



# COMPOSITIO MATHEMATICA

## On the Bateman–Horn conjecture for polynomials over large finite fields

Alexei Entin

Compositio Math. **152** (2016), 2525–2544.

[doi:10.1112/S0010437X16007570](https://doi.org/10.1112/S0010437X16007570)



FOUNDATION  
COMPOSITIO  
MATHEMATICA



LONDON  
MATHEMATICAL  
SOCIETY  
EST. 1865



# On the Bateman–Horn conjecture for polynomials over large finite fields

Alexei Entin

## ABSTRACT

We prove an analogue of the classical Bateman–Horn conjecture on prime values of polynomials for the ring of polynomials over a large finite field. Namely, given non-associate, irreducible, separable and monic (in the variable  $x$ ) polynomials  $F_1, \dots, F_m \in \mathbf{F}_q[t][x]$ , we show that the number of  $f \in \mathbf{F}_q[t]$  of degree  $n \geq \max(3, \deg_t F_1, \dots, \deg_t F_m)$  such that all  $F_i(t, f) \in \mathbf{F}_q[t]$ ,  $1 \leq i \leq m$ , are irreducible is

$$\left( \prod_{i=1}^m \frac{\mu_i}{N_i} \right) q^{n+1} (1 + O_{m, \max \deg F_i, n}(q^{-1/2})),$$

where  $N_i = n \deg_x F_i$  is the generic degree of  $F_i(t, f)$  for  $\deg f = n$  and  $\mu_i$  is the number of factors into which  $F_i$  splits over  $\overline{\mathbf{F}}_q$ . Our proof relies on the classification of finite simple groups. We will also prove the same result for non-associate, irreducible and separable (over  $\mathbf{F}_q(t)$ ) polynomials  $F_1, \dots, F_m$  not necessarily monic in  $x$  under the assumptions that  $n$  is greater than the number of geometric points of multiplicity greater than two on the (possibly reducible) affine plane curve  $C$  defined by the equation

$$\prod_{i=1}^m F_i(t, x) = 0$$

(this number is always bounded above by  $(\sum_{i=1}^m \deg F_i)^2/2$ , where  $\deg$  denotes the total degree in  $t, x$ ) and

$$p = \text{char } \mathbf{F}_q > \max_{1 \leq i \leq m} N_i,$$

where  $N_i$  is the generic degree of  $F_i(t, f)$  for  $\deg f = n$ .

## 1. Introduction

The classical Bateman–Horn conjecture [BH62] predicts the frequency at which a set of irreducible polynomials over the integers attains simultaneously prime values at integer points. Namely, let  $F_1, \dots, F_m \in \mathbf{Z}[x]$ ,  $\deg F_i > 0$ , be non-associate (i.e. no two differ just by a sign) irreducible polynomials over the integers and suppose that for each prime  $p$  there exists  $a \in \mathbf{Z}$  such that  $p \nmid F_1(a) \cdots F_k(a)$ . Then

$$\#\{1 \leq a < x \mid F_1(a), \dots, F_k(a) \text{ are prime}\} \sim \frac{C(F_1, \dots, F_m)}{\prod_{i=1}^m \deg F_i} \frac{x}{\log^m x},$$

Received 13 April 2015, accepted in final form 16 March 2016, published online 21 September 2016.

2010 Mathematics Subject Classification 11T55 (primary).

Keywords: function fields, irreducible polynomials, Bateman–Horn.

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 no. 320755.

This journal is © Foundation Compositio Mathematica 2016.

where

$$C(F_1, \dots, F_m) = \prod_{p \text{ prime}} \frac{1 - \nu(p)/p}{(1 - 1/p)^m},$$

$\nu(p)$  being the number of solutions to  $F_1(x) \cdots F_m(x) \equiv 0 \pmod{p}$  in  $\mathbf{Z}/p$ . The only proved case of the conjecture is the case of a single linear polynomial, which is the prime number theorem for arithmetic progressions.

In the present paper we establish an analogue of this conjecture for polynomials over large finite fields. Let  $q$  be a power of a prime  $p$  and  $\mathbf{F}_q$  the field with  $q$  elements. We will consider an analogue of the Bateman and Horn problem with the ring  $\mathbf{Z}$  replaced by the one-variable polynomial ring  $\mathbf{F}_q[t]$ . For polynomials in several variables we will denote by  $\text{deg}$  the total degree and by  $\text{deg}_t, \text{deg}_x$ , etc., the degree in the respective variable. Let  $F_1, \dots, F_m \in \mathbf{F}_q[t][x]$ ,  $\text{deg}_x F_i > 0$ , be non-associate (i.e. not differing by a multiplicative constant in  $\mathbf{F}_q^\times$ ), irreducible and separable over  $\mathbf{F}_q(t)$ , i.e.  $F_i \notin \mathbf{F}_q[t][x^p]$ . Let  $n$  be a natural number. We ask for how many of the polynomials  $f \in \mathbf{F}_q[t]$  with  $\text{deg} f = n$  all the values  $F_i(t, f) \in \mathbf{F}_q[t]$  for  $i = 1, \dots, m$  are irreducible. We are interested in the asymptotics of this quantity for fixed  $m, \text{deg} F_i, n$  and  $q \rightarrow \infty$ . We will attack this problem by two different methods, each applicable under different additional conditions on  $F_1, \dots, F_m, n$  and  $p = \text{char} \mathbf{F}_q$ , and obtain two sets of results. The first method requires the classification of finite simple groups for its strongest form, while the second method is more direct and does not use any non-elementary facts from group theory.

Our first set of results applies to  $F_i$  which are all monic in  $x$ . The second set of results, which applies also to the non-monic case, will be given as Theorem 1.4 at the end of the present section. To state our results, we define the *slope* of a polynomial

$$P(t, x) = \sum_{j=0}^r c_j(t)x^j, c_r = 1$$

which is monic in  $x$  to be

$$\text{sl} P = \max_{1 \leq j \leq r} \frac{\text{deg} c_{r-j}(t)}{j} \tag{1}$$

(the degree of the zero polynomial is  $-\infty$ ). The slope has the property that  $\text{sl} PQ \leq \max(\text{sl} P, \text{sl} Q)$  and  $\text{sl} PQ = \text{sl} P$  if  $\text{sl} P = \text{sl} Q$  (to see the latter, observe that if  $\text{sl} P = \text{sl} Q, \text{deg}_x P = r_P, \text{deg}_x Q = r_Q$  and  $j_P, j_Q$  are the largest indices for which the maximum in (1) is attained for  $P$  and  $Q$ , respectively, then the degree of the coefficient of  $x^{r_P+r_Q-j_P-j_Q}$  in  $PQ$  is  $(j_P + j_Q)\text{sl} P$ ). Also we always have  $\text{sl} P \leq \text{deg}_t P$ .

Our main result for the monic case is the following.

**THEOREM 1.1.** *Let  $F_1, \dots, F_m \in \mathbf{F}_q[t][x], \text{deg}_x F_i = r_i > 0$ , be non-associate irreducible polynomials which are separable over  $\mathbf{F}_q(t)$  (i.e.  $F_i \notin \mathbf{F}_q[t][x^p]$ ) and monic in  $x$ . Let  $n$  be a natural number satisfying  $n \geq 3$  and  $n \geq \text{sl} F_i$  for  $1 \leq i \leq m$ . Denote  $N_i = r_i n$ . Denote by  $\mu_i$  the number of irreducible factors into which  $F_i(t, x)$  splits over  $\overline{\mathbf{F}}_q$ . Then*

$$\begin{aligned} & \#\{f \in \mathbf{F}_q[t], \text{deg} f = n \mid F(t, f) \in \mathbf{F}_q[t] \text{ is irreducible}\} \\ &= \left( \prod_{i=1}^m \frac{\mu_i}{N_i} \right) q^{n+1} (1 + O_{m, \text{deg} F_i, n}(q^{-1/2})), \end{aligned} \tag{2}$$

the implicit constant in the  $O$ -notation depending only on  $m, \text{deg} F_i$  for  $1 \leq i \leq m$  and  $n$ .

The assumption that  $n \geq \text{sl } F_i$  implies that  $N_i = r_i n$  is the generic degree of  $F_i(t, f)$  for  $\deg f = n$ , i.e. if  $a_0, \dots, a_n$  are free variables, then

$$\deg_t F_i \left( t, \sum_{j=0}^n a_j t^j \right) = N_i.$$

We note that the implied constant can be made explicit, but we do not concern ourselves with tracking it. We conjecture that Theorem 1.1, as well as Theorems 1.2 and 1.3 stated below, hold more generally without the monicity condition on the  $F_i$ , without any conditions on  $n$  except  $n \geq 3$  and for even  $q$  as well. The separability condition on the  $F_i$  generally cannot be omitted; it is not difficult to construct inseparable polynomials violating the Bateman–Horn statistics or even not assuming any irreducible values on  $\mathbf{F}_q[t]$ . See [CCG08] and [Swa62].

The proof of Theorem 1.1 requires the classification of finite simple groups. If we content ourselves with a result valid for  $n \geq 3 \deg_x F_i, 1 \leq i \leq m$  (as well as  $n \geq \text{sl } F_i$ ), then a much more elementary result from group theory is sufficient. See the discussion following the statement of Theorem 3.2. The same applies to Theorem 1.2 below.

Several related, less general, results have been known previously. Bary-Soroker and Jarden [BJ12] established (2) for polynomials  $F_1, \dots, F_k$  which are characteristic-0-like and nodal (see [BJ12, §1] for the precise definitions) in the case  $n = 1$ . Bary-Soroker [Bar12] and Pollack [Pol08] treated the case of  $F_i$  independent of  $t$ , i.e.  $F_i \in \mathbf{F}_q[x]$ . Bary-Soroker also treated the case  $F_i = x + h_i, h_i \in \mathbf{F}_q[t], n \geq \deg h_i$  (analogue of the Hardy–Littlewood prime tuples conjecture) in [Bar14]. The special case  $m = 2$  of the latter result was previously established by Bender and Pollack [BP09]. Bank *et al.* [BB15] treated the case of a single linear polynomial and  $n \geq 3$  ( $n \geq 2$  for odd  $q$ ).

We follow the general strategy used in most of the cited work above, which reduces the Bateman–Horn conjecture to computing the Galois group of the set of polynomials  $F_i(t, a_n t^n + \dots + a_0)$  over the field  $\mathbf{F}_q(a_0, \dots, a_n)$ ,  $a_0, \dots, a_n$  being free variables, using a version of the Chebotarev density theorem (see the next section). The Galois group computation is the novel part of the present work. Unlike the previous results described above where the Galois group was computed directly by algebraic means, we will use the arithmetic significance of the Galois group provided by the density theorem to prove a strong transitivity property, after which we will invoke results about multiply transitive groups. For the proof of Theorem 1.4, we will use a more direct algebraic method.

In the setting of Theorem 1.1, we will not just compute the probability of all the  $F_i(t, f)$  being irreducible but the probability of any possible decomposition. For simplicity we state here our result just for the case of absolutely irreducible polynomials, by which we mean polynomials irreducible over  $\overline{\mathbf{F}}_q$ .

