Proc. Indian Acad. Sci. (Math. Sci.), Vol. 111, No. 2, May 2001, pp. 139–149.
Printed in India

Descent principle in modular Galois theory

SHREERAM S ABHYANKAR and PRADIPKUMAR H KESKAR*

Mathematics Department, Purdue University, West Lafayette, IN 47907, USA
*Mathematics Department, University of Pune, Pune 411 007, India
E-mail: ram@cs.purdue.edu; keskar@math.unipune.ernet.in

MS received 24 July 2000; revised 1 December 2000

Abstract. We propound a descent principle by which previously constructed equations over $\mathrm{GF}(q^n)(X)$ may be deformed to have incarnations over $\mathrm{GF}(q)(X)$ without changing their Galois groups. Currently this is achieved by starting with a vectorial (= additive) q-polynomial of q-degree m with Galois group $\mathrm{GL}(m, q)$ and then, under suitable conditions, enlarging its Galois group to $\mathrm{GL}(m, q^n)$ by forming its generalized iterate relative to an auxiliary irreducible polynomial of degree n. Elsewhere this was proved under certain conditions by using the classification of finite simple groups, and under some other conditions by using Kantor's classification of linear groups containing a Singer cycle. Now under different conditions we prove it by using Cameron-Kantor's classification of two-transitive linear groups.

Keywords. Galois group; iteration; transitivity.

1. Introduction

In this paper we make some progress towards understanding which finite groups are Galois groups of coverings of the affine line over a ground field of characteristic $p \neq 0$, having at most one branch point other than the point at infinity. We are specially interested in the case when the ground field is not algebraically closed. In particular we realize some of the matrix groups $\mathrm{GL}(m, q^n)$, where $q = p^u > 1$ is a power of p and $m > 0$ and $n > 0$ are integers, over smaller fields of characteristic p than had previously been accomplished. For a tie-up with the geometric case of an algebraically closed ground field and the arithmetic case of a finite ground field see Remark 5.1 at the end of the paper. Likewise, for a tie-up with Drinfeld module theory see Remark 5.2 at the end of the paper.

To describe the contents of the paper in greater detail, henceforth let $q = p^u > 1$ be a power of a prime p, let $m > 0$ and $n > 0$ be integers, and let $\mathrm{GF}(q) \subset k_q \subset K \subset \Omega$ be fields where Ω is an algebraic closure of K; note that there are no assumptions on the field k_q other than for it to contain $\mathrm{GF}(q)$. Also let $E = E(Y)$ be a monic separable vectorial q-polynomial of q-degree m in Y over K, i.e.,

$$E = E(Y) = Y^{q^m} + \sum_{i=1}^{m} X_i Y^{q^{m-i}} \quad \text{with} \quad X_i \in K \text{ and } X_m \neq 0, \tag{1.1}$$

where the elements $X_1, \ldots, X_m$ need not be algebraically independent over k_q. When we want to assume that, for a subset J^* of $\{1, \ldots, m\}$, the elements $\{X_i : i \in J^*\}$ are algebraically independent over k_q and $K = k_q(\{X_i : i \in J^*\})$ with $X_i = 0$ for all $i \notin J^*$, we may express this by saying that we are in the *generic* case of type J^*, and we may indicate it by writing $E^*_{m,q}$ for E and K^* for K. When J^* is the singleton $J^\flat = \{m\}$

we may say that we are in the *binomial* case. When J^* is the pair $J_\mu^\dagger = \{m-\mu, m\}$ with $1 \le \mu < m$ we may say that we are in the μ-*trinomial* case. When J^* is the set $J^\ddagger = \{m-\nu : \nu = 0$ or $\nu =$ a divisor of $m\}$, we may say that we are in the *divisorial* case. Note that the Y-derivative of $E(Y)$ is X_m and hence if $m \in J^*$ then in the generic case of type J^*, the equation $E(Y) = 0$ gives a covering of the affine line over $k_q(\{X_i : m \ne i \in J^*\})$ having $X_m = 0$ as the only possible branch point other than the point at infinity.

In the *general* (= not necessarily generic) case, let V be the set of all roots of E in Ω, and note that then V is an m-dimensional GF(q)-vector-subspace of Ω. Moreover, since GF(q) is assumed to be a subfield of k_q and hence of K, every K-automorphism of the splitting field $K(V)$ of E over K induces a GF(q)-linear transformation of V. Consequently $\mathrm{Gal}(E, K) < \mathrm{GL}(V)$, i.e., the Galois group of E over K may be regarded as a subgroup of GL(V) (see [Ab3]). If we do not assume GF(q) $\subset k_q$ then we only get $\mathrm{Gal}(E, K) < \Gamma\mathrm{L}(V)$, where $\Gamma\mathrm{L}(V)$ is the group of all semilinear transformations of V (see [Ab6]). By fixing a basis of V we may identify GL(V) with GL(m, q), and $\Gamma\mathrm{L}(V)$ with $\Gamma\mathrm{L}(m, q)$. If $J_1^\dagger \subset J^*$ then in the generic case of type J^*, as shown in [Ab2] to [Ab4], we have $\mathrm{Gal}(E^*_{m,q}, K^*) = \mathrm{GL}(m, q)$ but over GF(p), as shown in [Ab6], we have $\mathrm{Gal}(E^*_{m,q}, \mathrm{GF}(p)(\{X_i : i \in J^*\})) = \Gamma\mathrm{L}(m, q)$; for applications of these results see [Ab1] and [Ab5]. To mitigate this bloating we take recourse to generalized iteration as defined in Remark 3.30 of [Ab7] and repeated below. Here bloating refers to the fact that a more direct approach would give a Galois group which is larger than desired, when working over a smaller ground field, and the goal is to modify the covering in order to shrink the group from semilinear to general linear.

