The Möbius function and the residue theorem

Brian Conrad
Department of Mathematics,
University of Michigan, Ann Arbor, MI 48109-1109
Email: bdconrad@umich.edu

Keith Conrad
Department of Mathematics,
University of Connecticut, Storrs, CT 06269-3009
Email: kconrad@math.uconn.edu

Manuscript correspondence to: Keith Conrad at above address.
Tel: (860)486-3207, Fax: (860) 486-4238

Abstract

A classical conjecture of Bouniakowsky says that a non-constant irreducible polynomial in $\mathbf{Z}[T]$ has infinitely many prime values unless there is a local obstruction. Replacing $\mathbf{Z}[T]$ with $\kappa[u][T]$, where κ is a finite field, the obvious analogue of Bouniakowsky's conjecture is false. All known counterexamples can be explained by a new obstruction, and this obstruction can be used to fix the conjecture. The situation is more subtle in characteristic 2 than in odd characteristic. Here we illustrate the general theory for characteristic 2 in some examples.

Keywords: Möbius function, residue theorem, Bouniakowsky conjecture.

In memory of Arnold E. Ross

1 Introduction

When $f(T) \in \mathbf{Z}[T]$ is a non-constant irreducible polynomial, a classical conjecture of Bouniakowsky [1] asserts that $f(n)$ is prime for infinitely many integers n unless there is a *local* obstruction, *i.e.*, unless $f(n)$ is divisible by a common prime for every $n \in \mathbf{Z}$. For example, $T^2 - T + 6$ is irreducible, but has a local obstruction at 2: $n^2 - n + 6$ is always even. Bouniakowsky's conjecture is proved when $\deg f = 1$ (this is the qualitative form of Dirichlet's theorem), but no example has been settled when $\deg f > 1$. The key point we wish to emphasize is that the philosophy underlying Bouniakowsky's conjecture (and its quantitative refinements, as in work of Hardy–Littlewood) is that statistics on prime specializations should be governed by local considerations. A survey on connections between diophantine equations and Bouniakowsky's conjecture (more precisely, the broader conjecture of Schinzel which treats primality of several polynomials rather than a single polynomial) is in [8].

While Bouniakowsky did not discuss a conjecture in $\kappa[u][T]$, where κ is a finite field, there is an obvious formulation, as follows. Let $f(T) \in \kappa[u][T]$ be any irreducible with positive T-degree. Then there should be infinitely many irreducible values $f(g)$, as g runs over $\kappa[u]$, unless there is a local obstruction, *i.e.*, unless some irreducible in $\kappa[u]$ divides $f(g)$ for every $g \in \kappa[u]$. As in the classical case, this conjecture is a theorem when $\deg_T(f) = 1$.

Surprisingly, this obvious analogue of Bouniakowsky's conjecture is not generally true! For example, $T^8 + u^3$ is irreducible in $\mathbf{F}_2[u][T]$ and has no local obstructions (the values at $T = 0$ and $T = 1$ are relatively prime in $\mathbf{F}_2[u]$). Yet for every $g \in \mathbf{F}_2[u]$, $g^8 + u^3$ is reducible. The reducibility is clear when $g(0) = 0$, but not when $g(0) = 1$. In the latter case, Swan [7] proved the reducibility by showing $g^8 + u^3$ has an even number of irreducible factors. We find it convenient to write this in the form $\mu_{\mathbf{F}_2[u]}(g^8 + u^3) = 1$, where $\mu_{\mathbf{F}_2[u]}$ is the analogue on $\mathbf{F}_2[u]$ of the classical Möbius function. Note that whereas the property of being squarefree (*i.e.*, having non-zero Möbius-value) is local, the Möbius function is inherently global.

A general analysis of the preceding phenomenon is tied up with inseparability in positive characteristic and extends to higher genus curves ($\kappa[u]$ being the case of "genus 0"), and is joint work with R. Gross that will be presented in [2] and [3]. The main discovery we have found is that there is a *global* obstruction to irreducible values of polynomials in $\kappa[u][T]$, having no classical analogue in $\mathbf{Z}[T]$. This new obstruction is related to unusual statistics for $\mu_{\kappa[u]}(f(g))$ as g varies. By "unusual statistics" we mean: the average value of $\mu_{\kappa[u]}(f(g))$, in a sense which is made precise in [2], provably does not always tend to 0. In contrast, for any non-constant $f(T) \in \mathbf{Z}[T]$, one expects (in agreement with all numerical evidence) that the average $(1/x)\sum_{n \le x} \mu(f(n))$ tends to 0, although this has not been proved in any case where $\deg f > 1$. (When $\deg f = 1$, H. N. Shapiro [5] proved the condition $(1/x)\sum_{n \le x} \mu(f(n)) \to 0$ is equivalent to the quantitative form of Dirichlet's theorem.)

The extent to which the average value of $\mu_{\kappa[u]}(f(g))$, for some $f(T) \in \kappa[u][T]$, has non-zero limiting behavior turns out to be linked not only to a corrected $\kappa[u]$-analogue of Bouniakowsky's conjecture for $f(T)$, but to a quantiative refinement, *i.e.*, to a $\kappa[u]$-analogue

of the Hardy–Littlewood conjecture on the frequency of prime values of polynomials.

Example 1.1. Let $f(T) = T^{12} + (u+1)T^6 + u^4 \in \mathbf{F}_3[u][T]$. Numerical data suggest that, as g varies over $\mathbf{F}_3[u]$, $f(g)$ is irreducible about 1.33 times as often as is predicted by naive probabilistic arguments (based on an analogy with the classical situation in $\mathbf{Z}[T]$). We are unable to prove the asymptotic relation suggested by the data, but we can rigorously produce a number close to 1.33 in the statisics for non-zero values of $\mu_{\mathbf{F}_3[u]}(f(g))$, as follows. In [2], $\mu_{\mathbf{F}_3[u]}(f(g))$ is proved to be periodic in g, with period $u(u-1)$, for $\deg g \geq 1$. A consequence, as explained in [2], is that the average non-zero value of $\mu_{\mathbf{F}_3[u]}(f(g))$, in a certain sense, is exactly $-1/3$. This differs from 1 by $4/3 = 1.333\ldots$. (Consider an agreement between naive predictions and numerical data as corresponding to a trivial correction factor of 1.)

