic polynomials in T. The last example is particularly interesting, since $\widetilde{\Lambda}_{\mathbf{F}_3}(f;n)$ and $\Lambda_{\mathbf{F}_3}(f;n)$ lie on opposite sides of 1.

Numerically, in each example where $\widetilde{\Lambda}_\kappa(f;n) \neq \Lambda_\kappa(f;n)$ for $n \gg 0$, data for Conjecture 9.12 has been an excellent fit with $\Lambda_\kappa(f;n)$. Moreover, in the examples where $\widetilde{\Lambda}_\kappa(f;n) = \Lambda_\kappa(f;n)$ for $n \gg 0$, we have found a common explanation for this equality: $\mu(f(g)) = 0$ when $(f(g), M_{f,\kappa}^{\min}) \neq 1$ since, for every irreducible π dividing $M_{f,\kappa}^{\min}$, any root

$f(T)$	$\widetilde{\Lambda}_{\mathbf{F}_3}(f;n)$	$\Lambda_{\mathbf{F}_3}(f;n)$
Example 3.4	$1\ (n \geq 2)$	$0, 1, 2, 1, \ldots\ (n \geq 1)$
Example 3.5	$20/21\ (n \geq 3)$	$2/3\ (n \geq 3)$
$(u^2+2u+1)T^6+(u^2+2u)T^3+2u^2$	$1\ (n \geq 2)$	$0, 2, 0, 2, \ldots\ (n \geq 1)$
$(u+2)T^{12}+u^2T^6+u^3+2$	$6/7\ (n \geq 3)$	$6/5\ (n \geq 4)$

Table 9.2. Examples where $\widetilde{\Lambda}_{\mathbf{F}_3}(f;n) \neq \Lambda_{\mathbf{F}_3}(f;n)$ for $n \gg 0$

of $f(T)$ in $\kappa[u]/\pi$ is a multiple root (that is, it is also a root of $\partial_u(f(T)) = (\partial_u f)(T)$). This includes the vacuous case of those π dividing $M_{f,\kappa}^{\min}$ with $\omega_f(\pi) = 0$. It would be interesting to know if this is always the explanation.

Remark 9.15. The sequence $\Lambda_\kappa(f;n)$ is sensitive to our choice of sampling regions, taken to be the locus of all polynomials of degree n for increasing n. (The classical Hardy–Littlewood conjecture also uses a specific family of sampling regions: $\mathbf{Z} \cap [1,x]$.) If we instead sample over monic g of each degree, then we need a monic version of $\Lambda_\kappa(f;n)$. This is an effectively computable, but possibly new, periodic sequence (with mod 4 periodicity, *etc.*, by the same arguments). For example, if $\kappa = \mathbf{F}_5$ then $\mu(g^5+u)$ has both 1 and -1 as values as g runs over polynomials with odd degree, but only -1 is a value if we restrict to monic g of odd degree. Thus, we expect a change in the distribution of irreducibility counts for g^5+u if we restrict attention to monic g, and numerical data support this (in agreement with a monic version of Λ_κ).

Appendix: Convergence of Hardy–Littlewood constants

We want to discuss, in a general context, how products like $C(f)$ in (2.4) and (3.7) can be computed accurately. Some elementary representation theory will help us write down rapidly-converging product formulas.

Rather than restrict attention to polynomials in $\mathbf{Z}[T]$ or $\kappa[u][T]$, we allow polynomials to lie in $\mathscr{O}_{K,S}[T]$, where $\mathscr{O}_{K,S}$ is a ring of S-integers for a global field K, with S containing the set S_∞ of archimedean places in the number-field case. Let f be the product of r elements $f_1, \ldots, f_r \in \mathscr{O}_{K,S}[T]$ that are irreducible in $K[T]$, pairwise coprime in $K[T]$, and have no local obstructions at places on $\mathscr{O}_{K,S}$. (The last condition means each f_j defines a non-zero function on the residue field of each place). Set

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

where $\operatorname{Res}(\mathscr{O}_{K,S})$ denotes the residue at $s=1$ for the zeta-function $\zeta_{K,S}$ of $\operatorname{Spec}(\mathscr{O}_{K,S})$. Such numbers are called *Hardy–Littlewood constants*, and agree with (2.4) and (3.7) for $\mathscr{O}_{K,S} = \mathbf{Z}$ and $\mathscr{O}_{K,S} = \kappa[u]$.

