Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Algebra 297 (2006) 234–253
JOURNAL OF Algebra
www.elsevier.com/locate/jalgebra

A result of Hermite and equations of degree 5 and 6

Hanspeter Kraft [1]

Mathematisches Institut der Universität Basel, Rheinsprung 21, CH-4051 Basel, Switzerland

Received 27 January 2005

Available online 25 May 2005

Communicated by Corrado de Concini

Abstract

A classical result from 1861 due to HERMITE says that every separable equation of degree 5 can be transformed into an equation of the form $x^5 + bx^3 + cx + d = 0$. Later, in 1867, this was generalized to equations of degree 6 by JOUBERT. We show that both results can be understood as an explicit analysis of certain covariants of the symmetric groups S_5 and S_6. In case of degree 5, the classical invariant theory of binary forms of degree 5 comes into play whereas in degree 6 the existence of an outer automorphism of S_6 plays an essential rôle.

1. Introduction

Let L/K be a finite separable field extension of degree n. A classical problem is to find a generating element $x \in L$ whose equation

$$x^n + a_1 x^{n-1} + a_2 x^{n-2} + \cdots + a_{n-1} x + a_n = 0 \tag{1}$$

is as simple as possible. For example, a quadratic extension in characteristic $\neq 2$ has always a generator with equation $x^2 + b = 0$. One easily shows that for a separable extension of

E-mail address: hanspeter.kraft@unibas.ch.

[1] The author is supported by the Swiss National Science Foundation (Schweizerischer Nationalfonds).

0021-8693/$ – see front matter
doi:10.1016/j.jalgebra.2005.04.015

degree 3 there is a generator with equation $x^3 + bx + b = 0$ and similarly for an extension of degree 4 (see Section 7).

In degree 5 and 6 we have the following classical results which go back to HERMITE [Her61] and JOUBERT [Jou67].

Main Theorem. (a) *For any separable field extension L/K of degree* 5 *there is a generator $x \in L$ whose equation has the form*

$$x^5 + bx^3 + cx + c = 0,$$

except for $K = \mathbb{F}_2$ where the equation has the form $x^5 + x^3 + 1 = 0$.

(b) *Let L/K be a separable extension of degree* 6. *If* $\operatorname{char} K \neq 2$ *then there is a generator $x \in L$ whose equation has the form*

$$x^6 + bx^4 + cx^2 + dx + d = 0.$$

The arguments given by HERMITE and JOUBERT work in characteristic zero. They are short and elegant and both are based on some classical invariant theory. The idea is to construct "universal" TSCHIRNHAUS transformations which, applied to any generator of L/K, produce elements of L whose Eqs. (1) satisfy the required properties, i.e. $a_1 = a_3 = 0$. From this it is not difficult to obtain the Main Theorem (at least in characteristic zero) although this is not explicitly formulated in their papers. To get the result also in positive characteristic and, in particular, for finite fields needs a little more work and some explicit computations.

The aim of this note is to give a "modern" approach to these results, following (and explaining) the classical ideas. We will show that HERMITE's and JOUBERT's method can be considered as a careful and explicit analysis of certain covariants of the symmetric groups S_5 and S_6. In degree 5 the classical invariant theory of binary forms of degree 5 comes into play whereas in degree 6 the existence of an outer automorphism of S_6 plays an essential rôle. Another modern approach was given by CORAY in [Cor87]; it is based on rationality questions for cubic hypersurfaces.

There is the obvious question to generalize these results to higher degree. However, it was shown by REICHSTEIN in [Rei99] that, in general, this is not possible (see Example 1 in Section 3).

It should be pointed out here that the relation between equations of degree n and covariants of the symmetric group S_n has be studied in detail by BUHLER and REICHSTEIN in [BuR97] (see also [BuR99]).

2. Covariants and Tschirnhaus transformations

Let K be field and $n \in \mathbb{N}$ a positive integer. To any $x = (x_1, x_2, \ldots, x_n) \in K^n$ we associate the polynomial $\pi(x) := \prod_i (X - x_i) \in K[X]$. This defines a polynomial map

$$\pi : \mathbb{A}^n \to P_n$$

where P_n denotes the unitary polynomials of degree n:

$$P_n(K) := \{ f = X^n + a_1 X^{n-1} + a_2 X^{n-2} + \cdots + a_{n-1} X + a_n \mid a_i \in K \}.$$

The morphism π is defined over $\mathbb{Z}$ and corresponds to the *algebraic quotient* with respect to the natural action of the symmetric group S_n on $\mathbb{A}^n$ by permutations. This means that the polynomial functions on P_n are identified, via π, with the symmetric functions on $\mathbb{A}^n$,

$$\pi^* : \mathbb{Z}[P_n] = \mathbb{Z}[a_1, a_2, \ldots, a_n] \xrightarrow{\sim} \mathbb{Z}[\mathbb{A}^n]^{S_n}, \quad a_k \mapsto (-1)^k s_k(x_1, x_2, \ldots, x_n),$$

where s_k denotes the kth elementary symmetric function. If $f \in P_n(K)$ and $\xi_1, \ldots, \xi_n$ the roots of f (with multiplicities) in some field extension L/K then $\pi^{-1}(f)$ is the S_n-orbit of $\xi = (\xi_1, \ldots, \xi_n) \in \mathbb{A}^n(L) = L^n$.

Let $\Phi = (\varphi_1, \varphi_2, \ldots, \varphi_n) : \mathbb{A}^n \to \mathbb{A}^n$ be an S_n-equivariant polynomial map. By definition, Φ induces a morphism $\overline{\Phi} : P_n \to P_n$ such that the following diagram commutes:

$$\begin{array}{ccc} \mathbb{A}^n & \xrightarrow{\Phi} & \mathbb{A}^n \\ \downarrow{\scriptstyle \pi} & & \downarrow{\scriptstyle \pi} \\ P_n & \xrightarrow{\overline{\Phi}} & P_n. \end{array}$$

Such an S_n-equivariant morphism Φ is classically called a *covariant*. More generally, we have the following definition.

Definition 1. Let V, W be finite-dimensional K-representations of a finite group G. Then a G-equivariant K-morphism $\Phi : V \to W$ is called a *covariant of V of type W*. Moreover, Φ is said to be *faithful* if G acts faithfully on the image $\Phi(V_{\overline{K}})$ where $\overline{K}$ is the algebraic closure of K.

Clearly, covariants $\Phi, \Psi : \mathbb{A}^n \to \mathbb{A}^n$ can be added and multiplied by invariants p (i.e. by symmetric polynomials):

$$\Phi + \Psi := (\varphi_1 + \psi_1, \varphi_2 + \psi_2, \ldots, \varphi_n + \psi_n),$$
$$p\Phi := (p\varphi_1, p\varphi_2, \ldots, p\varphi_n).$$

Thus the covariants form a module over the ring of invariants. Moreover, we can form the "transvection" $(\Phi, \Psi) := (\varphi_1\psi_1, \varphi_2\psi_2, \ldots, \varphi_n\psi_n)$ of two covariants Φ and Ψ. It is obtained by composing the product $\Phi \times \Psi$ with the bilinear covariant $\mathbb{A}^n \times \mathbb{A}^n \to \mathbb{A}^n$, $(x_1, \ldots, x_n, y_1, \ldots, y_n) \mapsto (x_1 y_1, \ldots, x_n y_n)$.

All this also makes sense if Φ and Ψ are covariants of type $\mathbb{A}^n_{\text{sign}}$ where $\mathbb{A}^n_{\text{sign}}$ denotes the standard representation of S_n multiplied with the sign character. E.g., if $\Psi : \mathbb{A}^n \to \mathbb{A}^n_{\text{sign}}$ is a covariant (of type $\mathbb{A}^n_{\text{sign}}$) and $\Delta := \prod_{i<j}(x_i - x_j)$ then $\Delta\Psi$ is a covariant of type $\mathbb{A}^n$.

Remark 1. It is easy to see that if a covariant $\Phi = (\varphi_1, \dots, \varphi_n) : \mathbb{A}^n \to \mathbb{A}^n$ is not faithful then $\varphi_1 = \varphi_2 = \cdots = \varphi_n$, and this polynomial is an S_n-invariant. For example, if $\varphi_1 + \varphi_2 + \cdots + \varphi_n = 0$ and $\operatorname{char} K$ does not divide n, then Φ is faithful if $\Phi \neq 0$.