The pattern illustrated by Example 1.1 fits into the following more general picture. For all $f(T) \in \kappa[u][T]$ that we have found to have a noticeable excess or defect of irreducible values for $f(g)$ compared to naive (local) probabilistic predictions, $f(T)$ is a polynomial in T^p. Moreover, the excess or defect of irreducible values agrees numerically with a correction factor which is related to averages of non-zero Möbius values $\mu_{\kappa[u]}(f(g))$. The definition and analysis of this correction factor is given in [2], where the general case turns out to be more complicated than what is suggested by Example 1.1 alone. In particular, the case $p = 2$ is not as well understood as the case $p \neq 2$.

Having illustrated why the behavior of $\mu_{\kappa[u]}(f(g))$ is of interest in connection with a $\kappa[u]$-analogue of Bouniakowsky's conjecture, we turn to the main focus of this paper: the extra difficulties encountered in understanding $\mu_{\kappa[u]}(f(g))$ when κ has characteristic 2 rather than odd characteristic. Let us illustrate the difference between odd characteristic and characteristic 2 by considering $f(T) = T^p + u$. When κ has characteristic $p \neq 2$ and $g \in \kappa[u]$ is non-constant, $\mu_{\kappa[u]}(g^p + u)$ admits a very simple formula:

$$\mu_{\kappa[u]}(g^p + u) = (-1)^n \chi(-1)^{n(n+1)/2} \chi(c)^n, \tag{1.1}$$

where χ is the quadratic character on $\kappa^\times$, n is the degree of g, and c is the leading coefficient of g. In particular, $\mu_{\kappa[u]}(g^p + u)$ is determined by $\chi(c)$ and n mod 4 (or just n mod 2 if -1 is a square in κ). The proof of (1.1), which is discussed in [2], is an easy application of Swan's work.

The analogue of (1.1) in characteristic 2 is more subtle:

Example 1.2. Let κ be a finite field of characteristic 2. For any $g \in \kappa[u]$ with $\deg g \geq 1$,

$$\mu_{\kappa[u]}(g^2 + u) = (-1)^{\lfloor \frac{\deg g + 1}{2} \rfloor [\kappa:\mathbf{F}_2]} (-1)^{\mathrm{Tr}_{\kappa/\mathbf{F}_2}(s_2(\omega_g))}, \tag{1.2}$$

where $\lfloor \cdot \rfloor$ is the greatest integer function, $\mathrm{Tr}_{\kappa/\mathbf{F}_2}$ is the trace,

$$\omega_g := \frac{g}{g^2 + u} \, \mathrm{d}g \tag{1.3}$$

is a rational 1-form on $\mathbf{P}^1_\kappa$, and $s_2(\omega)$ (for any rational 1-form ω on $\mathbf{P}^1_\kappa$) denotes the second elementary symmetric function of the residues of ω at its geometric poles:

$$s_2(\omega) = \sum_{y_1 \neq y_2} \mathrm{Res}_{y_1}(\omega) \, \mathrm{Res}_{y_2}(\omega) \in \kappa,$$

the sum running over unordered pairs $\{y_1, y_2\}$ of distinct geometric poles of ω on $\mathbf{P}^1_\kappa$. The proof of (1.2) is given in §3.

As an illustration of (1.2), consider $g = u + \gamma$, where $\gamma \in \kappa$. Then $g^2 + u = u^2 + u + \gamma^2$, so $\mu_{\kappa[u]}(g^2 + u) = 1$ if $u^2 + u + \gamma^2$ has a root in κ, and $\mu_{\kappa[u]}(g^2 + u) = -1$ otherwise. Whether or not $u^2 + u + \gamma^2$ splits over κ is equivalent to whether or not $\mathrm{Tr}_{\kappa/\mathbf{F}_2}(\gamma^2) = 0$, so $\mu_{\kappa[u]}(g^2 + u) = (-1)^{\mathrm{Tr}_{\kappa/\mathbf{F}_2}(\gamma^2)}$. On the other hand, the differential form ω_g has poles at the roots r_1 and r_2 of $u^2 + u + \gamma^2$, and at ∞, with respective residues $r_1 + \gamma$, $r_2 + \gamma$, and 1. These three residues have second elementary symmetric function $\gamma + 1$, so the right side of (1.2) is $(-1)^{\mathrm{Tr}_{\kappa/\mathbf{F}_2}(\gamma)}$. This agrees with our direct calculation of $\mu_{\kappa[u]}(g^2 + u)$, since γ and γ^2 have the same trace to $\mathbf{F}_2$.

Our next characteristic 2 example has no residues in its statement, but they show up in its proof.

Example 1.3. Let κ be a finite field of characteristic 2. For any $g \in \kappa[u]$ with $\deg g \geq 3$,

$$\mu_{\kappa[u]}(g^8 + (u^3 + u)g^4 + u) = 1. \tag{1.4}$$

A proof of (1.4) is given in §2, where we also show the restriction $\deg g \geq 3$ is sharp: for some $c \in \kappa^\times$, the right side of (1.4) is -1 when $g = cu^2$.

The meaning of Examples 1.2 and 1.3 in the context of a Bouniakowsky-type conjecture over $\kappa[u]$ is the following. Examples 1.2 and 1.3 involve $T^2 + u$ and $T^8 + (u^3 + u)T^4 + u$. Both are irreducible in $\kappa[u][T]$ and have no local obstructions (each has relatively prime values at $T = 0$ and $T = 1$). The Möbius formula in Example 1.3 implies that $g^8 + (u^3 + u)g^4 + u$ is reducible when $\deg g \geq 3$, and thus $T^8 + (u^3 + u)T^4 + u$ is a counterexample to the obvious analogue of Bouniakowsky's conjecture over $\kappa[u]$. On the other hand, our formula for $\mu_{\kappa[u]}(g^2 + u)$ does not immediately rule out the possibility of $g^2 + u$ being irreducible for infinitely many g, and numerical testing for $\kappa = \mathbf{F}_2$ and $\mathbf{F}_4$ supports this possibility. In fact, Example 1.2 is a case where we believe (but we are not able to prove) that the obvious $\kappa[u]$-analogue of Bouniakowsky's conjecture is true.

