The autocorrelation of the Möbius function and Chowla’s conjecture for the rational function field in characteristic 2

Dan Carmon

Raymond and Beverly Sackler School of Mathematical Sciences, Tel Aviv University, Tel Aviv 69978, Israel

We prove a function field version of Chowla’s conjecture on the autocorrelation of the Möbius function in the limit of a large finite field of characteristic 2, extending previous work in odd characteristic.

1. Introduction

Let \( \mathbb{F}_q \) be a finite field of \( q \) elements, and let \( \mathbb{F}_q[x] \) be the polynomial ring over \( \mathbb{F}_q \). The Möbius function of a non-zero polynomial \( F \in \mathbb{F}_q[x] \) is defined to be \( \mu(F) = (-1)^r \) if \( F = cP_1 \cdots P_r \) with \( 0 \neq c \in \mathbb{F}_q \) and \( P_1, \ldots, P_r \) are distinct monic irreducible polynomials, and \( \mu(F) = 0 \) otherwise. Let \( M_n \subset \mathbb{F}_q[x] \) be the set of monic polynomials of degree \( n \) over \( \mathbb{F}_q \), which is of size \( \#M_n = q^n \).

For \( r > 0 \), distinct polynomials \( \alpha_1, \ldots, \alpha_r \in \mathbb{F}_q[x] \) with deg \( \alpha_i < n \), and \( \epsilon_i \in \{1, 2\} \), not all even, set

\[
C(\alpha_1, \ldots, \alpha_r; n) := \sum_{F \in M_n} \mu(F + \alpha_1)^{\epsilon_1} \cdots \mu(F + \alpha_r)^{\epsilon_r}.
\] (1.1)

In Carmon & Rudnick [1], an upper bound on \( |C(\alpha_1, \ldots, \alpha_r; n)| \) was established for fields of odd characteristic, demonstrating that \( \lim_{q \to \infty} (1/\#M_n) \sum_{F \in M_n} C(\alpha_1, \ldots, \alpha_r; n) = 0 \) holds for any fixed \( n > 1, r > 1 \). This is analogous to Chowla’s conjecture over function fields, in the limit of a large finite field.

This result has since found further applications. Bary-Soroker [2] uses a result similar to a part of the proof, named ‘square independence’, and computes a certain Galois group to be \( S_n^r \). This computation then implies many equidistribution and independence results, proving function field analogues, in the limit of a large base field, to myriad classical problems, such as...
as the Hardy–Littlewood conjecture, and the additive and Titchmarsh divisor problems; see Andrade et al. [3] for more details and examples. We stress that Bary-Soroker’s computation, and any implications thereof, were only valid in odd characteristic, owing to square independence having been established only in odd characteristic.

In this paper, we shall provide a bound on \(|C(\alpha_1, \ldots, \alpha_r; n)|\) in the case of characteristic 2, yielding the analogue to Chowla’s conjecture in this setting. We shall also verify square independence in characteristic 2, thus extending the validity of Bary-Soroker’s computation and all its implications.

Henceforth, we shall assume that \(q\) is even. As in odd characteristic, for \(r = 1\) and \(n > 1\), we have \(\sum_{F \in M_n} \mu(F) = 0\). For \(n = 1\), we have \(\mu(F) \equiv -1\) and the sum equals \((-1)^{\sum \epsilon_i}\). The case \(n = 2\) is a new special case in characteristic 2 and will be handled separately in §5. For \(n > 2, r > 1\), we show:

**Theorem 1.1.** Fix \(r > 1\) and assume that \(n > 2\) and that \(q\) is even. Then for any choice of distinct polynomials \(\alpha_1, \ldots, \alpha_r \in \mathbb{F}_q[x]\) with \(\max \deg \alpha_j < n\), and \(\epsilon_i \in \{1, 2\}\), not all even,

\[
|C(\alpha_1, \ldots, \alpha_r; n)| \leq r n q^{n-1/2} + \frac{3}{4} (r + 3) n^2 q^{n-1}. \tag{1.2}
\]

2. Analogues in characteristic 2

The starting point in [1] was Pellet’s formula, expressing the Möbius function in terms of the quadratic character of the discriminant

\[
\mu(F) = (-1)^{\deg F} \chi_2(\text{disc } F).
\]

For even \(q\), Pellet’s formula does not hold; indeed, even the usual quadratic character \(\chi_2\) itself is meaningless, as every element of \(\mathbb{F}_q\) is the square of another. There is, however, a similar formula which uses Berlekamp’s discriminant (first defined in [4]). We shall repeat here the definitions and required properties of Berlekamp’s discriminant.

(a) Definition of Berlekamp’s discriminant

Given a polynomial \(F(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_0, a_n \neq 0\) with coefficients in \(\mathbb{F}_q\), let \(r_1, \ldots, r_n\) be its roots in some algebraic extension of \(\mathbb{F}_q\). The Berlekamp discriminant of \(F\) is defined in terms of its roots as

\[
\text{Berl}(F) = \sum_{i < j} \frac{r_i r_j}{r_i^2 + r_j^2}. \tag{2.1}
\]

The expression \(\text{Berl}(F)\) is symmetric in the roots of \(F\), hence it is in \(\mathbb{F}_q\), and its value is independent of the extension used. Furthermore, taking a common denominator, we may write

\[
\frac{a_n^{2n-2} \sum_{i < j} (r_i r_j) \prod_{i' < j'} (r_i^2 + r_j^2))}{a_n^{2n-2} \prod_{i < j} (r_i^2 + r_j^2)}. \tag{2.2}
\]

Note that both the denominator and the numerator are symmetric polynomials in the roots of \(F\). Hence, they are homogeneous polynomials\(^1\) (over \(\mathbb{F}_2\)) in the coefficients of \(F\) of degree \(2n - 2\).

\(^1\)It is perhaps not trivial that they are indeed polynomials, rather than rational functions with a power of \(a_n\) in their denominators. We will see that they are indeed polynomials as a by-product of their computation. Furthermore, in all of our applications, \(F\) will be monic.
Furthermore, the denominator is in fact the discriminant of $F$, for, in characteristic 2, $r_i^2 + r_j^2 = (r_i - r_j)^2$. Following Berlekamp, we denote the numerator of Berl($F$) by $\xi(F)$, that is,

$$\text{Berl}(F) = \frac{\xi(F)}{\text{disc } F}.$$  

(2.3)

Note also that, in characteristic 2, $\text{disc } F = \delta(F)^2$, where

$$\delta(F) = a_{n-1}^n \prod_{i<j}(r_i + r_j)$$

is a polynomial in the coefficients of $F$ with total degree $n - 1$, and degree at most $d(n) = \lfloor (n - 1)/2 \rfloor$ in $a_0$—its leading term, as a polynomial in $a_0$, is $a_{n/2}^{(n-1)/2}a_0^{(n-1)/2}$ for odd $n$, and $a_{n/2}^{n/2}a_0^{(n-2)/2}$ for even $n$. The formulae for Berlekamp’s discriminant for degrees up to 3 are

$$\text{Berl}(ax + b) = \frac{0}{1},$$

$$\text{Berl}(ax^2 + bx + c) = \frac{ac}{b^2}$$

and

$$\text{Berl}(ax^3 + bx^2 + cx + d) = \frac{a^2d^2 + abcd + b^3d + ac^3}{(ad + bc)^2}.$$  