Assume now that the covariant $\Phi : \mathbb{A}^n \to \mathbb{A}^n$ is defined over K. Let $f \in P_n(K)$ and denote by $\xi = (\xi_1, \xi_2, \dots, \xi_n)$ the roots of f in some field extension L/K. Then $\Phi(\xi) =: (\bar{\xi}_1, \bar{\xi}_2, \dots, \bar{\xi}_n)$ are the roots of the transformed polynomial $\bar{f} := \overline{\Phi}(f)$:

$$f = \prod_{i=1}^{n} (X - \xi_i) \mapsto \bar{f} = \prod_{i=1}^{n} (X - \bar{\xi}_i).$$

Lemma 1. *Let $\Phi = (\varphi_1, \varphi_2, \dots, \varphi_n) : \mathbb{A}^n \to \mathbb{A}^n$ be a covariant defined over K.*

(1) *φ_1 is invariant under $S_{n-1} \subset S_n$ acting on the last $n-1$ variables, and $\varphi_k = (1k)\varphi_1$ where $(1k)$ denotes the transposition of 1 and k. Conversely, every S_{n-1}-invariant polynomial φ_1 defines a unique covariant Φ whose first component is φ_1.*
(2) *There is a uniquely defined polynomial $\varphi = \varphi(f, X) = \varphi(a_1, \dots, a_n, X) \in K[P_n][X]$ of degree $< n$ in X such that*

$$\varphi_i(x_1, x_2, \dots, x_n) = \varphi(a_1, a_2, \dots, a_n, x_i), \quad i = 1, 2, \dots, n,$$

where $a_k := (-1)^k s_k(x_1, \dots, x_n)$. Conversely, every such polynomial φ defines a covariant $\Phi : \mathbb{A}^n \to \mathbb{A}^n$.

Proof. Part (1) is clear, since $S_{n-1} \subset S_n$ is the stabilizer of $(1, 0, \dots, 0) \in \mathbb{A}^n$. The assertion (2) follows from (1) because

$$K[x_1, x_2, \dots, x_n]^{S_{n-1}} = \bigoplus_{j=0}^{n-1} K[x_1, x_2, \dots, x_n]^{S_n} x_1^j.$$

This is well known and even holds over $\mathbb{Z}$. □

A way to express this result is by saying that the covariants $\mathbb{A}^n \to \mathbb{A}^n$ form a *free module of rank n* over the invariants, with a basis given by the covariants

$$(x_1, \dots, x_n) \mapsto \left(x_1^j, \dots, x_n^j\right), \quad j = 0, 1, \dots, n-1.$$

For our purposes the following interpretation of Lemma 1 will be important. If the polynomial

$$f = x^n + a_1 x^{n-1} + a_2 x^{n-2} + \cdots + a_n$$

with coefficients $a_k \in K$ has a root ξ in some field extension L/K then $\bar{\xi} := \varphi(f, \xi)$ belongs to L, and $\bar{\xi}$ is a root of the transformed polynomial $\bar{f} := \overline{\Phi}(f)$. Following the classics we therefore make the following definition.

Definition 2. The polynomial $\varphi(f, X) \in K[P_n][X]$ from Lemma 1(2) will be called the TSCHIRNHAUS *transformation* associated to the covariant Φ.

Thus, by Lemma 1, the covariants $\Phi : \mathbb{A}^n \to \mathbb{A}^n$ correspond bijectively to TSCHIRNHAUS transformations $\varphi = \varphi(a_1, \ldots, a_n, X)$ by

$$\Phi(x_1, \ldots, x_n) = \big(\varphi(a_1, \ldots, a_n, x_1), \ldots, \varphi(a_1, \ldots, a_n, x_n)\big)$$

where $a_k := (-1)^k s_k(x_1, \ldots, x_n)$. So the general problem can be formulated as follows.

Problem. Given a field extension L/K of degree n and a generator $\xi \in L$ with equation $f(\xi) = 0$, find a covariant $\Phi : \mathbb{A}^n \to \mathbb{A}^n$ defined over K such that $\bar{f} := \overline{\Phi}(f)$ is as simple as possible.

Of course, the transformed equation $\bar{f}$ has to be irreducible so that $\bar{\xi} = \varphi(f, \xi)$ is again a generator of the extension L/K.

Our main goal is to prove the following two theorems which are due to HERMITE and JOUBERT. They imply the Main Theorem from the introduction about the special form of equations of degree 5 and 6. This will be shown in the following Section 3 for infinite fields K and in Section 6 for finite fields K. The proofs of the following two theorems will be given in Sections 4 and 5.

Theorem A *(*HERMITE*). There is a homogeneous S_5-covariant $\Phi : \mathbb{A}^5 \to \mathbb{A}^5$ of degree* 25*, defined over $\mathbb{Z}$ and faithful for every field K, with the property that*

$$s_1(\varphi_1, \varphi_2, \ldots, \varphi_5) = s_3(\varphi_1, \varphi_2, \ldots, \varphi_5) = 0.$$

The covariant Φ has the form $\Phi = (\Psi, \Delta\Omega)$ where $\Psi : \mathbb{A}^5 \to \mathbb{A}^5$ is homogeneous of degree 9*, $\Omega : \mathbb{A}^5 \to \mathbb{A}^5_{\text{sign}}$ homogeneous of degree* 6 *and $\Delta = \prod_{i<j}(x_i - x_j)$.*

Remark 2. We used the computer program Mathematica to show that, for every prime p, we have $s_4(\varphi_1, \ldots, \varphi_5) \not\equiv 0 \bmod p$.

For the case of degree 6 we need a slight modification. It is well known that the group S_6 has an outer automorphism τ which is unique up to inner automorphisms. We will denote by $\mathbb{A}^6_\tau$ the standard representation of S_6 twisted with τ, i.e. $\sigma \cdot_\tau x := \tau(\sigma)x$. Clearly, $\mathbb{A}^6_\tau$ has the same invariants and the same algebraic quotient $\pi : \mathbb{A}^6_\tau \to P_6$ as $\mathbb{A}^6$. Therefore, every covariant $\Phi : \mathbb{A}^6 \to \mathbb{A}^6_\tau$ gives rise to a commutative diagram

$$\begin{array}{ccc} \mathbb{A}^6 & \xrightarrow{\Phi} & \mathbb{A}^6_\tau \\ \downarrow{\scriptstyle \pi} & & \downarrow{\scriptstyle \pi} \\ P_6 & \xrightarrow{\overline{\Phi}} & P_6. \end{array}$$

It is easy to see that the results obtained so far carry over to this case with only minor modifications in the formulation, e.g. in Lemma 1(1) the component φ_1 is invariant under $\tau(S_5) \subset S_6$, and $\varphi_k = (1k) \cdot_\tau \varphi_1 = \tau((1k))\varphi_1$.

Theorem B *(*JOUBERT*). There is a homogeneous S_6-covariant $\Phi : \mathbb{A}^6 \to \mathbb{A}^6_\tau$ of degree* 18*, defined over $\mathbb{Z}$ and faithful for every field K of characteristic $\neq 2$, with the property that*

$$s_1(\varphi_1, \ldots, \varphi_6) = s_3(\varphi_1, \ldots, \varphi_6) = 0 \quad \text{and} \quad s_5(\varphi_1, \ldots, \varphi_6) = -2^5 \Delta^6.$$

The covariant Φ has the form $\Phi = \Delta\Psi$ where $\Psi : \mathbb{A}^6 \to (\mathbb{A}^6_\tau)_{\text{sign}}$ is of degree 3 *and $\Delta = \prod_{i<j}(x_i - x_j)$.*

Remark 3. It is interesting to remark that Ψ is the covariant of type $(\mathbb{A}^6_\tau)_{\text{sign}}$ of lowest possible degree (see Section 5 for more details). This observation will provide us with a conceptual proof of Theorem B (in characteristic zero), without any explicit calculations.

3. Equations of degree 5 and 6

We will now use the results of the previous section to deduce the Main Theorem about the special form of equations of degree 5 and 6 for infinite fields K. For the case of finite fields, we will need the explicit description of the covariants of HERMITE and JOUBERT; this will be done in Section 6.