Our convention is that

$$\prod_{v \notin S} \frac{1-\omega_f(v)/\operatorname{N}v}{(1-1/\operatorname{N}v)^r} := \prod_{n \geq 1} \prod_{\substack{\operatorname{N}v = n \\ v \notin S}} \frac{1-\omega_f(v)/\operatorname{N}v}{(1-1/\operatorname{N}v)^r}. \tag{A.1}$$

The convergence of the right side will usually only be conditional. If we are working over a number field, we can order the terms either by increasing value of the norm or according to the rational prime below each place. These both converge (with the same value) if either does since the subproduct over places with degree > 1 is absolutely convergent and the factors at places with degree 1 have the same order of appearance in both such orderings of the product. More generally, if K/K_0 is a finite extension of global fields, then we can write (A.1) as an infinite product indexed by places of K_0 in order of increasing norm.

Definition A.1. Let L be a field and $h \in L[T]$ be nonzero. If L has characteristic $p > 0$ and $h(T) = H(T^{p^m})$ with $m \geq 0$ maximal, then $H(T)$ is the *p-free part* of $h(T)$. If L has characteristic 0, the p-free part of h is defined to be h.

The p-free part of any irreducible in $L[T]$ is separable and irreducible, but the p-free part of a reducible polynomial may be inseparable.

Theorem A.2. *The infinite product* (A.1) *converges. Convergence is absolute if and only if the p-free part of each f_j is linear, where p is the characteristic of K.*

Proof. Note

$$\prod_{v \notin S} \frac{1 - \omega_f(v)/\,\mathrm{N}v}{(1 - 1/\,\mathrm{N}v)^r} = \prod_{v \notin S} \left(1 + \frac{r - \omega_f(v)}{\mathrm{N}v} + O\left(\frac{1}{\mathrm{N}v^2}\right)\right). \tag{A.2}$$

Taking logarithms termwise in the product and stripping away absolutely convergent subsums reduces the proof of convergence to a check that the series $\sum_{v \notin S}(r - \omega_f(v))/\mathrm{N}v$ converges, where the terms are added in the same sense that terms in (A.1) are multiplied.

Since the f_j's are pairwise coprime in $K[T]$,

$$\omega_f(v) = \omega_{f_1}(v) + \cdots + \omega_{f_r}(v) \tag{A.3}$$

for all but finitely many v. (The lack of any error term in (A.3) is due to the fact that we are considering polynomials in one variable. In the multivariable analogue of Theorem A.2 there are error terms and these can be estimated by using the Lang–Weil estimate.) By (A.3) we are reduced to checking convergence of

$$\sum_{v \notin S} \frac{1 - \omega_{f_j}(v)}{\mathrm{N}v}, \tag{A.4}$$

where we remind the reader that f_j is irreducible.