(b) Effective computation of Berlekamp’s discriminant

The formulae above do not lend themselves immediately to computations of Berl($F$) or $\xi(F)$ in terms of the coefficients of $F$. A computational method can be obtained by first lifting the coefficients of $F$ from $\mathbb{F}_q$ to a field of characteristic 0. To do so, choose an algebraic extension $K$ of $\mathbb{Q}$ such that $K$ becomes isomorphic to $\mathbb{F}_q$ when reduced modulo 2, and choose any lifting $F_0(x) = a_{0,n}x^n + a_{0,n-1}x^{n-1} + \cdots + a_{0,0}$ with coefficients in $K$ such that $F_0 \equiv F$ (mod 2). If the roots of $F_0$ in the algebraic closure are $r_{0,1}, \ldots, r_{0,n}$ (such that $r_{0,i} \equiv r_i$ (mod 2)) we have $\text{disc } F_0 = a_{0,n}^{2n-2} \prod_{i<j}(r_{0,i} - r_{0,j})^2$. Define similarly $\text{disc}_+ F_0 = a_{0,n}^{2n-2} \prod_{i<j}(r_{0,i} + r_{0,j})^2$. Note that $\text{disc}_+ F_0$ is a symmetric polynomial in the roots of $F_0$, and therefore an integral polynomial in the coefficients of $F_0$—in fact, it is the square of such a polynomial. Furthermore, it is clear that $\text{disc } F_0 \equiv \text{disc}_+ F_0$ (mod 4), as polynomials in either the roots or coefficients. Therefore, the expression $\xi_0(F_0) = (\text{disc}_+ F_0 - \text{disc } F_0)/4$ is also an integral polynomial in the coefficients of $F_0$. It is now easy to verify that, when $\xi_0$ is reduced modulo 2, the obtained polynomial must indeed be equal to $\xi(F)$, as given by the numerator of formula (2.2). In particular, the result of this process is independent of the lifting.

In fact, in our computations we will use a simpler lifting. We shall always compute the discriminants in general settings, where all of the coefficients of $F$ are either 0, 1 or a symbol from a set of variables $V$, never any explicit value in $\mathbb{F}_q$. Thus we need only lift the coefficients from $\mathbb{F}_2[V]$ to $\mathbb{Z}[V]$, which can be done, for example, by lifting 0,1 $\in \mathbb{F}_2$ to 0,1 $\in \mathbb{Z}$. The discriminants computed in this manner will yield the appropriate polynomials in $\mathbb{F}_2[V]$, and hence the same formulae will also be valid for any substitution of values from $\mathbb{F}_q$ into the variables of $V$.

Finally, we note that $\text{disc}_+(F_0)$, like the discriminant, can be expressed as a resultant, and is therefore easily computable,

$$\text{disc}_+(f(x)) = \text{Res}(f(x),f(-x)) \quad \text{Res}\left(\frac{f(x) - f(-x)}{2x},\frac{f(x) + f(-x)}{2}\right).$$  

(2.4)

Splitting $f$ into its even and odd parts as $f(x) = g(x^2) + x h(x^2)$, we may rewrite formula (2.4) as

$$\text{disc}_+(f(x)) = \text{Res}(h(x^2),g(x^2)) = \text{Res}(h(x),g(x))^2.$$  

(2.5)

Using these methods, we show the following technical lemma:

**Lemma 2.1.** For any $n > 2$, the degree of $\xi(F)$ in $a_0$ is at most $2d(n)$. 


Proof. For all \( n \), both \( \text{disc}(F_0) \), \( \text{disc}_+(F_0) \) have degree at most \( n - 1 \) in \( a_0 \). For odd \( n \), \( 2d(n) = n - 1 \), and we are done. For even \( n \), \( 2d(n) = n - 2 \), and so we must check that the coefficient of \( a_0^{-1} \) in \( \xi(F) \) vanishes. And, indeed, \( \text{disc}_+(F_0) \) is a square, hence its degree in \( a_0 \) must be even, and, therefore, less than \( n - 1 \). On the other hand, the leading coefficient of \( \text{disc}(F_0) \) is known to be \( \pm n! a_0^{n-1} a_0^{n-1} \), which is clearly 0 mod 8 for any even \( n > 2 \).

Note that the consequent of lemma 2.1 is false for \( n = 2 \), as \( \xi(F) = a_2 a_0 \) while \( 2d(2) = 0 \).

(c) An analogue to Pellet’s formula

The main theorem in [4] provides an analogue to Pellet’s formula in characteristic 2. We restate it here in more familiar terms. Let \( \chi_2 : \mathbb{F}_q \rightarrow \{ \pm 1 \} \) be defined by \( \chi_2(x) = 1 \) if \( x = y^2 + y \) for some \( y \in \mathbb{F}_q \), and \( \chi_2(x) = -1 \) otherwise. Note that the map \( y \mapsto y^2 + y \) is linear over \( \mathbb{F}_2 \), and its kernel is the set \( \{0, 1\} \). Therefore, its image is an \( \mathbb{F}_2 \)-linear subspace of \( \mathbb{F}_q \) with co-dimension 1, and thus \( \chi_2 \) is a group homomorphism, i.e. \( \chi_2(x + y) = \chi_2(x) \chi_2(y) \). We are interested in evaluating \( \chi_2(\text{Ber} F) \). From (2.1), we may write \( \text{Ber} F = \beta^2 + \beta \), where \( \beta = \sum_{i \leq j} (r_i / (r_i + r_j)) \in \mathbb{F}_q \), so we need only determine whether \( \beta \in \mathbb{F}_q \). Note that any odd permutation of the roots \( r_i \) changes \( \beta \) to \( \beta + 1 \), and that \( \beta \) is fixed under any even permutation. As \( \beta \in \mathbb{F}_q \) iff \( \beta \) is fixed under the Frobenius endomorphism, the value of \( \chi_2 \) is determined by the sign of the permutation on \( r_i \) given by the Frobenius endomorphism. The following analogue to Pellet’s formula is now immediate:

\[
\mu(F) = (-1)^{\deg F} \chi_2(\text{Ber} F).
\]

Note that this formula is only valid when \( F \) is square-free. Otherwise, \( \text{disc} F = 0 \), in which case \( \text{Ber} F \) is not properly defined. Somewhat informally, we may correct this by assigning \( \chi_2(\infty) = 0 \). However, this difficulty is more easily avoided by assuming that \( \text{disc}(F + a_i) \neq 0 \) for all \( i \). Indeed, as there are exactly \( q^{n-1} \) polynomials in \( M_n \) with \( \text{disc} F = 0 \), this assumption fails for at most \( q^{n-1} \) \( n \)-tuples \( a_0, \ldots, a_{n-1} \). This negligible error will be collected into the \( \frac{\chi_2(\text{Ber} F)}{d} \) term further along. We may henceforth assume that all \( \epsilon_i = 1 \), as terms with \( \epsilon_i = 2 \) do not affect the remaining non-vanishing summands at all.