**THEOREM 1.2.** *Let  $F_1, \dots, F_m \in \mathbf{F}_q[t][x]$  and  $n$  satisfy all the conditions of Theorem 1.1 and moreover assume that the  $F_i$  are absolutely irreducible. Denote  $N_i = r_i n$ , where  $r_i = \deg_x F_i$ .*

*Fix partitions*

$$N_i = \sum_{k=1}^{M_i} s_{ik}, s_{ik} \geq 1$$

*of each  $N_i$  for  $1 \leq i \leq m$ . Then the number of  $f \in \mathbf{F}_q[t], \deg f = n$ , such that each  $F_i(t, f)$  decomposes into  $M_i$  irreducible factors of degrees  $s_{ik}, 1 \leq k \leq M_i$  is*

$$\left( \prod_{i=1}^m \mathcal{P}(s_{i1}, \dots, s_{iM_i}) q^{n+1} \right) (1 + O_{k, \deg F_i, n}(q^{-1/2})),$$

where  $\mathcal{P}(s_{i1}, \dots, s_{iM_i})$  is the probability of a random permutation in  $S_{N_i}$  having the cycle structure  $(s_{i1}, \dots, s_{iM_i})$ .

The connection between decompositions of polynomials and cycle structures of permutations will be made clear in the following section. Meanwhile, note that for absolutely irreducible  $F_1, \dots, F_m$ , Theorem 1.2 implies Theorem 1.1, since the probability of a permutation in  $S_{N_i}$  being a full (i.e. length  $N_i$ ) cycle is  $1/N_i$ . For not absolutely irreducible  $F_i$  a similar result can be obtained with the usual permutation groups replaced by certain permutational wreath products. See Theorem 2.3 in the next section.

As a byproduct, we will also obtain the following result, which is a generalised Chowla conjecture for polynomials (the original Chowla conjecture for integers appears in [Cho65]).

**THEOREM 1.3.** *Let  $F_1, \dots, F_m$  and  $n$  satisfy the conditions of Theorem 1.1, except the condition  $n \geq 3$  is not required if  $q$  is odd (if  $q$  is even, we still require  $n \geq 3$ ). Then for every sequence  $s_1, \dots, s_m = \pm 1$  of signs, the number of  $f \in \mathbf{F}_q[t], \deg f = n$ , such that  $\mu(F_i(t, f)) = s_i$  for  $1 \leq i \leq m$  is*

$$\frac{q^{n+1}}{2^m} (1 + O_{k, \deg F_i, n}(q^{-1/2})).$$

Here  $\mu$  denotes the Möbius function for  $\mathbf{F}_q[t]$ .

It is easy to see that in the case of absolutely irreducible polynomials Theorem 1.3 follows from Theorem 1.2, but in fact if  $q$  is odd it follows directly from the much simpler Proposition 5.1 on the multiplicative independence of the discriminants  $\text{Disc}_t F_i(t, a_0 + \dots + a_n t^n) \in \overline{\mathbf{F}}_q(a_0, \dots, a_n)$  ( $a_0, \dots, a_n$  being free variables) modulo squares in  $\overline{\mathbf{F}}_q(a_0, \dots, a_n)^\times$ . Similarly, in the even characteristic case it follows from Proposition 5.2, which is a similar statement about the linear independence of Berlekamp discriminants. The proof of Theorem 1.3 or Propositions 5.1 and 5.2 does not require the classification of finite simple groups or any other non-elementary fact from group theory. The equivalence of Theorem 1.3 and Propositions 5.1 and 5.2 is shown by the method of Carmon and Rudnick [CR14], which they applied to the special case of Theorem 1.3 with  $F_i = x + h_i$  for  $h_i \in \mathbf{F}_q[t]$  with  $n \geq \deg h_i$  and  $q$  odd. Carmon recently generalised the result to even characteristic [Car15]. The only part of their proof which requires this special form is proving Proposition 5.1 for this case. In fact the only part of the present work that requires the conditions that all the  $F_i$  are monic in  $x$  and that  $n \geq \text{sl } F_i$  for all  $i$  is in the proof of Propositions 5.1 and 5.2 so if one can prove Theorem 1.3 or Propositions 5.1 and 5.2 without these conditions then one can dispense with them in all of our results. We conjecture that for  $n \geq 3$  these conditions are not required.

Next we state our second set of results obtained by a different method, which apply to not necessarily monic  $F_1, \dots, F_m$ . We will need to consider the possibly reducible affine plane curve  $C$  over  $\mathbf{F}_q$  defined by the equation  $F(t, x) = 0$ , where  $F = \prod_{i=1}^m F_i$ . For a point  $P \in C(\overline{\mathbf{F}}_q)$ , we denote by  $m_P$  its multiplicity in  $C$ , i.e. the degree of the lowest-degree form appearing in the Taylor expansion of  $F(t, x)$  around  $P$ .

**THEOREM 1.4.** *Let  $F_1, \dots, F_m \in \mathbf{F}_q[t][x]$  with  $\deg_x F_i > 0$  be non-associate irreducible polynomials which are separable over  $\mathbf{F}_q(t)$ , i.e.  $F_i \notin \mathbf{F}_q[t][x^p]$ . Set  $F = \prod_{i=1}^m F_i$ . Let  $C/\overline{\mathbf{F}}_q$  be the (possibly reducible) affine plane curve defined by  $F(t, x) = 0$ . Let  $n \geq 3$  be a natural number satisfying*

$$n > \#\{P \in C(\overline{\mathbf{F}}_q) \mid m_P > 2\}.$$

Denote  $N_i = \deg_t F_i(t, a_n t^n + \dots + a_0)$ , where  $a_0, \dots, a_n$  are free variables. Denote by  $\mu_i$  the number of irreducible factors into which  $F_i(t, x)$  splits over  $\overline{\mathbf{F}}_q$ . Assume that

$$p > \max_{1 \leq i \leq m} N_i.$$

Then the assertions of Theorems 1.1–1.3 hold in this case as well.

The proof of Theorem 1.4 is also based on a Galois group computation, like the proof of Theorems 1.1–1.3, but the computation is more direct and does not use any non-elementary results from group theory. It will be carried out in §§ 6 and 7.

## 2. Galois groups, Frobenius classes and equidistribution

Let  $p$  be a prime number and  $q$  a power of  $p$ . Let  $g \in \mathbf{F}_q[t]$  be a separable polynomial (i.e. having distinct roots over  $\overline{\mathbf{F}}_q$ ) of degree  $N$ . The Frobenius map  $\text{Fr}_q$  (given by  $y \mapsto y^q$ ) defines a permutation of the roots of  $g$ , which gives a well-defined conjugacy class  $\Theta(g)$  of the symmetric group  $S_N$ . The degrees of the prime factors of  $g$  correspond to the cycle lengths of  $\Theta(g)$ . In particular,  $g$  is irreducible if and only if  $\Theta(g)$  is (the conjugacy class of) a full cycle. We call  $\Theta(g)$  the Frobenius class of  $g$ . If  $g_1, \dots, g_m \in \mathbf{F}_q[t]$  are separable polynomials with  $\deg g_i = N_i$ , we get a conjugacy class  $\Theta(g_1, \dots, g_m)$  in  $S_{N_1} \times \dots \times S_{N_m}$  by taking the product of the individual Frobenius classes  $\Theta(g_i)$ . We call  $\Theta(g_1, \dots, g_m)$  the Frobenius class of  $g_1, \dots, g_m$ .

Let  $F_1, \dots, F_m \in \mathbf{F}_q[t][x]$ ,  $\deg_x F_i > 0$  be non-associate, irreducible and separable over  $\mathbf{F}_q(t)$  and  $n$  a natural number. Set  $F = \prod_{i=1}^m F_i$ . Note that by our assumptions  $F$  is separable over  $\mathbf{F}_q(t)$ . Let  $a_0, \dots, a_n$  be free variables,  $\mathbf{f} = a_n t^n + \dots + a_0 \in \mathbf{F}_q[\mathbf{a}, t]$  ( $\mathbf{a}$  is a shorthand for  $a_0, \dots, a_n$ ) and  $N_i = \deg_t F_i(t, \mathbf{f})$ .

*Convention.* For the rest of the paper the asymptotic big- $O$  notation always implies a constant depending only on  $m$ ,  $\deg F_i$  for  $1 \leq i \leq m$  and  $n$ .

To proceed further, we need the following lemma.

LEMMA 2.1. *Under the above assumptions on the  $F_i$  and  $F = \prod F_i$ , the polynomial  $F(t, \mathbf{f}) \in \mathbf{F}_q[\mathbf{a}][t]$  is separable over  $\mathbf{F}_q(\mathbf{a})$ . For all but  $O(q^n)$  of the polynomials  $f \in \mathbf{F}_q[t]$  such that  $\deg f = n$  (the total number of such polynomials is  $q^n(q-1)$ ), the polynomials  $F_i(t, f) \in \mathbf{F}_q[t]$  have degree  $N_i$  and  $F(t, f) \in \mathbf{F}_q[t]$  is separable.*

*Proof.* This is proved in [Rud14]. □

The second part of the lemma implies that for all but  $O(q^n)$  of the polynomials  $f \in \mathbf{F}_q[t]$ ,  $\deg f = n$  (again note that the total number of such polynomials is  $q^n(q-1)$ ), the Frobenius class  $\Theta(F_1(t, f), \dots, F_m(t, f))$  is a well-defined conjugacy class in  $S_{N_1} \times \dots \times S_{N_m}$ .

Let  $L$  be the splitting field of  $F(t, \mathbf{f})$  over  $\mathbf{F}_q(\mathbf{a})$ . Denote by  $G$  its Galois group. It can be viewed as a subgroup of  $S_{N_1} \times \dots \times S_{N_m}$  by its action on the roots of each  $F_i(t, \mathbf{f})$ . These roots are all distinct by the first assertion of Lemma 2.1. Let  $\mathbf{F}_{q^\nu}$  be the algebraic closure of  $\mathbf{F}_q$  in  $L$ . Denote by  $G_1$  the set of  $\sigma \in G$  such that  $\sigma$  acts as the Frobenius map  $\text{Fr}_q$  on  $\mathbf{F}_{q^\nu}$ . It is a coset of the normal subgroup  $\text{Gal}(L/\mathbf{F}_{q^\nu}(\mathbf{a})) \subset G$ .

The fundamental tool we will use in the present work is the following equidistribution result.

THEOREM 2.2. *Let  $F_1, \dots, F_m \in \mathbf{F}_q[t][x]$ ,  $\deg_x F_i > 0$ , be non-associate, irreducible and separable over  $\mathbf{F}_q(t)$ ,  $F = \prod_{i=1}^m F_i$ . Here we do not assume that the  $F_i$  are monic in  $x$ . Let  $n$  be a natural*

number,  $a_0, \dots, a_n$  be free variables,  $\mathbf{a} = (a_0, \dots, a_n)$ ,  $\mathbf{f} = \sum_{j=0}^n a_j t^j$  and  $N_i = \deg_t F_i(t, \mathbf{f})$ . Denote by  $L$  the splitting field of  $F(t, \mathbf{f})$  over  $\mathbf{F}_q(\mathbf{a})$ ,  $G = \text{Gal}(L/\mathbf{F}_q(\mathbf{a}))$  its Galois group and  $\mathbf{F}_{q^\nu}$  the algebraic closure of  $\mathbf{F}_q$  in  $L$ . Denote by  $G_1$  the set of  $\sigma \in G$  acting as  $\text{Fr}_q$  on  $\mathbf{F}_{q^\nu}$ . Consider  $G$  as a subgroup of  $S_{N_1} \times \dots \times S_{N_m}$  via its action on the roots of  $F_i(t, \mathbf{f})$ .