Let $F_j \in \mathscr{O}_{K,S}[T]$ denote the p-free part of f_j, so $\omega_{f_j}(v) = \omega_{F_j}(v)$ for all $v \notin S$. (When K is a number field, $F_j = f_j$.) Since each F_j is irreducible in $K[T]$, for all but finitely many v the number $\omega_{F_j}(v)$ of solutions to $F_j = 0$ in the residue field $\mathscr{O}_v/\mathfrak{m}_v$ equals the number of relative places of degree 1 lying over v in the field $K[T]/(F_j(T))$. Thus,

$$\sum_{\substack{\mathrm{N}v \leq x \\ v \notin S}} \frac{\omega_{f_j}(v)}{\mathrm{N}v} = \sum_{\mathrm{N}w \leq x} \frac{1}{\mathrm{N}w} + \text{const.} + o(1),$$

where v runs over places of K outside S and w runs over places of $K[T]/(F_j)$. For w running over places in any global field E,

$$\sum_{\mathrm{N}w \leq x} \frac{1}{\mathrm{N}w} = \log\log x + c_E + o(1)$$

for some constant c_E. Applying this to $E = K[T]/(F_j(T))$ and to $E = K$, we subtract and see that (A.4) converges.

Now it remains to check that the product (A.2) converges absolutely if and only if each (irreducible) f_j has linear p-free part F_j. Absolute convergence of a product, by definition, means absolute convergence of its related series of logarithms, and in the case of (A.2) this translates into convergence of

$$\sum_{v \notin S} \frac{|r - \omega_f(v)|}{\mathrm{N}v}. \tag{A.5}$$

When all F_j's are linear, by (A.3) we have $\omega_f(v) = \sum \omega_{F_j}(v) = r$ for all but finitely many v, so all but finitely many terms in (A.5) are 0. Conversely, assume some F_j is nonlinear. By the Chebotarev density theorem, for a positive proportion of v all F_j's split completely modulo v. (Here we need that the F_j's are separable, or equivalently that the field $K[T]/(F_j)$ is separable over K.) Since $\omega_{F_j}(v) = \deg F_j$ for all but finitely many of these v, we see that if some F_j is nonlinear then $\omega_f(v) = \sum_{j=1}^r \deg F_j > r$ for such v. This makes the v-th term in (A.5) at least as large as $1/\mathrm{N}v$ for a positive proportion of v, so (A.5) diverges. ■

For numerical work, we need absolutely convergent products for Hardy–Littlewood constants. To convert (A.2) into an absolutely convergent product, we will multiply each factor by an additional term so the v-th factor is $1 + O(1/\mathrm{N}v^2)$. This has been discussed in the literature when $K = \mathbf{Q}$ [4], [10], [26], [27], [28], [29].

First we set up notation. Let $f_1, \ldots, f_r \in \mathscr{O}_{K,S}[T]$ satisfy the hypotheses of (3.6): they are irreducible and pairwise coprime in $K[T]$, and their product f does not have a local obstruction at any $v \notin S$. For any finite place v, let $\zeta_K(v,s) = 1/(1 - \mathrm{N}v^{-s})$ be the v-th Euler factor of $\zeta_K(s)$. Set $K_j = K[T]/(f_j)$. Writing $\zeta_{K_j}(s)$ as a product over finite places of K (rather than of K_j), let $\zeta_{K_j}(v,s)$ be the v-th factor: $\zeta_{K_j}(s) = \prod_v \zeta_{K_j}(v,s)$. Note that $\zeta_{K_j}(v,s)$ is the reciprocal of a polynomial in $\mathrm{N}v^{-s}$ with degree $[K_j : K]$ for all but finitely many v. The number $\zeta_{K_j}(v,1) = \zeta_{K_j}(v,s)|_{s=1}$ is what matters in the next theorem.

Theorem A.3. *With notation as above, the Hardy–Littlewood constant equals*

$$\frac{1}{\mathrm{Res}(\mathscr{O}_{K_1,S_1}) \cdots \mathrm{Res}(\mathscr{O}_{K_r,S_r})} \prod_{v \notin S} \left(1 - \frac{\omega_f(v)}{\mathrm{N}v}\right) \zeta_{K_1}(v,1) \cdots \zeta_{K_r}(v,1),$$

where the product is absolutely convergent. Here v runs over the places of K not in S, and S_j is the set of places of K_j that lie over S.