Theorem 1. *Let K be an infinite field.*

(1) *If L/K is a separable field extension of degree* 5*, then there is a generator x of L/K with equation*

$$x^5 + bx^3 + cx + d = 0.$$

(2) *If L/K is a separable field extension of degree* 6 *and* $\operatorname{char} K \neq 2$*, then there is a generator x of L/K with equation*

$$x^6 + bx^4 + cx^2 + dx + e = 0.$$

Example 1 *(*REICHSTEIN [Rei99]*).* Let k be a field of characteristic $\neq 3$ containing a primitive third root of unity. Define

$$L := k(x_1, x_2, \ldots, x_r) \supset K := k\big(x_1^3, x_2^3, \ldots, x_r^3\big).$$

If $x \in L$, $x \neq 0$, then $\operatorname{Tr}_{L/K}(x^3) \neq 0$. In particular, there is no generator x of L/K whose equation has the form

$$x^n + a_2 x^{n-2} + a_4 x^{n-4} + a_5 x^{n-5} + \cdots + a_n$$

where $n := [L : K] = 3^r$. (In fact, if $x = \sum_{i_1,i_2,\dots,i_r} a_{i_1 i_2 \dots i_r} x_1^{i_1} x_2^{i_2} \cdots x_r^{i_r}$ then $\mathrm{Tr}_{L/K}\, x = 0$ if and only if $a_{i_1 i_2 \dots i_r} = 0$ whenever $i_1, i_2, \dots, i_r \in 3\mathbb{N}$. From this observation the claim follows immediately.) This example also shows that for $n = 3^r$ there is no faithful covariant $\Phi : \mathbb{A}^n \to \mathbb{A}^n$ such that $s_1(\varphi_1, \dots, \varphi_n) = s_3(\varphi_1, \dots, \varphi_n) = 0$.

In view of Theorems A and B from the previous section, the theorem above is an immediate consequence of the next proposition. In fact, applying the TSCHIRNHAUS transformation of HERMITE, respectively JOUBERT, to a "general" element of L/K we get a new generator of L/K whose equation has the form

$$x^5 + bx^3 + cx + d = 0, \quad \text{respectively } x^6 + bx^4 + cx^2 + dx + e = 0.$$

In order to reduce further to the form of the equations claimed in the Main Theorem of the introduction (i.e. $d = c$, respectively $e = d$) we have to show that $c \neq 0$ respectively $d \neq 0$. This will be done in Section 6.

Proposition 1. *Let K be an infinite field and L/K a separable field extension of degree n. If $\Phi : \mathbb{A}^n \to \mathbb{A}^n$ is a faithful covariant defined over K and φ the corresponding* TSCHIRNHAUS *transformation, then there is a generator ξ of L/K such that $\bar{\xi} := \varphi(\xi)$ also generates L over K, i.e. if $f(x) = 0$ is the equation of ξ then the transformed equation $\bar{f} := \overline{\Phi}(f)$ is again irreducible. Moreover, if $s \in K[x_1, \dots, x_n]$ is a non-zero polynomial function on $\mathbb{A}^n$ then $\xi \in L$ can be chosen in such a way that $s(\xi_1, \xi_2, \dots, \xi_n) \neq 0$ where $(\xi_1, \dots, \xi_n)$ are the conjugates of ξ.*

For the proof, we need the next two lemmas. (In the first one, the field K is arbitrary.)

Lemma 2. *Let $\Phi : \mathbb{A}^n \to \mathbb{A}^n$ be a covariant defined over an arbitrary field K and $f \in P_n(K)$ an irreducible separable polynomial with splitting field L/K and Galois group $G = \mathrm{Gal}(L/K)$. Then the image $\bar{f} = \overline{\Phi}(f)$ has the form $\bar{f} = h^m$ with an irreducible polynomial $h \in P_k(K)$ of degree $k = n/m$ whose splitting field $\bar{L} \subset L$ is G-invariant. In particular, $\mathrm{Gal}(\bar{L}/K) \simeq G/N$ where $N := \mathrm{Gal}(L/\bar{L})$.*

Proof. This is an exercise in Galois theory. If $f = \prod_{i=1}^n (X - x_i)$ and $\Phi(x_1, \dots, x_n) = (\bar{x}_1, \dots, \bar{x}_n)$, i.e. $\bar{x}_i = \varphi(f, x_i)$, then $\bar{f} = \prod_{i=1}^n (X - \bar{x}_i)$. Since Φ is defined over K we see that $\Phi : L^n \to L^n$ is G-equivariant. Hence, the splitting field $\bar{L} = K[\bar{x}_1, \dots, \bar{x}_n]$ of $\bar{f}$ is G-invariant and G acts transitively on the set $\Lambda := \{\bar{x}_1, \dots, \bar{x}_n\}$. If $[K[\bar{x}_1] : K] = k$ then the stabilizer of $\bar{x}_1$ has index k in G and the orbit Λ consists of k elements, say $\Lambda = \{\bar{x}_1, \dots, \bar{x}_k\}$. It follows that $h := \prod_{i=1}^k (X - \bar{x}_i) \in K[X]$ is irreducible and $\bar{f} = h^m$ where $m = n/k$. □

Example 2. If $f \in P_n(K)$ has Galois group S_n then either $\bar{f}$ is irreducible with Galois group S_n or $\bar{f} = (X - a)^n$ with $a \in K$. In fact, if x is a root of f and L the splitting field then $\mathrm{Gal}(L/K[x]) \simeq S_{n-1}$ which is a maximal subgroup of S_n. Thus every element $y \in K[x] \setminus K$ generates $K[x]/K$.

Lemma 3. *Let K be an infinite field and L/K a finite separable field extension of degree n. Then the subset*

$$\mathrm{Irr}_{L/K} := \{ f \in P_n(K) \mid f \text{ irreducible and } f(x) = 0 \text{ for some } x \in L \}$$

is Zariski-dense in $P_n(\overline{K})$ where $\overline{K}$ denotes the algebraic closure of K.

Proof. Let $L = K(x)$ with equation $f(x) = 0$. A linear combination $y = \sum_{i=0}^{n-1} a_i x^i$ ($a_i \in K$) is a generator for L/K if and only if the powers $(1, y, y^2, \ldots, y^{n-1})$ are linearly independent over K. It follows that the corresponding subset

$$A := \left\{ (a_0, a_1, \ldots, a_{n-1}) \in K^n \;\middle|\; \sum_{i=0}^{n-1} a_i x^i \text{ generates } L/K \right\}$$

is Zariski-dense in $\overline{K}^n$. As a consequence, if $x_1 := x, x_2, \ldots, x_n$ are the roots of f in some splitting field $L' \supset L$ of f, then

$$B := \left\{ \left(\sum_{i=0}^{n-1} a_i x_1^i, \sum_{i=0}^{n-1} a_i x_2^i, \ldots, \sum_{i=0}^{n-1} a_i x_n^i \right) \;\middle|\; (a_0, a_1, \ldots) \in A \right\} \subset L'^n$$

is Zariski-dense in $\overline{K}^n$. Hence its image $\pi(B) \subset P_n(\overline{K})$ is Zariski-dense, too. By construction, $\pi(B)$ is the set considered in the lemma. □

Now we can prove Proposition 1.

Proof of Proposition 1. By Lemma 3 the set $\pi^{-1}(\mathrm{Irr}_{L/K})$ is Zariski-dense in $\overline{K}^n$. Therefore, the subset

$$I := \{ \xi = (\xi_1, \ldots, \xi_n) \in \pi^{-1}(\mathrm{Irr}_{L/K}) \mid \bar{\xi} := \Phi(\xi) \text{ has trivial stabilizer in } S_n \}$$

is Zariski-dense, too, because Φ is faithful. This means that the $\bar{\xi}_i$'s are all different and so, by Lemma 2, the image $\bar{f} := \overline{\Phi}(f) = \prod_i (X - \bar{\xi}_i)$ is irreducible. Also, since I is Zariski-dense, the function s does not vanish on I. □

4. Proof of Theorem A

For any covariant $\Phi = (\varphi_1, \ldots, \varphi_n) : \mathbb{A}^n \to \mathbb{A}^n$ the functions $s_k(\varphi_1, \ldots, \varphi_n)$ are symmetric, hence can be regarded as functions on $P_n(K)$, as we have already seen above. Denote by W_n the vector space of binary forms of degree n:

$$W_n(K) := K[X, Y]_n = \left\{ f = \sum_{i=0}^{n} a_i X^{n-i} Y^i \;\middle|\; a_i \in K \right\}.$$

We will identify $P_n(K)$ with the binary forms f with leading coefficient $a_0 = 1$ by setting $Y = 1$. Then every polynomial $q = q(a_1, \dots, a_n)$ on $P_n(K)$ of degree d defines, by homogenizing, a homogeneous polynomial $\tilde{q}(a_0, a_1, \dots, a_n) := a_0^d q(a_1/a_0, a_2/a_0, \dots, a_n/a_0)$ on $W_n(K)$ of the same degree. Thus, from every covariant $\Phi : \mathbb{A}^n \to \mathbb{A}^n$ we obtain n homogeneous functions $\tilde{s}_k \in K[W_n]$ defined by

$$\tilde{s}_k(a_0, a_1, \dots, a_n) := a_0^{d_k} s_k(\varphi_1, \varphi_2, \dots, \varphi_n), \quad k = 1, \dots n,$$

where d_k is the degree of $s_k(\varphi_1, \dots, \varphi_n)$ considered as a function on P_n which is the same as the degree of s_k in each variable x_j. Now recall that the group $\mathrm{SL}_2(K)$ acts linearly on $W_n(K)$ by considering a binary form f as a function on the standard representation K^2 of $\mathrm{SL}_2(K)$: $gf(v) := f(g^{-1}v)$ for $g \in \mathrm{SL}_2(K)$ and $v \in K^2$.