Then, for every conjugacy class  $C$  in  $S_{N_1} \times \dots \times S_{N_m}$ , we have

$$\begin{aligned} & \#\{f \in \mathbf{F}_q[t], \deg f = n \mid \Theta(F_1(t, f), \dots, F_m(t, f)) = C\} \\ &= \frac{\#(C \cap G_1)}{\#G_1} q^{n+1} (1 + O_{m, \deg F_i, n}(q^{-1/2})). \end{aligned}$$

This result is now quite standard and can be seen for example as a special case of [Bar12, Proposition 2.2]. A variant of Theorem 2.2 appears as [ABR15, Theorem 3.1]. There only the case of monic  $F_i$  and  $\nu = 1$  is considered. However, the result is deduced from a more general explicit Chebotarev theorem [ABR15, Theorem A.4], which implies Theorem 2.2 in the same way. Theorem 2.2 can also be viewed as a 0-dimensional case of Deligne–Katz equidistribution [KS99], with the Galois group acting as the monodromy group. Our results (especially Theorem 1.2) can then be seen as a 0-dimensional disconnected fiber analogue of the much deeper equidistribution results of [KS99].

Theorem 2.2 reduces the study of the factorisation statistics of

$$F_1(t, f), \dots, F_m(t, f)$$

to the computation of the Galois group  $G = \text{Gal}(L/\mathbf{F}_q(\mathbf{a}))$  as a permutation group on the roots of  $F_i(t, \mathbf{f})$  over  $\mathbf{F}_q(\mathbf{a})$ . It is this computation which is the heart of the present work. One of the novelties in our work is that unlike all the previous work cited in the introduction we actually use Theorem 2.2 in the computation of the Galois group (in the proof of multiple transitivity, see the next section) and not only apply it after the Galois group has been computed directly.

Theorem 1.2 would follow from Theorem 2.2 if we can show that for  $F_1, \dots, F_m$  satisfying the conditions of Theorem 1.2, the Galois group of  $F(t, \mathbf{f})$  ( $F = \prod F_i$ ) over  $\mathbf{F}_q(\mathbf{a})$  is the maximal possible, i.e.  $S_{N_1} \times \dots \times S_{N_m}$ . In the general case (not necessarily absolutely irreducible  $F_i$ ), the answer is a little bit more complicated and will be stated next.

Let  $P \in \mathbf{F}_q[t][x]$  be irreducible and separable. There is some minimal field  $\mathbf{F}_{q^\mu}$  over which  $P$  splits into absolutely irreducible factors. It is not difficult to see that the number of these factors is  $\mu$  and they are transitively permuted by the Galois group  $\text{Gal}(\mathbf{F}_{q^\mu}/\mathbf{F}_q)$ , which is cyclic of order  $\mu$ .

**THEOREM 2.3.** *Let  $F_1, \dots, F_m \in \mathbf{F}_q[t][x]$ ,  $\deg_x F_i = r_i > 0$  and  $n \geq 3$  satisfy the conditions of Theorem 1.1. Let  $a_0, \dots, a_n$  be free variables,  $\mathbf{a} = (a_0, \dots, a_n)$ ,  $\mathbf{f} = \sum_{j=0}^n a_j t^j$  and  $N_i = \deg_t F_i(t, \mathbf{f}) = r_i n$  (the last equality holds because  $n \geq \text{sl } F_i$ ).*

Let

$$F_i = \prod_{j=1}^{\mu_i} P_{ij}, P_{ij} \in \mathbf{F}_{q^{\mu_i}}[t][x]$$

be the decomposition of  $F_i$  into absolutely irreducible factors. It is easy to see that  $\deg_t P_{ij}(t, \mathbf{f}) = N_i/\mu_i$ . Let  $\mathbf{F}_{q^\mu}$  be the composite of all the  $\mathbf{F}_{q^{\mu_i}}$ .

Denote by  $L$  the splitting field of  $F(t, \mathbf{f})$  over  $\mathbf{F}_q(\mathbf{a})$  and let  $G = \text{Gal}(L/\mathbf{F}_q(\mathbf{a}))$  be its Galois group, which can be considered a subgroup of  $S_{N_1} \times \dots \times S_{N_m}$  through its action on the roots of  $F_i(t, \mathbf{f})$ . Denote by  $\Omega_{ij}$  the set of roots of  $P_{ij}$  in  $L$ . The Galois group  $H_\mu = \text{Gal}(\mathbf{F}_{q^\mu}/\mathbf{F}_q) =$

$\text{Gal}(\mathbf{F}_{q^\mu}(\mathbf{a})/\mathbf{F}_q(\mathbf{a}))$  is isomorphic to  $\mathbf{Z}/\mu$  and for each  $i$  it acts on the set  $\{P_{ij}\}_{1 \leq j \leq \mu_i}$  transitively; the action factors through a principal action of  $H_{\mu_i} = \text{Gal}(\mathbf{F}_{q^{\mu_i}}/\mathbf{F}_q) \cong \mathbf{Z}/\mu_i$ . For  $\sigma \in H_\mu$ , we denote by  $\Omega_{ij}^\sigma$  the set of roots of  $P_{ij}^\sigma$ . The Galois action of  $\sigma$  on the roots sends  $\Omega_{ij}$  into  $\Omega_{ij}^\sigma$ .

The following holds:  $G$  consists of all the permutations  $\pi$  on  $\bigcup_{ij} \Omega_{ij}$  for which there exists some  $\sigma \in H_\mu$  such that  $\pi(\Omega_{ij}) = \Omega_{ij}^\sigma$  for each  $i$  and  $j$ .

Note that if all the  $F_i$  are absolutely irreducible, i.e.  $\mu_i = 1$  for  $1 \leq i \leq m$ , then we get  $G = S_{N_1} \times \dots \times S_{N_m}$ . Theorem 2.5 gives a complete description of the Galois group of  $\mathbf{F}_i(t, \mathbf{f})$  over  $\mathbf{F}_q(\mathbf{a})$  as a permutation group together with its map to  $H_\mu = \text{Gal}(\mathbf{F}_{q^\mu}/\mathbf{F}_q) \cong \mathbf{Z}/\mu$ . It is now an elementary exercise on permutation groups to deduce Theorems 1.1–1.3 from Theorems 2.2 and 2.3. For example, Theorem 1.1 follows from the fact that the probability of a random permutation  $\pi$  of  $\bigcup_{ij} \Omega_{ij}$  satisfying  $\pi(\Omega_{ij}) = \Omega_{ij}^\sigma$  for all  $i, j$  for a given generator  $\sigma$  of  $H_\mu$  (considered as a cyclic permutation on  $\{1, \dots, \mu_i\}$  for each  $i$ ) being a full cycle on each  $\bigcup_{j=1}^{\mu_i} \Omega_{ij}$  is  $\prod_{i=1}^m (\mu_i/N_i)$ . The elementary proof of this fact is carried out in [Bar12]. Theorem 1.2 follows at once from Theorems 2.2 and 2.3 and the remark immediately following the statement of Theorem 2.3.

The essential result from which Theorem 2.3 will follow is the following.

**THEOREM 2.4.** *Let  $p > 2$  be a prime number and  $k = \overline{\mathbf{F}}_p$ . Let*

$$F_1, \dots, F_m \in k[t][x], \quad \deg_x F_i = r_i > 0,$$

*be non-associate, irreducible, monic in  $x$  and separable over  $\overline{\mathbf{F}}_p(t)$ . Let  $n \geq 3$  be a natural number such that  $n \geq \text{sl } F_i$  for  $1 \leq i \leq m$ . Let  $a_0, \dots, a_n$  be free variables,  $\mathbf{a} = (a_0, \dots, a_n)$ ,  $\mathbf{f} = \sum_{j=0}^n a_j t^j$  and  $N_i = \deg_t F_i(t, \mathbf{f}) = r_i n$ . Denote by  $L$  the splitting field of  $\prod_{i=1}^m F(t, \mathbf{f})$  over  $k(\mathbf{a})$  and  $G = \text{Gal}(L/k(\mathbf{a}))$  its Galois group. Then  $G$  is the full permutation group  $S_{N_1} \times \dots \times S_{N_m}$  acting on the roots of  $F_i(t, \mathbf{f})$  over  $k(\mathbf{a})$ .*

The proof of Theorem 2.4 will occupy the next three sections and is the heart of the present work. Theorem 2.3 follows directly from Theorem 2.4 applied to the factors of the  $F_i$  over  $\mathbf{F}_{q^\mu}$  (we use the notation of Theorem 2.3). To see that it can be applied to the factors, it needs to be verified that  $n \geq \text{sl } F_i$  implies the same for the factors of  $F_i$  over  $\mathbf{F}_{q^\mu}$ . But these factors have the same slope as  $F_i$  since  $\text{sl } PQ = \text{sl } P$  whenever  $\text{sl } P = \text{sl } Q$ . So, Theorem 2.4 applies to the factors of the  $F_i$  over  $\mathbf{F}_{q^\mu}$ . Now in the notation of Theorem 2.3 every  $\sigma \in H_\mu = \text{Gal}(\mathbf{F}_{q^\mu}/\mathbf{F}_q) = \text{Gal}(\mathbf{F}_{q^\mu}(\mathbf{a})/\mathbf{F}_q(\mathbf{a}))$  lifts to some  $\sigma' \in \text{Gal}(L/\mathbf{F}_q(\mathbf{a}))$ . It can then be composed with some element of  $\text{Gal}(L/\mathbf{F}_{q^\mu})$  (which can be chosen to permute each  $\Omega_{ij}$  as we please) to obtain any permutation of the form described in the assertion of Theorem 2.3.

Thus, Theorem 2.4 implies Theorem 2.3, which in turn implies Theorems 1.1–1.3. By the same considerations, Theorem 1.4 follows from the following result.

**THEOREM 2.5.** *Let  $p$  be prime. Let  $F_1, \dots, F_m \in \overline{\mathbf{F}}_p[t][x]$ ,  $\deg_x F_i > 0$ , be non-associate irreducible polynomials which are separable over  $\overline{\mathbf{F}}_p(t)$  and not necessarily monic in  $x$ . Denote  $F = \prod_{i=1}^m F_i$ . Let  $C$  be the affine plane curve defined by  $F(t, x) = 0$ . Let  $n \geq 3$  be a natural number such that*

$$n > \#\{P \in C(\overline{\mathbf{F}}_p) \mid m_P > 2\}$$

*(the definition of the multiplicity  $m_P$  is given just before the statement of Theorem 1.4). Denote  $\mathbf{f} = \sum_{j=0}^n a_j t^j \in \overline{\mathbf{F}}_p(a_0, \dots, a_n)[t]$ , where  $\mathbf{a} = (a_0, \dots, a_n)$  are free variables and  $N_i = \deg_t F_i(t, \mathbf{f})$ . Assume that*

$$p > \max_{1 \leq i \leq m} N_i.$$



Let  $L$  be the splitting field of  $F(t, \mathbf{f})$  over  $\overline{\mathbf{F}}_p(\mathbf{a})$  and let  $G = \text{Gal}(L/\overline{\mathbf{F}}_p(\mathbf{a}))$  be its Galois group, which we view as a subgroup of  $S_{N_1} \times \cdots \times S_{N_m}$  via its action on the roots of  $F_i(t, \mathbf{f})$ . Then in fact  $G = S_{N_1} \times \cdots \times S_{N_m}$  is the full product of permutation groups.

The proof of Theorem 2.5 will be given in § 6.