DEFINITION 1.2

For every nonnegative integer j we inductively define the jth *iterate* $E^{[[j]]}$ of E by putting $E^{[[0]]} = E^{[[0]]}(Y) = Y$, $E^{[[1]]} = E^{[[1]]}(Y) = E(Y)$, and $E^{[[j]]} = E^{[[j]]}(Y) = E(E^{[[j-1]]}(Y))$ for all $j > 1$. Next we define the *generalized* rth *iterate* $E^{[r]}$ of E for any $r = r(T) = \sum r_i T^i \in \Omega[T]$ with $r_i \in \Omega$ (and $r_i = 0$ for all except a finite number of i), where T is an indeterminate, by putting $E^{[r]} = E^{[r]}(Y) = \sum r_i E^{[[i]]}(Y)$. Note that, for the Y-derivative $E_Y^{[r]}(Y)$ of $E^{[r]}(Y)$ we clearly have

$$E_Y^{[r]}(Y) = E_Y^{[r]}(0) = r(X_m) \tag{1.2.1}$$

and hence if $r(X_m) \ne 0$ then $E^{[r]}$ is a separable vectorial q-polynomial over Ω whose q-degree in Y equals m times the T-degree of r. Also note that the definition of $E^{[r]}$ remains valid for any vectorial E without assuming it to be monic or separable. Moreover, in such a general set-up, this makes the additive group of all vectorial q-polynomials $E = E(Y)$ in Y over Ω into a $\Omega[T]$-*premodule* having all the properties of a module except the left distributive law and the associativity of multiplication, i.e., for all $r, r' \in \Omega[T]$ we have $E^{[r+r']} = E^{[r]} + E^{[r']}$, but for all E, E' over Ω we need not have $(E+E')^{[r]} = E^{[r]} + E'^{[r]}$, and in general $E^{[rr']}$ need not be equal to $(E^{[r]})^{[r']}$. Reverting to the fixed monic separable vectorial E exhibited in (1.1), the said premodule structure makes Ω into a GF(q)[T]-*module* when for every $r \in \mathrm{GF}(q)[T]$ and $z \in \Omega$ we define the 'product' of r and z to be $E^{[r]}(z)$; we denote this GF(q)[T]-module by Ω_E. Now let us fix

$$s = s(T) \in R = \mathrm{GF}(q)[T] \text{ of } T\text{-degree } n \text{ with } s(X_m) \ne 0 \tag{1.2.2}$$

and note that then $E^{[s]}$ is a separable vectorial q-polynomial of q-degree mn in Y over K, and the coefficient of its highest degree term equals the coefficient of the highest degree

of $s(T)$. Let $V^{[s]}$ be the set of all roots of $E^{[s]}$ in Ω, and note that then $V^{[s]}$ is an (mn)-dimensional GF(q)-vector-subspace of Ω. Let GF$(q, s) = R/sR$ where sR is the ideal generated by s in $R = \mathrm{GF}(q)[T]$, and let $\omega : R \to \mathrm{GF}(q, s)$ be the canonical epimorphism. Now $V^{[s]}$ is a submodule of Ω_E and as such it is annihilated by sR and hence we may regard it as a GF(q, s)-module; note that then, for every $r \in R$ and $z \in \Omega$, the 'product' of $\omega(r)$ and z is given by $\omega(r)z = E^{[r]}(z) = \sum r_i E^{[i]}(z)$, and for every $g \in \mathrm{Gal}(K(V^{[s]}), K)$ we have $g(\omega(r)z) = \sum g(r_i)E^{[i]}(g(z)) = (\omega(r))g(z)$; also note that for all $r \in R$ and $z \in \Omega$ we have $rz = \omega(r)z = E^{[r]}(z) = \theta(r, z)$ with $\theta(r, z) \in (\mathrm{GF}(q)[X_1, \ldots, X_m])[z]$. It follows that, in a natural manner,

$$\mathrm{Gal}(E^{[s]}, K) < \mathrm{GL}(V^{[s]}), \tag{1.2.3}$$

where GL$(V^{[s]})$ is the *group of all* GF(q, s)*-linear automorphisms* of $V^{[s]}$, by which we mean all additive isomorphisms $\sigma : V^{[s]} \to V^{[s]}$ such that for all $\eta \in \mathrm{GF}(q, s)$ and $z \in V^{[s]}$ we have $\sigma(\eta z) = \eta\sigma(z)$. Note that

$$s \text{ irreducible in } R \Rightarrow \mathrm{GL}(V^{[s]}) \approx \mathrm{GL}(m, q^n), \tag{1.2.4}$$

where $\approx$ denotes isomorphism. Also note that the Y-derivative of $E^{[s]}(Y)$ is $s(X_m)$ and hence if $m \in J^*$ and s is irreducible in R then in the generic case of type J^*, the equation $E^{[s]}(Y) = 0$ gives a covering of the affine line over $k_q(\{X_i : m \neq i \in J^*\})$ having $s(X_m) = 0$ as the only possible branch point other than the point at infinity; this branch point is rational if and only if $n = 1$.

Now part of what was proved in [Ab7] can be stated as follows:

Trinomial Lemma 1.3. *If* $J_1^\dagger \subset J^*$ *then in the generic case of type* J^* *we have* $\mathrm{Gal}(E^*_{m,q}, K^*) = \mathrm{GL}(m, q)$.

In Note 3.37 of [Ab7] the following problem about generalized iterations was posed.

Problem. Show that if $J^* = \{1, 2, \ldots, m\}$ then in the generic case of type J^* we have $\mathrm{Gal}(E^{*[s]}_{m,q}, K^*) = \mathrm{GL}(V^{[s]})$.

In [AS1] this was proved when $s = T^n$ and in Theorem 3.25 of [Ab7] that result was semilinearized. Likewise in [AS2] it was proved under the assumptions that s is irreducible and m is a square-free integer with $\mathrm{GCD}(m, n) = 1$ and $\mathrm{GCD}(mnu, 2p) = 1$, where we recall that u is the exponent of p in q, i.e., u is the positive integer defined by $q = p^u$. Actually, what was proved in (1.18) of [AS2] was the following slightly more general result.