3. Reduction to a counting problem

Continuing analogously to [1], we may write

\[
C(\alpha_1, \ldots, \alpha_r; n) = (-1)^m \sum_{F \in M_n} \chi_2(\text{Ber} F + \alpha_1) + \cdots + \text{Ber} F + \alpha_r) \chi_2(\text{Ber} F + \alpha_r)).
\]

We single out the constant term \( t := F(0) \) of \( F \in M_n \) and write \( F(x) = f(x) + t \), with

\[
f(x) = x^n + a_{n-1} x^{n-1} + \cdots + a_1 x
\]

and set

\[
B_f(t) := \text{Ber} f(x) + t = \frac{\xi(f(x) + t)}{\text{disc}(f(x) + t)} = \frac{\xi(t)}{D_f(t)} = \frac{\xi(t)}{\delta_f^2(t)}
\]

which is a rational function of height\(^2\) at most \( n - 1 \) in \( t \). Therefore, we have

\[
|C(\alpha_1, \ldots, \alpha_r; n)| \leq \sum_{a \in \mathbb{F}_q} \left| \sum_{t \in \mathbb{F}_q} \chi_2(B_{f + a \alpha_1}(t) + \cdots + B_{f + a \alpha_r}(t)) \right|.
\]

In order to bound the character sum, we apply Weil’s theorem to the appropriate Artin–Schreier curve. See [5, theorem 1] for the general claim and proof; we state it here for characteristic 2.

\^2The height \( h(t) \) of a rational function is the maximum of the degrees of its numerator and denominator. It is equal to the total order of its poles (resp. zeros), including poles (resp. zeros) at infinity.
Theorem 3.1. Let $\mathbb{F}_q$ be a field of characteristic 2, and let $p \in \mathbb{F}_q(t)$ be a rational function which is not of the form $H^2(t) + H(t) + c$ for any $H(t) \in \mathbb{F}_q(t), c \in \mathbb{F}_q$. Starting with the projective curve $y^2 + y = p(t)$, using translations of the variable $y$ by appropriate rational functions in $\mathbb{F}_q(t)$, we may obtain an isomorphic curve $y^2 + y = \tilde{p}(t)$, satisfying:

1. $\tilde{p}(t) = p(t) + Q^2(t) + Q(t)$ for some rational function $Q(t) \in \mathbb{F}_q(t)$ with $\text{ht}(Q) \leq \frac{1}{2} \text{ht}(p)$;
2. the poles $P_0, \ldots, P_s$ of $\tilde{p}$ are all poles of $p$;
3. the order $d_i$ of the pole $P_i$ in $\tilde{p}$ is less than or equal to its order in $p$;
4. the orders $d_i$ are all odd.

The following bound then holds:

$$\left| \sum_{t \in \mathbb{F}_q} \chi_2(\tilde{p}(t)) \right| \leq 2gq^{1/2} + 1,$$  \hspace{1cm} (3.5)

where $g$ is the genus of the (isomorphic) curves, given by

$$g = \frac{\sum_{i=0}^{s} (d_i + 1) - 2}{2}.$$  \hspace{1cm} (3.6)

Note that the condition $p \neq H^2(t) + H(t) + c$ was necessary (and sufficient) in order to ensure that $\tilde{p}$ is not a constant function.

In our case, we want a bound for the character sum of $p$, not of $\tilde{p}$. Note that $\chi_2(Q^2(t) + Q(t)) = 1$ whenever $t$ is not a pole of $Q$, and equals 0 at poles. Therefore, $\chi_2(\tilde{p}(t))$ and $\chi_2(p(t))$ may differ only at the poles of $Q$, and, if they do, they differ by at most 1. Thus,

$$\left| \sum_{t \in \mathbb{F}_q} \chi_2(p(t)) \right| \leq \left| \sum_{t \in \mathbb{F}_q} \chi_2(\tilde{p}(t)) \right| + \text{ht}(Q(t)).$$  \hspace{1cm} (3.7)

We will use the following easy corollary of (3.6) to estimate the genus:

Corollary 3.2. Let $\mathbb{F}_q$ be a field of characteristic 2, and let $p \in \mathbb{F}_q(t)$ be a rational function not of the form $H^2(t) + H(t) + c$. Suppose that the order of $p$ in all of its poles is even, except in at most one pole. Then $g \leq (\text{ht}(p) - 1)/2$, where $g$ is the genus of the curve $y^2 + y = p(t)$.

For us, the relevant rational function is $p(t) = B_{f+\alpha_1}(t) + \cdots + B_{f+\alpha_r}(t)$, which has height at most $r(n - 1)$. The denominator of $p$ is $(\delta_{f+\alpha_1} \cdots \delta_{f+\alpha_r})^2$, which indeed shows that all its poles have even order, with the sole possible exception of a pole at infinity. Hence $p$ satisfies the conditions of corollary 3.2, provided $p(t) \neq H^2(t) + H(t) + c$. Combining (3.5), (3.7) and corollary 3.2, we obtain

$$\left| \sum_{t \in \mathbb{F}_q} \chi_2(p(t)) \right| \leq 2gq^{1/2} + 1 + \text{ht}(Q) < \text{ht}(p)q^{1/2} + \frac{\text{ht}(p)}{2} + 1 < r(nq)^{1/2},$$  \hspace{1cm} (3.8)

which, when applied to (3.4), provides the major term in (1.2).

We need now only find a way to bound the size of the set $G_n^c$ of ‘bad’ $a$’s where $p(t) = H^2(t) + H(t) + c$. As in odd characteristic, we cover $G_n^c$ by simpler, algebraic varieties.

(a) Covering $G_n^c$

For simplicity, and without loss of generality, assume $\alpha_1 = 0$.

Proposition 3.3. We can write $G_n^c \subset A_n \cup B_n \cup C_n$ where:

- $A_n$ is the set of those $a \in \mathbb{F}_q^{n-1}$ for which $\deg \delta_j = d(n)$ and $\delta_j$ is not coprime to $\xi_j \delta_j^2 - \xi_j^2$, that is,
  $$A_n = \{a \in \mathbb{F}_q^{n-1} : \deg \delta_j = d(n), \text{Res}(\delta_j, \xi_j \delta_j^2 - \xi_j^2) = 0\};$$  \hspace{1cm} (3.9)
— \( B_n = \bigcup_{j \neq 1} B(j) \) where \( B(j) \) is the set of those \( a \in \mathbb{F}_q^{n-1} \) for which \( \deg \delta_f = \deg \delta_{f+ \alpha_j} = d(n) \) and \( \delta_{f+ \alpha_j}(t) \) have a common zero, that is,

\[
B(j) = \{ a \in \mathbb{F}_q^{n-1} : \deg \delta_f = \deg \delta_{f+ \alpha_j} = d(n), \text{Res}(\delta_f(t), \delta_{f+ \alpha_j}(t)) = 0 \};
\]

— \( C_n = \bigcup_j C(j) \), where

\[
C(j) = \{ a \in \mathbb{F}_q^{n-1} : \deg \delta_{f+ \alpha_j} < d(n) \}.
\]

Henceforth, let us denote \( \mathcal{E}_f = \xi_f \delta_f^2 - \xi_f^2 \).