The basic idea of HERMITE is to arrange the covariant Φ in such a way that the homogeneous polynomials $\tilde{s}_k$ become SL_2-*invariant* functions on W_n and then to use our knowledge about SL_2-invariants and, in particular, the fact that there are no SL_2-invariants in certain degrees.

In order to achieve this we will use the following classical result (see [Sch68, II.4 Satz 2.10]).

Proposition 2. *Assume that K is algebraically closed of characteristic 0. Let $q \in K[x_1, \dots, x_n]$ be a symmetric polynomial which is of degree d in each variable and let $\tilde{q} = \tilde{q}(a_0, \dots, a_n) \in K[W_n]$ be the corresponding homogeneous polynomial of degree d. Then $\tilde{q}$ is an SL_2-invariant if and only if the following two conditions hold*:

(T) $q(x_1 + t, x_2 + t, \dots, x_n + t) = q(x_1, x_2, \dots, x_n)$ *for $t \in K$, i.e., q only depends on the differences $x_i - x_j$;*

(R) $n \cdot d$ *is even and* $(x_1 x_2 \cdots x_n)^d q(1/x_1, 1/x_2, \dots, 1/x_n) = (-1)^{nd/2} q(x_1, x_2, \dots, x_n)$.

It then follows that q is homogeneous of degree $nd/2$.

Outline of proof. The group $\mathrm{SL}_2(K)$ is generated by the matrices $\left[\begin{smallmatrix} 1 & t \\ & 1 \end{smallmatrix}\right]$ ($t \in K$) and $\left[\begin{smallmatrix} & i \\ i & \end{smallmatrix}\right]$ ($i := \sqrt{-1}$). Therefore, a (homogeneous) function $\tilde{q} \in K[W_n]$ is $\mathrm{SL}_2(K)$-invariant if and only if $\tilde{q}(f)$ does not change under the following substitutions:

$$f(X, Y) \mapsto f(X - t, Y) \quad \text{and} \quad f(X, Y) \mapsto f(iY, iX).$$

Write $f(X, Y) = a_0 \prod_{i=1}^n (X - x_i Y)$ so that $\tilde{q}(f) = q(x_1, x_2, \dots, x_n)$. Since

$$f(X - t, Y) = a_0 \prod_{i=1}^n \big((X - t) - x_i Y\big) = a_0 \prod_{i=1}^n \big(X - (x_i + t)Y\big)$$

the invariance under $f(X, Y) \mapsto f(X - t, Y)$ is equivalent to (T). Since

$$f(iY, iX) = a_0 \prod_{k=1}^n (iY - x_k iX) = a_0 (-i)^n x_1 x_2 \cdots x_n \prod_{k=1}^n \left(X - \frac{1}{x_k} Y\right)$$

the invariance under $f(X,Y) \mapsto f(iY, iX)$ is equivalent to the condition

$$q(x_1, x_2, \ldots, x_n) = (-i)^{nd}(x_1 x_2 \cdots x_n)^d \cdot q\left(\frac{1}{x_1}, \frac{1}{x_2}, \ldots, \frac{1}{x_n}\right)$$

which is (R). Moreover, $\tilde{q}$ is also invariant under $f(X,Y) \mapsto f(tX, t^{-1}Y)$ for $t \in K^*$ which implies that q is homogeneous of degree $nd/2$. □

Later on we will use the following easy fact: If two arbitrary homogeneous polynomials q_1, q_2 satisfy one of the conditions (T) or (R) then the same holds for the product $q_1 q_2$.

Example 3. Start with $\Delta := \prod_{i<j}(x_i - x_j)$. It is easy to see that the symmetric polynomial Δ^2 satisfies the conditions (T) and (R) with $d := 2(n-1)$. The corresponding homogeneous invariant of degree d is the *discriminant* D of a binary form of degree n. The polynomial Δ itself satisfies the conditions (T) and (R), but is skew-symmetric, i.e. $\sigma\Delta = \operatorname{sign}(\sigma) \cdot \Delta$ for $\sigma \in S_n$.

Now we are ready to prove Theorem A from Section 3. In his note [Her61] HERMITE considers the following polynomial in $\mathbb{Z}[x_1, x_2, \ldots, x_5]$:

$$\begin{aligned}\psi_1 := &\big[(x_1 - x_2)(x_1 - x_5)(x_4 - x_3) + (x_1 - x_3)(x_1 - x_4)(x_2 - x_5)\big] \\ &\cdot \big[(x_1 - x_2)(x_1 - x_3)(x_5 - x_4) + (x_1 - x_4)(x_1 - x_5)(x_2 - x_3)\big] \\ &\cdot \big[(x_1 - x_2)(x_1 - x_4)(x_5 - x_3) + (x_1 - x_3)(x_1 - x_5)(x_4 - x_2)\big]. \qquad (2)\end{aligned}$$

One easily checks that ψ_1 is symmetric in $x_2, x_3, \ldots, x_5$, hence, by Lemma 1, defines a covariant

$$\Psi = (\psi_1, \psi_2, \ldots, \psi_5) : \mathbb{A}^5 \to \mathbb{A}^5$$

of degree 9, defined over $\mathbb{Z}$, where $\psi_k := (1k)\psi_1$. The functions ψ_i obviously satisfy the condition (T) of Proposition 2, and one finds

$$x_1^3 (x_1 x_2 \cdots x_5)^3 \psi_1\left(\frac{1}{x_1}, \frac{1}{x_2}, \ldots, \frac{1}{x_5}\right) = -\psi_1(x_1, \ldots, x_5).$$

Therefore, the polynomial

$$\varphi_1 := \psi_1 \cdot \prod_{1<i<j} (x_i - x_j) \cdot \Delta \qquad (3)$$

has the properties (T) and (R) with $d = \deg_{x_i} \varphi_1 = 10$, and φ_1 is symmetric in $x_2, \ldots, x_5$. By construction, the corresponding covariant

$$\Phi = (\varphi_1, \varphi_2, \ldots, \varphi_5) : \mathbb{A}^5 \to \mathbb{A}^5$$

is again defined over $\mathbb{Z}$ and has degree $25 = 9 + 6 + 10$. We claim that Φ satisfies the properties of Theorem A.

In fact, it follows from Proposition 2, choosing for K the algebraic closure of $\mathbb{Q}$, that the homogeneous polynomials $\tilde{s}_k \in K[W_5]$ corresponding to the symmetric functions $s_k(\varphi_1, \ldots, \varphi_5)$ are SL_2-invariants of W_5 of degree $10k$. Since a symmetric polynomial which is divisible by Δ is automatically divisible by Δ^2 we see that $\tilde{s}_1$ is divisible by the discriminant D and $\tilde{s}_3$ by D^2, and we get $\deg \tilde{s}_1/D = 10 - 8 = 2$ and $\deg \tilde{s}_3/D^2 = 3 \cdot 10 - 2 \cdot 8 = 14$. On the other hand, the SL_2-invariants of W_5 are generated by invariants I_4, I_8, I_{12}, and I_{18} of degree 4, 8, 12, and 18 (see [Sch68, II.9 Satz 2.26]). Hence, there are no invariants in degree 2 and 14 and so $\tilde{s}_1 = \tilde{s}_3 = 0$ which proves Theorem A.