For characteristic 2, the main result in [2] is the following theorem that has Examples 1.2 and 1.3 as special cases.

Theorem 1.4. *Let κ be a finite field of characteristic 2 and $f(T) \in \kappa[u][T]$ be a polynomial in T^2, say $f(T) = a(T^2)$. Assume $f(T)$ is squarefree with positive T-degree and has no irreducible factor in $\kappa[u]$. For non-zero $g \in \kappa[u]$, set*

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

There exists a non-zero $M_f \in \kappa[u]$ such that for $g_1, g_2 \in \kappa[u]$ with sufficiently large degrees, the congruences $\deg g_1 \equiv \deg g_2 \bmod 4$ and $g_1 \equiv g_2 \bmod M_f$ imply

$$(-1)^{\mathrm{Tr}_{\kappa/\mathbf{F}_2}(s_2(\omega_{a,g_1}))}\mu_{\kappa[u]}(f(g_1)) = (-1)^{\mathrm{Tr}_{\kappa/\mathbf{F}_2}(s_2(\omega_{a,g_2}))}\mu_{\kappa[u]}(f(g_2)).$$

If $4|\deg_T a$ or $[\kappa : \mathbf{F}_2]$ is even, then the congruence condition on $\deg g_i$ can be dropped.

Remark 1.5. If $f(T)$ is a polynomial in T^4, so $a(T)$ is itself a polynomial in T^2, then using M_a in place of M_f and noting $\mathrm{d}(g^2) = 0$ for any g, all g_1 and g_2 of large degree in $\kappa[u]$ satisfy

$$\deg g_1 \equiv \deg g_2 \bmod 2, \quad g_1 \equiv g_2 \bmod M_a \Longrightarrow \mu_{\kappa[u]}(f(g_1)) = \mu_{\kappa[u]}(f(g_2)).$$

However, if $f(T)$ is only a polynomial in T^2, the sign $(-1)^{\mathrm{Tr}_{\kappa/\mathbf{F}_2}(s_2(\omega_{a,g}))}$ does not seem to behave in a simple manner in general. This accounts for our current inability to formulate a completely satisfactory characteristic 2 analogue of Bouniakowsky's conjecture.

Theorem 1.4 explains part of Examples 1.2 and 1.3, using $f(T) = T^2 + u$ and $f(T) = T^8 + (u^3 + u)T^4 + u$. (In the second case, $f(T)$ is a polynomial in T^4. That is why Example 1.3 has a simpler appearance than Example 1.2.) Indeed, the proof of Theorem 1.4 in [2] turns out to imply that $M_f = 1$ in Example 1.2 and $M_a = 1$ in Example 1.3. Therefore, according to Theorem 1.4, when g has sufficiently large degree,

$$(-1)^{\mathrm{Tr}_{\kappa/\mathbf{F}_2}(s_2(\omega_g))}\mu_{\kappa[u]}(g^2 + u)$$

only depends on $\deg g \bmod 4$ (ω_g is as in (1.3)) and

$$\mu(g^8 + (u^3 + u)g^4 + u)$$

is independent of g. What the proof of Theorem 1.4 does not easily tell us is the effective lower bound on $\deg g$ in the two examples.

The proof of Theorem 1.4 in [2] is long and involves a mixture of algebraic and 2-adic arguments, and the proof of the higher-genus version of Theorem 1.4 in [3] uses rigid analytic geometry and deformation theory, together with a technique for pulling up results from the genus-zero case. In the present paper, we illustrate some of the general ideas in the proof of Theorem 1.4 by proving (1.2) and (1.4) in a self-contained way, including the effective lower bounds. The methods we use in these specific examples are, for the most part, specializations of the methods used to analyze the general case in genus zero. Our hope is that working out these examples here will make the general proof of Theorem 1.4 in [2] easier to follow.

Terminology. Write W for the Witt vectors of κ (*e.g.*, $W = \mathbf{Z}_2$ when $\kappa = \mathbf{F}_2$), and write K for the fraction field of W. We will be working with polynomials in $W[u]$, and want to fix the meaning of two terms in the context of this ring. A polynomial in $W[u]$ is called *unitary* when its leading coefficient is a unit. For a non-zero polynomial $h(u) \in \kappa[u]$, a *lift* of h to $W[u]$ is any $H(u) \in W[u]$ which reduces to h and satisfies $\deg H = \deg h$. The degree condition is equivalent to requiring that H is unitary. For instance, $2u^2 + u + 3 \in W[u]$ reduces to $u + 1$ in $\kappa[u]$ but it is not considered to be a lift of $u + 1$.

We thank M. Larsen for some suggestions related to Example 1.3 and the referee for some comments on an earlier version of this paper. The first author thanks the NSF and the Sloan Foundation for financial support.

2 Example 1.3

Since polynomials in T^4 are easier to treat, we discuss Example 1.3 before Example 1.2.

Our strategy for proving (1.4) has three steps. Define

$$F(T) := T^2 + (u^3 + u)T + u \in \mathbf{Z}[u][T],$$

so (1.4) is equivalent to: $\mu_{\kappa[u]}(F(g^4)) = 1$ when $g \in \kappa[u]$ has degree at least 3.

Our first step in the direction of (1.4) will use a formula of Swan to show

$$\mu_{\kappa[u]}(F(g^4)) = \chi(R_W(F(G^4), F(G^4)')), \tag{2.1}$$

where $g \in \kappa[u]$ is non-constant, $G \in W[u]$ is any lift of g ($\deg G = \deg g$, so G is unitary), R_W is the resultant on $W[u]$ with respect to W, and χ is a certain quadratic character on $W^\times$. The derivative $F(G^4)'$ is a u-derivative.

Our second step will simplify the right side of (2.1). The resultant $R_W(F(G^4), F(G^4)')$ is difficult to compute symbolically, since $F(G^4)'$ depends on G'. We will use the residue theorem to show, that $F(G^4)'$ can be replaced with $(\partial_u F)(G^4)$ in (2.1):

$$\mu_{\kappa[u]}(F(g^4)) = \chi(R_W(F(G^4), (\partial_u F)(G^4))). \tag{2.2}$$

(The resultant in (2.1) and (2.2) is computed in characteristic 0. The characteristic 0 polynomials $F(G^4)'$ and $(\partial_u F)(G^4)$ are usually not equal, although their reductions to characteristic 2 agree.) Up to this stage, g can be any non-constant polynomial in $\kappa[u]$.