Proof. We will assume \( a \in \mathbb{C}_n \setminus (B_n \cup C_n) \), and show that \( a \in A_n \). By \( a \not\in B(j) \cup C(j) \cup C(1) \), \( \delta_f \) is coprime to \( \delta_{f+ \alpha_j} \) for all \( j \neq 1 \). From \( a \in \mathbb{C}_n \), we obtain

\[
p(t) = \frac{\xi_{f+ \alpha_j}(t)}{D_{f+ \alpha_j}(t)} + \cdots + \frac{\xi_{f+ \alpha_j}(t)}{D_{f+ \alpha_j}(t)} = H_f^2(t) + H(t) + c.
\]

Note that, from \( a \not\in C(1) \) and lemma 2.1, we have \( \deg D_f = 2d(n) \geq \deg \xi_f \). Consider a root of \( \delta_f \) with multiplicity \( m \). Then, it is a pole of \( p \) with multiplicity at most \( 2m \). Hence, it is a pole of \( H \) with multiplicity at most \( m \), its multiplicity in \( \delta_f \). As this is true for every root of \( \delta_f \), it follows that we may write \( H = H_n/\delta_f H_d \), where \( H_n, H_d \) are polynomials and \( H_d \) is coprime to \( \delta_f \). Thus, there exists a unique polynomial \( H_f \) with \( \deg H_f < \deg \delta_f \) such that \( H_f \equiv (H_n/H_d) \pmod{\delta_f} \). We may then write \( H_n = H_f H_d + H_d \delta_f \) for some polynomial \( H_r \), or, equivalently, \( H = H_f/\delta_f + H_r/H_d \). Substituting this relation in (3.12), we obtain from \( \deg H_f < \deg \delta_f \) and \( \deg \xi_f \leq \deg D_f \), as well as \( \delta_f \) being coprime to \( D_{f+ \alpha_j} \) and \( H_d \), that

\[
\frac{\xi_f(t)}{\delta_f(t)} = \frac{(H_f/\delta_f)^2 + H_f/\delta_f + c_2}{\delta_f(t)}. \tag{3.13}
\]

Multiplying by \( \delta_f^2 \), we obtain

\[
\xi_f = H_f^2 + \delta_f H_f + c_2 \delta_f^2. \tag{3.14}
\]

Differentiating the last formula, we get

\[
\xi'_f = \delta_f H_f + \delta_f H'_f. \tag{3.15}
\]

Reducing equations (3.14) and (3.15) modulo \( \delta_f \), we find

\[
\xi_f \equiv H_f^2 \pmod{\delta_f} \tag{3.16}
\]

and

\[
\xi'_f \equiv \delta_f H_f \pmod{\delta_f}. \tag{3.17}
\]

From which we easily derive

\[
\xi_f^2 \equiv \delta_f^2 H_f^2 \equiv \delta_f^2 \xi_f \pmod{\delta_f}. \tag{3.18}
\]

Congruence (3.18) states that \( \delta_f \) must divide \( \mathcal{E}_f = \xi_f \delta_f^2 - \xi_f^2 \), \( a \not\in C(1) \) implies in particular that \( \delta_f \) is not constant, and, therefore, \( \delta_f \) and \( \mathcal{E}_f \) are not coprime—thus \( a \in A_n \) by definition.

(b) Bounding degrees and sizes

In order to complete the proof we need to provide bounds for the degrees of the polynomials defining \( A_n \) and \( B_n \), and show that these polynomials are not identically zero. We shall first compute the bounds assuming the polynomials do not vanish, which we will prove in the next section. We must also bound the size of \( C_n \). We begin with the following lemma.

**Lemma 3.4.** Let \( A = A^\prime \cup \{ t \} \) be a set of variables. Let \( f, g \in \mathbb{F}_q[A] \) be homogeneous polynomials in the variables \( A \) of degrees \( d_f, d_g \), respectively. Let \( n_f, n_g \) be their respective degrees as polynomials in the variable \( t \) with coefficients in \( \mathbb{F}_q[A^\prime] \). Set \( R = \text{Res}_t(f, g) \in \mathbb{F}_q[A^\prime] \). Then \( R \) is a homogeneous polynomial in the variables \( A^\prime \) of degree \( d_f n_g + d_g n_f - d_f d_g \).
Proof. Write \( f = \sum_{k=0}^{n_f} a_k t^k, g = \sum_{k=0}^{n_g} b_k t^k \), where \( a_k, b_k \) are homogeneous polynomials in the variables \( A' \) of degrees \( d_f - k, d_g - k \), respectively. Let \( \prod a_k^{r_k} \prod b_k^{s_k} \) be any arbitrary monomial appearing in \( R \). By well-known properties of the resultant, we have
\[
\sum r_k = n_g, \quad \sum s_k = n_f \quad \text{and} \quad \sum k(r_k + s_k) = n_f n_g.
\]
It follows that the total degree of the monomial in the variables \( A' \) is
\[
\sum (d_f - k)r_k + \sum (d_g - k)s_k = d_f \sum r_k + d_g \sum s_k - \sum k(r_k + s_k) = d_f n_g + n_f d_g - n_f n_g
\]
as claimed. \(\blacksquare\)

In order to obtain bounds on the sizes of \( A_n, B_n \) from the degrees of their defining polynomials, we will use the following elementary lemma [6, §4, lemma 3.1]:

**Lemma 3.5.** Let \( h(X_1, \ldots, X_m) \in \mathbb{F}_q[X_1, \ldots, X_m] \) be a non-zero polynomial of total degree at most \( d \). Then the number of zeros of \( h(X_1, \ldots, X_m) \) in \( \mathbb{F}_q^m \) is at most
\[
\#\{ x \in \mathbb{F}_q^m : h(x) = 0 \} \leq d q^{m-1}.
\]

(c) **Bounding \( C_n \)**

We note that, for odd \( n > 2 \), \( \delta_f + \alpha_j(t) \) is always of degree exactly \( (n-1)/2 \). For even \( n > 2 \), \( \deg \delta_f + \alpha_j(t) = (n-2)/2 \) iff the coefficient of \( x^{n-1} \) in \( f + \alpha_j \) is non-zero. This is true simultaneously for every \( j \), for all but at most \( r q^{n-2} \) tuples \( a \), where \( a_{n-1} \in \{ \alpha_j, n-1 \} \). Hence \( \# C_n \leq r q^{n-2} \). This contribution will be merged into the other bounds.

(d) **Bounding \( B_n \)**

\( \delta_f(t) \) and \( \delta_f + \alpha_j(t) \) have total degree \( n-1 \) in \( a_{n-1}, \ldots, a_1, t \), and, by definition of \( B(j) \), they have degree \( d(n) \) as polynomials in \( t \). Hence by lemma 3.4, \( \text{Res}(\delta_f(t), \delta_f + \alpha_j(t)) \) has total degree \( 2(n-1)d(n) - d(n)^2 \) in the coefficients \( a_{n-1}, \ldots, a_1 \), which equals \( \frac{3}{4}(n-1)^2 \) for odd \( n \) and \((3n-2)(n-2))/4 for even \( n \); in either case, we may round this up to \( \frac{3}{4}n^2 \) and obtain\(^3\)
\[
\# B_n < \frac{3}{4} (r-1) n^2 q^{n-2}.
\]

(e) **Bounding \( A_n \)**

We have seen that the total degree of \( \xi_f \) in \( a_0, a_{n-1}, \ldots, a_1, t \) is \( 2(n-1) \), and that the degree of \( \xi_f \) in \( t \) is at most \( 2d(n) \). \( \delta_f \) has total degree \( n-1 \) and degree \( d(n) \) in \( t \). Thus, we find that the degree of \( \Sigma_f = \xi_f \delta_f^2 - \xi_f^2 \) in \( t \) is at most \( 4d(n) - 2 \) (i.e. \( 2n - 4 \) for odd \( n \) and \( 2n - 6 \) for even \( n \)), and its total degree in \( a_0, a_{n-1}, \ldots, a_1, t \) is exactly \( 4n - 6 \). As the degree of \( \delta_f \) is constant, we may assume that the polynomial given by \( \text{Res}(\delta_f, \Sigma_f) \) is fixed by always assuming \( \Sigma_f \) is of degree exactly \( 4d(n) - 2 \). This is valid, as adding leading zeros to only one of the polynomials multiplies the resultant by a non-zero factor. We now have s
\[
\deg_x \delta_f = d(n), \quad \text{tot.} \deg \delta_f = n-1
\]
and \( \deg_x \Sigma_f = 4d(n) - 2 \), \( \text{tot.} \deg \Sigma_f = 4n - 6 \).