Finally, the covariant Φ is faithful. First of all, $\Phi \not\equiv 0$ modulo p for all primes p. In fact, one easily sees that the leading term of the polynomial ψ_1 is $-x_1^6 x_2^3$ and so the leading term of φ_1 has coefficient ± 1. Now the faithfulness follows from Remark 1 for $\operatorname{char} K \neq 5$. If $\operatorname{char} K \neq 2$ and Φ were not faithful for K then φ_1 is an invariant and so $\varphi_1 \cdot \Delta^{-1} = \psi_1 \cdot \prod_{1<i<j}(x_i - x_j)$ a semi-invariant, hence divisible by Δ. This is not possible since ψ_1 does not vanish for $x_1 = x_2$. □

Remark 4. By construction, the covariant Φ has the form $\Phi = \Delta(\Psi, \Omega) = (\Psi, \Delta\Omega)$ where

$$\Omega = (\omega_1, \ldots, \omega_5) : \mathbb{A}^5 \to \mathbb{A}^5_{\mathrm{sign}}$$

is the homogeneous covariant of degree 6 defined by $\omega_1 := \prod_{1<i<j}(x_i - x_j)$ (and $\omega_k := -(1k)\omega_1$ for $k \geqslant 2$). The representation $\mathbb{A}^5_{\mathrm{sign}}$ of S_5 contains the subrepresentation

$$U := \{x = (x_1, \ldots, x_5) \in \mathbb{A}^5 \mid x_1 + \cdots + x_5 = 0\},$$

and the image of the covariant $\Omega : \mathbb{A}^5 \to \mathbb{A}^5_{\mathrm{sign}}$ is contained in U. (The last statement is clear since $\omega_1 + \cdots + \omega_5$ is skew symmetric of degree 9, hence equal to 0, because every skew symmetric polynomial is divisible by Δ.)

It is interesting to note that Ω is the covariant of type U of smallest possible degree because the representation U occurs in $K[\mathbb{A}^5]$ for the first time in degree 6. (In fact, U is the irreducible representation corresponding to the partition $(2, 1, 1, 1)$, and $K[\mathbb{A}^5]_6$ contains the induced representation $\operatorname{Ind}_{S_2}^{S_5} K$ because the stabilizer of $x_1 x_2^2 x_3^3 \in K[\mathbb{A}^5]_6$ is S_2.)

Remark 5. Using a computer program like SINGULAR [GPS01] or MATHEMATICA, one can check directly that $s_1(\varphi_1, \ldots, \varphi_5) = s_3(\varphi_1, \ldots, \varphi_5) = 0$. So the ingenious part of HERMITE's short note is the discovery of the functions ψ_i above. In fact, his remark is the following, see [Her61]. He was trying to write out the invariant of degree 18 of the binary forms of degree 5 in terms of the roots $x_1, \ldots, x_5$. Thus, he was looking for a polynomial expression ψ in the differences $(x_i - x_j)$ which satisfies the conditions (T) and (R) of Proposition 2 where $d = 18$. He discovered that $\psi := \psi_1\psi_2\psi_3\psi_4\psi_5$ has this property, i.e. that ψ can be written as a product of 5 terms where each one is invariant with respect to one of the standard subgroups $S_4 \subset S_5$. And, of course, he immediately realized that this can be used to transform and simplify equations of degree 5.

5. Proof of Theorem B

We will give two proofs for Theorem B. The first one is more conceptual, but only works in characteristic zero. The second follows the explicit calculations given by JOUBERT and is valid in all characteristics $\neq 2$.

First Proof. Here the base field is $\mathbb{Q}$. If $\lambda = (\lambda_1, \lambda_2, \ldots)$ is a partition of 6, we will denote by V_λ the irreducible representation of S_6 associated to λ (see [FuH91, §4.1]). So $V_{(6)}$ is the trivial representation and $V_{(1,1,\ldots,1)}$ is the sign representation. It is not hard to see that twisting $V_{(5,1)}$ with the outer automorphism τ we obtain the representation $V_{(2,2,2)}$ which is isomorphic to $V_{(3,3)} \otimes \text{sign}$.

Let V denote the standard representation of S_6, i.e. $V \simeq V_{(6)} \oplus V_{(5,1)}$. Then, as we just said, $V_\tau \supset V_{(2,2,2)}$. One easily sees that the third symmetric power S^3V contains the representation $V_{(3,3)}$. In fact, all symmetric powers S^iV are permutation representation. Since the stabilizer of $e_1e_2e_3 \in S^3V$ is $S_3 \times S_3$, we see that S^3V contains the induced representation $\text{Ind}_{S_3\times S_3}^{S_6} \mathbb{Q}$ which contains $V_{(3,3)}$.

It follows that

$$V_{(3,3)} \simeq V_{(2,2,2)} \otimes \text{sign} \subset V_\tau \otimes \text{sign},$$

which implies that there is a non-trivial covariant $\Psi : V \to V_\tau \otimes \text{sign}$ of degree 3. Multiplying Ψ with Δ, we finally get a covariant

$$\Phi := \Delta\Psi : \mathbb{A}^6 \to \mathbb{A}^6_\tau$$

of degree $3 + \deg \Delta = 18$. We claim that Φ satisfies the properties of Theorem B. In fact, for every k the function $s_k(\varphi_1, \ldots, \varphi_6) = s_k(\psi_1, \ldots, \psi_6)\Delta^k$ is symmetric and so $s_{2k+1}(\psi_1, \ldots, \psi_6)$ is skew-symmetric of degree $6k + 3$, hence is divisible by Δ. Since $\deg \Delta = 15$, we get $s_1 = s_3 = 0$. To see that Φ is faithful, we simply remark that otherwise $\Phi = 0$ because $s_1 = \varphi_1 + \cdots + \varphi_6 = 0$ (see Remark 1). □

Second Proof. This proof needs some preparation. We will consider the elements of the symmetric group S_6 as permutations of the projective line

$$\mathbb{PF}_5 = \mathbb{F}_5 \cup \{\infty\} = \{\infty, 0, 1, 2, 3, 4\},$$

so that $H := \text{PGL}_2(\mathbb{F}_5)$ becomes a subgroup of S_6 isomorphic to S_5. This subgroup is the image of the standard $S_5 \subset S_6$ under an outer automorphism τ. Let S_6 act on the set of subsets of $\mathbb{PF}_5$ consisting of 2 elements, and define $N \subset S_6$ to be the normalizer of the subset

$$\mathcal{M} := \big\{\{\infty, 0\}, \{1, 4\}, \{2, 3\}\big\}.$$

We obtain a surjective homomorphism $\rho : N \to S_3$ with kernel isomorphic to $(\mathbb{Z}_2)^3$ generated by the transpositions $(\infty\,0)$, $(1\,4)$, $(2\,3)$. The following result is known and easy to prove.

Lemma 4. *Set $N_0 := N \cap H$, the normalizer of $\mathcal{M}$ in H, and $\eta := \left[\begin{smallmatrix} 1 & 1 \\ & 1 \end{smallmatrix}\right] \in H$.*

(1) $\rho(N_0) = S_3$ *and* $\ker \rho|_{N_0} \simeq (\mathbb{Z}_2)^2$.
(2) N_0 *is isomorphic to* S_4.
(3) $H = N_0 \cup \eta N_0 \cup \eta^2 N_0 \cup \eta^3 N_0 \cup \eta^4 N_0$.