We will study the 2-adic valuations of roots of certain auxiliary polynomials in order to show $R_W(F(G^4), (\partial_u F)(G^4))$ is a square in $W^\times$ when $G \in W[u]$ is unitary and $\deg G \geq 3$. Therefore (2.2) is equal to 1 for all g with degree at least 3, which proves (1.4). A more careful study of the case $\deg G = 2$ will show us the right side of (2.2) is -1 for at least one g of degree 2.

Now we carry out this strategy.

Step 1: Derive (2.1).

We begin by recalling a formula of Swan which describes the Möbius function on separable polynomials in $\kappa[u]$ in terms of a polynomial lifting into characteristic 0. This will suffice for our intended application, since $g^8 + (u^3 + u)g^4 + u$ is separable for any g in $\kappa[u]$. Indeed, suppose an irreducible π in $\kappa[u]$ divides both $g^8 + (u^3 + u)g^4 + u$ and its derivative $(u^2 + 1)g^4 + 1$:

$$g^8 + (u^3 + u)g^4 + u \equiv 0 \bmod \pi, \quad (u^2 + 1)g^4 + 1 \equiv 0 \bmod \pi.$$

Feeding the second congruence into the first, we get $g^8 \equiv 0 \bmod \pi$, so $\pi | g$. Thus, the second congruence becomes $1 \equiv 0 \bmod \pi$, a contradiction.

For a separable $h \in \kappa[u]$, let H be any lift of h to $W[u]$. A formula of Swan [7] expresses $\mu_{\kappa[u]}(h)$ in terms of the discriminant of H:

$$\mu_{\kappa[u]}(h) = (-1)^{\deg h}\chi(\mathrm{disc}_W H), \tag{2.3}$$

where χ is a quadratic character on $W^\times$ which we define in the next paragraph. The discriminant of H, for us, is defined as

$$\mathrm{disc}_W H := \frac{(-1)^{d(d-1)/2}\prod_{H(\alpha)=0} H'(\alpha)}{(\mathrm{lead}\, H)^d}, \tag{2.4}$$

where $d = \deg H$. (This definition is unaffected when H is scaled by a non-zero constant, which is not the case for the usual definition of the polynomial discriminant in the literature.) Since h is separable, $\mathrm{disc}_W H$ is in $W^\times$. In fact, by an easy extension of Stickelberger's congruence modulo 4 for discriminants over $\mathbf{Z}$, $\mathrm{disc}_W H \in \mu_{\mathrm{odd}}(W)\cdot(1+4W)$, where $\mu_{\mathrm{odd}}(W)$ is the group of odd-order roots of unity in W.

The quadratic character χ in (2.3) will be defined in terms of the product decomposition

$$W^\times = \mu_{\mathrm{odd}}(W)\cdot(1+2W).$$

The squares $(W^\times)^2$ have index 2 in the subgroup $\mu_{\mathrm{odd}}(W)\cdot(1+4W)$ [4, p. 47]. Initially define χ as the quadratic character on $\mu_{\mathrm{odd}}(W)\cdot(1+4W)$ whose kernel is $(W^\times)^2$. Explicitly, for $\zeta \in \mu_{\mathrm{odd}}(W)$ and $w \in W$,

$$\chi(\zeta(1+4w)) = (-1)^{\mathrm{Tr}_{\kappa/\mathbf{F}_2}(w \bmod 2W)}. \tag{2.5}$$

In particular, χ is trivial on $1+8W$; this will be crucial later. To avoid the tedium of verifying that every element of $W^\times$ to which we will apply χ lies in $\mu_{\mathrm{odd}}(W)\cdot(1+4W)$, extend χ arbitrarily to a character on $W^\times$. The extended character is quadratic since $(W^\times)^2$ lies in (in fact, equals) the kernel of the original χ. We can consider χ as a quadratic character on $W^\times/(1+8W)$.

(When $\kappa = \mathbf{F}_2$, χ is the quadratic Legendre symbol $(\frac{\cdot}{2})$ on $\mathbf{Z}_2^\times/(1+8\mathbf{Z}_2)$ and (2.3) says $\mu_{\mathbf{F}_2[u]}(h) = (-1)^{\deg h}(\frac{\mathrm{disc}\, H}{2})$, which is a formula going back to Stickelberger [6].)

We return to the intended application. Choose $g \in \kappa[u]$ with degree $n \geq 1$, and let $h = F(g^4) = g^8 + (u^3+u)g^4 + u$. Since h is separable and $\deg h = 8n$, (2.3) implies

$$\mu_{\kappa[u]}(F(g^4)) = \chi(\mathrm{disc}_W(F(G^4))), \tag{2.6}$$

where $G \in W[u]$ is any lift of g. (Then $F(G^4)$ is a lift of $F(g^4)$, and G and $F(G^4)$ are unitary.)

Now we change (2.6) into an equation involving resultants. We use the standard definition of resultants: when D is a domain and H_1 and H_2 are non-zero in $D[u]$,

$$R_D(H_1,H_2) := (\mathrm{lead}\, H_1)^{\deg H_2}\prod_{H_1(\alpha)=0} H_2(\alpha), \tag{2.7}$$

the product running over the roots of H_1 (with multiplicity) in a splitting field. In steps 2 and 3 below, we will use the following three properties of resultants:

a) $R_D(H_1,H_2) = (-1)^{(\deg H_1)(\deg H_2)}R_W(H_2,H_1)$. (Thus $R_D(H_1,H_2) = R_D(H_2,H_1)$ when one of the H_j's has even degree.)

b) Resultants are bimultiplicative in each argument.

c) When $H_1 \equiv H_2 \bmod H_3$, $R_D(H_3, H_1) = (\operatorname{lead} H_3)^{\deg H_1 - \deg H_2} R_D(H_3, H_2)$.