*Remark.* Theorems 2.4 and 2.5 hold in fact for any algebraically closed field  $k$  and not just  $k = \overline{\mathbf{F}}_p$ , provided the required conditions on the characteristic are satisfied or the characteristic is 0. This is because for any fixed  $m, n, \deg F_i$  the assertion can be formulated in the first-order language of fields and all the algebraically closed fields with a given characteristic are elementarily equivalent for this language. Furthermore, if a statement in the first-order language of fields holds in algebraically closed fields of arbitrarily large characteristic, it must also hold in characteristic 0. See [Mar02, § 3.2].

### 3. Computing the Galois group: an outline

In the present section we outline the proof of Theorem 2.4, from which all our other results follow (see § 2). We will show that under the assumptions of Theorem 2.4 for each  $F_i$  the Galois group  $G_i$  of each  $F_i(t, \mathbf{f})$  over  $k(\mathbf{a})$  is  $S_{N_i}$ . Furthermore, we will show that the permutation sign map  $G \rightarrow \{\pm 1\}^m$  is onto. This will finish the proof of Theorem 2.4 by the following elementary lemma on permutation groups.

LEMMA 3.1. *Let  $G \subset S_{N_1} \times \cdots \times S_{N_m}$  be such that the projections  $G \rightarrow S_{N_i}$  are onto and the permutation sign map*

$$G \rightarrow \prod_{i=1}^m S_{N_i}/A_{N_i} \cong \{\pm 1\}^m$$

*is also onto. Then  $G = S_{N_1} \times \cdots \times S_{N_m}$ .*

*Proof.* See [Bar12, Lemma 3.2]. □

In § 4, we will prove that each  $G_i$  is  $(n + 1)$ -transitive (as a permutation group) on the roots of  $\mathbf{F}_i(t, \mathbf{f})$  (Proposition 4.1). Since we always assume that  $n \geq 3$ , this implies that each  $G_i$  is 4-transitive. While the result is stated over  $\overline{\mathbf{F}}_p$ , our proof will use the arithmetic significance of the Galois group, namely we will use Theorem 2.2 in this step as well.

In § 5, we will show that the sign map  $G \rightarrow \{\pm 1\}^m$  is onto. We call this the sign-independence property. In particular,  $G_i \not\subset A_{N_i}$  for each  $i$ . Next we will use the following deep fact from group theory (see [Cam99, Theorem 4.11] or [DM96, § 7.3]).

THEOREM 3.2. *Let  $G \subset S_N$  be a 4-transitive permutation group not contained in  $A_N$ . Then  $G = S_N$ .*

In fact it is known that except for  $S_N$  and  $A_N$  the only 4-transitive groups are simple Mathieu groups (which then must be contained in  $A_N$ ). The proof of this fact requires the classification of finite simple groups, more precisely the Schreier conjecture (that the outer automorphism group of each finite simple group is solvable) which follows from it.

Theorem 3.2 combined with the 4-transitivity of the  $G_i$  and the fact that  $G_i \not\subset A_{N_i}$  implies that  $G_i = S_{N_i}$ . Combined with the sign-independence property of  $G$  and Lemma 3.1, this shows that  $G = S_{N_1} \times \cdots \times S_{N_m}$  as asserted in Theorem 2.4.

*Remark.* By much more elementary means it can be shown that a  $\lceil 3\sqrt{N} - 2 \rceil$ -transitive group  $G \subset S_N$  is either  $S_N$  or  $A_N$ ; see [Hal63, §5.7]. Therefore, if in addition to the assumptions of Theorem 1.1 we assume that  $n \geq 3 \deg_x F_i, 1 \leq i \leq m$ , our results can be proved without using the classification of finite simple groups.

#### 4. Computing the Galois group: multiple transitivity

In the present section we will prove the  $(n + 1)$ -transitivity property for the Galois groups of the individual  $F_i(t, \mathbf{f})$  asserted after the statement of Lemma 3.1. Let  $p$  be a prime number. Let  $F_1 \in \overline{\mathbf{F}}_p[t][x], \deg_x F_1 > 0$ , be irreducible and separable over  $\overline{\mathbf{F}}_p(t)$ . In the present section we do not assume that  $F_1$  is monic in  $x$ , nor that  $p$  is odd. Neither do we require any conditions on  $n$ . Proposition 4.1 below is valid in this generality.

Let  $n$  be a natural number,  $\mathbf{a} = (a_0, \dots, a_n)$  free variables and  $f = \sum_{j=0}^n a_j t^j \in \overline{\mathbf{F}}_p[\mathbf{a}][t]$ . Denote  $N = N_1 = \deg_t F_1(t, \mathbf{f})$ . Let  $\alpha_1, \dots, \alpha_N$  be the roots of  $F_1(t, \mathbf{f})$  in the algebraic closure of  $\overline{\mathbf{F}}_p(\mathbf{a})$ . By Lemma 2.1, they are distinct. Let  $G = G(\overline{\mathbf{F}}_p(\mathbf{a}, \alpha_1, \dots, \alpha_N) / \overline{\mathbf{F}}_p(\mathbf{a}))$  be the Galois group of  $F_1(t, \mathbf{f})$ . We view  $G$  as a permutation group on  $\alpha_1, \dots, \alpha_N$ .

In the present section we prove the following transitivity property.

**PROPOSITION 4.1.** *The action of  $G$  on  $\alpha_1, \dots, \alpha_N$  is  $(n + 1)$ -transitive, i.e. every sequence of  $n + 1$  distinct roots of  $F_1(t, \mathbf{f})$  can be mapped to any other such sequence by some element of  $G$ .*

Although the proposition is formulated over an algebraically closed field and one might expect a purely algebraic or algebro-geometric proof for it, our approach is actually to use its arithmetic significance implied by Theorem 2.2.

Denote by  $L$  the smallest extension of  $\mathbf{F}_p(\mathbf{a})$  containing the coefficients of  $F_1$  and  $\alpha_1, \dots, \alpha_N$ . It is finitely generated over  $\mathbf{F}_p$  and so the algebraic closure of  $\mathbf{F}_p$  in  $L$  is a finite field  $\mathbf{F}_q$ . Replacing  $q$  by a large enough power and replacing  $L$  by  $L\mathbf{F}_q$ , we may assume that  $\text{Gal}(L/\mathbf{F}_q(\mathbf{a})) = \text{Gal}(L\overline{\mathbf{F}}_q/\overline{\mathbf{F}}_q(\mathbf{a})) = G$ . The field  $L$  is the splitting field of  $F_1(t, \mathbf{f})$  over  $\mathbf{F}_q(\mathbf{a})$  and by our assumptions  $\mathbf{F}_q$  is algebraically closed in  $L$ . These properties persist if we replace  $q$  by any power of it. This will be used later.

In the present section we continue using the convention of § 2 that the asymptotic  $O$ -notation has an implied constant depending on  $n, \deg F_1$ . For a polynomial  $g \in \mathbf{F}_q[t]$  and a natural number  $e$ , we denote by  $\ell_e(g)$  the number of length- $e$  sequences of distinct roots of  $g$  in the coefficient field  $\mathbf{F}_q$ . We have of course  $\ell_e(g) = \prod_{i=0}^{e-1} (\ell_1(g) - i)$ . For a permutation  $\sigma \in S_N$ , we will also denote by  $\ell_e(\sigma)$  the number of length- $e$  sequences of distinct fixed points of  $\sigma$ . This is well defined on conjugacy classes in  $S_N$ . Observe that for a separable  $g$  with  $\deg g = N$ , we have  $\ell_e(g) = \ell_e(\Theta(g))$ , where  $\Theta(g)$  is the Frobenius class of  $g$ . In the present section the Frobenius classes are defined via the action of  $\text{Fr}_q$  (not  $\text{Fr}_p$ ).

**PROPOSITION 4.2.** *We have*

$$\sum_{\substack{f \in \mathbf{F}_q[t] \\ \deg f = n}} \ell_{n+1}(F_1(t, f)) = q^{n+1}(1 + O(q^{-1/2})).$$

*Proof.* Denote by  $C$  the affine plane curve defined by  $F_1(t, x) = 0$ . It is absolutely irreducible since  $F_1$  is absolutely irreducible. Let  $C^{n+1}$  be the  $(n + 1)$ -fold product of  $C$  with itself and  $V \subset C^{n+1}$  the open subset of  $(n + 1)$ -tuples of points with distinct  $t$ -coordinates,  $X = C^{n+1} \setminus V$  its closed complement. The variety  $C^{n+1}$  is irreducible and defined by equations of degree  $O(1)$ .

The proper subvariety  $X$  is also defined by equations of degree  $O(1)$ . Therefore, by the Lang–Weil estimates we have  $\#V(\mathbf{F}_q) = q^{n+1}(1 + O(q^{-1/2}))$ . For every sequence of points  $(\tau_i, \xi_i) \in C(\mathbf{F}_q)$ ,  $i = 1, \dots, n + 1$ , with distinct  $\tau_i$ , there is a unique polynomial  $f \in \mathbf{F}_q[t]$ ,  $\deg f \leq n$  such that  $f(\tau_i) = \xi_i, 1 \leq i \leq n + 1$ .

Now, by the definition of  $\ell_e$ , we have

$$\begin{aligned} \sum_{\substack{f \in \mathbf{F}_q[t] \\ \deg f \leq n}} \ell_{n+1}(F_1(t, f)) &= \sum_{\substack{\tau_1, \dots, \tau_{n+1} \in \mathbf{F}_q \\ \text{distinct}}} \#\{f \in \mathbf{F}_q[t], \deg f \leq n, F_1(\tau_i, f(\tau_i)) = 0\} \\ &= \sum_{\substack{(\tau_1, \xi_1), \dots, (\tau_{n+1}, \xi_{n+1}) \in C(\mathbf{F}_q) \\ \tau_i \text{ distinct}}} \#\{f \in \mathbf{F}_q[t], \deg f \leq n, f(\tau_i) = \xi_i\} \\ &= \#V(\mathbf{F}_q) = q^{n+1}(1 + O(q^{-1/2})). \end{aligned} \tag{3}$$

Since the number of  $f$  with  $\deg f < n$  is  $q^n$ , we may replace the condition  $\deg f \leq n$  in the summation with  $\deg f = n$ , introducing an error of  $O(q^n)$ .  $\square$

Recall that  $G = \text{Gal}(L/\mathbf{F}_q(\mathbf{a})) \subset S_N$  and  $\mathbf{F}_q$  is algebraically closed in  $L$ . For a random variable  $X$  on a finite probability space  $S$ , we will denote by  $\langle X(s) \rangle_{s \in S}$  its expected value. We always assume the probability measure to be uniform on the space.

PROPOSITION 4.3. *We have*

$$\langle \ell_{n+1}(\sigma) \rangle_{\sigma \in G} = 1.$$

*Proof.* By Theorem 2.2, the Frobenius elements of  $F_1(t, f)$  for  $f \in \mathbf{F}_q[t], \deg f = n$  are equidistributed in the  $S_N$ -conjugacy classes of  $G$  up to  $O(q^{-1/2})$  (note that  $\nu = 1$  in the notation of Theorem 2.2 since  $\mathbf{F}_q$  is algebraically closed in  $L$ ). Using Proposition 4.2, we see that

$$\langle \ell_{n+1}(\sigma) \rangle_{\sigma \in G} = \langle \ell_{n+1}(\Theta(F_1(t, f))) \rangle_{f \in \mathbf{F}_q[t], \deg f = n} + O(q^{-1/2}) = 1 + O(q^{-1/2})$$

(we may disregard those  $f$  with  $F_1(t, f)$  non-separable by Lemma 2.1). The implicit constant in the error term depends only on  $\deg F, n$  and not on  $q$ . We have observed in the beginning of the section that  $q$  may be replaced by any power of  $q$  with all of our assumptions remaining valid. Since  $\langle \ell_e(\sigma) \rangle_{\sigma \in G}$  is a rational number with denominator dividing  $N!$ , replacing  $q$  with a large enough power of it we see that we must have an equality  $\langle \ell_{n+1}(\sigma) \rangle_{\sigma \in G} = 1$ .  $\square$

To complete the proof of Proposition 4.1, we need the following elementary lemma from group theory.