Weak divisorial Theorem 1.4. *Assume that* s *is irreducible in* R, *and* $J^\ddagger \subset J^*$. *Also assume that* m *is a square-free integer with* $\mathrm{GCD}(m, n) = 1$, *and* $\mathrm{GCD}(mnu, 2p) = 1$. *Then in the generic case of type* J^* *we have* $\mathrm{Gal}(E^{*[s]}_{m,q}, K^*) = \mathrm{GL}(V^{[s]}) \approx \mathrm{GL}(m, q^n)$.

Now CPT (= the classification of projectively transitive permutation groups, i.e., subgroups of GL acting transitively on nonzero vectors) is a remarkable consequence of CT (= the classification theorem of finite simple groups). The implication CT $\Rightarrow$ CPT was mostly proved by Hering [He1, He2]; it is also discussed by Cameron [Cam], Kantor [Ka2], and Liebeck [Lie]. The proof of (1.4) given in [AS2] makes essential use of the following weaker version of CPT, which follows by scanning the list of projectively transitive permutation groups given in [Ka2] or [Lie].

Weak CPT 1.5. Let d be an odd positive integer, and let $G < \mathrm{GL}(d, p)$ be transitive on the nonzero vectors GF$(p)^d \setminus \{0\}$. Then there exist positive integers b, c with $bc = d$ and a group G_0 with $\mathrm{SL}(b, p^c) < G_0 < \Gamma\mathrm{L}(b, p^c)$ such that $G \approx G_0$.

The $m = 1$ case of (1.4), without the hypothesis $\mathrm{GCD}(mnu, 2p) = 1$, was proved by Carlitz [Car] (also see Hayes [Hay]) in connection with his explicit class field theory. In our proof of (1.4) we used the following variation of Carlitz's result which we reproved as Theorem 1.20 in [AS2]; recall that a univariate polynomial $\widetilde{F}(Y) = \sum_{i=0}^{N} \widetilde{F}_i Y^i$ of positive degree N in Y is said to be *Eisenstein* relative $(\widetilde{R}, \widetilde{M})$, where $\widetilde{M}$ is a prime ideal in a ring $\widetilde{R}$, if $\widetilde{F}_N \in \widetilde{R} \setminus \widetilde{M}$, $\widetilde{F}_i \in \widetilde{M}$ for $1 \le i \le N-1$, and $\widetilde{F}_0 \in \widetilde{M} \setminus \widetilde{M}^2$.

Carlitz irreducibility lemma 1.6. *Assume that s is irreducible in R, and $J^\flat \subset J^*$. Let $s^*(T)$ be a nonconstant irreducible factor of $s(T)$ in $k_q[T]$, and let M^* be the ideal in $R^* = k_q[\{X_i : i \in J^*\}]$ generated by $\{X_i : i \in J^* \setminus J^\flat\} \cup \{s^*(X_m)\}$. Then, for $m = 1$, in the generic case of type J^* we have that $M^* = s^*(X_m)R^*$ is a maximal ideal in $R^* = k_q[X_m]$, $Y^{-1}E^{*[s]}_{1,q}(Y)$ is Eisenstein relative to (R^*, M^*), $Y^{-1}E^{*[s]}_{1,q}(Y)$ is irreducible in $K^*[Y]$, and $\mathrm{Gal}(E^{*[s]}_{1,q}, K^*) = \mathrm{GL}(V^{[s]}) \approx \mathrm{GL}(1, q^n)$. Moreover, without assuming $m = 1$, but assuming $\mathrm{GCD}(m, n) = 1$, in the generic case of type J^* we have that M^* is a maximal ideal in R^*, $Y^{-1}E^{*[s]}_{m,q}(Y)$ is Eisenstein relative to (R^*, M^*), $Y^{-1}E^{*[s]}_{m,q}(Y)$ is irreducible in $K^*[Y]$, and $\mathrm{Gal}(E^{*[s]}_{m,q}, K^*)$ has an element of order $q^{mn} - 1$.*

In proving (1.4), in addition to items (1.5) and (1.6), we also used the first part of the following well-known versatile lemma which was initiated by Singer in [Sin] and which was stated as Lemma 1.23 in [AS2]; for an elementary proof of a supplemented version of this see Lemma 5.13 and §6 of [Ab8].

Singer cycle lemma 1.7. *Let $A \in \mathrm{GL}(m, q)$ have order $e = q^m - 1$. Then $\det(A)$ has order $\epsilon = q - 1$, and A acts transitively on the nonzero vectors $\mathrm{GF}(q)^m \setminus \{0\}$, i.e., it is an e-cycle in the symmetric group S_e (and as such it is called a Singer cycle). Moreover, in $\mathrm{GL}(m, q)$ all subgroups generated by such elements, i.e., all cyclic subgroups of order e, form a nonempty complete set of conjugates.*

Now the last assertion of (1.6) says that if s is irreducible in R and $J^\flat \subset J^*$ with $\mathrm{GCD}(m, n) = 1$ then $\mathrm{Gal}(E^{*[s]}_{m,q}, K)$, as a subgroup of $\mathrm{GL}(m, q^n)$, contains a Singer cycle. In his 1980 paper [Ka1], without using CT, Kantor proved the following variation (1.8) of (1.5) by replacing the hypothesis of G acting transitively on nonzero vectors by the stronger hypothesis that G contains a Singer cycle.

Kantor's Singer cycle theorem 1.8. *If $G < \mathrm{GL}(m, q^n)$ contains an element of order $q^{mn} - 1$ then for some divisor m' of m we have $\mathrm{GL}(m', q^{nm/m'}) \triangleleft G$, where $\mathrm{GL}(m', q^{nm/m'})$ is regarded as a subgroup of $\mathrm{GL}(m, q)$ in a natural manner.*

As a consequence of (1.6) and (1.8), but without using (1.5), and hence without using CT, in (5.18) of [Ab8] we proved the following stronger version (1.9) of (1.4) in which the assumption $\mathrm{GCD}(mnu, 2p) = 1$ is replaced by the weaker assumption $\mathrm{GCD}(m, p) = 1$.