Now we can prove Theorem B. For the polynomial functions on $\mathbb{A}^6$ we use the variables $x_\infty, x_0, x_1, x_2, x_3, x_4$. Define, as in JOUBERT's paper,

$$\begin{aligned} h &:= (x_\infty - x_4)(x_1 - x_3)(x_2 - x_0) + (x_0 - x_1)(x_4 - x_2)(x_3 - x_\infty) \\ &= x_\infty x_0(x_2 + x_3 - x_1 - x_4) + x_1 x_4(x_\infty + x_0 - x_2 - x_3) + x_2 x_3(x_1 + x_4 - x_\infty - x_0). \end{aligned}$$

It is easy to see that h is semi-invariant with respect to the subgroup N defined above. For $\sigma \in N_0 \simeq S_4$, we have $\sigma h = \operatorname{sign}(\rho(\sigma)) \cdot h = \operatorname{sign}_{N_0}(\sigma) \cdot h$. Therefore, by Lemma 4(3), the function

$$h + \eta(h) + \eta^2(h) + \eta^3(h) + \eta^4(h)$$

is semi-invariant with respect to H. We claim that the coefficients of this polynomial are all ± 3. In fact,

$$\begin{aligned} h = (x_1x_2x_3 + x_2x_3x_4 + x_4x_0x_1) - (x_1x_2x_4 + x_2x_3x_0 + x_3x_4xx_1) \\ + x_\infty(x_0x_2 + x_3x_0 + x_4x_1) - x_\infty(x_0x_1 + x_2x_3 + x_4x_0) \end{aligned}$$

and each bracket expression is a sum of three monomials from a single orbit under the group generated by the cyclic permutation $\eta = (12340) \in H$. Denoting by o_{ijk} the sum of the monomials in the orbit of $x_i x_j x_k$ under the group $\langle \eta \rangle \subset H$, e.g.

$$o_{123} = x_1x_2x_3 + x_2x_3x_4 + x_3x_4x_0 + x_4x_0x_1 + x_0x_1x_2,$$

we see that

$$h + \eta(h) + \eta^2(h) + \eta^3(h) + \eta^4(h) = 3(o_{123} - o_{124} + x_\infty o_{02} - x_\infty o_{01}).$$

Thus

$$\psi_1 := \frac{1}{3}\big(h + \eta(h) + \eta^2(h) + \eta^3(h) + \eta^4(h)\big)$$

is the sum of all square-free monomials $x_i x_j x_k$ $(i \neq j \neq k \neq i)$ with coefficients ± 1. Since ψ_1 is semi-invariant with respect to H, we see that $\varphi_1 := \Delta \cdot \psi_1$ is invariant with respect to H and defines, by Lemma 1, a homogeneous covariant

$$\Phi = (\varphi_1, \varphi_2, \ldots, \varphi_6) : \mathbb{A}^6 \to \mathbb{A}^6_\tau$$

of degree $3 + \deg \Delta = 18$, defined over $\mathbb{Z}$. The same degree argument as in the first proof shows that $s_1(\varphi_1, \ldots, \varphi_6) = s_3(\varphi_1, \ldots, \varphi_6) = 0$. Moreover, in characteristic $\neq 2$, the polynomial ψ_1 is not a semi-invariant with respect to the whole group S_6, hence φ_1 is not an S_6-invariant, and so Φ is faithful (see Remark 1). □

Remark 6. We have seen above that ψ_1 is the sum of all square-free monomials $x_i x_j x_k$ ($i \neq j \neq k \neq i$) with coefficients ± 1. Thus, for any field K of characteristic 2, we have $\psi_1 = s_3$. Hence neither Ψ nor Φ is faithful in characteristic 2. We do not know if the Main Theorem for extensions of degree 6 also holds in characteristic 2.

6. The case of finite fields

In this section we will show that the methods of HERMITE and JOUBERT also work for finite fields thus completing the proof of the Main Theorem. For extensions of degree 6 in characteristic $\neq 2$, this will follows from what we have done in Sections 3 and 5. Recall that JOUBERT's covariant

$$\Phi : \mathbb{A}^6 \to \mathbb{A}^6_\tau$$

has the form $\Phi = \Psi \cdot \Delta$ where $\Psi : \mathbb{A}^6 \to (\mathbb{A}^6_\tau)_{\text{sign}}$ is of degree 3. If follows that $s_5(\psi_1, \ldots, \psi_6)$ is a semi-invariant of degree 15, hence an integral multiple of Δ. We claim that

$$s_5(\psi_1, \ldots, \psi_6) = \pm 2^s \cdot \Delta \quad \text{for some } s \in \mathbb{N}.$$

In fact, if $s_5(\psi_1, \ldots, \psi_6) \equiv 0 \bmod p$ for a prime $p \neq 2$, then it follows from Theorem 1 of Section 3 that for any infinite field K of characteristic p and any extension L/K of degree 6 there is a generator ξ whose equation has the form $x^6 + a_2 x^4 + a_4 x^2 + a_6 = 0$. But this implies that L contains a subfield $L' := K(\xi^2)$ of degree 3 over K which clearly does not hold for generic extensions of degree 6.

Remark 7. An explicit calculation shows that

$$s_5(\psi_1, \ldots, \psi_6) = -2^5 \cdot \Delta.$$

The next proposition shows that the covariant of JOUBERT, applied to any separable irreducible polynomial of degree 6 over any field of characteristic $\neq 2$, always gives an irreducible polynomial. In particular, this proves the Main Theorem for extensions of degree 6.

Proposition 3. *Let K be any field of characteristic $\neq 2$ and let $f \in K[x]$ be an irreducible separable polynomial of degree* 6. *If $\Phi : \mathbb{A}^6 \to \mathbb{A}^6_\tau$ is the covariant constructed by* JOUBERT *then $\bar{f} = \overline{\Phi}(f)$ is irreducible. Moreover, the linear term of $\bar{f}$ has a non-zero coefficient.*

Proof. By Theorem B, we have $\bar{f} := \overline{\Phi}(f) = x^6 + bx^4 + cx^2 + dx + e$ If $\xi = (\xi_1, \xi_2, \ldots, \xi_6)$ are the (distinct) roots of f, then $\Delta(\xi) \neq 0$ and so

$$d = s_5\big(\varphi_1(\xi), \ldots, \varphi_6(\xi)\big) = \pm 2^s \Delta^6 \neq 0.$$

On the other hand, if $\bar{f}$ were reducible then, by Lemma 2, $\bar{f} = h^k$ where h is irreducible and $k = 2, 3$ or 6. The case $\bar{f} = h^2$ cannot occur since then h should have the form $x^3 + ax$. In the other two cases a short calculation shows that h is either $(x^2 - a)$ or x. But then the coefficient d of the linear term of $\bar{f}$ is zero. □

A similar result does not hold for the covariant Φ of HERMITE. In fact, if we start with an irreducible polynomial of the form $f(x) = x^5 - a$ then $\overline{\Phi}(f) = 0$. (This can verified by using the explicit form (3) of Φ given in Section 4; in fact, $\psi_1(1, \zeta, \zeta^2, \zeta^3, \zeta^4) = 0$ for any fifth root of unity ζ.) However, we have the following result.

Proposition 4. *Let $\Phi = (\varphi_1, \varphi_2, \ldots, \varphi_5) : \mathbb{A}^5 \to \mathbb{A}^5$ be the covariant of* HERMITE.

(1) *For every prime p the symmetric polynomial $s_4(\varphi_1, \ldots, \varphi_5) \in \mathbb{Z}[x_1, \ldots, x_5]$ is non-zero modulo p.*
(2) *$s_4(\varphi_1, \ldots, \varphi_5)$ is divisible by Δ^6 in $\mathbb{Z}[x_1, \ldots, x_5]$, and the quotient $S_4 := s_4(\varphi_1, \ldots, \varphi_5)/\Delta^6$ is homogeneous of degree* 40.
(3) *If L/K is a separable extension of degree* 5 *and $\xi \in L$ a generator with equation $f(x) = 0$ such that $S_4(\xi_1, \ldots, \xi_5) \neq 0$, then $\overline{\Phi}(f)$ is irreducible.* ($\xi_1, \ldots, \xi_5$ *are the conjugates of* ξ.)

Proof. Recall the definition of HERMITE's covariant $\Phi : \mathbb{A}^5 \to \mathbb{A}^5$:

$$\varphi_1 = \psi_1 \cdot \prod_{1<i<j} (x_i - x_j) \cdot \Delta$$

where

$$\begin{aligned}\psi_1 := {} & \big[(x_1 - x_2)(x_1 - x_5)(x_4 - x_3) + (x_1 - x_3)(x_1 - x_4)(x_2 - x_5)\big] \\ & \cdot \big[(x_1 - x_2)(x_1 - x_3)(x_5 - x_4) + (x_1 - x_4)(x_1 - x_5)(x_2 - x_3)\big] \\ & \cdot \big[(x_1 - x_2)(x_1 - x_4)(x_5 - x_3) + (x_1 - x_3)(x_1 - x_5)(x_4 - x_2)\big]\end{aligned}$$

and $\Delta = \prod_{i<j}(x_i - x_j)$. Put $\tilde{\psi}_1 := \psi_1 \cdot \prod_{1<i<j}(x_i - x_j)$. It is easy to see that the leading term of $\tilde{\psi}_1$ and $\tilde{\psi}_2$ is $\pm x_1^6 x_2^5 x_3^4$ and that the leading term of $\tilde{\psi}_4$ and $\tilde{\psi}_5$ is $\pm x_1^6 x_2^6 x_3^2 x_4$. Moreover, one finds that $\psi_3(t^4, t^3, t^2, t, 1) = 0$. This implies that

$$\begin{aligned}& s_4\big(\tilde{\psi}_1, \ldots, \tilde{\psi}_5\big)\big(t^4, t^3, t^2, t, 1\big) \\ & \quad = \tilde{\psi}_1\big(t^4, \ldots, t, 1\big) \cdot \tilde{\psi}_2\big(t^4, \ldots, t, 1\big) \cdot \tilde{\psi}_4\big(t^4, \ldots, t, 1\big) \cdot \tilde{\psi}_5\big(t^4, \ldots, t, 1\big)\end{aligned}$$

and the leading term of this product is $\pm t^{188}$. Thus $s_4(\tilde{\psi}_1, \ldots, \tilde{\psi}_5)$ and hence $s_4(\varphi_1, \ldots, \varphi_5) = s_4(\tilde{\psi}_1, \ldots, \tilde{\psi}_5)\Delta^4$ is non-zero modulo p for every prime p, proving (1).