Comparing (2.4) and (2.7),

$$\operatorname{disc}_W H = \frac{(-1)^{d(d-1)/2} R_W(H(u), H'(u))}{(\operatorname{lead} H)^{2d-1}}, \tag{2.8}$$

where $d = \deg H$. Using $H = G^8 + (u^3 + u)G^4 + u = F(G^4)$ in (2.8), we get

$$\operatorname{disc}_W(F(G^4)) = \frac{R_W(F(G^4), F(G^4)')}{(\operatorname{lead} G)^{8(16n-1)}}. \tag{2.9}$$

Since χ is quadratic, (2.6) and (2.9) imply (2.1).

Step 2: Derive (2.2).

Reduction $W[u] \to \kappa[u]$ commutes with differentiation, but does not generally commute with the calculation of resultants. The reason is that resultants depend on degrees and leading coefficients, and a leading coefficient in characteristic 0 may vanish under reduction (causing the degree to drop). For example, when $c \in W^\times$, $R_W(cu^2+u, 2u+1) = c-2 \in W^\times$ and the resultant of the reduced polynomials is $R_\kappa(\overline{c}u^2 + u, 1) = 1 \in \kappa^\times$. Usually, $c - 2 \neq 1$ in $W/2W = \kappa$.

Nevertheless, reduction and calculation of resultants can behave well together. For example, when H_1 and H_2 are both unitary polynomials in $W[u]$, the reduction of $R_W(H_1, H_2)$ equals $R_\kappa(\overline{H}_1, \overline{H}_2)$. More importantly for us, when at least one of H_1 or H_2 is unitary,

$$R_\kappa(\overline{H}_1, \overline{H}_2) \in \kappa^\times \iff R_W(H_1, H_2) \in W^\times. \tag{2.10}$$

Indeed, by property (a) of resultants, it suffices to show this equivalence when H_1 is unitary. In that case, the reason $R_W(H_1, H_2)$ may not reduce to $R_\kappa(\overline{H}_1, \overline{H}_2)$ is that the degree of $\overline{H}_2$ may be smaller than that of H_2. The effect of a degree drop in $\kappa[u]$ (in other words, computing a resultant with an artificially inflated degree assigned to one of the polynomials) is a scaling of the actual resultant by a power of the reduction of $\operatorname{lead} H_1$. This is a unit factor, which does not affect the property of a resultant lying in $\kappa^\times$ or not.

Since $F(g^4)$ is separable, so $\operatorname{disc}_\kappa(F(g^4)) \neq 0$, $R_W(F(G^4), F(G^4)') \in W^\times$ by (2.10). Since $(\partial_u F)(G^4)$ and $F(G^4)'$ have the same reduction in $\kappa[u]$, (2.10) implies the resultant $R_W(F(G^4), (\partial_u F)(G^4))$ is also in $W^\times$. Therefore (2.2) will follow from

$$\frac{R_W(F(G^4), F(G^4)')}{R_W(F(G^4), (\partial_u F)(G^4))} \in (W^\times)^2. \tag{2.11}$$

Equation (2.11) is what we will prove in the rest of Step 2.

To simplify notation, for non-zero $w_1, w_2 \in W$ we will write $w_1 \sim w_2$ to denote equality up to unit square factor (*i.e.*, $w_1/w_2 \in (W^\times)^2$).

Let $c = \operatorname{lead} G \in W^\times$, so $\operatorname{lead}(F(G^4)) = c^8 \in (W^\times)^2$. Then

$$R_W(F(G^4), F(G^4)') \sim \prod_\alpha F(G^4)'|_{u=\alpha}, \tag{2.12}$$

$$R_W(F(G^4), (\partial_u F)(G^4)) \sim \prod_\alpha (\partial_u F)(G^4)|_{u=\alpha}, \tag{2.13}$$

where α runs over the roots of $F(G^4)$ in an algebraic closure $\overline{K}$ (K is the fraction field of W). The α's in fact lie in the valuation ring of $\overline{K}$, which we write as $\overline{W}$. The suppressed square factors on the right sides of (2.12) and (2.13) are $c^{8(8n-1)}$ and $c^{8(4n+2)}$, respectively.

By the Chain Rule,

$$F(G^4)' = (\partial_u F)(G^4) + (\partial_T F)(G^4) \cdot 4G^3G'. \tag{2.14}$$

Feeding this into (2.12) and (2.13) gives

$$\begin{aligned}\frac{R_W(F(G^4), F(G^4)')}{R_W(F(G^4), (\partial_u F)(G^4))} &\sim \prod_\alpha \left(1 + 4\frac{(\partial_T F)(G^4)G^3G'}{(\partial_u F)(G^4)}\bigg|_{u=\alpha}\right) \\ &\equiv 1 + 4\sum_\alpha \frac{(\partial_T F)(G^4)G^3G'}{(\partial_u F)(G^4)}\bigg|_{u=\alpha} \bmod 8W.\end{aligned}$$

(The sum over α is in W since it is Galois-invariant and $(\partial_u F)(G^4)|_{u=\alpha} \in \overline{W}^\times$.)

Let $P = F(G^4) \in W[u]$. Since P is unitary and its reduction in $\kappa[u]$ is separable, its roots in $\overline{K}$ lie in $\overline{W}$ and are simple; different roots have different reductions in the residue field of $\overline{W}$. Hence, we arrive at the key observation: P is a local parameter at each of its roots, so we can write each term in the sum over α as a residue at α:

$$\frac{(\partial_T F)(G^4)G^3G'}{(\partial_u F)(G^4)}\bigg|_{u=\alpha} = \operatorname{Res}_\alpha\left(\frac{(\partial_T F)(G^4)G^3G'}{(\partial_u F)(G^4)}\,\frac{\mathrm{d}P}{P}\right).$$

Computing $\mathrm{d}P$, the differential form on the right side can be written as

$$\begin{aligned}\frac{(\partial_T F)(G^4)G^3G'}{(\partial_u F)(G^4)}\,\frac{\mathrm{d}P}{P} &= \frac{(\partial_T F)(G^4)G^3G'}{(\partial_u F)(G^4)}\,\frac{(\partial_u F)(G^4) + (\partial_T F)(G^4)4G^3G'}{F(G^4)}\,\mathrm{d}u \\ &= \frac{(\partial_T F)(G^4)G^4}{F(G^4)}\,\frac{\mathrm{d}G}{G} + \frac{4((\partial_T F)(G^4)G^3G')^2}{(\partial_u F)(G^4)F(G^4)}\,\mathrm{d}u.\end{aligned}$$