Strong divisorial theorem 1.9. *Assume that s is irreducible in R, and $J^\ddagger \subset J^*$. Also assume that m is a square-free integer with $\mathrm{GCD}(m, n) = 1$, and $\mathrm{GCD}(m, p) = 1$. Then in the generic case of type J^* we have $\mathrm{Gal}(E^{*[s]}_{m,q}, K^*) = \mathrm{GL}(V^{[s]}) \approx \mathrm{GL}(m, q^n)$.*

In (1.14) of [Ab9] we settled another case of the above Problem by proving the following Theorem without using the above results (1.4) to (1.9).

Two step theorem 1.10. *Assume that s is irreducible in R, and $J_1^\dagger = J^*$. Also assume that $m = n = 2$. Then in the generic case of type J^* we have* $\mathrm{Gal}(E_{m,q}^{*[s]}, K^*) = \mathrm{GL}(V^{[s]}) \approx \mathrm{GL}(m, q^n)$.

The proof of (1.10) was based on the following lemma which was stated as Lemma 1.16 in [Ab9] and established in §3 of that paper.

Packet throwing lemma 1.11. *Let $\widetilde{M}$ be the maximal ideal in a regular local domain $\widetilde{R}$ of dimension $d > 0$ with quotient field $\widetilde{K}$. Let $\widetilde{F}(Y) = \sum_{0\le i\le N} \widetilde{F}_i Y^i$ be a polynomial of degree $N > 0$ in Y which is Eisenstein relative to $(\widetilde{R}, \widetilde{M})$.* [Note that then for some elements $F_2, \ldots, F_d$ in $\widetilde{R}$ we have $(\widetilde{F}_0, F_2, \ldots, F_d)\widetilde{R} = \widetilde{M}$.] *Let $\widehat{K} = \widetilde{K}(\eta)$ where η is an element in an overfield of $\widetilde{K}$ with $\widetilde{F}(\eta) = 0$, and let $\widehat{R} = \widetilde{R}[\eta]$ and $\widehat{M} = \eta\widehat{R} + \widetilde{M}\widehat{R}$. Then $\widehat{R}$ is the integral closure of $\widetilde{R}$ in $\widehat{K}$, $\widehat{R}$ is a d dimensional regular local domain with maximal ideal $\widehat{M}$, $\widehat{M} \cap \widetilde{R} = \widetilde{M}$, and for any $\widehat{\eta} \in \widehat{K}$ with $\widetilde{F}(\widehat{\eta}) = 0$ and any $F_2, \ldots, F_d$ in $\widetilde{R}$ with $(\widetilde{F}_0, F_2, \ldots, F_d)\widetilde{R} = \widetilde{M}$ we have $(\widehat{\eta}, F_2, \ldots, F_d)\widehat{R} = \widehat{M}$, and hence for any $\widehat{\eta} \in \widehat{K}$ with $\widetilde{F}(\widehat{\eta}) = 0$ we have $\widehat{\eta} \in \widehat{M} \setminus \widehat{M}^2$. Moreover, if for some positive integer $D < N - 1$ we have $\widetilde{F}_D \notin \widetilde{M}^2 + \widetilde{F}_0\widetilde{R}$ and $\widetilde{F}_i \in \widetilde{M}^{D+2-i} + \widetilde{F}_0\widetilde{R}$ for $1 \le i \le D-1$, and $\eta_1, \ldots, \eta_D$ are pairwise distinct elements in $\widehat{K}$ with $\widetilde{F}(\eta_j) = 0$ for $1 \le j \le D$, then $\widetilde{F}(Y) = \widehat{F}(Y)\prod_{1\le j\le D}(Y - \eta_j)$ where $\widehat{F}(Y)$ is a polynomial of degree $N - D$ in Y which is Eisenstein relative to $(\widehat{R}, \widehat{M})$.*

In proving (1.10), the following consequence of (1.11) was implicitly used; in §2 we shall explicitly deduce it from (1.11).

Two transitivity lemma 1.12. *Assume that s is irreducible in R, and we are in the generic case of type J^* with $J^\flat \subset J^*$ and $m > 1$.* [*Note that by* (1.2) *we know that then* $\mathrm{Gal}(E_{q,m}^{*[s]}, K^*) < \mathrm{GL}(V^{[s]}) \approx \mathrm{GL}(m, q^n)$ *and hence we may regard* $\mathrm{Gal}(E_{q,m}^{*[s]}, K^*)$ *to be acting on the $(m-1)$-dimensional projective space $\mathcal{P}(m-1, q^n)$ over* $\mathrm{GF}(q^n)$ (*where the action is not faithful unless $q^n = 2$*).] *Let $N = q^{mn} - 1$ and $\widetilde{F}(Y) = Y^{-1}E_{q,m}^{*[s]}(Y) = \sum_{0\le i\le N} \widetilde{F}_i Y^i$ with $\widetilde{F}_i \in R^* = k_q[\{X_j : j \in J^*\}]$. Assume that the localization of R^* at some nonzero prime ideal in it is a regular local domain $\widetilde{R}$ with maximal ideal $\widetilde{M}$ such that $\widetilde{F}(Y)$ is Eisenstein relative to $(\widetilde{R}, \widetilde{M})$. Let $D = q^n - 1$ and assume that $\widetilde{F}_D \notin \widetilde{M}^2 + \widetilde{F}_0\widetilde{R}$ and $\widetilde{F}_i \in \widetilde{M}^{D+2-i} + \widetilde{F}_0\widetilde{R}$ for $1 \le i \le D-1$. Then* $\mathrm{Gal}(E_{q,m}^{*[s]}, K^*)$ *is two transitive on the $(m-1)$-dimensional projective space $\mathcal{P}(m-1, q^n)$ over* $\mathrm{GF}(q^n)$.

In Theorem I of [CKa], Cameron–Kantor proved the following:

Cameron-Kantor's two transitivity theorem 1.13. *If $m > 2$ and $G < \Gamma\mathrm{L}(m, q)$ is two transitive on the projective space $\mathcal{P}(m-1, q)$, then either* $\mathrm{SL}(m, q) < G$ *or $G =$ the alternating group A_7 inside* $\mathrm{SL}(4, 2)$.