\(^3\)Note that, as \( n \geq 3 \), the process of rounding up to \( \frac{3}{4}n^2 \) adds at least \( \frac{2}{4}n^2 \). This, together with the rounding of the bound on \( \# A_n \), covers the two instances where we neglected an error of \( r q^{n-1} \): the first in the assumption that \( \delta(F + a) \neq 0 \), the second in bounding \( C_n \).
Hence by lemma 3.4, the degree of $\text{Res}(\delta_f, \Sigma_f)$ in $a_{n-1}, \ldots, a_1$ is $(n - 1)(4d(n) - 2) + d(n)(4n - 6) - d(n)(4d(n) - 2)$, which is equal to $(n - 1)(3n - 5)$ for odd $n$ and $3n^2 - 10n + 6$ for even $n$. In either case, we may round this up to $3n^2$ and obtain

$$\#A_n < 3n^2q^{n-2};$$

combining this with (3.20) we get

$$\#G_n^c < \frac{3}{4}(r + 3)n^2q^{n-2},$$

proving theorem (1.1).

4. Non-vanishing of the resultants

(a) Non-vanishing of the polynomials defining $B_n$

Proposition 4.1. Given a non-zero polynomial $\alpha \in \mathbb{F}_q[x]$ with $\deg \alpha < n$, the function $a \mapsto \text{Res}(\delta_f(t), \delta_{f+a}(t))$ is not the zero polynomial, that is, the polynomial function

$$R(a) := \text{Res}(\delta_f(t), \delta_{f+a}(t)) \in \mathbb{F}_2[\bar{a}]$$

(4.1)

is not identically zero.

We note that the proof of the analogous proposition in odd characteristic [1, proposition 3.1] did not in fact rely on the characteristic being odd. We may follow the same arguments to see again that $R(a)$ cannot be identically zero. More accurately, the proof in [1] referred to the polynomial $\text{Res}(D_f(t), D_{f+a}(t))$, which, in characteristic 2, is equivalent to $\text{Res}(\delta_f^2(t), \delta_{f+a}^2(t)) = R(a)^q$. The main observation behind the proof was that the roots of $D_f(t)$, which are the same as the roots of $\delta_f(t)$, are exactly those $t$ for which there exists some $\rho$ (in some fixed algebraic closure of $\mathbb{F}_q$) that satisfies $f'(\rho) = 0$ and $t = -f(\rho)$. This observation is just as valid in characteristic 2 as are the calculations that followed. This completes the proof of inequality (3.20).

(b) Non-vanishing of the polynomials defining $A_n$

We wish to show that the algebraic condition for being in $A_n$, i.e. $\text{Res}(\delta_f, \Sigma_f) = 0$, is not always satisfied. We will demonstrate this by giving explicit examples of $f$ that do not satisfy the equation.\footnote{A different approach was used in the case of odd characteristic. The approach used here could have been applied there partially. For example, the polynomial $f + t = x^n + ax^2 + t$ yields $D_f(t) = (-1)^{(n-3)/2}a^{n-1} + (1 - n)t^{-1}a^n$, which satisfies $\text{disc}(D_f(t)) \neq 0$ given $\gcd(q, n(n - 1)) = 1, a \neq 0$. While this covers many cases, the remaining cases are not as easily dealt with. The algebraic approach managed to avoid this division into cases completely.} We will construct two generic examples, depending on the parity of $n$. The relevant computations are given in more detail in §7.

Consider first $n \geq 3$ odd, and take $f + t = x^n + ax^2 + t$. An easy computation then yields

$$\delta_f(t) = t^{n-1},$$

$$\xi_f(t) = \frac{(-1)^{(n-1)/2}n^n - 1}{4}t^{n-1} + a^n t$$

and

$$\xi'_f(t) = a^n.$$
For the case of $n \geq 4$ even, we will consider the polynomial $f + t = x^n + ax^{n-1} + bx + t$, with $a, b \neq 0$. Let us write $n = 2m$. An easy computation yields $\delta_f(t) = a^{m+1}m^{-1} + b^m$ and $\delta'_f(t) = (m - 1) a^m t^{m-2}$, and a longer computation yields

$$
\xi_f = \begin{cases} 
0 & m \equiv 0, 1 \pmod{4} \\
\delta^2_f & m \equiv 2, 3 \pmod{4} \\
m \equiv 0, 2 \pmod{4} \\
\delta^2_f & m \equiv 1, 3 \pmod{4}
\end{cases}
$$

and hence

$$
\Xi_f = \xi_f \delta^2_f - \xi^2_f = \begin{cases} 
a^{2m+1}b^{m+1}3m^{-6} & m \equiv 0 \pmod{4} \\
na^{2m+1}b^{m+1}3m^{-6} + a^m t^{m-4} \delta^2_f & m \equiv 2 \pmod{4} \\
a^{n-2}b^{n-2}t^{n-2} & m \equiv 1 \pmod{2}.
\end{cases}
$$

As $b \neq 0$, clearly $t = 0$ is not a root of $\delta_f$, but, in all cases above, it is either the sole root of $\Xi_f$ or of a combination of $\xi_f$ and $\delta_f$. Either way, it is clear that $\delta_f$ and $\Xi_f$ can have no common roots, and $\text{Res}(\delta_f, \Xi_f) \neq 0$. Thus, we have shown inequality (3.21).

5. The case $n = 2$

For $n = 2$, the inequality (1.2) is not always valid—sometimes there are correlations in the Möbius function. The following proposition covers all cases where $n = 2$.

**Proposition 5.1.** Let $\alpha_1, \ldots, \alpha_r \in \mathbb{F}_q[x]$ be distinct linear polynomials $\alpha_i = a_ix + b_i$, and let $\epsilon_1, \ldots, \epsilon_r \in \{1, 2\}$. Set

$$
A = \{a_i : 1 \leq i \leq r\},
$$

$$
b_a = \sum_{i : a_i = a} \epsilon_i b_i, \quad A_b = \{a \in A : b_a \neq 0\} \quad \text{and} \quad \gamma_a = \sum_{i : a_i = a} \epsilon_i \text{ mod } 2, \quad A_\gamma = \{a \in A : \gamma_a \neq 0\}.
$$

One of the following relations then holds:

$$
|C(\alpha_1, \ldots, \alpha_r; 2)| \leq rq \quad \text{if } A_\gamma \neq \emptyset
$$

$$
|C(\alpha_1, \ldots, \alpha_r; 2)| \leq rq^{3/2} \quad \text{if } A_\gamma = \emptyset, \ A_b \neq \emptyset
$$

and

$$
|C(\alpha_1, \ldots, \alpha_r; 2)| \geq q^2 - rq \quad \text{if } A_\gamma = \emptyset, \ A_b = \emptyset.
$$