Since $F(g^4)$ is separable in characteristic 2, the residue (at each α) of the second differential form on the right side lies in $4\overline{W}$, so the sum is in $4W$, and hence is in $2W$. Therefore

$$\sum_\alpha \frac{(\partial_T F)(G^4)G^3G'}{(\partial_u F)(G^4)}\bigg|_{u=\alpha} \equiv \sum_\alpha \operatorname{Res}_\alpha\left(\frac{(\partial_T F)(G^4)G^4}{F(G^4)}\,\frac{\mathrm{d}G}{G}\right) \bmod 2W.$$

The roots α of $F(G^4)$ include all the poles of $((\partial_T F)(G^4)G^3/F(G^4))\,\mathrm{d}G$ except perhaps ∞. Moreover, reduction $\alpha \mapsto \overline{\alpha}$ gives a bijection between the geometric roots of $F(G^4)$ and $F(g^4)$. Therefore by the residue theorem over κ,

$$\begin{aligned}\sum_{\overline{\alpha}} \operatorname{Res}_{\overline{\alpha}}\left(\frac{(\partial_T F)(g^4)g^4}{F(g^4)}\,\frac{\mathrm{d}g}{g}\right) &= -\operatorname{Res}_\infty\left(\frac{(\partial_T F)(g^4)g^4}{F(g^4)}\,\frac{\mathrm{d}g}{g}\right) \\ &= -\deg_T(F)\operatorname{ord}_\infty(g) \\ &= 0\end{aligned}$$

since $\deg_T(F) = 0$ in κ. This establishes (2.11), and therefore (2.2).

Step 3: Let $G = cu^n + \cdots$ in $W[u]$ with $c \in W^\times$ and $n \geq 1$. We will prove the resultant $R_W(F(G^4), (\partial_u F)(G^4))$ is a square in $W^\times$ when $n \geq 3$ and that it is not in the kernel of χ for some G of degree 2.

We write out the resultant more fully:

$$R_W(F(G^4), (\partial_u F)(G^4)) = R_W(G^8 + (u^3 + u)G^4 + u, (3u^2 + 1)G^4 + 1).$$

Using some resultant algebra, we are going to show

$$R_W(F(G^4), (\partial_u F)(G^4)) \sim R_W(6u^5 + 2u^3 + 1, G^4 - 2u^3), \tag{2.15}$$

where $\sim$ has the same meaning as in the discussion of Step 2. We then will use 2-adic algebra to prove $R_W(6u^5 + 2u^3 + 1, G^4 - 2u^3)$ is a square in $W^\times$ when $n \geq 3$.

Using properties (a), (b), and (c) of resultants as listed above (2.8), we get

$$\begin{aligned}
R_W(F(G^4), (\partial_u F)(G^4)) &= R_W((3u^2 + 1)G^4 + 1, G^8 + (u^3 + u)G^4 + u) \\
&= R_W((3u^2 + 1)G^4 + 1, G^8 + u(-2u^2G^4 - 1) + u) \\
&= R_W((3u^2 + 1)G^4 + 1, G^8 - 2u^3G^4) \\
&= R_W((3u^2 + 1)G^4 + 1, G)^4 \cdot R_W((3u^2 + 1)G^4 + 1, G^4 - 2u^3) \\
&= R_W(G, (3u^2 + 1)G^4 + 1)^4 \cdot R_W(G^4 - 2u^3, (3u^2 + 1)G^4 + 1) \\
&= c^{4(4n+2)} \cdot c^{4(4n-3)} R_W(G^4 - 2u^3, (3u^2 + 1)2u^3 + 1) \\
&\sim R_W(6u^5 + 2u^3 + 1, G^4 - 2u^3),
\end{aligned}$$

which is (2.15).

Let β run over the roots of $6u^5 + 2u^3 + 1$ in $\overline{K}$. Clearly $|\beta|_2 > 1$, so $G(\beta) \neq 0$ (roots of G are integral over W). Therefore

$$\begin{aligned}
R_W(6u^5 + 2u^3 + 1, G^4 - 2u^3) &= 6^{4n} \prod_\beta (G(\beta)^4 - 2\beta^3) \\
&= \left(6^n \prod_\beta G(\beta)\right)^4 \prod_\beta \frac{G(\beta)^4 - 2\beta^3}{G(\beta)^4} \\
&= R_W(6u^5 + 2u^3 + 1, G)^4 \prod_\beta \left(1 - \frac{2\beta^3}{G(\beta)^4}\right).
\end{aligned}$$

Normalizing the 2-adic valuation to be the usual one on $\mathbf{Q}_2$, each β has 2-adic valuation $-1/5$ ($1/\beta$ is the root of an Eisenstein polynomial) and G is unitary, so $G(\beta)$ has 2-adic valuation $-n/5$ and $2\beta^3/G(\beta)^4$ has 2-adic valuation $(4n+2)/5 > 1$. Therefore the product over β, which is a field norm down to K, lies in $W^\times$.

Since $R_W(6u^5 + 2u^3 + 1, G) = 6^n \prod_\beta G(\beta)$ and there are five β's, the 2-adic valuation of $R_W(6u^5 + 2u^3 + 1, G)$ is $n + 5(-n/5) = 0$. Thus, $R_W(6u^5 + 2u^3 + 1, G) \in W^\times$, so its fourth power is in $(W^\times)^2$. Therefore by (2.2),

$$\mu_{\kappa[u]}(F(g^4)) = \chi\left(\prod_\beta \left(1 - \frac{2\beta^3}{G(\beta)^4}\right)\right) \tag{2.16}$$

as long as $n \geq 1$.

We now look at $\prod_\beta (1 - 2\beta^3/G(\beta)^4)$. The valuation of $2\beta^3/G(\beta)^4$, which is $(4n+2)/5$, is ≥ 3 for $n \geq 4$. When $n = 3$, the valuation of $2\beta^3/G(\beta)^4$ is $14/5 > 2$, so

$$\prod_\beta \left(1 - \frac{2\beta^3}{G(\beta)^4}\right) \equiv 1 \bmod 2^{14/5}. \tag{2.17}$$

The left side of (2.17) lies in $W^\times$ and W is unramified over $\mathbf{Z}_2$, so (2.17) actually holds modulo 2^3. Thus $\prod_\beta (1 - 2\beta^3/G(\beta)^4)$ lies in $1 + 8W \subset (W^\times)^2$ for $n \geq 3$, not just for $n \geq 4$. This concludes the explanation of Example 1.3 when $n \geq 3$.