The original infinite-product definition of these constants is theoretically important; it shows up, for example, in work on upper bounds related to the classical Hardy–Littlewood conjecture [3, Lemma 3]. Before proving Theorem A.3, we give examples in $\mathbf{Q}$ and in $\mathbf{F}_2(x)$.

Example A.4. We write the Hardy–Littlewood constant for primes of the form n^2+1 in $\mathbf{Z}$ as an absolutely convergent product. Here $K = \mathbf{Q}$, $S = S_\infty$, $r = 1$, $f_1(T) = T^2 + 1$, $K_1 = \mathbf{Q}(i)$, $\mathrm{Res}(\mathscr{O}_{K_1}) = \pi/4$, $\omega_{f_1}(p) = 1 + \chi_4(p)$, and $\zeta_{K_1}(p,s) = (1 - p^{-s})^{-1}(1 - \chi_4(p)p^{-s})^{-1}$.

Theorem A.3 says, after some algebra,

$$\prod_p \frac{1 - \omega_f(p)/p}{1 - 1/p} = \frac{4}{\pi} \prod_p \left(1 - \frac{\chi_4(p)}{p-1}\right) \frac{1}{1 - \chi_4(p)/p}. \tag{A.6}$$

Concretely, we have interlaced the Euler product for $\pi/4 = \prod_p (1 - \chi_4(p)/p)^{-1}$ into the product defining $C(T^2+1)$. Using PARI, we collect in Table A.1 approximations to both sides of (A.6). The improvement on the right side is clear.

n	Left side of (A.6)	Right side of (A.6)
10^2	1.351546	1.372739
10^3	1.370454	1.372814
10^4	1.371023	1.372813
10^5	1.372350	1.372813
10^6	1.372811	1.372813

Table A.1. Comparison of (A.6) using $p \leq n$

Example A.5. Let $K = \mathbf{F}_2(u)$ and $f(T) = T^2 + (u^4 + u^2)T + u^5 + u^3 + u^2 + u + 1$, viewed in $\mathbf{F}_2[u][T]$ (here $S = \{\infty\}$ and $r = 1$). The field $K_1 = K(\theta)$, where $f(\theta) = 0$, ramifies over the places u and $u + 1$ of $\mathbf{F}_2(x)$. For all primes π in $\mathbf{F}_2[u]$, $\omega_f(\pi) = 1 + \chi(\pi)$, where χ is the nontrivial character on $\mathrm{Gal}(K_1/K)$. Since $\mathrm{Res}(\mathbf{F}_2[u]) = 1/\log 2$, Theorem A.3 says

$$\log 2 \prod_{\pi} \frac{1 - \omega_f(\pi)/\mathrm{N}\pi}{1 - 1/\mathrm{N}\pi} = \frac{1}{\mathrm{Res}(\mathscr{O}_{K_1,S_1})} \prod_{\pi} \left(1 - \frac{\chi(\pi)}{\mathrm{N}\pi - 1}\right) \frac{1}{1 - \chi(\pi)/\mathrm{N}\pi}. \tag{A.7}$$

To compute $\mathrm{Res}(\mathscr{O}_{K_1,S_1})$, write $\zeta_{K_1}(s) = L(z)/(1 - z)(1 - 2z)$, where $z = 2^{-s}$. Since ∞ splits in K_1, $\mathrm{Res}(\mathscr{O}_{K_1,S_1}) = L(1/2)/2\log 2$. We write $L(1/2)$ rather than $L(1)$ since we are viewing L as a function of 2^{-s}. Since K_1 has genus 3 by the Hurwitz formula (a little care is needed since ramification over u and $u + 1$ is wild), $L(z)$ must have degree 6. A calculation shows $L(z) = 1 + z + 4z^5 + 8z^6$. The factor outside the product on the right side of (A.7) is thus $(8/7)\log 2$. In Table A.2 we compare the two sides of (A.7).

n	Left side of (A.7)	Right side of (A.7)
8	$1.184185 \log 2$	$1.211391 \log 2$
9	$1.193363 \log 2$	$1.211409 \log 2$
10	$1.201499 \log 2$	$1.211417 \log 2$
11	$1.213269 \log 2$	$1.211423 \log 2$
12	$1.211185 \log 2$	$1.211422 \log 2$

Table A.2. Comparison of (A.7) using $\deg \pi \leq n$

Now we prove Theorem A.3, inspired by the abelian case over $\mathbf{Q}$ in [4, p. 124].

