JOUBERT'S THEOREM FAILS IN CHARACTERISTIC 2

ZINOVY REICHSTEIN

ABSTRACT. Let L/K be a separable field extension of degree 6. An 1867 theorem of P. Joubert asserts that if $\mathrm{char}(K) \neq 2$ then L is generated over K by an element whose minimal polynomial is of the form $t^6 + at^4 + bt^2 + ct + d$ for some $a, b, c, d \in K$. We show that this theorem fails in characteristic 2.

I. Introduction. The starting point for this note is the following classical theorem.

Theorem 1. (P. Joubert, 1867) *Let L/K be a separable field extension of degree* 6. *Assume* $\mathrm{char}(K) \neq 2$. *Then there is a generator $y \in L$ for L/K (i.e., $L = K(y)$) whose minimal polynomial is of the form*

$$t^6 + at^4 + bt^2 + ct + d \tag{2}$$

for some $a, b, c, d \in K$.

Joubert [Jo] gave a formula, which associates to an arbitrary generator x for L/K another generator $y \in L$ whose minimal polynomial is of the form (2). He did not state Theorem 1 in the above form, did not investigate under what assumptions on L, K and x his formula applies, and, most likely, only considered fields of characteristic zero. A proof of Theorem 1 based on an enhanced version of Joubert's argument has been given by H. Kraft [Kr, Main Theorem (b)]. A different (earlier) modern proof of Theorem 1, based on arithmetic properties of cubic hypersurfaces, is due to D. Coray [Co, Theorem 3.1]. (Coray assumed that $\mathrm{char}(K) \neq 2, 3$.) Since both of these proofs break down in characteristic 2, Kraft [Kr, Remark 6] asked if Theorem 1 remains valid when $\mathrm{char}(K) = 2$. In this paper we will show that the answer is "no" in general but "yes" under some additional assumptions on L/K.

II. **Notational conventions.** Suppose L/K is a field extension of of degree n. Every $y \in L$, defines a K-linear transformation $L \to L$ given by $z \mapsto yz$. We will denote the characteristic polynomial of this linear transformation by $t^n - \sigma_1(y)t^{n-1} + \cdots + (-1)^n \sigma_n(y)$. It is common to write $\mathrm{tr}(y)$ in place of $\sigma_1(y)$. The minimal and the characteristic polynomial of y coincide if and only if y is a generator for L/K.

If L/K is separable, then $\sigma_i(y) = s_i(y_1, \ldots, y_n)$, where $y_1, \ldots, y_n$ are the Galois conjugates of y and s_i is the ith elementary symmetric polynomial. Furthermore, if $[L : K] = 6$, then condition (2) of Theorem 1 is equivalent to $\sigma_1(y) = \sigma_3(y) = 0$.

We will be particularly interested in the "general" field extension L_n/K_n of degree n constructed as follows. Let F be a field and $x_1, \ldots, x_n$ be independent variables over F. The symmetric group S_n acts on $F(x_1, \ldots, x_n)$ by permuting $x_1, \ldots, x_n$. Set $K_n := F(x_1, \ldots, x_n)^{\mathrm{S}_n} = F(a_1, \ldots, a_n)$, where $a_i = s_i(x_1, \ldots, x_n)$, and $L_n := F(x_1, \ldots, x_n)^{\mathrm{S}_{n-1}} = K_n(x_1)$, where S_{n-1} permutes $x_2, \ldots, x_n$. Note that by construction L_n/K_n is a separable extension of degree n.

2010 *Mathematics Subject Classification.* 12E10, 12F10.

The author was partially supported by National Sciences and Engineering Research Council of Canada Discovery grant 250217-2012.

We remark that since S_n has no proper subgroups containing S_{n-1}, there are no proper subextensions between K_n and L_n. Thus for $n \geqslant 2$, $y \in L_n$ generates L_n/K_n if and only if $y \notin K_n$.

III. Main results.

Theorem 3. *Let F be a field of characteristic 2, $m \geq 1$ be an integer, and $n := 2 \cdot 3^m$. Then there is no $y \in L_n - K_n$ such that $\sigma_1(y) = \sigma_3(y) = 0$.*

In particular, setting $m = 1$, we see that Theorem 1 fails in characteristic 2. We will deduce Theorem 3 from the following more general

Theorem 4. *Let F be a field of characteristic 2, $m \geqslant 1$ be an integer, p be an odd prime, and $n := 2p^m$. Then there is no $y \in L_n - K_n$ such that $\mathrm{tr}(y) = \mathrm{tr}(y^2) = \cdots = \mathrm{tr}(y^p) = 0$.*

By Newton's formulas, $\mathrm{tr}(y^3) = \mathrm{tr}(y)^3 - 3\,\mathrm{tr}(y)\sigma_2(y) + 3\sigma_3(y)$. Thus in characteristic $\neq 3$,

$$\sigma_1(y) = \sigma_3(y) = 0 \quad \Longleftrightarrow \quad \mathrm{tr}(y) = \mathrm{tr}(y^3) = 0\,.$$

Moreover, in characteristic 2, $\mathrm{tr}(z^2) = \mathrm{tr}(z)^2$ for any $z \in L_n$ and thus

$$\mathrm{tr}(y) = \mathrm{tr}(y^2) = \cdots = \mathrm{tr}(y^p) = 0 \quad \Longleftrightarrow \quad \mathrm{tr}(y) = \mathrm{tr}(y^3) = \cdots = \mathrm{tr}(y^{p-2}) = \mathrm{tr}(y^p) = 0\,.$$

In particular, for $p = 3$, Theorem 4 reduces to Theorem 3.