Now consider the case $n = 2$. We will show $\mu_{\kappa[u]}(F(g^4)) = -1$ for some g in $\kappa[u]$ which is a $\kappa^\times$-multiple of u^2. Considering (2.16), our task is equivalent to finding $c \in W^\times$ such that the product

$$\prod_\beta \left(1 - \frac{2\beta^3}{(c\beta^2)^4}\right) = \prod_\beta \left(1 - \frac{2}{c^4\beta^5}\right),$$

which lies in $W^\times$, is not in the kernel of χ. Noting $2\beta^5$ has valuation 0, we compute the product modulo 2^4 as

$$\begin{aligned} \prod_\beta \left(1 - \frac{2}{c^4\beta^5}\right) &= \prod_\beta \left(1 - \frac{4}{c^4(2\beta^5)}\right) \\ &\equiv 1 - 4\sum_\beta \frac{1}{c^4(2\beta^5)} \bmod 2^4. \end{aligned}$$

Let β_0 denote one of the β's, and let $L = K(\beta_0)$. The sum over β is $c^{-4}\operatorname{Tr}_{L/K}(1/(2\beta_0)^5)$, which lies in W, so (2.16) and the definition (2.5) of χ tell us

$$\mu_{\kappa[u]}(F((\overline{c}u^2)^4)) = (-1)^{\operatorname{Tr}_{W/\mathbf{Z}_2}(c^{-4}\operatorname{Tr}_{\mathscr{O}_L/W}(1/(2\beta_0^5)))}. \tag{2.18}$$

Since L/K is totally ramified, $\operatorname{Tr}_{\mathscr{O}_L/W}(x) \equiv [\mathscr{O}_L : W]x \bmod \mathfrak{m}_L$ for any $x \in \mathscr{O}_L$. Therefore, since $[\mathscr{O}_L : W] = 5$, $\operatorname{Tr}_{\mathscr{O}_L/W}(1/(2\beta_0^5))$ lies in $W^\times$. Looking again at (2.18), it is now immediate that c exists such that $\mu(F((\overline{c}u^2)^4)) = -1$. (Running through this norm and trace argument with a unitary lift G of any g with degree $n \geq 2$, $\mu_{\kappa[u]}(F(g^4))$ equals $(-1)^{\operatorname{Tr}_{W/\mathbf{Z}_2}(\beta_0^3/(2G(\beta_0)^4))}$. The 2-adic valuation of $\beta_0^3/(2G(\beta_0)^4)$ is $(4n-8)/5 \geq 0$, which gives an alternate conclusion to the proof of (1.4) for $n \geq 3$.)

3 Example 1.2

The explanation of Example 1.2 follows ideas similar to those of Example 1.3, but we meet some new complications because $T^2 + u$ is not a polynomial in T^4. (The convenience of polynomials in T^4 is due to calculations like (2.14), which introduce a factor of 4 in undesirable terms involving G', thus simplifying mod 8 computations to mod 2 computations, *i.e.*, to computations in the residue field of an unramified extension of W.)

Let $g \in \kappa[u]$ have degree $n \geq 1$. Since $g^2 + u$ has degree $2n$, (2.3) says

$$\mu_{\kappa[u]}(g^2 + u) = \chi(\operatorname{disc}_W(G^2 + u)), \tag{3.1}$$

where G is a lift of g to $W[u]$. In particular, G has degree n. Set $c = \operatorname{lead} G \in W^\times$.

Letting $\alpha \in \overline{W}$ run over the roots of $G^2 + u$ in $\overline{K}$,

$$\operatorname{disc}_W(G^2 + u) = \frac{(-1)^n}{c^{2n}} \prod_\alpha (1 + 2G(\alpha)G'(\alpha)). \tag{3.2}$$

Since $Q := G^2 + u$ is a local parameter at each α, $G(\alpha)G'(\alpha) = \operatorname{Res}_\alpha(GG'(\mathrm{d}Q/Q))$. Writing

$$\begin{aligned} GG' \frac{\mathrm{d}Q}{Q} &= \frac{GG'(1 + 2GG')}{G^2 + u} \,\mathrm{d}u \\ &= \frac{G}{G^2 + u} \,\mathrm{d}G + 2 \cdot (GG')^2 \frac{\mathrm{d}u}{G^2 + u}, \end{aligned}$$

we have

$$\prod_\alpha (1 + 2G(\alpha)G'(\alpha)) = \prod_\alpha \left(1 + 2 \operatorname{Res}_\alpha \left(\frac{G}{G^2 + u} \,\mathrm{d}G\right) + 4 \operatorname{Res}_\alpha \left((GG')^2 \frac{\mathrm{d}u}{G^2 + u}\right)\right). \tag{3.3}$$

The product only matters modulo 8, since in (3.1) we apply χ to the product. Since $g^2 + u$ is separable in characteristic 2, each α lies in an unramified extension of K. Thus, each of the residues on the right side of (3.3) is in W^{nr}, the valuation ring of the maximal unramified extension of K inside of $\overline{K}$.