Proof. Our goal is to introduce additional factors into (A.2) to kill $(r - \omega_f(v))/\mathrm{N}v$, making the v-th term $1 + O(1/\mathrm{N}v^2)$ and thus making convergence absolute.

In the function field case, the zeta function for K_j has the same Euler factors as the zeta function for the maximal separable subextension over K. Thus we can replace each f_j with its p-free part, so without loss of generality all f_j are separable.

Now we recall a convenient formula for $\omega_f(v) - r$ in terms of representation theory, following [17, p. 26]. By (A.3), we have

$$\omega_f(v) = \omega_{f_1}(v) + \cdots + \omega_{f_r}(v) \tag{A.8}$$

for all but finitely many v. Since $\omega_{f_j}(v)$ counts solutions to $f_j = 0$ in $\mathscr{O}_v/\mathfrak{m}_v$, we can express $\omega_{f_j}(v)$ in terms of group characters at a Frobenius element over v. Specifically, let K_j' be a Galois extension of K containing K_j, and set $G_j = \mathrm{Gal}(K_j'/K)$, $H_j = \mathrm{Gal}(K_j'/K_j)$. The G_j-action on the distinct roots of f_j is isomorphic to the left G_j-action on G_j/H_j.

Linearize this action to a permutation representation ρ_j of G_j on $\mathbf{C}[G_j/H_j] = \operatorname{Ind}_{H_j}^{G_j}(\mathbf{C})$, with character χ_j. Then $\omega_{f_j}(v) = \chi_j(\mathrm{Frob}_v)$, where Frob_v is any Frobenius over v in G_j, at least when v is unramified in K_j'. Only a finite number of v are excluded by this constraint.

There is a trivial subrepresentation of ρ_j, and it occurs only once since ρ_j is a quotient of the regular representation of G_j. Write $\rho_j = 1_{G_j} \oplus \rho_j'$ and $\chi_j = 1 + \chi_j'$, so ρ_j' has no trivial subrepresentation. Substituting $\omega_{f_j}(v) = \chi_j(\mathrm{Frob}_v)$ into (A.8),

$$\omega_f(v) = r + \chi_1'(\mathrm{Frob}_v) + \cdots + \chi_r'(\mathrm{Frob}_v) \tag{A.9}$$

for all but finitely many v. Therefore the $1/\mathrm{N}v$ part of the factor at v in (A.2) is killed when the factor at v is multiplied by

$$\prod_{j=1}^{r} \frac{1}{\det(I - \rho_j'(\mathrm{Frob}_v)\,\mathrm{N}v^{-1})} = \prod_{j=1}^{r}\left(1 + \frac{\chi_j'(\mathrm{Frob}_v)}{\mathrm{N}v} + O\left(\frac{1}{\mathrm{N}v^2}\right)\right), \tag{A.10}$$

at least for all but finitely many v.

We can write (A.10) in terms of Euler factors at $s = 1$ for the zeta functions of K and K_j, but this requires some notation as follows. For a Galois extension of global fields E/F and a finite-dimensional complex representation ρ of $\mathrm{Gal}(E/F)$, the Artin L-function of ρ is an Euler product over places of F. For any place v of the base F, let $L(v, \rho, s)$ be the v-Euler factor of $L(\rho, s)$. This is the reciprocal of a polynomial in $\mathrm{N}v^{-s}$.