LEMMA 4.4. *Let  $G$  be a finite group acting on a finite set  $X$ . For  $\sigma \in G$ , denote by  $\ell_e(\sigma)$  the number of length- $e$  sequences of distinct fixed points for the action of  $G$  on  $X$ . Then*

$$\langle \ell_e(\sigma) \rangle_{\sigma \in G} \geq 1,$$

*with equality if and only if  $G$  is  $e$ -transitive.*

*Proof.* First we prove the assertion for  $e = 1$ . For  $x \in X$ , we denote by  $O_x$  its orbit and by  $G_x$  its stabiliser. We have

$$\frac{1}{\#G} \sum_{\sigma \in G} \ell_1(\sigma) = \frac{1}{\#G} \#\{(\sigma, x) \in G \times X \mid \sigma x = x\} = \frac{1}{\#G} \sum_{x \in X} \#G_x = \sum_{x \in X} \frac{1}{\#O_x} \geq 1,$$

since  $\#O_x \leq \#X$  for all  $x$ . Equality holds if and only if  $O_x = X$  for all  $x \in X$ , i.e. if the action of  $G$  is transitive.

To prove the assertion for general  $e$ , consider the set  $X^{(e)}$  of  $e$ -sequences of distinct elements in  $X$  with the  $G$ -action defined by  $\sigma(x_1, \dots, x_e) = (\sigma x_1, \dots, \sigma x_e)$ . For  $\sigma \in G$ , the number of fixed points for this action is exactly  $\ell_e(\sigma)$ . The action of  $G$  on  $X^{(e)}$  is transitive if and only if the action of  $G$  on  $X$  is  $e$ -transitive. Now applying the case  $e = 1$  to the action of  $G$  on  $X^{(e)}$ , we obtain our assertion.  $\square$

Combining Proposition 4.3 and Lemma 4.4, we see that the action of  $G$  on  $\alpha_1, \dots, \alpha_N$  is  $(n + 1)$ -transitive, which finishes the proof of Proposition 4.1.

### 5. Computing the Galois group: sign independence

Let  $p$  be a prime number. Denote  $k = \overline{\mathbf{F}}_p$ . Let  $F_1, \dots, F_m \in k[t][x]$ ,  $\deg_x F_i = r_i > 0$  be non-associate, irreducible and separable over  $k(t)$ . Denote  $F = \prod_{i=1}^m F_i$ . Let  $n$  be a natural number,  $\mathbf{a} = (a_0, \dots, a_n)$  free variables and  $f = \sum_{j=0}^n a_j t^j \in k[\mathbf{a}][t]$ . Assume that  $n \geq \text{sl } F_i, 1 \leq i \leq m$ . Denote  $N_i = r_i n = \deg_t F_i(t, \mathbf{f})$  (the last equality follows from the assumption  $n \geq \text{sl } F_i$ ). Let  $L$  be the splitting field of  $F(t, \mathbf{f})$  over  $k(\mathbf{a})$ ,  $G = \text{Gal}(L/k(\mathbf{a}))$  its Galois group, which we view as a subgroup of  $S_{N_1} \times \dots \times S_{N_m}$  via its action on the roots of  $F_i(t, \mathbf{f})$ .

From § 3, we know that to complete the proof of Theorem 2.4 it is enough to show that the sign projection map  $G \rightarrow \prod_{i=1}^m S_{N_i}/A_{N_i} \cong \{\pm 1\}^m$  is onto. In the case of odd  $p$ , by a well-known fact from Galois theory this is equivalent to the discriminants  $\text{Disc}_t F_i(t, \mathbf{f})$  being linearly independent as elements of  $k(\mathbf{a})^\times/k(\mathbf{a})^{\times 2}$  (note that they are non-zero by Lemma 2.1). This fact follows from the expression of the discriminant of a polynomial  $g = \sum_{i=0}^N b_i t^i, b_N \neq 0$  over an arbitrary field  $K$  as  $\text{Disc } g = \Delta^2$ , where

$$\Delta = b_N^{N-1} \prod_{i < j} (\rho_i - \rho_j),$$

$\rho_i$  being the roots of  $g$ . The expression  $\Delta$  is fixed by even permutations but not by odd ones (in odd characteristic), so  $\Delta \in K$  if and only if the Galois group of  $g$  over  $K$  is contained in  $A_N$  (this argument applies to a single polynomial, but is easy to extend to the case of several  $g_1, \dots, g_k$ ). This is valid only in odd characteristic. A similar criterion can be formulated in characteristic 2 using Berlekamp discriminants [Ber76].

#### 5.1 Sign independence: odd $p$

Assume that  $p > 2$ . We will need the following basic facts about the discriminant (see [GKZ14, § 12]). For every natural number  $N$ , there is a universal polynomial  $\mathcal{D}_N(b_0, \dots, b_N) \in \mathbf{Z}[b_0, \dots, b_N]$  such that over any field  $K$  and for any  $g = \sum_{j=0}^N B_j t^j \in K[t], B_N \neq 0$ , we have

$$\text{Disc } g = \mathcal{D}_N(B_0, \dots, B_N).$$

Furthermore, if we assign to each variable  $b_j$  the weight  $j$ , the polynomial  $\mathcal{D}_N(b_0, \dots, b_N)$  is homogeneous of (weighted) degree  $N(N - 1)$ .

For polynomials  $g_1, g_2$  with non-zero discriminants, we have

$$\text{Disc } g_1 g_2 \equiv \text{Disc } g_1 \text{Disc } g_2 \pmod{K^{\times 2}}.$$

Therefore, to show the multiplicative independence of the discriminants  $\text{Disc}_t F_i(t, \mathbf{f})$  (modulo squares) it is enough to show that the discriminant of any partial product of the  $F_i$  is not a

square. Without loss of generality, we may assume that this partial product is  $F = \prod_{i=1}^m F_i$  (otherwise repeat the argument with a subset of the  $F_i$ ). Note that since  $\text{sl } PQ \leq \max(\text{sl } P, \text{sl } Q)$ , we have  $n \geq \text{sl } F$ . So, it is enough to prove the following.

**PROPOSITION 5.1.** *Let  $F \in k[t][x]$  with  $\deg_x F = r > 0$  be separable over  $k(t)$  and monic in  $x$ . Assume that  $n \geq \text{sl } F$  and as usual  $\mathbf{a} = (a_0, \dots, a_n)$  are free variables,  $\mathbf{f} = \sum_{j=0}^n a_j t^j$ . Then  $\text{Disc}_t F(t, \mathbf{f})$  is not a square in  $k(\mathbf{a})$ .*

*Proof.* We may assume without loss of generality that  $F(0, x)$  is separable. Otherwise find an  $\alpha \in k$  such that  $F(\alpha, x)$  is separable and replace  $t$  with  $t - \alpha$  (such an  $\alpha$  exists because  $\text{Disc}_x F(t, x) \neq 0$  since  $F$  is separable over  $k(t)$ ). This does not change the discriminant of  $F(t, \mathbf{f})$ . By Lemma 2.1,  $F(0, \mathbf{f})$  is separable over  $k(\mathbf{a})$ .

Let us assign weights to the variables  $a_j$  by  $w(a_j) = j$  (and to monomials by additivity). For a polynomial  $H \in k[\mathbf{a}]$ , we will denote by  $\deg_w H$  the highest weight of a monomial appearing in it. The fact that  $F$  is monic in  $x$  and  $n \geq \text{sl } F$  imply that  $N = \deg_t F(t, \mathbf{f}) = rn$ . Write

$$F(t, \mathbf{f}) = \sum_{j=0}^N C_j(\mathbf{a})t^j, C_j \in k[\mathbf{a}],$$

$$F(0, \mathbf{f}) = \sum_{j=0}^N D_j(\mathbf{a})t^j, D_j \in k[\mathbf{a}].$$

We have  $\deg_w C_j \leq j$ . Moreover, the degree  $j$  form of each  $C_j$  with respect to  $w$  is exactly  $D_j$  (since a polynomial of the form  $t^\mu \mathbf{f}^\nu$  has a coefficient of weight  $j - \nu$  at  $t^j$ ), which is homogeneous of degree  $j$  by construction. We have  $\deg_t F(t, \mathbf{f}) = \deg_t F(0, \mathbf{f}) = rn = N$ . It follows from the homogeneity of the discriminant with weight  $j$  for the coefficient of  $t^j$  and the fact that  $\text{Disc}_t F(0, \mathbf{f}) \neq 0$  that the degree  $N(N - 1)$  form of  $\text{Disc}_t F(t, \mathbf{f})$  is exactly  $\text{Disc}_t F(0, \mathbf{f})$ , which is homogeneous of degree  $N(N - 1)$ . This is the leading (highest weight) form of  $\text{Disc}_t F(t, \mathbf{f})$ . It is therefore enough to show that  $\text{Disc}_t F(0, \mathbf{f})$  is not a square. But this is just a special case of the proposition for a polynomial with constant coefficients (i.e. independent of  $t$ ) and this has been proved in [Bar12, Proposition 1.7] for odd  $q$ . □

**5.2 Sign independence:  $p = 2$**

Assume that  $p = 2$ . We recall the definition and basic facts about the Berlekamp discriminant. See [Ber76] and [Car15] for more details. Let  $N$  be a natural number and  $b_0, \dots, b_N$  free variables. In the case of even characteristic, the discriminant  $\mathcal{D}_N(b_0, \dots, b_N) \in \mathbf{F}_2[b_0, \dots, b_N]$  (reduced modulo 2) is in fact the square of a polynomial  $\delta_N(b_0, \dots, b_N) \in \mathbf{F}_2[b_0, \dots, b_N]$ . For a field  $K \supset \mathbf{F}_2$  and a polynomial  $g = \sum_{j=0}^N B_j t^j \in K[t]$ ,  $B_N \neq 0$ , we will denote  $\delta(g) = \delta_N(B_0, \dots, B_N)$ . If  $\rho_1, \dots, \rho_N$  are the roots of  $g$  in an algebraic closure of  $K$ , the *Berlekamp discriminant* of  $g$  is defined to be

$$\text{BDisc}(g) = \sum_{1 \leq i < j \leq N} \frac{\rho_i \rho_j}{\rho_i^2 + \rho_j^2}.$$

It can be written as

$$\text{BDisc}(g) = \frac{\xi(B_0, \dots, B_N)}{\delta(g)^2},$$

where  $\xi(b_0, \dots, b_N) \in \mathbf{F}_2(b_0, \dots, b_N)$  is a universal polynomial depending on  $N$ . We will denote  $\xi(g) = \xi_N(g)$  when  $\deg g = N$ . If we assign the weights  $w(b_i) = i$  to the variables  $b_0, \dots, b_N$ , then  $\delta_N$  is homogeneous of degree  $N(N - 1)/2$  and  $\xi_N$  is homogeneous of degree  $N(N - 1)$ .