Let $\omega_G = (G/(G^2 + u)) \,\mathrm{d}G$ and $\omega_g = (g/(g^2 + u)) \,\mathrm{d}g$. These are rational 1-forms on $\mathbf{P}^1_K$ and $\mathbf{P}^1_\kappa$. The right side of (3.3) is congruent modulo $8W^{\mathrm{nr}}$ to

$$1 + 2 \sum_\alpha \operatorname{Res}_\alpha(\omega_G) + 4 \sum_\alpha \operatorname{Res}_\alpha \left((GG')^2 \frac{\mathrm{d}u}{G^2 + u}\right) + 4 \sum_{\alpha_1 \neq \alpha_2} \operatorname{Res}_{\alpha_1}(\omega_G) \operatorname{Res}_{\alpha_2}(\omega_G). \tag{3.4}$$

The first sum in (3.4) is $-\operatorname{Res}_\infty(\omega_G)$. The second sum only matters mod $2W^{\mathrm{nr}}$, so we compute its reduction mod 2. For each $\overline{\alpha}$, $g^2 + u$ is a local parameter at $\overline{\alpha}$. Thus, since the residue characteristic is 2, we have:

$$\begin{aligned} \operatorname{Res}_{\overline{\alpha}} \left((gg')^2 \frac{\mathrm{d}u}{g^2 + u}\right) &= \operatorname{Res}_{\overline{\alpha}} \left((gg')^2 \frac{\mathrm{d}(g^2 + u)}{g^2 + u}\right) \\ &= \left(\operatorname{Res}_{\overline{\alpha}} \left(gg' \frac{\mathrm{d}(g^2 + u)}{g^2 + u}\right)\right)^2 \\ &= \left(\operatorname{Res}_{\overline{\alpha}} \left(g \frac{g' \mathrm{d}u}{g^2 + u}\right)\right)^2 \\ &= (\operatorname{Res}_{\overline{\alpha}} \omega_g)^2. \end{aligned}$$

Since $\alpha \mapsto \overline{\alpha}$ is a bijection between geometric roots of $G^2 + u$ and $g^2 + u$,

$$\sum_\alpha \operatorname{Res}_{\overline{\alpha}} \left(\frac{(gg')^2}{g^2 + u} \,\mathrm{d}u\right) = \sum_{\overline{\alpha}} (\operatorname{Res}_{\overline{\alpha}} \omega_g)^2 = \left(\sum_{\overline{\alpha}} \operatorname{Res}_{\overline{\alpha}} \omega_g\right)^2 = (\operatorname{Res}_\infty \omega_g)^2 \tag{3.5}$$

by the residue theorem over κ.

We conclude from preceding calculations that $\prod_\alpha(1+2G(\alpha)G'(\alpha))$ is congruent modulo $8W^{\mathrm{nr}}$ (even modulo $8W$, by Galois invariance) to

$$1 - 2\,\mathrm{Res}_\infty(\omega_G) + 4(\mathrm{Res}_\infty\,\omega_G)^2 + 4\sum_{\alpha_1\neq\alpha_2}\mathrm{Res}_{\alpha_1}(\omega_G)\,\mathrm{Res}_{\alpha_2}(\omega_G). \tag{3.6}$$

The second term in (3.6) is $-2\,\mathrm{Res}_\infty(\omega_G) = 2n$ and the third is $-4\sum_\alpha \mathrm{Res}_\alpha(\omega_G)\,\mathrm{Res}_\infty(\omega_G)$. Since we are working modulo 8, so -4 may be replaced with 4, equations (3.1), (3.2), (3.3), and (3.6) give

$$\mu_{\kappa[u]}(g^2+u) = \chi((-1)^n(1+2n+4s_2(\omega_g))) = \chi((-1)^n(1+2n)+4s_2(\omega_g)).$$

Since $(-1)^n(1+2n) \equiv 1 + 4\lfloor\frac{1+n}{2}\rfloor \bmod 8$, (2.5) gives

$$\mu_{\kappa[u]}(g^2+u) = (-1)^{\mathrm{Tr}_{\kappa/\mathbf{F}_2}(\lfloor\frac{1+n}{2}\rfloor + s_2(\omega_g))} = (-1)^{\lfloor\frac{1+n}{2}\rfloor[\kappa:\mathbf{F}_2]}(-1)^{\mathrm{Tr}_{\kappa/\mathbf{F}_2}(s_2(\omega_g))},$$

concluding the verification of Example 1.2.

Remark 3.1. In the general proof of Theorem 1.4, the 1-form $(gg')^2\mathrm{d}u/(g^2+u)$ is replaced with $\eta_g := s_g^2\mathrm{d}r_g/r_g$ where r_g equals $f(g)$ and s_g is a certain rational function (depending on g and g'), and the role of ω_g is played by $\theta_g := s_g\mathrm{d}r_g/r_g$. In such generality, a calculation much like (3.5) equates the sum of the residues of η_g at poles of θ_g on $\mathbf{A}^1_\kappa$ with the sum of $(\mathrm{Res}_\infty\theta_g)^2$ and the residues of η_g at its poles $x \neq \infty$ that are *not* poles of θ_g.

The miracle is that η_g has vanishing residue at such x because in characteristic 2

$$\mathrm{Res}_x\eta_g = \mathrm{Res}_x\left(s_g^2\,\frac{\mathrm{d}r_g}{r_g}\right) = \left(\mathrm{Res}_x\left(s_g\,\frac{\mathrm{d}r_g}{r_g}\right)\right)^2 = (\mathrm{Res}_x\theta_g)^2 = 0.$$

Thus, the sum of the residues of η_g at the finite poles of θ_g is equal to $\mathrm{Res}_\infty\eta_g = (\mathrm{Res}_\infty\theta_g)^2$, as in (3.5), even when η_g has finite poles away from the finite poles of θ_g. This is crucial in the proof of Theorem 1.4.

References

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

[2] B. Conrad, K. Conrad, R. Gross, Irreducible specialization in genus 0 (submitted).

[3] B. Conrad, K. Conrad, R. Gross, Irreducible specialization in higher genus (in preparation).

[4] S. Lang, "Algebraic Number Theory," 2nd ed., Springer-Verlag, New York, 1994.

[5] H. N. Shapiro, Some assertions equivalent to the prime number theorem for arithmetic progressions, Comm. Pure Appl. Math. 2 (1949) 293–308.

[6] L. Stickelberger, Über eine neue Eigenschaft der Diskriminanten algebraischer Zahlkörper, in: *Verhandlungen des Ersten Internationalen Mathematiker-Kongresses* (Leipzig, 1898), Krause Reprint Limited, Nendeln, 1967, pp. 182–193.

[7] R. Swan, Factorization of polynomials over finite fields, Pacific J. Math. 12 (1962) 1099–1106.

[8] P. Swinnerton-Dyer, Some applications of Schinzel's hypothesis to Diophantine equations, in: K. Győry, H. Iwaniec, J. Urbanowicz (Eds.), Number Theory in Progress (Vol. 1), de Gruyter, Berlin, 1999, pp. 503–530.