Since ρ_j' is nearly a permutation representation, the behavior of local factors of Artin L-functions under induction implies

$$\frac{1}{\det(I - \rho_j'(\mathrm{Frob}_v)\,\mathrm{N}v^{-1})} = \frac{\zeta_{K_j}(v,1)}{\zeta_K(v,1)}$$

for all but finitely many places v of K. (The notations $\zeta_K(v,s)$ and $\zeta_{K_j}(v,s)$ were defined just before Theorem A.3.) Therefore (A.9) and (A.10) imply

$$\frac{1 - \omega_f(v)/\mathrm{N}v}{(1 - 1/\mathrm{N}v)^r} \prod_{j=1}^{r} \frac{\zeta_{K_j}(v,1)}{\zeta_K(v,1)} = 1 + O\left(\frac{1}{\mathrm{N}v^2}\right). \tag{A.11}$$

Since $\zeta_K(v,1) = 1/(1 - 1/\mathrm{N}v)$, the denominators on the left side of (A.11) cancel each other.

Using (A.11), write

$$\prod_{v \notin S} \frac{1 - \omega_f(v)/\mathrm{N}v}{(1 - 1/\mathrm{N}v)^r} = \prod_{v \notin S}\left(1 - \frac{\omega_f(v)}{\mathrm{N}v}\right)\zeta_{K_1}(v,1)\cdots\zeta_{K_r}(v,1) \;\cdot\; \prod_{v \notin S}\prod_{j=1}^{r} \frac{\zeta_K(v,1)}{\zeta_{K_j}(v,1)}. \tag{A.12}$$

On the right, the first product over v is absolutely convergent by (A.11) and the second product over v converges since the other two products over v converge. The product on the left side of (A.12), and thus the second product on the right side, is usually only conditionally convergent.

We now evaluate the second product on the right side using a method that will not require *a priori* knowledge of its convergence. The partial products there involve the Euler factors of the zeta functions of K and the K_j's at $s = 1$. Asymptotics for such products are governed by the generalization of the Mertens' asymptotic $\prod_{p \le x}(1 - 1/p)^{-1} \sim e^{\gamma}\log x$: as w runs over the places of any global field F, and S is a finite set of places (with S containing

S_∞ in the number-field case),

$$\prod_{\substack{\mathrm{N}w \leq x \\ w \notin S}} \frac{1}{1 - 1/\mathrm{N}w} \sim \mathrm{Res}(\mathscr{O}_{F,S}) e^\gamma \log x \tag{A.13}$$

as $x \to \infty$. This goes back to [35, p. 274] when F is a number field and $S = S_\infty$. For a proof over global fields, see [22]. Using (A.13), the second product on the right side of (A.12) equals $\mathrm{Res}(\mathscr{O}_{K,S})^r / \mathrm{Res}(\mathscr{O}_{K_1,S_1}) \cdots \mathrm{Res}(\mathscr{O}_{K_r,S_r})$, where S_j is the set of places of K_j lying over S. (The powers of $e^\gamma \log x$ that arise in the numerator and denominator asymptotics from (A.13) exactly cancel each other.) Now divide both sides of (A.12) by $\mathrm{Res}(\mathscr{O}_{K,S})^r$. ■

When $K = \mathbf{Q}$, the formula in Theorem A.3 agrees with a formula of Davenport and Schinzel [10, p. 182], but the structure of our formula is clearer for our purposes.

Using representation theory more fully, Kurokawa [17, Theorem A2] obtained another rapidly convergent product that we have found to be useful in some calculations.