A fundamental property of the Berlekamp discriminant is that the Galois group of  $g$  over  $K$  contains an odd permutation of the roots of  $g$  if and only if there exists a  $\tau \in K$  such that  $\text{BDisc}(g) = \tau^2 + \tau$ . We now need to prove the following analogue of Proposition 5.1 for even characteristic.

PROPOSITION 5.2. *Let  $F \in k[t][x]$  with  $\deg_x F = r > 0$  be separable over  $k(t)$  and monic in  $x$ . Assume that  $n \geq \max(3, \text{sl } F)$  and as usual  $\mathbf{a} = (a_0, \dots, a_n)$  are free variables,  $\mathbf{f} = \sum_{j=0}^n a_j t^j$ . Then  $\text{BDisc}_t F(t, \mathbf{f})$  is not of the form  $\tau^2 + \tau$  for  $\tau \in k(\mathbf{a})$ .*

*Proof.* Assume to the contrary that there exists  $\tau \in k(\mathbf{a})$  such that

$$\text{BDisc}_t F(t, \mathbf{f}) = \tau^2 + \tau.$$

Since  $\text{BDisc}_t F(t, \mathbf{f}) = \xi(F(t, \mathbf{f})) / \delta(F(t, \mathbf{f}))^2$  and  $k(\mathbf{a})$  is a unique factorisation domain, we can write  $\tau = u / \delta(F(t, \mathbf{f}))$ ,  $u \in k(\mathbf{a})$ , and we have

$$\xi(F(t, \mathbf{f})) = u^2 + \delta(F(t, \mathbf{f}))u. \tag{4}$$

Now let us assign the weights  $w(a_i) = i$  to the variables and denote by  $\text{lf } H$  the leading form of a polynomial  $H \in k(\mathbf{a})$  with respect to this weight. By the homogeneity properties of  $\delta_N, \xi_N$  and the fact that

$$\deg F(t, \mathbf{f}) = \deg F(0, \mathbf{f}) = n \deg_x F$$

(since  $n \geq \text{sl } F$ ), we have that (denoting  $N = n \deg_x F$ )

$$\begin{aligned} \text{lf } \delta(F(t, \mathbf{f})) &= \delta(F(0, \mathbf{f})), \text{ lf } \xi(F(t, \mathbf{f})) = \xi(F(0, \mathbf{f})), \\ \deg_w \delta(F(t, \mathbf{f})) &= \deg_w \delta(F(0, \mathbf{f})) = N(N - 1)/2, \\ \deg_w \xi(F(t, \mathbf{f})) &= \deg_w \xi(F(0, \mathbf{f})) = N(N - 1) \end{aligned}$$

and, from (4), we also have  $\deg_w u = N(N - 1)/2$ . Using these facts and taking leading forms in (4), we deduce that

$$\xi(F(0, \mathbf{f})) = (\text{lf } u)^2 + \delta(F(0, \mathbf{f})) \cdot \text{lf } u,$$

so, taking  $\tau_1 = \text{lf } u / \delta(F(0, \mathbf{f}))$ , we have

$$\text{BDisc}(F(0, \mathbf{f})) = \tau_1^2 + \tau_1.$$

Now assuming, as we may after a shift in the variable  $t$ , that  $F(0, \mathbf{f}) \in k(\mathbf{a})[t]$  is separable, we have reduced our problem to the constant-coefficient case (i.e. the case when  $F(t, x)$  is independent of  $t$ ). But this is a special case of [Car15, Lemma 6.3]. □

### 6. Proof of Theorem 1.4

The proof of Theorem 2.5, from which Theorem 1.4 follows, will occupy the present section as well as the next one. Let  $p$  be a prime number. Denote  $k = \overline{\mathbf{F}}_p$ . Let  $F_1, \dots, F_m \in k[t][x]$ ,  $\deg_x F_i > 0$ , be non-associate, irreducible and separable over  $k(t)$ . Denote  $F = \prod_{i=1}^m F_i$ . Let  $n \geq 3$  be a natural number. At this point we impose no further restrictions on the  $F_i, n$  or  $p$ . They will be required later. Let  $\mathbf{a} = (a_0, \dots, a_n)$  be free variables over  $k$ ,

$$\mathbf{f} = \sum_{j=0}^n a_j t^j \in k[\mathbf{a}][t], \quad N_i = \deg_t F_i(t, \mathbf{f}).$$

By Lemma 2.1, the polynomial  $F(t, \mathbf{f}) \in k[\mathbf{a}][t]$  is separable over  $k(\mathbf{a})$ . Let  $L$  be the splitting field of  $F(t, \mathbf{f})$  over  $k(\mathbf{a})$ ,  $G = \text{Gal}(L/k(\mathbf{a}))$  its Galois group. We view  $G$  as a subgroup of  $S_{N_1} \times \cdots \times S_{N_m}$  via its action on the roots of each  $F_i(t, \mathbf{f})$ . Our aim is to show that in fact  $G = S_{N_1} \times \cdots \times S_{N_m}$  under the assumptions of Theorem 2.5. Denote by  $G_i \subset S_{N_i}$  the Galois group of  $F_i(t, \mathbf{f})$  over  $k(\mathbf{a})$ . This is the projection of  $G$  to  $S_{N_i}$  defined by restricting its action to the roots of  $F_i(t, \mathbf{f})$ .

By Proposition 4.1, each  $G_i$  acts 2-transitively (in fact  $(n + 1)$ -transitively) on the roots of  $F_i(t, \mathbf{f})$ . Suppose that we could show that for each  $i = 1, \dots, m$ , there exists an element  $\sigma \in G$  which transposes two roots of  $F_i(t, \mathbf{f})$  and leaves all the other roots of  $F(t, \mathbf{f})$  fixed. Then  $G_i = S_{N_i}$ , since  $G_i$  is 2-transitive and  $S_{N_i}$  is generated by transpositions. Furthermore, the sign projection map  $G \rightarrow \prod_{i=1}^m S_{N_i}/A_{N_i} \cong \{\pm 1\}^m$  is onto. Lemma 3.1 would then imply that  $G = \prod_{i=1}^m S_{N_i}$ .

It is therefore sufficient to prove the existence of transpositions as above. Our plan is to construct a discrete valuation ring in  $k(\mathbf{a})$  which ramifies in  $L$  such that its inertia group contains the required transposition. An important ingredient in the proof is the following technical claim, the main idea behind its proof having been suggested to the author by U. Zannier.

PROPOSITION 6.1. *Let  $H \in k[\mathbf{a}] \setminus k[a_n]$  be an irreducible polynomial. Let  $K$  be the algebraic closure of the field of fractions of  $k[\mathbf{a}]/H$ . Let  $b_i$  be the image of  $a_i$  in  $K$ ,  $g = \sum_{j=0}^m b_j t^j \in K[t]$ . Then one of the following holds for  $F(t, g) \in K[t]$ .*

(i)  $F(t, g)$  is separable, i.e. has only simple roots in  $K$ .

(ii)  $F(t, g)$  has one root of multiplicity two and the other roots are simple.

(iii) There exists a point  $(\tau, \xi)$  on the affine plane curve  $C$  defined by  $F(t, x) = 0$  with multiplicity  $m_P$  such that  $H \sim \mathbf{f}(\tau) - \xi \in k[\mathbf{a}]$  (by  $\sim$ , we denote association),  $\tau$  is a root of  $F(t, g)$  of exact multiplicity  $m_P$  and the other roots of  $F(t, g)$  are simple.

The proof of Proposition 6.1 will be given in § 7. We now proceed to the proof of Theorem 2.5. Assume that

$$n > \#\{P \in C \mid m_P > 2\}.$$

Also assume that  $p > \max N_i$ . We have seen that it is enough to produce for each  $1 \leq i \leq m$  an element  $\sigma \in G$  which transposes two roots of  $F_i(t, \mathbf{f})$  and fixes the other roots of  $F(t, \mathbf{f})$ . By symmetry, it is enough to show this for  $i = 1$ . Let  $(\tau_l, \xi_l), l = 0, \dots, n - 2$ , be distinct points on  $C$  including all the points  $P$  such that  $m_P > 2$ . Denote

$$\begin{aligned} a'_l &= \mathbf{f}(\tau_l) - \xi_l, \\ \mathcal{K} &= k(a'_0, \dots, a'_{n-2}, a_n), \end{aligned}$$

$\alpha$  a root of  $F_1(t, \mathbf{f})$  in  $L$  (recall that  $L$  is the splitting field of  $F(t, \mathbf{f})$  over  $k(\mathbf{a})$ ). The variables  $a'_0, \dots, a'_{n-2}, a_{n-1}, a_n$  are obtained from  $a_0, \dots, a_n$  by an invertible affine transformation (it is invertible because of the non-vanishing of Vandermonde determinants with distinct second column entries). We have  $k(\mathbf{a}) = \mathcal{K}(a_{n-1})$ . The field  $k(\mathbf{a}, \alpha) = \mathcal{K}(a_{n-1}, \alpha)$  can now be viewed as a one-variable function field over  $\mathcal{K}$ .

LEMMA 6.2. *The polynomial  $F_1(t, \mathbf{f})$  considered as an element of  $\mathcal{K}[t, a_{n-1}]$  is irreducible over  $\overline{\mathcal{K}}$ .*

*Proof.* We may write

$$\mathbf{f} = u(t) + a_{n-1}v(t), \quad \deg_t u, \deg_t v \leq n, \tag{5}$$

where  $u \in \mathcal{K}[t]$  has coefficients which are affine forms in  $a'_0, \dots, a'_{n-2}, a_n$  and  $v(t) \in k[t]$ . We have  $v \neq 0$  again by the non-vanishing of the Vandermonde determinant with second column entries  $\tau_0, \dots, \tau_{n-2}, t$ . By (5), we have

$$a_{n-1} = (\mathbf{f} - u(t))/v(t).$$

The elements  $t, \mathbf{f} \in \mathcal{K}(a_{n-1}, t)$  are algebraically independent over  $\overline{\mathcal{K}}$ . Suppose that  $F_1(t, \mathbf{f}) = U(t, a_{n-1})V(t, a_{n-1})$ , where  $U, V \in \overline{\mathcal{K}}[t, a_{n-1}]$  are non-constant. Then

$$F_1(t, \mathbf{f}) = U\left(t, \frac{\mathbf{f} - u(t)}{v(t)}\right)V\left(t, \frac{\mathbf{f} - u(t)}{v(t)}\right).$$

But this is impossible since  $F_1(t, \mathbf{f})$  is irreducible in  $\overline{\mathcal{K}}[t, \mathbf{f}]$  and, by Gauss’s lemma, also in  $\overline{\mathcal{K}}(t)[\mathbf{f}]$ . We have obtained a contradiction.  $\square$

**PROPOSITION 6.3.** *There exists an irreducible polynomial  $H \in k[\mathbf{a}]$  which when viewed as a polynomial in  $\mathcal{K}[a_{n-1}]$  is non-constant and defines a place which is ramified in the extension*

$$k(\mathbf{a}) = \mathcal{K}(a_{n-1}) \subset \mathcal{K}(a_{n-1}, \alpha) = k(\mathbf{a}, \alpha)$$

*of one-variable function fields over  $\mathcal{K}$  (recall that  $\alpha$  is a root of  $F_1(t, \mathbf{f})$ ).*

*Proof.* By the previous lemma, the extension  $\mathcal{K}(a_{n-1}) \subset \mathcal{K}(a_{n-1}, \alpha)$  is geometric, i.e.  $\mathcal{K}(a_{n-1})$  is algebraically closed in  $\mathcal{K}(a_{n-1}, \alpha)$ . We assumed that  $p > N_1$ , so the Galois closure of  $\mathcal{K}(a_{n-1}, \alpha)$  over  $\mathcal{K}(a_{n-1})$  is a Galois extension of degree prime to  $p$  and so must ramify at some finite place of  $\mathcal{K}(a_{n-1})$  (the tame fundamental group of the affine line is zero; see [SGA1, § XIII, Corollary 2.12]), which can be defined by some irreducible polynomial  $H \in k[\mathbf{a}] = k[a'_0, \dots, a'_{n-2}, a_n][a_{n-1}]$ . Of course  $H$  ramifies in  $\mathcal{K}(a_{n-1}, \alpha)$  as well. This proves the proposition.  $\square$

We remark that the proof of Proposition 6.3 is the only place where we use the condition  $p > \max N_i$ . Now let  $H$  be as asserted in the proposition. Note that since  $H$  is non-constant as a polynomial in  $\mathcal{K}[a_{n-1}]$ , we have  $H \notin k[a_n]$  and also  $H$  is not associate to a polynomial of the form  $f(\tau) - \xi$  for any point  $P = (\tau, \xi) \in C$  with  $m_P > 2$  (since these elements are in  $\mathcal{K}$  by construction). The image of the polynomial  $F(t, \mathbf{f}) \in k[\mathbf{a}][t]$  modulo  $H$  has degree  $N = \deg_t F(t, \mathbf{f})$  (since the leading coefficient of  $F(t, \mathbf{f})$  is in  $k[a_n]$  and so is prime to  $H$ ) and, by Proposition 6.1, it has at most one double root over the algebraic closure of the fraction field of  $k[\mathbf{a}]/H$ , the other roots being simple, and no root of multiplicity 3 or higher.