**Proof.** One may easily see that, for a quadratic polynomial,

$$
\text{Berl}(x^2 + ax + b) = \frac{b}{a^2}.
$$

In particular, $\mu(x^2 + ax + b) = 0 \iff a = 0$, and otherwise

$$
\mu(x^2 + ax + b) = \chi_2 \left( \frac{b}{a^2} \right).
$$
Clearly for \( f = x^2 + sx + t, \prod_i \mu(f + a_i)^{\varepsilon_i} = 0 \iff s \in A, \) so we may take our sum only over \( s \notin A. \) There is no further contribution to the product from \( a_i \) where \( \varepsilon_i = 2. \) We compute:

\[
C(\alpha_1, \ldots, \alpha_r; 2) = \sum_{f \in M_2} \prod_i \mu(f + \alpha_i)^{\varepsilon_i} = \sum_{s \notin A} \sum_{t \in F_q} \chi_2 \left( \sum_i \varepsilon_i \frac{b_{i} + t}{s + a_{i}^{2}} \right) \\
= \sum_{s \notin A} \sum_{t \in F_q} \chi_2 \left( \sum_{a \in A} \frac{b_{a} + s}{s + a^{2}} \right) \\
= \sum_{s \notin A} \sum_{t \in F_q} \chi_2 \left( \sum_{a \in A} \frac{b_{a}}{s + a^{2}} \right) \chi_2 \left( \sum_{a \in A} \frac{\varepsilon_{a}}{s + a^{2}} \right) \\
= \sum_{s \notin A} \chi_2 \left( \sum_{a \in A} \frac{b_{a}}{s + a^{2}} \right) \left( \sum_{a \in A} \frac{1}{s + a^{2}} \right) \left( \sum_{t \in F_q} \left( \sum_{a \in A} \frac{1}{s + a^{2}} \right) t \right). \tag{5.4}
\]

Note that, for any constant \( c, \)

\[
\sum_{t \in F_q} \chi_2(ct) = \begin{cases} 0 & c \neq 0 \\ q & c = 0. \end{cases}
\]

We now have two cases. If \( A_{\gamma} \neq \emptyset, \) then \( \sum_{a \in A_{\gamma}} (1/(s^2 + a^2)) = 0 \) for at most \( \#A_{\gamma} - 1 < r \) values of \( s. \) Hence, in this case, we have

\[
|C(\alpha_1, \ldots, \alpha_r; 2)| \leq (\#A_{\gamma} - 1)q < rq, \tag{5.5}
\]

which is the first case of proposition 5.1. On the other hand, if \( A_{\gamma} \) is empty, then (5.4) becomes

\[
C(\alpha_1, \ldots, \alpha_r; 2) = q \sum_{s \notin A} \chi_2 \left( \sum_{a \in A} \frac{b_{a}}{s + a^{2}} \right). \tag{5.6}
\]

Once again, we have two cases. If \( A_{b} = \emptyset, \) then clearly

\[
C(\alpha_1, \ldots, \alpha_r; 2) = q(q - \#A) \geq q^2 - rq, \tag{5.7}
\]

i.e. there is full correlation—every term in the sum is either 0 or 1. This is the third case of proposition 5.1. Finally, we are left with the case \( A_{b} \neq \emptyset. \) By the change of variables \( y = y + \sum_{a \in A_{b}} (\sqrt{\frac{b_{a}}{s}}) \) we see that the curve \( y^2 + y = \sum_{a \in A_{b}} (b_{a}/(s^2 + a^2)) \) is equivalent to the curve \( y^2 + y = \sum_{a \in A_{b}} (\sqrt{\frac{b_{a}}{s}}) \). The rational function \( \sum_{a \in A_{b}} (\sqrt{\frac{b_{a}}{s}}) \) has exactly \( \#A_{b} \) distinct simple poles; hence, by theorem 3.1, the genus of these curves is exactly \( \#A_{b} - 1. \) Note that \( A_{b} \leq r/2; \) indeed, \( A_{\gamma} = \emptyset \) implies that each \( a \in A_{b} \) is represented at least twice in the sequence \( \{a_i\}. \) Applying theorem 3.1 to equation (5.6) then yields

\[
|C(\alpha_1, \ldots, \alpha_r; 2)| \leq (2(\#A_{b} - 1)q^{1/2} + 1)q < rq^{3/2}. \tag{5.8}
\]

Completing the proof of proposition 5.1.

\[\begin{align*}
\end{align*}\]

### 6. Square independence

In [2, proposition 3.1], Bary-Soroker computes the following Galois group:

**Proposition 6.1 (Bary-Soroker).** Let \( q \) be an odd prime power, let \( n, r \) be positive integers, let \( \alpha = (\alpha_1, \ldots, \alpha_r) \in \mathbb{F}_q[x]^r \) be an \( r \)-tuple of distinct polynomials each of degree \( < n, \) let \( U = (U_0, \ldots, U_{n-1}) \) be an \( n \)-tuple of variables over \( \mathbb{F}_q, \) and let \( F = x^n + U_0 x^{n-1} + \cdots + U_0 \in \mathbb{F}_q[U, x]. \) For each \( i = 1, \ldots, r, \) let \( F_i = F + \alpha_i. \) Let \( \mathbb{F}_q \) be an algebraic closure of \( \mathbb{F}_q, \) let \( E = \mathbb{F}_q(U), \) let \( F_i \) be the splitting field of \( F_i \) over \( E, \) and let \( F \) be the splitting field of \( \prod_{i=1}^r F_i \) over \( E. \) Then \( \text{Gal}(F/E) \cong S_n^r. \)

We wish to extend this computation also to even \( q. \) The reliance on odd \( q \) lies in the following lemma [2, lemma 3.3]:

\[
\begin{align*}
\end{align*}\]
Lemma 6.2. For a separable polynomial $f \in E[x]$, denote by $\delta(f)$ the square class of its discriminant $\text{disc}_x(f)$ in the $\mathbb{F}_2$-vector space $E^2/(E^2)^2$. The square classes $\delta_x(F_1), \ldots, \delta_x(F_r)$ are linearly independent.

The lemma is proved using some of the arguments from [1]. It then sets the ground for the application of a final lemma [7, lemma 3.4]:

Lemma 6.3. If the square classes $\delta_x(F_1), \ldots, \delta_x(F_r)$ are linearly independent, then $F_1, \ldots, F_r$ are linearly disjoint over $E$.

Lemma 6.3, together with $F$ being the compositum of $F_1, \ldots, F_r$ and with the classical fact that $\text{Gal}(F_1/E) \cong S_r$, easily yields proposition 6.1; see [2, §3] for the full details.

We note that lemma 6.3 was proved in [7] also in characteristic 2. In this case, $\delta(f)$ needs to be defined in terms of Berlekamp’s discriminant, as the residue class $[\text{Berl}(f)]$ in $E/\wp(E)$, where $\wp(y) = y^2 + y$. We shall prove lemma 6.2 in this context, for $n > 2$, in analogy to the proof in [2]. This is the last required piece in the proof of proposition 6.1 for characteristic 2.