Remark A.6. It is tempting to evaluate the second product over v on the right side of (A.12) by using

$$\prod_{j=1}^{r} \frac{\zeta_{K,S}(s)}{\zeta_{K_j,S_j}(s)} = \frac{\zeta_{K,S}(s)^r}{\zeta_{K_1,S_1}(s) \cdots \zeta_{K_r,S_r}(s)} \to \frac{\mathrm{Res}(\mathscr{O}_{K,S})^r}{\mathrm{Res}(\mathscr{O}_{K_1,S_1}) \cdots \mathrm{Res}(\mathscr{O}_{K_r,S_r})}$$

as $s \to 1^+$ and invoking the analogue of Abel's theorem for Euler products. This argument gives the correct answer and is the method in [10, p. 183], but it is invalid because there is no direct analogue of Abel's theorem for Euler products. Consider [35, pp. 279–280] the identity $\zeta(2s-1)/\zeta(s) = \prod_p (1 - p^{-s})/(1 - p^{1-2s})$ for $\mathrm{Re}(s) > 1$ (with the right side absolutely convergent in this open half-plane). The left side has an analytic continuation to $\mathbf{C}$. If we formally let $s = 1$ then the product on the right is (conditionally) convergent with value 1 when we multiply terms in order of increasing p, but the analytic continuation has value $1/2$ at $s = 1$.

To test (1.8) numerically for a specific f, it is necessary to accurately estimate the associated Hardy–Littlewood constant, but this is rather time-consuming; in fact, this estimation is usually the most delicate part of numerical testing of (1.8). We therefore conclude this appendix by giving a consequence of the combination of (1.2) in the separable case and (1.8) in the inseparable case that involves no Hardy–Littlewood constants and so is much easier to check in practice.

Suppose $p \neq 2$ and $f(T) \in \kappa[u][T^p]$ satisfies the Bouniakowsky conditions (we can also take $p = 2$ if $f(T) \in \kappa[u][T^4]$). We have $f(T) = F(T^{p^m})$ for a maximal $m \geq 1$, and this p-free part $F(T)$ of $f(T)$ clearly satisfies the Bouniakowsky conditions and $F(T)$ is separable in T. We expect that $f(T)$ satisfies (1.8) and $F(T)$ satisfies (1.2). The proof of Theorem A.2 shows $C(f) = C(F)$, so dividing (1.8) for f by (1.2) for F cancels out the contribution of the mysterious Hardy–Littlewood constants and leads to the prediction

$$\frac{\#\{g \in \kappa[u] : \deg g = n, f(g) \text{ prime}\}}{\#\{g \in \kappa[u] : \deg g = n, F(g) \text{ prime}\}} \to \frac{\Lambda_\kappa(f; n)}{p^m} \tag{A.14}$$

as $n \to \infty$; the right side of (A.14) is periodic in $n \bmod 4$ for $n \gg 0$, so this limit is understood to be taken for (large) n running through a fixed congruence class modulo 4. The two sides of (A.14) can be computed independently for increasing n (using (4.12) to compute Möbius values in $\Lambda_\kappa(f;n)$), and there are no Hardy–Littlewood constants on either side.

References

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

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

[3] P. T. Bateman and R. M. . Stemmler, *Waring's problem for algebraic number fields and primes of the form* $(p^r-1)/(p^d-1)$, Illinois J. Math. **6** (1962), 142–156. MR 25 #2059

[4] P. T. Bateman and R. A. Horn, *Primes represented by irreducible polynomials in one variable*, Proc. Sympos. Pure Math., Vol. VIII, Amer. Math. Soc., Providence, R.I., 1965, pp. 119–132. MR 31 #1234

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

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

[7] J.-L. Colliot-Thélène and P. Swinnerton-Dyer, *Hasse principle and weak approximation for pencils of Severi-Brauer and similar varieties*, J. Reine Angew. Math. **453** (1994), 49–112. MR 95h:11060

[8] B. Conrad and K. Conrad, *The Möbius function and the residue theorem*, *Journal of Number Theory* (to appear).

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

[10] H. Davenport and A. Schinzel, *A note on certain arithmetical constants*, Illinois J. Math. **10** (1966), 181–185. MR 32 #5632

[11] L. E. Dickson, *A new extension of Dirichlet's theorem on prime numbers*, Mess. Math. (1904), 155–161.