We have $s_4(\varphi_1, \ldots, \varphi_5) = s_4(\tilde{\psi}_1, \ldots, \tilde{\psi}_5)\Delta^4$. In addition, $\tilde{\psi}_i\tilde{\psi}_j\tilde{\psi}_k$ is divisible by Δ for $i \neq j \neq k \neq i$, and so $s_4(\tilde{\psi}_1, \ldots, \tilde{\psi}_5)$ is divisible by Δ^2 since it is symmetric. Now (2) follows because $\deg \tilde{\psi}_i = 15$ and $\deg \Delta = 10$.

Finally, let $f \in K[x]$ be an irreducible separable polynomial of degree 5 with roots $\xi_1, \ldots, \xi_5 \in \overline{K}$. If $\bar{f} := \overline{\Phi}(f)$ is reducible then $\bar{f} = (x-a)^5$ by Lemma 2. If $\operatorname{char} K \neq 5$ then $a = 0$, because $5a = s_1(\varphi_1(\xi), \ldots, \varphi_5(\xi)) = 0$. For $\operatorname{char} K = 5$, we get $\bar{f} = x^5 - a^5$. In both cases we see that $S_4(\xi_1, \ldots, \xi_5) = 0$ which contradicts the assumption. Thus we get (3). □

A crucial step in the proof of the Main Theorem for infinite fields K was Proposition 1 which says that for a faithful covariant $\Phi : \mathbb{A}^n \to \mathbb{A}^n$ defined over K and a separable extension L/K we can always find a generator $\xi \in L$ such that $\varphi(\xi)$ is also a generator for L/K, or, equivalently, that $\overline{\Phi}(f)$ is irreducible where $f \in K[x]$ is the minimal polynomial of ξ. However, if K is finite it is not clear that such a $\xi \in L$ exists. One expects that this is the case if K is large enough. In fact, we have the following more precise result. (For our proof we will only need the second part.)

Proposition 5. *Let K be a finite field and L/K a separable extension of degree n. Let $\Phi : \mathbb{A}^n \to \mathbb{A}^n$ be a faithful homogeneous covariant defined over K and*

$$\varphi = p_0 + p_1 X + p_2 X^2 + \cdots + p_{n-1} X^{n-1}$$

the corresponding TSCHIRNHAUS *transformation (see Definition 2). If $\varphi(\xi_1, \ldots, \xi_n) \in K$ for all generators ξ of L/K where $\xi_1, \ldots, \xi_n$ are the conjugates of ξ, then*

$$|K| \leqslant \min\{\deg p_j \mid j > 0 \text{ and } p_j \neq 0\} < \deg \Phi.$$

Moreover, if $S \in K[x_1, \ldots, x_n]$ is a homogeneous symmetric polynomial such that $S(\xi_1, \ldots, \xi_n) = 0$ for all generators ξ of L/K then

$$|K| \leqslant \deg S.$$

Proof. If $\varphi(\xi) \in K$ for a generator ξ of L/K then $p_1(\xi) = p_2(\xi) = \cdots = p_{n-1}(\xi) = 0$ because $1, \xi, \xi^2, \ldots, \xi^{n-1}$ are linearly independent over K. Now fix a generator θ of L/K and consider the following linear change of coordinates

$$x_i = y_0 + y_1\theta_i + y_2\theta_i^2 + \cdots + y_{n-1}\theta_i^{n-1}, \quad i = 1, 2, \ldots, n,$$

where $\theta_1 := \theta, \theta_2, \ldots, \theta_n$ are the conjugates of θ. Each p_j and also S are transformed into homogeneous polynomials $\tilde{p}_j(y_0, \ldots, y_{n-1})$ and $\tilde{S}(y_0, \ldots, y_{n-1})$ of the same degree. In addition, $\tilde{p}_j$ and $\tilde{S}$ have their coefficients in K, because p_j and S do and are symmetric.

If $\xi = a_0 + a_1\theta + \cdots + a_{n-1}\theta^{n-1}$ is a generator of L/K then, by assumption, we have $\tilde{p}_j(a_0, \ldots, a_{n-1}) = 0$ for $j \geqslant 1$. Thus each $\tilde{p}_j$ vanishes on $K^n \setminus F$ where F is the finite

union of all subspaces corresponding to proper subfields $L' \subset L$ containing K. The following Lemma 5 shows that F is contained in a proper linear subspace of K^n and so Lemma 6 implies that $|K| \leqslant \deg p_j$ if $p_j \neq 0$, and also $|K| \leqslant \deg S$, hence the claim. □

Lemma 5. *Let L/K be an extension of finite fields of degree $n > 1$ and $p_1, p_2, \ldots, p_k$ be the prime factors of n. Then the sum of the proper subfields $M \subset L$ containing K has codimension*

$$\frac{n}{p_1 p_2 \cdots p_k}(p_1 - 1)(p_2 - 1)\cdots(p_k - 1) \geqslant 1.$$

The following proof was communicated to me by MIHAELA POPOVICIU and JAN DRAISMA.

Proof. For every divisor d of n, we denote by L_d the (unique) subfield of L with $[L : L_d] = d$. Then the span of the proper subfields of L containing K is given by

$$L_{p_1} + L_{p_2} + \cdots + L_{p_k}.$$

We can therefore assume that $K = L_{p_1 p_2 \cdots p_k}$, i.e. that n is square-free. We proceed by induction on the number k of prime factors of n, the case $n = p_1$ being trivial. By relabeling the p_i's, we can assume that p_k is not equal to the characteristic of K. Then we claim that

$$(L_{p_1} + \cdots + L_{p_{k-1}}) \cap L_{p_k} = L_{p_1} \cap L_{p_k} + \cdots + L_{p_{k-1}} \cap L_{p_k}. \tag{4}$$

The inclusion $\supseteq$ is clear. For the converse, suppose that $\alpha \in L_{p_k}$ can be written as

$$\alpha = \alpha_1 + \alpha_2 + \cdots + \alpha_{k-1} \quad \text{where } \alpha_i \in L_{p_i}.$$

Let $F : L \to L$ be the FROBENIUS operator of the extension L/L_{p_k} and put

$$H := \frac{1}{p_k}\bigl(\mathrm{Id} + F + F^2 + \cdots + F^{p_k - 1}\bigr).$$

The linear operator H is the projection onto the fixed points $L^F = L_{p_k}$ and stabilizes all L_{p_i}. Thus $\alpha = H(\alpha) = H(\alpha_1) + H(\alpha_2) + \cdots + H(\alpha_k)$ and $H(\alpha_i) \in L_{p_i} \cap L_{p_k}$ which proves our claim (4). Using this we get

$$\begin{aligned}
\operatorname{codim}_K(L_{p_1} + \cdots + L_{p_k}) &= \operatorname{codim}_K(L_{p_1} + \cdots + L_{p_{k-1}}) + \operatorname{codim} L_{p_k} \\
&\quad - \operatorname{codim}_K(L_{p_1} + \cdots + L_{p_{k-1}}) \cap L_{p_k} \\
&= \operatorname{codim}_K(L_{p_1} + \cdots + L_{p_{k-1}}) + \operatorname{codim}_K L_{p_k} \\
&\quad - \operatorname{codim}_K(L_{p_1} \cap L_{p_k} + \cdots + L_{p_{k-1}} \cap L_{p_k}).
\end{aligned}$$