Proof. Consider an arbitrary extension $\mathbb{F}_{q^r}$, write $t = U_0$ and consider specializations $(U_t, \ldots, U_{t-1}) \mapsto (a, a_1, \ldots, a_{n-1}) \in \mathbb{F}_{q^{n-1}}$. The specialization of the polynomial $F$ is $f = x^n + a_{n-1}x^{n-1} + \cdots + a_1x + t \in \mathbb{F}_{q^n}[x, t]$. $E/\wp(E)$ is specialized into $\mathbb{F}_{q^n}(t)/\wp(\mathbb{F}_{q^n}(t))$, and we shall work in this quotient. Examine the proof of proposition 3.3. In its course we have in fact shown that if $[\text{Berl}(f + a_1)] \neq 0$ (i.e. $a \notin A_n \cup C(1)$), and $\text{disc}_c(f + a_1)$ is coprime to $\text{disc}_c(f + a_i)$ for each $i \neq 1$ (i.e. $a \notin B_n \cup C_n$), then $[\text{Berl}(f + a_1)] \cdots [\text{Berl}(f + a_r)] \neq 0$ (i.e. $a \in G_n$). We have also demonstrated that this must be the case for all but at most $\frac{3}{4}(r + 3)\tau_a q^{(n-2)}$ specializations. The same arguments imply that if we further require that $[\text{Berl}(f + a_1)] \neq 0$ for each $i = 1, \ldots, r$, and that $\text{disc}_c(f + a_i)$ is coprime to $\text{disc}_c(f + a_j)$ for all $j \neq i$, then for any non-empty set of indices $S \subset \{1, \ldots, r\}$, we have $\sum_{i \in S}[\text{Berl}(f + a_i)] \neq 0$. It is an equivalent statement to $[\text{Berl}(f + a_1)], \ldots, [\text{Berl}(f + a_r)]$ being linearly independent in $\mathbb{F}_{q^n}(t)/\wp(\mathbb{F}_{q^n}(t))$. Furthermore, this must be the case for all but at most $\frac{3}{4}r(r + 3)\tau_a q^{(n-2)}$ of the $q^{(n-1)}$ specializations $a \in \mathbb{F}_{q^{n-1}}$. In particular, for large enough exponent $e$, it is clear that there must exist at least one such specialization. But since a linear dependence in $E/\wp(E)$ would survive specialization, it follows that the set of $\delta_x(F_1) = [\text{Berl}(F + a_i)]$ must indeed be linearly independent in $E/\wp(E)$.

Remark. Lemma 6.2 is not true in general for $n = 2$, where linear dependence of the classes is possible, as demonstrated in the third case of proposition 5.1. The result is still valid under more specific conditions that rule out the possibility of dependence occurring in any subset of $a_1, \ldots, a_r$. Specifically, we require that there is no non-empty subset $S \subset \{1, \ldots, r\}$ such that $\{a_i : i \in S\}$ all share the same linear term and $\sum_{i \in S}(x^2 + a_i) = 0$.

7. Computations

In this section, we calculate $\delta(f), \xi(f)$ for the polynomials $f = x^n + ax^2 + t$ with $n$ odd, and $f = x^n + ax^{n-1} + bx + t$ with $n$ even. To do so, we will compute $\text{disc}(f)$ and $\text{disc}_+(f)$. Writing $f(x) = g(x^2) + xh(x^2)$, we see from (2.5) that (working modulo 2)

$$\delta^2(f) = \text{disc}(f) = \text{disc}_+(f) = \text{Res}(h(x), g(x))^2. \quad (7.1)$$

Hence $\delta(f) = \text{Res}(h(x), g(x))$. The formula for resultant of binomials is well known (see [3, lemma 3]) and we obtain

$$\delta(x^n + ax^2 + t) = \text{Res}(x^{(n-1)/2}, ax + t) = t^{(n-1)/2}$$

and

$$\delta(x^n + ax^{n-1} + bx + t) = \text{Res}(ax^{n/2-1} + bx^{n/2} + t) = a^{n/2}b^{n/2-1} + b^{n/2}. $$
We now need to compute \( \text{disc}(f) \):
\[
\text{disc}(x^n + ax \gamma + b) = \text{Res}(x^{(n-1)/2}, ax \gamma + b) = t^{n-1}
\]
and
\[
\text{disc}(x^n + ax^{n-1} + bx + t) = \text{Res}(ax^{(n-1)/2} + b, x^{n/2} + t)^2
= a^n t^{n-2} + b^n + 2a^{n/2}b^{n/2} + 1. \quad (7.2)
\]

We now need to compute \( \text{disc}(f) \) in characteristic 0.

For odd \( n \), we apply the general formula for trinomial discriminants [8, theorem 2] and obtain
\[
\text{disc}(x^n + ax^2 + t) = (-1)^{(n-1)/2}(n^{n-1} + 4(n-2)^n - a^n t).
\]

It is now immediate to compute
\[
\xi(x^n + ax^2 + t) = \frac{\text{disc} - \text{disc}}{4} \mod 2 = \frac{(-1)^{(n-1)/2}n^{n-1} - 1}{4} t^{n-1} + a^n t.
\]

The computation for even \( n \) is somewhat longer and more complicated:
\[
\text{disc}(x^n + ax^{n-1} + bx + t)
= (-1)^{(n-1)/2} \text{Res}(x^n + ax^{n-1} + bx + t, nx^{n-1} + (n-1)ax^{n-2} + b)
= (-1)^{(n-1)/2} \text{Res}(x^n + ax^{n-1} + bx + t, nx^n + (n-1)ax^{n-1} + bx)
= (-1)^{(n-1)/2} \text{Res}(x^n + ax^{n-1} + bx + t, ax^{n-1} + (n-1)bx + nt)
= (-1)^{(n-1)/2} \text{Res}(x^n + (2-n)bx + (1-n)t, ax^{n-1} + (n-1)bx + nt)
= (-1)^{(n-1)/2} \frac{1}{t} a^n \text{Res}
\left(-x^n + (n-2)bx + (n-1)t, x^{n-1} + \frac{(n-1)b}{a} + \frac{nt}{a}\right)
\quad \ldots \quad (7.2)
\]
\[
= (-1)^{(n-1)/2} \frac{1}{(n-1)t^2} a^n \text{Res}
\left(-x^n + (n-2)bx + (n-1)t, x^{n-1} + \frac{(n-1)b}{a} x + \frac{nt}{a}\right)
\quad \ldots \quad (7.2)
\]
\[
= (-1)^{(n-1)/2} \frac{1}{(n-1)t^2} a^n \text{Res}
\left(-x^n + (n-2)bx + (n-1)t, \frac{(n-1)b}{a} x^2 + \frac{(n-2)b}{a} x + (n-1)t\right)
\quad \ldots \quad (7.2)
\]
\[
= (-1)^{(n-1)/2} \frac{1}{(n-1)t^2} \text{Res}((n-1)bx^2 + ((n-2)ab + nt)x + (n-1)at, x^n - (n-2)bx + (n-1)t)
\quad \ldots \quad (7.2)
\]
\[
= (-1)^{(n-1)/2} \frac{1}{(n-1)t^2} \text{Res}(ax^2 + bx + \gamma, x^n - ux - v)
\quad \ldots \quad (7.2)
\]

where \( a = (n-1)b, \beta = (n-2)ab + nt, \gamma = (n-1)at, u = (n-2)b \) and \( v = (n-1)t \) are defined by the last equivalence. Note that \( a, \gamma, v \) are odd and \( \beta, u \) are even. The number of terms in the last resultant is unbounded as \( n \) increases. However, we are only interested in \( \text{disc}(f) \) mod 8, and, as \( \beta \) is even, only finitely many of these terms will be non-zero modulo 8. Therefore, we will continue our computation modulo 8.