As a consequence of (1.6), (1.7), (1.12), (1.13), and the coefficient computations of §3, but *without using* (1.5) or (1.8) to (1.10), in §4 we shall prove the following theorem. With an eye on further applications, the computations of §3 are more extensive than what we need here.

Main theorem 1.14. *Assume that s is irreducible in R, and $n < m$ with* $\mathrm{GCD}(m, n) = 1$ *and $J_n^\dagger \subset J^*$. Then in the generic case of type J^* we have* $\mathrm{Gal}(E_{m,q}^{*[s]}, K^*) = \mathrm{GL}(V^{[s]}) \approx \mathrm{GL}(m, q^n)$.

In §5 we shall make some motivational and philosophical remarks.

2. Proof of two transitivity lemma

To continue with the discussion of (1.2), for a moment assume that s is irreducible in R with $s(X_m) \neq 0$ and $m > 1$. Then by (1.2.3) and (1.2.4) we have $\mathrm{Gal}(E^{[s]}, K) < \mathrm{GL}(V^{[s]}) \approx \mathrm{GL}(m, q^n)$ and hence we may regard $\mathrm{Gal}(E^{[s]}, K)$ to be acting on the $(m-1)$-dimensional projective space $\mathcal{P}(m-1, q^n)$ over $\mathrm{GF}(q^n)$ (where the action is not faithful unless $q^n = 2$). Let $N = q^{mn} - 1$ and $F(Y) = Y^{-1}E^{[s]}(Y)$. Then $F(Y) \in K[Y]$ is of Y-degree N. For a moment assume that $F(Y)$ is irreducible in $K[Y]$ and let $\widehat{K} = K(\eta)$ where η is a root of $F(Y)$ in Ω. Then $[\widehat{K} : K] = N$ and $\mathrm{Gal}(E^{[s]}, K)$ is transitive on $\mathcal{P}(m-1, q^n)$. Let R_0 be the set of all nonzero members of R of T-degree less than n. Then, in the notation of (1.2), $(\omega(r)\eta)_{r\in R_0}$ are all the distinct 'nonzero scalar multiples' of η in the (R/s)-vector space $V^{[s]}$, and clearly R_0 is the set of all $\alpha_0 + \alpha_1 T + \cdots + \alpha_{n-1}T^{n-1}$ with $(\alpha_0, \alpha_1, \ldots, \alpha_{n-1}) \in \mathrm{GF}(q)^n \setminus \{(0, 0, \ldots, 0)\}$. This gives us D distinct roots of $F(Y)$ in $\widehat{K}$ where $D = q^n - 1$. Therefore $F(Y) = \widehat{F}^*(Y) \prod_{r\in R_0}(Y - \omega(r)\eta)$ where $\widehat{F}^*(Y) \in \widehat{K}[Y]$ is of Y-degree $N - D = q^{mn} - q^n > 1$. Now $(\omega(r)\eta)_{r\in R_0}$ is the inverse image of a point in $\mathcal{P}(m-1, q^n)$ under the natural surjection $\mathrm{GF}(q^n)^m \setminus \{0\} \to \mathcal{P}(m-1, q^n)$ obtained by identifying $V^{[s]}$ with $\mathrm{GF}(q^n)^m$ via a basis. It follows that if $\widehat{F}^*(Y)$ is irreducible in $\widehat{K}$ then $\mathrm{Gal}(E^{[s]}, K)$ is two transitive on $\mathcal{P}(m-1, q^n)$. It is also clear that if $F(Y) = \widehat{F}(Y) \prod_{1\leq i\leq D}(Y - \eta_i)$ where $\eta_1, \ldots, \eta_D$ are distinct roots of $F(Y)$ in $\widehat{K}$ and $\widehat{F}(Y) \in \widehat{K}[Y]$ is irreducible then we must have $\widehat{F}^*(Y) = \widehat{F}(Y)$. Therefore we get the following:

Projective action lemma 2.1. *In the situation of* (1.2) *assume that s is irreducible in R with $s(X_m) \neq 0$ and $m > 1$. Let $F(Y) = Y^{-1}E^{[s]}(Y)$ and note that then $F(Y) \in K[Y]$ is of Y-degree $N = q^{mn} - 1$. Assume that $F(Y)$ is irreducible in $K[Y]$ and let $\widehat{K} = K(\eta)$ where η is a root of $F(Y)$ in Ω. Then $[\widehat{K} : K] = N$ and $\mathrm{Gal}(E^{[s]}, K)$ is transitive on $\mathcal{P}(m-1, q^n)$. Moreover, if upon letting $D = q^n - 1$ we have $F(Y) = \widehat{F}(Y) \prod_{i\leq i\leq D}(Y - \eta_i)$ where $\eta_1, \ldots, \eta_D$ are distinct roots of $F(Y)$ in $\widehat{K}$ and $\widehat{F}(Y) \in \widehat{K}[Y]$ is irreducible then* $\mathrm{Gal}(E^{[s]}, K)$ *is two transitive on $\mathcal{P}(m-1, q^n)$.*

Since Eisenstein polynomials are irreducible, upon taking $E = E^*_{m,q}$ with $F = \widetilde{F}$ and $K = K^* = \widetilde{K}$ in (2.1), by (1.11) we get (1.12).