Denote by  $R$  the discrete valuation ring  $\mathcal{K}[a_{n-1}]_H$  in  $\mathcal{K}(a_{n-1})$  and by  $S$  any discrete valuation ring lying over it in  $L$ . The field  $R/HR$  is isomorphic to the field of fractions of  $k[\mathbf{a}]/H$  as a  $k[\mathbf{a}]$ -module (recall that  $k[\mathbf{a}] = \mathcal{K}[a_{n-1}]$ ). Denote by  $\eta$  a prime element of  $S$ . Since  $H$  is ramified in the extension  $\mathcal{K}[a_{n-1}] = k[\mathbf{a}] \subset L$ , the inertia group of  $S$  relative to  $R$  is non-empty, so there exists a non-trivial  $\sigma \in G$  satisfying  $\sigma\alpha_j \equiv \alpha_j \pmod{\eta S}$  for any root  $\alpha_j$  of  $F(t, \mathbf{f})$ . By the previous paragraph, there can be at most two roots  $\alpha_1, \alpha_2$  of  $F(t, \mathbf{f})$  for which  $\alpha_1 \equiv \alpha_2 \pmod{\eta S}$  and they must be roots of  $F_1(t, \mathbf{f})$  since  $R$  ramifies in the extension defined by  $F_1(t, \mathbf{f})$ . Therefore,  $\sigma$  transposes  $\alpha_1, \alpha_2$  and leaves the other roots of  $F(t, \mathbf{f})$  fixed, which is exactly what we needed to complete the proof of Theorem 2.5.



**7. Proof of Proposition 6.1**

We keep the setting and notation of the previous section, but we assume no restrictions on  $n$  and  $p$  other than  $n \geq 3$ . Let  $H \in k[\mathbf{a}] \setminus k[a_n]$  be an irreducible polynomial. Let  $K$  be the algebraic closure of the field of fractions of  $k[\mathbf{a}]/H$ . Let  $b_i$  be the image of  $a_i$  in  $K$ ,  $g = \sum_{j=0}^m b_j t^j \in K[t]$ .

We begin by noting that  $H(b_0, \dots, b_n) = 0$  and, up to a constant, this is the only relation satisfied by  $b_0, \dots, b_n$ . By our assumption that  $H \notin k[a_n]$ , we have

$$\deg_t F_i(t, g) = N_i = \deg_t F_i(t, \mathbf{f}),$$

since the leading coefficient of  $F_i(t, f)$  is in  $k[a_n]$  and so is prime to  $H$ .

Denote  $d = \deg_x F$ . Over a finite separable extension  $E$  of  $k(t)$ , we may factor

$$F(t, x) = c(t) \prod_{i=1}^d (x - \zeta_i/c(t)),$$

where  $\zeta_i \in E$  are integral over  $k[t]$  and  $c(t)$  is the leading coefficient of  $F$  as a polynomial in  $x$ . The  $\zeta_i$  are distinct since the  $F_i$  are distinct, irreducible and separable. We will denote  $\phi_i = \zeta_i/c(t) \in E$ .

A *place* on a field  $\mathcal{F}$  with values in a field  $\mathcal{E}$  is a map  $\pi : \mathcal{F} \rightarrow \mathcal{E} \cup \{\infty\}$  such that  $\mathcal{R} = \pi^{-1}(\mathcal{E})$  is a valuation ring in  $\mathcal{F}$  and  $\pi|_{\mathcal{R}}$  is a ring homomorphism. See [Jac89, §§9.7 and 9.8] for the definition and basic properties of valuation rings and places. The most important fact we will use is that if  $\mathcal{E}$  is algebraically closed and  $R \subset \mathcal{F}$  is any subring, then any homomorphism  $\pi : R \rightarrow \mathcal{F}$  can be extended to an  $\mathcal{E}$ -valued place on  $\mathcal{F}$ . If  $\pi$  is a place on  $\mathcal{F}$  which is regular on  $R$ , i.e. does not assume  $\infty$ , then it is also regular on the integral closure of  $R$  in  $\mathcal{K}$ .

Recall that  $K$  is the algebraic closure of the field of fractions of  $k[\mathbf{a}]/H$ . Let  $\alpha \in K$  be any element. Evaluation at  $\alpha$  defines a place  $\pi_\alpha : k(t) \rightarrow K \cup \{\infty\}$  which is regular (i.e. does not assume  $\infty$ ) on  $k[t]$ . For each  $\alpha \in K$ ,  $\pi_\alpha$  can be extended to a place  $E \rightarrow K \cup \{\infty\}$ , which we will also denote by  $\pi_\alpha$ . We choose one such extension for each  $\alpha \in K$ . Since the  $\zeta_i$  are integral over  $k[t]$ , we have  $\pi_\alpha(\zeta_i) \in K$ . If  $c(\alpha) \neq 0$ , then  $\pi_\alpha(\phi_i) \in K$ . For  $h \in E$ , we will use the notation  $h(\alpha) = \pi_\alpha(h)$ . For  $h \in k(t)$ , this coincides with the usual meaning of  $h(\alpha)$ . Note that for  $h \in E$  and any algebraically closed field  $k \subset K' \subset K$ , we have  $h(\alpha) \in K' \cup \{\infty\}$  whenever  $\alpha \in K'$ . This is because  $h$  is algebraic over  $k(t)$ , so  $h(\alpha)$  is algebraic over  $k(\alpha)$  (since  $\pi_\alpha$  is a ring homomorphism on its valuation ring).

The usual derivative defines a derivation  $D : k(t) \rightarrow k(t)$  over  $k$ . Since  $E/k(t)$  is a finite separable extension,  $D$  extends uniquely to a derivation  $D : E \rightarrow E$ . For  $h \in E$ , we will denote  $h' = Dh$ . On  $k(t)$ , this coincides with the usual definition of the derivative. Similarly, the usual derivative on  $K(t)$  can be extended to  $K(t)E = K(t)[x]/F(t, x)$  and this extension is coherent with the extension from  $k(t)$  to  $E$  by uniqueness.

While we may assume that  $p > 2$  for our application, we will prove Proposition 6.1 for  $p = 2$  as well, as this only requires a slight modification and the proposition might be useful in full generality. To accommodate the case  $p = 2$ , we will need to use the second Hasse–Schmidt derivative (see [Gol03, § 1.3] for background on this notion). For a polynomial  $f = \sum_{j=0}^e u_j t^j \in \ell[t]$  over a field  $\ell$ , it is defined by

$$f^{(2)} = \sum_{j=2}^e u_j \binom{j}{2} t^{j-2}.$$

The second Hasse–Schmidt derivative is  $\ell$ -linear and satisfies

$$(f_1 f_2)^{(2)} = f_1^{(2)} f_2 + f_1' f_2' + f_1 f_2^{(2)}. \tag{6}$$

In characteristic  $\neq 2$ , we have  $f^{(2)} = \frac{1}{2}f''$ . An element  $\alpha \in \bar{\ell}$  is a triple (or higher multiplicity) root of  $f$  if and only if  $f(\alpha) = f'(\alpha) = f^{(2)}(\alpha) = 0$ . Like the usual derivative, the second Hasse–Schmidt derivative has a unique extension to  $\ell(t)$  and then to any finite separable extension of it. We extend the second Hasse–Schmidt derivative from  $k(t)$  to  $E$  and from  $K(t)$  to  $K(t)E$  (these are coherent extensions by uniqueness, i.e. the derivative on  $K(t)E$  when restricted to  $E$  coincides with the derivative on  $E$ ).

After setting up the notion of evaluation of elements of  $E$  (which we view as algebraic functions on  $K$ ) on elements of  $K$  and the notions of derivative and second Hasse–Schmidt derivative for elements (algebraic functions) in  $K(t)E$ , we proceed to the proof of Proposition 6.1.

First assume that there exists an element  $\tau \in k$  such that  $F(\tau, g(\tau)) = 0$ . Since  $k$  is algebraically closed, we have  $\xi = \mathbf{g}(\tau) \in k$ . The point  $(\tau, \xi) \in \mathbf{A}^2(k)$  lies on the curve  $C$  defined by  $F(t, x) = 0$ . We have  $g(\tau) = \sum_{j=0}^n b_j \tau^j = \xi$ , so

$$H \sim \sum_{j=0}^n a_j \tau^j - \xi$$

(this relation is irreducible since it is linear and so it is associate with  $H$  by uniqueness). We also see that  $\tau$  is the only root of  $F(t, g)$  contained in  $k$  (again by the uniqueness of the algebraic relation satisfied by  $b_0, \dots, b_n$ ). For simplicity, we assume that  $\tau = \xi = 0$ , otherwise we may shift the variables  $t, x$  by a constant without affecting either the assumptions or the conclusion of the proposition. Then  $H \sim a_0, b_0 = 0$  and  $b_1, \dots, b_n$  are free variables over  $k$ . We also have  $g = b_1 t + \dots + b_n t^n$ .

Let  $m_P$  be the multiplicity of  $(\tau, \xi) = (0, 0)$  as a point on  $C$ . We claim that the multiplicity of 0 as a root of  $F(t, g) \in K[t]$  is exactly  $m_P$ . Let  $F = \sum_{l=m_P}^d \mathcal{F}_l(t, x)$ ,  $\deg \mathcal{F}_l = l$ , be the decomposition of  $F$  into homogeneous forms. We have

$$F(t, g) = \sum_{l=m_P}^d \mathcal{F}_l(t, b_1 t + \dots + b_n t^n) = \mathcal{F}_{m_P}(1, b_1) t^{m_P} + \text{terms of degree } > m_P.$$

Since  $\mathcal{F}_{m_P} \neq 0$ , this proves our claim.