Let \( r_{1,2} = (-\beta \pm \sqrt{\beta^2 - 4\alpha \gamma})/2\alpha \) be the roots of \( ax^2 + bx + \gamma \). Let \( R_k = \alpha^k (r_1^k + r_2^k) \). It is easy to see that
\[
R_0 = 2, \quad R_1 = -\beta, \quad \forall k \geq 2, R_k = -\beta R_{k-1} - \alpha \gamma R_{k-2}.
\]
Thus, for all $k$, $R_k$ is a polynomial in $\beta, \alpha \gamma$. Using the recursion formula, one may show by induction that, for $k \geq 1$,

$$R_k = \sum_{0 \leq l \leq k/2} (-1)^{k-l} \left( \binom{k-l}{l} + \binom{k-l-1}{l-1} \right) (\alpha \gamma)^l \beta^{k-2l}. \quad (7.3)$$

Since we are only interested in $R_k \mod 8$, we may discard all monomials containing powers of $\beta$ greater than 3, and obtain

$$R_k \mod \beta^3 = \begin{cases} 
(-1)^m(2\alpha \gamma)^m - m^2(\alpha \gamma)^{m-1} \beta^2 & k = 2m \text{ even} \\
(-1)^m k(\alpha \gamma)^m & k = 2m - 1 \text{ odd}. 
\end{cases} \quad (7.4)$$

We now compute:

$$\text{Res}(ax^2 + \beta x + \gamma, x^n - ux - v) = \alpha^n(r_1^n - ur_1 - v)(r_2^n - ur_2 - v)$$

$$= \alpha^n((r_1 r_2)^n + u^2 r_1 r_2 + v^2 - ur_1 r_2(r_1^{n-1} + r_2^{n-1}) - v(r_1^n + r_2^n) + uv(r_1 + r_2))$$

$$= \gamma^n + u^2 \alpha^{n-1} \gamma + v^2 \alpha^n - u\gamma R_{n-1} - vR_n - u\alpha \gamma \beta.$$

We can now replace $\alpha, \beta, \gamma, u, v, R_{n-1}, R_n$ by their actual values. We first expand and simplify the part of the expression not involving $R_{n-1}, R_n$:

$$\gamma^n + u^2 \alpha^{n-1} \gamma + v^2 \alpha^n - u\alpha \gamma \beta = \gamma^n + \alpha^{n-1}(u^2 \gamma + v^2 \alpha - uv\beta)$$

$$= (n-1)^n(\alpha^n\beta^n + b^{n-1}((n-2)^2 \alpha^2 t + (n-1)^2 \beta t^2 - (n-2)bt((n-2)ab + nt)))$$

$$= (n-1)^n(\alpha^n\beta^n + b^n t^2).$$

Next, set $n = 2m$, and expand

$$(-1)^m(\alpha \gamma R_{n-1} + vR_n) = (n-1)\mu \alpha^{m-1} \gamma^m \beta + 2\nu \alpha^m \gamma^m - m^2 \nu \alpha^{m-1} \gamma^{m-1} \beta^2$$

$$\equiv (\alpha \gamma)^{m-1}(u\gamma \beta + 2\nu \alpha \gamma - m^2 \nu \beta^2)$$

$$= (n-1)^{m-1}(abt)^{m-1}((n-2)abt\beta + 2(n-1)^2 bt^2 - m^2 t^2 \beta^2)$$

$$\equiv 2(n-1)^m \alpha^m \beta^m t^{m+1} + (n-2)^2 \alpha^{m+1} \beta^{m+1} t^m - m^2 \nu \alpha^{m-1} \alpha^{m-1} \gamma \beta^{m-1} (\mod 8),$$

where the cancellations in the last congruence are due to identities such as $(n-1)^2 \equiv 1 (\mod 8)$, $n(n-2) \equiv 0 (\mod 8)$, $m^2(n-2)^2 \equiv 0 (\mod 16)$ and $ax \equiv a (\mod 8)$ when $a$ is divisible by 4 and $x$ is odd. Note that, of the last two terms, exactly one has coefficient $\equiv 0 (\mod 8)$, and the other coefficient $\equiv 4 (\mod 8)$, determined by the parity of $m$.

Combining all expansions, we obtain

$$\text{disc}(x^n + ax^{n-1} + bx + t) = (-1)^{n/2 + 1} \frac{1}{(n-1)^2} \text{Res}(ax^2 + \beta x + \gamma, x^n - ux - v)$$

$$\equiv (-1)^{m+1}(n-1)^{m-1}(a^{n-2} + b^n) + 2m \alpha^{m-1} \beta^{m-1}$$

$$+ 4 \begin{cases} 
\alpha^{m+1} \beta^{m+1} t^{m-2} & m \equiv 0 (\mod 2) \\
\alpha^{m-1} \beta^{m-1} t^m & m \equiv 1 (\mod 2). 
\end{cases} \quad (7.5)$$

Noting also that $a^n t^{n-2} + b^n = \delta_f^2$ and

$$(-1)^{m+1}(n-1)^{n-1} \equiv \begin{cases} 
1 & m \equiv 0, 1 (\mod 4) \\
5 & m \equiv 2, 3 (\mod 4), 
\end{cases} \quad (7.6)$$
we can now combine (7.2) and (7.5) to obtain
\[
\xi(x^n + ax^{n-1} + bx + t) = \frac{\text{disc}_+ - \text{disc}}{4} \mod 2
\]
\[
= \begin{cases} 
0 & m \equiv 0, 1 \pmod{4} \\
\delta^2_f & m \equiv 2, 3 \pmod{4} \\
\delta^2_f & m \equiv 0, 2 \pmod{4} \\
\delta^2_f & m \equiv 1, 3 \pmod{4}
\end{cases}
\]
as claimed.

Acknowledgements. The author thanks Lior Bary-Soroker for introducing him to Berlekamp’s discriminant and for helpful discussions, and Ze’ev Rudnick for guidance and many helpful suggestions.

Funding statement. The research leading to these results has received funding from the European Research Council under the European Union’s Seventh Framework Programme (FP7/2007-2013)/ERC grant agreement no. 320755.

References

1. Carmon D, Rudnick Z. 2014 The autocorrelation of the Möbius function and Chowla’s conjecture for the rational function field. Q. J. Math. 65, 53–61. (doi:10.1093/qmath/has047)
2. Bary-Soroker L. 2014 Hardy–Littlewood tuple conjecture over large finite fields. Int. Math. Res. Not. 2014, 568–575. (doi:10.1093/imrn/rns249)
3. Andrade JC, Bary-Soroker L, Rudnick Z. 2014 Shifted convolution and the Titchmarsh divisor problem over \( \mathbb{F}_q[t] \). (http://arxiv.org/abs/1407.2076)
4. Berlekamp ER. 1976 An analog to the discriminant over fields of characteristic two. J. Algebra 38, 315–317. (doi:10.1016/0021-8693(76)90222-2)
5. Perel’muter GI. 1963 On certain character sums. Uspekhi Mat. Nauk 18, 145–149.
6. Schmidt WM. 2004 Equations over finite fields: an elementary approach, 2nd edn. Heber City, UT: Kendrick Press.
7. Bary-Soroker L. 2012 Irreducible values of polynomials. Adv. Math. 229, 854–874. (doi:10.1016/j.aim.2011.10.006)
8. Swan RG. 1962 Factorization of polynomials over finite fields. Pac. J. Math. 12, 1099–1106. (doi:10.2140/pjm.1962.12.1099)