3. Coefficient computations

Let $R^\natural = \mathrm{GF}(q)[X_1, \ldots, X_m]$. Then clearly for every $\nu > 0$ we have

$$E^{[[\nu]]}(Y) = Y^{q^{m\nu}} + \sum_{i=1}^{m\nu} D_{\nu,i} Y^{q^{m\nu-i}} \text{ with } D_{\nu,i} \in R^\natural. \tag{3.1}$$

Also

$$E^{[[1]]}(Y) = E(Y) = Y^{q^m} + \sum_{i=1}^{m} X_i Y^{q^{m-i}} \tag{3.2}$$

and hence for every integer $\nu > 1$ we have

$$E^{[[\nu]]}(Y) = E(E^{[[\nu-1]]}(Y)) = \left(Y^{q^{m\nu-m}} + \sum_{i=1}^{m\nu-m} D_{\nu-1,i} Y^{q^{m\nu-m-i}}\right)^{q^m} + \sum_{v=1}^{m} X_v \left(Y^{q^{m\nu-m}} + \sum_{w=1}^{m\nu-m} D_{\nu-1,w} Y^{q^{m\nu-m-w}}\right)^{q^{m-v}}$$

$$= \left(Y^{q^{m\nu}} + \sum_{i=1}^{m\nu-m} D_{\nu-1,i}^{q^m} Y^{q^{m\nu-i}} \right) + \left(\sum_{v=1}^{m} X_v Y^{q^{m\nu-v}} + \sum_{v=1}^{m} \sum_{w=1}^{m\nu-m} X_v D_{\nu-1,w}^{q^{m-v}} Y^{q^{m\nu-v-w}} \right)$$

and therefore, for any positive integer i, upon letting

$$Q(i) = \left\{ \begin{array}{l} \text{the set of all pairs of integers } (v, w) \\ \text{with } 1 \le v \le m \text{ and } 1 \le w \le m\nu - m \\ \text{such that } v + w = i \end{array} \right. \tag{3.3}$$

we get

$$D_{\nu,i} = \sum_{(v,w)\in Q(i)} X_v D_{\nu-1,w}^{q^{m-v}} \quad \text{if} \quad m\nu - m < i \le m\nu \tag{3.4}$$

and

$$D_{\nu,i} = X_i + D_{\nu-1,i}^{q^m} + \sum_{(v,w)\in Q(i)} X_v D_{\nu-1,w}^{q^{m-v}} \quad \text{if} \quad 1 \le i \le m. \tag{3.5}$$

By induction we shall show that for every $\nu > 0$ we have

$$D_{\nu,m\nu} = X_m^\nu \tag{3.6}$$

and

$$\left\{ \begin{array}{l} \text{if } l \text{ is an integer with } 1 \le l < m \\ \text{such that } X_i = 0 \text{ whenever } m - l < i < m \\ \text{then } D_{\nu,i} = 0 \text{ whenever } m\nu - l < i < m\nu \\ \text{and } D_{\nu,m\nu-l} = X_{m-l} \sum_{\lambda=0}^{\nu-1} X_m^{(\nu-1)+\lambda(q^l-1)} \end{array} \right. \tag{3.7}$$

and

$$\left\{ \begin{array}{l} \text{if } j \text{ is an integer with } 1 \le j \le m \\ \text{such that } X_i = 0 \text{ whenever } 1 \le i < j \\ \text{then for } 1 \le i \le \min(m, 2j-1) \text{ we have} \\ D_{\nu,i} = \sum_{\lambda=0}^{\nu-1} X_i^{q^{m\lambda}} \\ \text{which we know to be zero if } 1 \le i < j. \end{array} \right. \tag{3.8}$$

By (3.2), this is obvious for $\nu = 1$. So let $\nu > 1$ and assume true for $\nu - 1$. Then clearly $Q(m\nu) = \{(m, m\nu - m)\}$, and hence by (3.4) and the $\nu - 1$ version of (3.6) we get

$$\begin{aligned} D_{\nu,m\nu} &= X_m D_{\nu-1,m\nu-m} \\ &= X_m X_m^{\nu-1} \\ &= X_m^\nu. \end{aligned}$$

Likewise, if l is an integer with $1 \le l < m$ such that $X_i = 0$ whenever $m - l < i < m$, then by (3.4) we get

$$D_{\nu,i} = \left\{ \begin{array}{ll} X_m D_{\nu-1,m\nu-m-i} & \text{if } m\nu - l < i < m\nu \\ X_m D_{\nu-1,m\nu-m-l} + X_{m-l} D_{\nu-1,m\nu-m}^{q^l} & \text{if } m\nu - l = i \end{array} \right.$$

and hence by the $\nu-1$ versions of (3.6) and (3.7) we get

$$D_{\nu,i}=0 \quad \text{if} \quad m\nu-l<i<m\nu$$

and

$$\begin{aligned} D_{\nu,m\nu-l} &= X_{m-l}\left(X_m^{(\nu-1)q^l}+\sum_{\lambda=0}^{\nu-1}X_m^{(\nu-2)+\lambda(q^l-1)}\right)\\ &= X_{m-l}\sum_{\lambda=0}^{\nu-1}X_m^{(\nu-1)+\lambda(q^l-1)}. \end{aligned}$$

Similarly, if j is an integer with $1\le j\le m$ such that $X_i=0$ whenever $1\le i<j$, then for all i,v,w with $1\le i\le 2j-1$ and $(v,w)\in Q(i)$ we have either $v<j$ or $w<j$, and hence by (3.5) and the $\nu-1$ version of (3.8) we see that for $1\le i\le \min(m,2j-1)$ we have

$$D_{\nu,i}=X_i+D_{\nu-1,i}^{q^m}=X_i+\left(\sum_{\lambda=0}^{\nu-2}X_i^{q^{m\lambda}}\right)^{q^m}=\sum_{\lambda=0}^{\nu-1}X_i^{q^{m\lambda}}.$$

4. Proof of main Theorem

To prove the Main Theorem 1.14, assume that s is irreducible in R and $n<m$ with $\mathrm{GCD}(m,n)=1$. Also assume that we are in the generic case of type J^* with $J_n^\dagger\subset J^*$. In view of (1.2.3) and (1.2.4), after identifying $V^{[s]}$ with $\mathrm{GF}(q^n)^m$ via a basis, we have $\mathrm{Gal}(E_{q,m}^{*[s]},K^*)<\mathrm{GL}(m,q^n)$ and we may regard $\mathrm{Gal}(E_{q,m}^{*[s]},K^*)$ as acting on $\mathcal{P}(m-1,q^n)$ (where the action is not faithful unless $q^n=2$). We want to show that $\mathrm{Gal}(E_{q,m}^{*[s]},K^*)=\mathrm{GL}(m,q^n)$.