Applying the induction hypothesis to the extensions $L/L_{p_1 p_2 \cdots p_{k-1}}$ and L_{p_k}/K we find

$$\operatorname{codim}_K(L_{p_1} + \cdots + L_{p_{k-1}}) = p_k(p_1 - 1)(p_2 - 1)\cdots(p_{k-1} - 1),$$
$$\operatorname{codim}_K(L_{p_1} \cap L_{p_k} + \cdots + L_{p_{k-1}} \cap L_{p_k}) = (p_1 - 1)\cdots(p_{k-1} - 1) + \operatorname{codim}_K L_{p_k},$$

hence

$$\operatorname{codim}_K(L_{p_1} + \cdots + L_{p_k}) = (p_k - 1)(p_1 - 1)(p_2 - 1)\cdots(p_{k-1} - 1). \qquad \square$$

Lemma 6. *Let K be a finite field and $f \in K[y_0, y_1, \ldots, y_m]$ a non-zero homogeneous polynomial. If f vanishes on $K^{m+1} \setminus W$ where W is a proper linear subspace then $|K| \leqslant \deg f$.*

Proof. By a linear change of coordinates we can assume that W is contained in the hyperplane given by $y_0 = 0$. Then the polynomial $\bar{f}(y_1, y_2, \ldots, y_m) := f(1, y_1, \ldots, y_m)$ is non-zero, has degree $\leqslant \deg f$ and vanishes on K^m. Now the claim follows by an easy induction on m, since a polynomial in one variable of degree d has at most d different roots. $\square$

Now we are ready to give a proof of the Main Theorem for extensions of degree 5. If K is infinite or $|K| \geqslant 40$ and L/K an extension of degree 5 then there is a generator ξ of L/K such that $s_4(\varphi_1(\xi), \ldots, \varphi_5(\xi)) \neq 0$, by Proposition 4(1) and (2) together with Proposition 5. It follows that the transformed equation $\bar{f}$ is irreducible (Proposition 4(3)) and has the form $x^5 + ax^3 + bx + c$ where $b \neq 0$ and the claim follows.

It remains to discuss the finite fields $K = \mathbb{F}_q$ where $q \leqslant 37$ and $q \neq 2$ and to show that in each case there is an irreducible polynomial of degree 5 of the required form. It clearly suffices to consider the fields $\mathbb{F}_q$ where $q = 2^2, 2^3, 2^5, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37$. In all these cases, there are the following irreducible polynomials of degree 5:

$$\begin{aligned}
&\mathbb{F}_{2^2}: \quad x^5 + ax + a \quad \text{where } a \in \mathbb{F}_{2^2} \setminus \mathbb{F}_2,\\
&\mathbb{F}_{2^3}: \quad x^5 + bx^3 + bx + b \quad \text{where } b^3 + b^2 + 1 = 0,\\
&\mathbb{F}_{2^5}: \quad x^5 + cx^3 + x + 1 \quad \text{where } c^5 + c^4 + c^3 + c^2 + 1 = 0,\\
&\mathbb{F}_3: \quad x^5 - x - 1,\\
&\mathbb{F}_5: \quad x^5 - x - 1,\\
&\mathbb{F}_7: \quad x^5 - 2x - 2,\\
&\mathbb{F}_{11}: \quad x^5 - x - 1,\\
&\mathbb{F}_{13}: \quad x^5 - x - 1,\\
&\mathbb{F}_{17}: \quad x^5 + 4x + 4,\\
&\mathbb{F}_{19}: \quad x^5 + 3x + 3,\\
&\mathbb{F}_{23}: \quad x^5 + 2x + 2,\\
&\mathbb{F}_{29}: \quad x^5 - 4x - 4,
\end{aligned}$$

$\mathbb{F}_{31}$: $x^5 + 3x + 3$,

$\mathbb{F}_{37}$: $x^5 - 3x - 3$.

This finishes the proof of the Main Theorem. □

7. Equations of degree 3 and 4

To complete the picture we want to describe the situation for equations of degree 3 and 4. First we have the following general result.

Lemma 7. *For $n > 2$ there is a faithful covariant $\Phi = (\varphi_1, \dots, \varphi_n) : \mathbb{A}^n \to \mathbb{A}^n_{\text{sign}}$ of degree $\binom{n-1}{2}$ such that $\varphi_1 + \cdots + \varphi_n = 0$.*

Proof. Define $\varphi_1 := \prod_{1<i<j\leqslant n}(x_i - x_j)$ and $\varphi_k := -(1k)\varphi_1$. Then $\Phi := (\varphi_1, \dots, \varphi_n)$ is a faithful covariant of type $\mathbb{A}^n_{\text{sign}}$. Since $\varphi_1 + \cdots + \varphi_n$ is skew symmetric of degree $< \deg \Delta$ the claim follows. □

Proposition 6. *Let L/K be a separable field extension of degree n where K is either infinite or* char K *is prime to n.*

(1) *If $[K : L] > 2$ there is a generator $x \in L$ with* $\operatorname{tr} x = 0$.
(2) *If $[L : K] = 3$ then there is a generator $x \in L$ which satisfies an equation of the form*

$$x^3 + ax + a = 0.$$

(3) *If $[L : K] = 4$ then there is a generator $x \in L$ which satisfies an equation of the form*

$$x^4 + ax^2 + bx + b = 0.$$

Proof. (1) This is well know if char K is prime to n. If K is infinite, then it follows from Lemma 7 together with Proposition 1.

(2) By (1) we can assume that there is a generator $x \in L$ with equation $x^3 + bx + c = 0$. If $b \neq 0$ the claim follows by replacing x by $\frac{b}{c}x$. If $b = 0$ (which can happen only if char $K \neq 3$) then $y := x + x^2$ satisfies the equation $y^3 + 3cy + c - c^2 = 0$ which reduces to the previous case.

(3) Again by (1) we can assume that there is a generator x of L/K such that $\operatorname{tr} x = 0$. Thus x satisfies an equation of the form $x^4 + bx^2 + cx + d = 0$. If $c \neq 0$ we are done as in (2). Otherwise, char $K \neq 2$ and the element $y := b/2 + x + x^2$ is again a generator of L/K. An easy calculation shows that the coefficient of the linear term of the equation of y is equal to $a^2 - 4d$ which is non-zero because $x^4 + ax^2 + d$ is irreducible, by assumption. □

Remark 8. Replacing in (2) the element x by $1/x$ we see that for a separable extension L/K of degree 3 there is always a generator x such that $x^3 + x^2 \in K$. This was mentioned to me by DAVID MASSER.

Acknowledgments

I thank Zinovy Reichstein and Jean-Pierre Serre for helpful discussions about the subject of this paper.

References

[BuR97] J. Buhler, Z. Reichstein, On the essential dimension of a finite group, Compositio Math. 106 (1997) 159–179.

[BuR99] J. Buhler, Z. Reichstein, On Tschirnhaus transformations, in: S.D. Ahlgren, et al. (Eds.), Topics in Number Theory, Kluwer Academic, 1999, pp. 127–142.

[Cor87] D.F. Coray, Cubic hypersurfaces and a result of Hermite, Duke Math. J. 54 (1987) 657–670.

[FuH91] W. Fulton, J. Harris, Representation Theory, Grad. Texts in Math., vol. 129, Springer-Verlag, Heidelberg, 1991.

[GPS01] G.-M. Greuel, G. Pfister, H. Schönemann, SINGULAR 2.0. A Computer Algebra System for Polynomial Computations, Centre for Computer Algebra, University of Kaiserslautern, 2001, http://www.singular.uni-kl.de.

[Her61] C. Hermite, Sur l'invariant du 18^{e} ordre des formes du cinquième degré et sur le rôle qu'il joue dans la résolution de l'équation du cinquième degré, extrait de deux lettres de M. Hermite à l'éditeur, J. Reine Angew. Math. 59 (1861) 304–305.

[Jou67] P. Joubert, Sur l'équation du sixième degré, C. R. Acad. Sci. Paris 64 (1867) 1025–1029.

[Rei99] Z. Reichstein, On a theorem of Hermite and Joubert, Canadian J. Math. 51 (1999) 69–95.

[Sch68] I. Schur, Vorlesungen über Invariantentheorie, Grundlehren Math. Wiss., vol. 143, Springer-Verlag, Heidelberg, 1968 (bearbeitet und herausgegeben von H. Grunsky).