Now we want to show that any other root  $\alpha \neq 0$  of  $F(t, g)$  is simple. Let  $\alpha \neq 0$  be such a root. We have observed that necessarily  $\alpha \notin k$ . In particular,  $c(\alpha) \neq 0$  (since  $k$  is algebraically closed and  $c(t) \in k(t)$ ). Therefore,  $\phi_i = \zeta_i/c$  are regular at  $\alpha$  (i.e.  $\pi_\alpha$  is regular at  $\phi_i$ ) and so are  $\phi'_i, \phi_i^{(2)}$ . We have

$$F(\alpha, g(\alpha)) = c(\alpha) \prod_{i=1}^d (g(\alpha) - \phi_i(\alpha)) = 0,$$

so for some  $i$  we must have  $g(\alpha) = \phi_i(\alpha)$ , i.e.

$$b_1 \alpha + b_2 \alpha^2 + \dots + b_n \alpha^n = \phi_i(\alpha). \tag{7}$$

We will assume that  $i = 1$ , so  $g(\alpha) = \phi_1(\alpha)$ . Now assume that  $\alpha$  is a double root of  $F(t, g)$ . Then  $F(t, g)'(\alpha) = 0$ . We have

$$F(t, g)' = c(t) \sum_{i=1}^d (g' - \phi'_i) \prod_{\substack{1 \leq j \leq d \\ j \neq i}} (g - \phi_j) + c'(t) \prod_{i=1}^d (g - \phi_i);$$

therefore, either  $g'(\alpha) = \phi'_1(\alpha)$ , or  $g(\alpha) = \phi_j(\alpha)$  for some  $j \neq 1$ . The latter cannot happen since then  $\phi_j(\alpha) = \phi_1(\alpha)$  and so  $\alpha \in k$  (because  $(\zeta_i - \zeta_1)(\alpha) = 0$ ,  $\zeta_i - \zeta_1$  divides some non-zero polynomial in  $k[t]$  and  $k$  is algebraically closed), which is a contradiction. So, we have

$$g'(\alpha) = b_1 + 2b_2\alpha + \dots + nb_n\alpha^{n-1} = \phi'_1(\alpha). \tag{8}$$

Multiplying (8) by  $\alpha$  and subtracting from (7), we obtain

$$-b_2\alpha^2 - 2b_3\alpha^3 - \dots - (n-1)b_n\alpha^n = \phi_1(\alpha) - \alpha\phi'_1(\alpha).$$

Since  $\alpha \neq 0$ , it follows that  $b_2$  lies in the algebraic closure of  $k(b_3, \dots, b_n, \alpha)$  and by (8) so does  $b_1$ . This implies that the transcendence degree of  $k(b_1, \dots, b_n)$  over  $k(b_3, \dots, b_n)$  is at most 1, which is a contradiction since  $b_1, \dots, b_n$  are algebraically independent over  $k$ .

Now we handle the case when  $F(t, g)$  has no roots in  $k$ . First we show that  $F(t, g)$  has no root of multiplicity 3 or higher. Assume to the contrary that  $\alpha \in K$  is such a root. Then  $F(t, g)(\alpha) = F(t, g)'(\alpha) = F(t, g)^{(2)}(\alpha) = 0$ . As above, this implies (using the product rule (6) and the fact that  $\alpha \notin k$ ) that for some  $i$  we have

$$g(\alpha) = \phi_i(\alpha), \quad g'(\alpha) = \phi'_i(\alpha), \quad g^{(2)}(\alpha) = \phi_i^{(2)}(\alpha).$$

We assume that this happens for  $i = 1$ . The relation

$$g^{(2)}(\alpha) = b_2 + 3b_3\alpha + \dots + \binom{n}{2}\alpha^{n-2} = \phi_1^{(2)}(\alpha)$$

implies that  $b_2$  is algebraic over  $k(b_3, \dots, b_n, \alpha)$ . The same then follows for  $b_1, b_0$  from the relations  $g'(\alpha) = \phi'_1(\alpha), g(\alpha) = \phi_1(\alpha)$ . This implies that  $k(b_0, \dots, b_n)$  has transcendence degree 1 over  $k(b_3, \dots, b_n, \alpha)$ , so the transcendence degree of  $k(b_0, \dots, b_n)$  over  $k$  is at most  $n - 1$ . This is a contradiction since  $b_0, \dots, b_n$  satisfy only one algebraic relation over  $k$ .

Finally, we want to exclude the possibility of two double roots  $\alpha, \beta \notin k, \alpha \neq \beta$  of  $F(t, f)$ . Assume to the contrary that such  $\alpha, \beta$  exist. Arguing as in the previous cases, we see that there must exist  $i \neq j$  such that

$$g(\alpha) = \phi_i(\alpha), \quad g'(\alpha) = \phi'_i(\alpha), \quad g(\beta) = \phi_j(\beta), \quad g'(\beta) = \phi'_j(\beta). \tag{9}$$

We assume that  $i = 1, j = 2$ . The relations (9) imply that

$$\begin{aligned} b_0 + b_1\alpha + b_2\alpha^2 + b_3\alpha^3, & \quad b_0 + b_1\beta + b_2\beta^2 + b_3\beta^3, \\ b_1 + 2b_2\alpha + 3b_3\alpha^2, & \quad b_1 + 2b_2\beta + 3b_3\beta^2 \end{aligned}$$

are all algebraic over  $k(b_4, \dots, b_n, \alpha, \beta)$  (recall that  $n \geq 3$ ). This gives an inhomogeneous linear system of equations for  $b_0, b_1, b_2, b_3$  over the algebraic closure of  $k(b_4, \dots, b_n, \alpha, \beta)$  with determinant

$$\det \begin{bmatrix} 1 & \alpha & \alpha^2 & \alpha^3 \\ 1 & \beta & \beta^2 & \beta^3 \\ & 1 & 2\alpha & 3\alpha^2 \\ & 1 & 2\beta & 3\beta^2 \end{bmatrix} = -(\alpha - \beta)^4 \neq 0.$$

This implies that  $b_0, b_1, b_2, b_3$  are algebraic over  $k(b_4, \dots, b_n, \alpha, \beta)$ , which is a contradiction since  $b_0, \dots, b_n$  satisfy only one algebraic relation over  $k$ . This concludes the proof of Proposition 6.1.

## ACKNOWLEDGEMENTS

The author would like to thank Zeév Rudnick for many useful discussions and for his encouragement during the research leading to this paper. The present work is part of the author's PhD studies at Tel-Aviv University under his supervision. The author would also like to thank Umberto Zannier for suggesting the main idea behind the proof of Proposition 6.1 and Dan Carmon for suggesting the way to treat the even characteristic case. The author would also like to thank Lior Bary-Soroker for some useful discussions, Brian Conrad for some useful remarks and suggestions and Keith Conrad for his careful reading of the paper and many helpful remarks about the exposition. The author would like to thank the MathOverflow community and especially Peter Müller for pointing out some of the results on multiply transitive groups used in the present work. Finally, the author would like to thank the anonymous referee of this paper for providing multiple expositional corrections and suggestions.

## REFERENCES

- ABR15 J. C. Andrade, L. Bary-Soroker and Z. Rudnick, *Shifted convolutions and the Titchmarsh divisor problem over  $\mathbf{F}_q[t]$* , *Philos. Trans. R. Soc. A* **373** (2015), doi: [10.1098/rsta.2014.0308](https://doi.org/10.1098/rsta.2014.0308).
- BB15 E. Bank and L. Bary-Soroker, *Prime polynomial values of linear functions in short intervals*, *J. Number Theory* **151** (2015), 263–275.
- Bar12 L. Bary-Soroker, *Irreducible values of polynomials*, *Adv. Math.* **229** (2012), 854–874.
- Bar14 L. Bary-Soroker, *Hardy–Littlewood tuple conjecture over large finite fields*, *Int. Math. Res. Not. IMRN* **2014** (2014), 568–575, doi: [10.1093/imrn/rns249](https://doi.org/10.1093/imrn/rns249).
- BJ12 L. Bary-Soroker and M. Jarden, *On the Bateman–Horn conjecture about polynomial rings*, *Münster J. Math.* **5** (2012), 41–58.
- BH62 P. T. Bateman and R. A. Horn, *A heuristic asymptotic formula concerning the distribution of prime numbers*, *Math. Comput.* **16** (1962), 363–367.
- BP09 A. O. Bender and P. Pollack, *On quantitative analogues of the Goldbach and twin prime conjectures over  $\mathbf{F}_q[t]$* , Preprint (2009), [arXiv:0912.1702](https://arxiv.org/abs/0912.1702) [math.NT].
- Ber76 E. R. Berlekamp, *An analog to the discriminant over fields of characteristic two*, *J. Algebra* **38** (1976), 315–317.
- Cam99 P. J. Cameron, *Permutation groups*, London Mathematical Society Student Texts, vol. 45 (Cambridge University Press, 1999).
- Car15 D. Carmon, *The autocorrelation of the Möbius function and Chowla's conjecture for the rational function field in characteristic 2*, *Phil. Trans. R. Soc. A* **373** (2015), doi: [10.1098/rsta.2014.0311](https://doi.org/10.1098/rsta.2014.0311).
- CR14 D. Carmon and Z. Rudnick, *The autocorrelation of the Möbius function and Chowla's conjecture for the rational function field*, *Q. J. Math.* **65** (2014), 53–61, doi: [10.1093/qmath/has047](https://doi.org/10.1093/qmath/has047).
- Cho65 S. Chowla, *The Riemann hypothesis and Hilbert's tenth problem*, *Mathematics and its Applications*, vol. 4 (Gordon & Breach, New York, 1965).
- CCG08 B. Conrad, K. Conrad and R. Gross, *Prime specialization in genus 0*, *Trans. Amer. Math. Soc.* **360** (2008), 2867–2908.
- DM96 J. D. Dixon and B. Mortimer, *Permutation groups*, Graduate Texts in Mathematics, vol. 163 (Springer, New York, 1996).
- GKZ14 I. M. Gelfand, M. M. Kapranov and A. V. Zelevinsky, *Discriminants, resultants and multidimensional determinants*, *Mathematics Theory and Applications* (Birkhäuser, Boston, MA, 2014).
- Gol03 D. M. Goldschmidt, *Algebraic functions and projective curves*, Graduate Texts in Mathematics, vol. 215 (Springer, New York, 2003).

- Hal63 M. Hall, *The theory of groups* (Macmillan, New York, 1963).
- Jac89 N. Jacobson, *Basic algebra II*, 2nd edn (Freeman, New York, 1989).
- KS99 N. M. Katz and P. Sarnak, *Random matrices, Frobenius eigenvalues, and monodromy*, American Mathematical Society Colloquium Publications, vol. 45 (American Mathematical Society, Providence, RI, 1999).
- Mar02 D. Marker, *Model theory: an introduction*, Graduate Texts in Mathematics, vol. 217 (Springer, New York, 2002).
- Pol08 P. Pollack, *Simultaneous prime specializations of polynomials over finite fields*, Proc. Lond. Math. Soc. (3) **97** (2008), 545–567.
- Rud14 Z. Rudnick, *Square-free values of polynomials over the rational function field*, J. Number Theory **135** (2014), 60–66.
- SGA1 A. Grothendieck *et al.*, *SGA 1 – Revêtements étales et groupe fondamental*, Lecture Notes in Mathematics, vol. 224 (Springer, Berlin, Heidelberg, 1971).
- Swa62 R. G. Swan, *Factorization of polynomials over finite fields*, Pacific J. Math. **12** (1962), 1099–1106.

Alexei Entin aentin@stanford.edu

Department of Mathematics, Stanford University, 450 Serra Mall,  
Stanford, CA 94305-2125, USA