Let $N=q^{mn}-1$ and $\widetilde{F}(Y)=Y^{-1}E_{q,m}^{*[s]}(Y)=\sum_{0\le i\le N}\widetilde{F}_iY^i$ with $\widetilde{F}_iY^i\in R^*=k_q[\{X_j:j\in J^*\}]$. Let $D=q^n-1$. Note that $s=s(T)=\sum_{0\le\nu\le n}s_\nu T^\nu$ with $s_\nu\in\mathrm{GF}(q)$ and $s_n\ne 0$. Let $\overline{k}_q$ be an algebraic closure of k_q in Ω, and let ζ be a root of $s(T)$ in $\overline{k}_q$. Since $s(T)$ is irreducible in R, we get $\zeta^{q^n-1}=1$ and $s'(\zeta)\ne 0$ where $s'(T)$ is the T-derivative of $s(T)$. Let $\widetilde{R}$ be the localization of $\overline{k}_q[X_n,X_m]$ at the maximal ideal generated by X_n and $X_m-\zeta$. Then $\widetilde{R}$ is two dimensional regular local domain with maximal ideal $\widetilde{M}=(X_n,X_m-\zeta)\widetilde{R}$.

For a moment suppose that $k_q=\overline{k}_q$ and $J_n^\dagger=J^*$, and let us write $K^\dagger$ for K^* and $E_{m,q}^\dagger$ for $E_{m,q}^*$. Now by (1.6) and (1.7) we see that $\widetilde{F}(Y)$ is Eisenstein relative to $(\widetilde{R},\widetilde{M})$, and the determinantal map $\mathrm{Gal}(E_{m,q}^{\dagger[s]},K^\dagger)\to\mathrm{GF}(q^n)\setminus\{0\}$ is surjective. By (1.2.1) we have

$$\widetilde{F}_0=s(X_m).$$

By taking $l=n$ in (3.7) we see that

$$\widetilde{F}_i=0 \text{ for } 1\le i\le D-1$$

and

$$\widetilde{F}_D=X_{m-n}\Theta(X_m),$$

where

$$\Theta(X_m) = \sum_{0\le\nu\le n} s_\nu \sum_{0\le\lambda\le\nu-1} X_m^{(\nu-1)+\lambda(q^n-1)}.$$

Since $\zeta^{q^n-1} = 1$, we get

$$\sum_{0\le\lambda\le\nu-1} \zeta^{(\nu-1)+\lambda(q^n-1)} = \nu\zeta^{\nu-1}$$

and therefore

$$\Theta(\zeta) = \sum_{0\le\nu\le n} s_\nu \nu\zeta^{\nu-1} = s'(\zeta) \neq 0.$$

It follows that

$$\widetilde{F}_D \notin \widetilde{M}^2 + \widetilde{F}_0\widetilde{R}$$

and hence by (1.12) we conclude that $\mathrm{Gal}(E_{m,q}^{\dagger[s]}, K^\dagger)$ is two transitive on $\mathcal{P}(m-1, q^n)$. If $n > 1$ then by (1.13) we see that $\mathrm{SL}(m, q^n) < \mathrm{Gal}(E_{m,q}^{\dagger[s]}, K^\dagger)$ and hence, because the determinantal map $\mathrm{Gal}(E_{m,q}^{\dagger[s]}, K^\dagger) \to \mathrm{GF}(q^n)\backslash\{0\}$ is surjective, we must have $\mathrm{Gal}(E_{m,q}^{\dagger[s]}, K^\dagger) = \mathrm{GL}(m, q^n)$. If $n = 1$ then by (1.3) we get $\mathrm{Gal}(E_{m,q}^{\dagger[s]}, K^\dagger) = \mathrm{GL}(m, q^n)$. Thus in both the cases we have $\mathrm{Gal}(E_{m,q}^{\dagger[s]}, K^\dagger) = \mathrm{GL}(m, q^n)$.

Now let us return to the case when the field k_q need not be algebraically closed. Since $\overline{k}_q$ is an overfield of k_q and $E_{m,q}^{\dagger[s]}$ is obtained from $E_{m,q}^{*[s]}$ by putting $X_i = 0$ for all $i \in J^* \setminus J_n^\dagger$, in view of the extension principle (cf. p. 93 of [Ab2]) and the specialization principle (cf. p. 1894 of [AbL]), see that $\mathrm{Gal}(E_{m,q}^{\dagger[s]}, K^\dagger) < \mathrm{Gal}(E_{m,q}^{*[s]}, K^*)$. Therefore $\mathrm{Gal}(E_{m,q}^{*[s]}, K^\dagger) = \mathrm{GL}(m, q^n)$.

5. Concluding remarks

Let us end with some remarks on motivation and philosophy.

Remark 5.1 (*Algebraic fundamental groups*). The algebraic fundamental group $\pi_A(L_k)$ of the affine line L_k over a field k is defined to be the set of all Galois groups of finite unramified Galois coverings of the affine line L_k over k. Similarly we define $\pi_A(L_{k,t})$ for $L_{k,t} = L_k$ punctured at t points, and more generally we define $\pi_A(C_{g,w})$ for a nonsingular projective genus g curve C over k punctured at $w+1$ points. Let $Q(p)$ be the set of all quasi-p groups, i.e., finite groups G such that $G = p(G)$ where $p(G)$ is the subgroup of G generated by all of its p-Sylow subgroups, and more generally let $Q_t(p)$ be the set of all quasi-(p,t) groups, i.e., those G for which $G/p(G)$ is generated by t generators. In [Ab1], as *geometric conjectures* it was predicted that if k is an algebraically closed field of characteristic p then $\pi_A(L_k) = Q(p)$, and more generally $\pi_A(L_{k,t}) = Q_t(p)$ and $\pi_A(C_{g,w}) = Q_{2g+w}(p)$. In 1994, these were settled affirmatively by Raynaud [Ray] and Harbater [Har]. For higher dimensional versions of the geometric conjectures see [Ab5]. Then, mostly inspired by Fried–Guralnick–Saxl [FGS] and Guralnick–Saxl [GuS], we turned our attention to coverings defined over finite fields. In [Ab6] this led to the *arithmetical question* asking whether $\pi_A(L_{\mathrm{GF}(q)}) = Q_1(p)$, the philosophy behind this being that dropping from an algebraically closed field to a finite field is somewhat like adding a branch point. In particular we may ask whether $\pi_A(L_{k,1})$ contains $Q_1(p)$ where

k is an overfield of GF(q). As indicated in the introduction, in doing this arithmetical problem, the linear groups got bloated towards their semilinear versions and the attempt to unbloat them led us to generalized iterations.