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

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

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

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

[16] G. H. Hardy and J. E. Littlewood, *Some problems of Partitio Numerorum III: on the expression of a number as a sum of primes*, Acta Math. **44** (1923), 1–70.

[17] N. Kurokawa, *Special values of Euler products and Hardy-Littlewood constants*, Proc. Japan Acad. Ser. A Math. Sci. **62** (1986), 25–28. MR 87j:11127

[18] S. Lang, *Algebra*, 3rd ed., Graduate Texts in Mathematics, vol. 211, Springer-Verlag, New York, 2002. MR 2003e:00003

[19] A. E. Pellet, *Sur la décomposition d'une fonction entière en facteurs irréductibles suivant un module premier p*, C. R. Acad. Sci. Paris **86** (1878), 1071–1072.

[20] B. Poonen, *Squarefree values of multivariable polynomials*, Duke Math. J. **118** (2003), 353–373. 1 980 998

[21] B. Poonen and Y. Tschinkel, *Arithmetic of Higher-Dimensional Varieties*, Progress in Mathematics, vol. 226, Birkhäuser, Boston, 2004.

[22] M. Rosen, *A generalization of Mertens' theorem*, J. Ramanujan Math. Soc. **14** (1999), 1–19. MR 2000e:11143

[23] W. M. Ruppert, *Reducibility of cubic polynomials mod p*, J. Number Theory **57** (1996), 198–206. MR 97a:11197

[24] A. Schinzel and W. Sierpiński, *Sur certaines hypothèses concernant les nombres premiers*, Acta Arith. 4 (1958), 185–208; erratum **5** (1958), 259. MR 21 #4936

[25] J -P. Serre, *Local fields*, Graduate Texts in Mathematics, vol. 67, Springer-Verlag, New York, 1979. MR82e:12016

[26] D. Shanks and M. Lal, *Bateman's constants reconsidered and the distribution of cubic residues*, Math. Comp. **26** (1972), 265–285. MR 46 #1734

[27] D. Shanks, *Calculation and applications of Epstein zeta functions*, Math. Comp. **29** (1975), 271–287, Collection of articles dedicated to Derrick Henry Lehmer on the occasion of his seventieth birthday. MR 53 #13114a

[28] ______, *Corrigenda: "Calculation and applications of Epstein zeta functions" (Math. Comp.* **29** *(1975), 271–287)*, Math. Comp. **29** (1975), 1167. MR 53 #13114b

[29] ______, *Corrigendum: "Calculation and applications of Epstein zeta functions" (Math. Comp.* **29** *(1975), 271–287)*, Math. Comp. **30** (1976), 900. MR 53 #13114c

[30] H. N. Shapiro, *Some assertions equivalent to the prime number theorem for arithmetic progressions*, Comm. Pure Appl. Math. **2** (1949), 293–308. MR 11,419d

[31] R. G. Swan, *Factorization of polynomials over finite fields*, Pacific J. Math. **12** (1962), 1099–1106. MR 26 #2432

[32] P. Swinnerton-Dyer, *Rational points on pencils of conics and on pencils of quadrics*, J. London Math. Soc. (2) **50** (1994), 231–242. MR 95i:14022

[33] A. R. Wadsworth, *Discriminants in characteristic two*, Linear and Multilinear Algebra **17** (1985), 235–263. MR 86m:12004

[34] L. C. Washington, *A family of cubic fields and zeros of* 3*-adic* L*-functions*, J. Number Theory **63** (1997), 408–417. MR 98e:11126

[35] A. Wintner, *A factorization of the densities of the ideals in algebraic number fields*, Amer. J. Math. **68** (1946), 273–284. MR 7,416b

Department of Mathematics, University of Michigan, Ann Arbor, MI 48109-1109
E-mail address: bconrad@umich.edu

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

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