Remark 5.2 (*Division points and Drinfeld modules*). The generalized iterations themselves came out of the theory of Drinfeld modules as developed in his paper [Dri]. This work of Drinfeld seems to have been inspired by Serre's work [Se1] on division points of elliptic curves which was later generalized by him [Se2] to abelian varieties. In turn, our description of the module $E^{[s]}$ in (1.2) is based on the ideas of Drinfeld modules. For a discussion of Drinfeld modules and their relationship with division points of elliptic curves and abelian varieties see Goss [Gos]. Very briefly, the roots of the separable vectorial q-polynomial E of q-degree $2m$ exhibited in (1.1) form a $2m$ dimensional GF(q)-vector-space on which the Galois group of E acts. The said Galois group also acts on the roots of $E^{[s]}$ discussed in (1.2) which are the analogues of 's-division points of E.' Indeed, we have used the letter E to remind ourselves of elliptic curves in case of $m = 1$ and more generally of $2m$ dimensional abelian varieties. We hope that the present descent principle can somehow be 'lifted' to characteristic zero. Before that it should be made to work in the symplectic situation, the bloated semilinear equations for which can be found in [Ab7]. Prior to that the GL work of this paper should be completed.

Acknowledgement

This work was partly supported by NSF Grant DMS 99-88166 and NSA grant MDA 904-97-1-0010.

References

[Ab1] Abhyankar S S, Coverings of algebraic curves, *Am. J. Math.* **79** (1957) 825–856
[Ab2] Abhyankar S S, Galois theory on the line in nonzero characteristic, *Bull. Am. Math. Soc.* **27** (1992) 68–133
[Ab3] Abhyankar S S, Nice equations for nice groups, *Israel J. Math.* **88** (1994) 1–24
[Ab4] Abhyankar S S, Projective polynomials, *Proc. Am. Math. Soc.* **125** (1997) 1643–1650
[Ab5] Abhyankar S S, Local fundamental groups of algebraic varieties, *Proc. Am. Math. Soc.* **125** (1997) 1635–1641
[Ab6] Abhyankar S S, Semilinear transformations, *Proc. Am. Math. Soc.* **127** (1999) 2511–2525
[Ab7] Abhyankar S S, Galois theory of semilinear transformations, Proceedings of the UF Galois Theory Week 1996 (ed.) Helmut Voelklein *et al, London Math. Soc., Lecture Note Series* **256** (1999) 1–37
[Ab8] Abhyankar S S, Desingularization and modular Galois theory (to appear)
[Ab9] Abhyankar S S, Two step descent in modular Galois theory, theorems of Burnside and Cayley, and Hilbert's thirteenth problem (to appear)
[AbL] Abhyankar S S and Loomis P A, Once more nice equations for nice groups, *Proc. Am. Math. Soc.* **126** (1998) 1885–1896
[AS1] Abhyankar S S and Sundaram G S, Galois theory of Moore– Carlitz–Drinfeld modules, *C. R. Acad. Sci. Paris* **325** (1997) 349–353
[AS2] Abhyankar S S and Sundaram G S, Galois groups of generalized iterates of generic vectorial polynomials (to appear)
[Cam] Cameron P J, Finite permutation groups and finite simple groups, *Bull. London Math. Soc.* **13** (1981) 1–22
[CKa] Cameron P J and Kantor W M, 2-Transitive and antiflag transitive collineation groups of finite projective spaces, *J. Algebra* **60** (1979) 384–422
[Car] Carlitz L, A class of polynomials, *Trans. Am. Math. Soc.* **43** (1938) 167–182
[Dri] Drinfeld V G, Elliptic Modules, *Math. Sbornik* **94** (1974) 594–627

[FGS] Fried M D, Guralnick R M and Saxl J, Schur covers and Carlitz's conjecture, *Israel J. Math.* **82** (1993) 157–225

[GuS] Guralnick R M and Saxl J, Monodromy groups of polynomials, Groups of Lie Type and their Geometries (eds) W M Kantor and L Di Marino (Cambridge University Press) (1995) 125–150

[Gos] Goss D, Basic Structures of Function Field Arithmetic (Springer-Verlag) (1996)

[Har] Harbater D, Abhyankar's conjecture on Galois groups over curves, *Invent. Math.* **117** (1994) 1–25

[Hay] Hayes D R, Explicit class field theory for rational function fields, *Trans. Am. Math. Soc.* **189** (1974) 77–91

[He1] Hering C, Transitive linear groups and linear groups which contain irreducible subgroups of prime order, *Geometriae Dedicata* **2** (1974) 425–460

[He2] Hering C, Transitive linear groups and linear groups which contain irreducible subgroups of prime order II, *J. Algebra* **93** (1985) 151–164

[Ka1] Kantor W M, Linear groups containing a Singer cycle, *J. Algebra* **62** (1980) 232–234

[Ka2] Kantor W M, Homogeneous designs and geometric lattices, *J. Combinatorial Theory* **A38** (1985) 66–74

[Lie] Liebeck M W, The affine permutation groups of rank three, *Proc. London Math. Soc.* **54** (1987) 477–516

[Ray] Raynaud M, Revêtment de la droit affine en charactéristic $p > 0$ et conjecture d'Abhyankar, *Invent. Math.* **116** (1994) 425–462

[Se1] Serre J-P, Propriétés galoisiennes des points d'ordre fini des courbes elliptiques, *Invent. Math.* **15** (1972) 259–331

[Se2] Serre J-P, Résumé des cours et travaux, *Annuaire du Collège de France* **85–86** (1985)

[Sin] Singer J, A theorem in finite projective geometry and some applications in number theory, *Trans. Am. Math. Soc.* **43** (1938) 377–385