Theorem 5. *Let L/K be a separable field extension of degree 6. Assume $\mathrm{char}(K) = 2$ and one of the following conditions holds:*

(a) there exists an intermediate extension $K \subset L' \subset L$ such that $[L' : K] = 3$,

(b) K is a C_1-field.

Then there is a generator $y \in L$ for L/K satisfying $\sigma_1(y) = \sigma_3(y) = 0$.

For background material on C_1-fields, see [Se$_1$, Sections II.3].

IV. Proof of Theorem 4: the overall strategy. It is easy to see that if Theorem 4 fails for a field F, it will also fail for the algebraic closure $\overline{F}$. We will thus assume throughout that F is algebraically closed.

Our proof of Theorem 4 will use the fixed point method, in the spirit of the arguments in [RY$_2$, Section 6]. The idea is as follows. Assume the contrary: $\mathrm{tr}(y) = \cdots = \mathrm{tr}(y^p) = 0$ for some $y \in L_n - K_n$. Based on this assumption we will construct a projective F-variety $\overline{X}$ with an S_n-action and an S_n-equivariant rational map $\varphi_y \colon \mathbb{A}^n \dashrightarrow \overline{X}$ defined over F. Here S_n acts on $\mathbb{A}^n$ by permuting coordinates in the usual way. The Going Down Theorem of J. Kollár and E. Szabó [RY$_1$, Proposition A.2] tells us that every abelian subgroup $G \subset \mathrm{S}_n$ of odd order has a fixed F-point in $\overline{X}$. On the other hand, we will construct an abelian p-subgroup G of S_n with no fixed F-points in $\overline{X}$. This leads to a contradiction, showing that y cannot exist. We will now supply the details of the proof, following this outline.

V. Construction of $\overline{X}$, φ_y, and the abelian subgroup $G \subset \mathrm{S}_n$. Every $y \in L_n$ gives rise to an S_n-equivariant rational map (i.e., a rational covariant)

$$\begin{gathered} f_y \colon \mathbb{A}^n \dashrightarrow \mathbb{A}^n \\ f_y(\alpha) = (h_1(y)(\alpha), \ldots, h_n(y)(\alpha))\,, \end{gathered}$$

where $\mathbb{A}^n$ is the n-dimensional affine space defined over F, $\alpha = (a_1, \ldots, a_n) \in \mathbb{A}^n$, elements of $F(x_1, \ldots, x_n)$ are viewed as rational functions on $\mathbb{A}^n$, and $h_1, \ldots, h_n$ are representatives of the left cosets of S_{n-1} in S_n, such that $h_i(1) = i$. Note that $h_1(y) = y, h_2(y), \ldots, h_n(y)$ are the conjugates of y in $F(x_1, \ldots, x_n)$. Since $y \in L_n := F(x_1, \ldots, x_n)^{\mathrm{S}_{n-1}}$, $h_i(y) \in F(x_1, \ldots, x_n)$

depends only on the coset $h_i\,\mathrm{S}_{n-1}$ (i.e., only on i) and not on the particular choice of h_i in this coset.

Recall that we are assuming that $\mathrm{tr}(y) = \cdots = \mathrm{tr}(y^p) = 0$. Thus the image of f_y is contained in the S_n-invariant subvariety $X \subset \mathbb{A}^n$ given by

$$a_1 + \cdots + a_n = a_1^2 + \cdots + a_n^2 = \cdots = a_1^p + \cdots + a_n^p = 0\,. \tag{6}$$

Because n is even and we are working in characteristic 2, if X contains $\alpha \in \mathbb{A}^n$ then X contains the linear span of α and $\alpha_0 := (1, \dots, 1)$. Using this observation, we define an S_n-equivariant rational map $\varphi_y \colon \mathbb{A}^n \dashrightarrow \overline{X}$ as a composition $\varphi_y \colon \mathbb{A}^n \overset{f_y}{\dashrightarrow} X \overset{\pi}{\dashrightarrow} \overline{X}$, where π denotes the linear projection $\mathbb{A}^n \dashrightarrow \mathbb{P}(F^n/D)$, $D := \mathrm{Span}_F(\alpha_0)$ is a 1-dimensional S_n-invariant subspace in F^n, and $\overline{X} \subset \mathbb{P}(F^n/D)$ is the image of X under π. Points in the projective space $\mathbb{P}(F^n/D) \simeq \mathbb{P}^{n-2}$ correspond to 2-dimensional linear subspaces $L \subset F^n$ containing D. Points in $\overline{X}$ correspond to 2-dimensional linear subspaces $L \subset F^n$, such that $D \subset L \subset X$. In particular, $\overline{X}$ is closed in $\mathbb{P}(F^n/D)$. The rational map π associates to a point $\alpha \in \mathbb{A}^n$ the 2-dimensional subspace spanned by α and α_0. Note that $\pi(\alpha)$ is well defined if and only if $\alpha \notin D$. Since we are assuming that $y \notin K_n$, the image of f_y is not contained in D. Thus the composition $\varphi_y = \pi \circ f_y \colon \mathbb{A}^n \dashrightarrow \overline{X}$ is a well-defined S_n-equivariant rational map.

Finally, the abelian p-subgroup $G \subset \mathrm{S}_n$ we will be working with, is defined as follows. Recall that $n = 2p^m$. The regular action of $(\mathbb{Z}/p\mathbb{Z})^m$ on itself allows us to view $(\mathbb{Z}/p\mathbb{Z})^m$ as a subgroup of S_{p^m}. Here we denote the elements of $(\mathbb{Z}/p\mathbb{Z})^m$ by $g_1, \dots, g_{p^m}$ and identify $\{1, \dots, p^m\}$ with $\{g_1, \dots, g_{p^m}\}$. We now set $G := (\mathbb{Z}/p\mathbb{Z})^m \times (\mathbb{Z}/p\mathbb{Z})^m \hookrightarrow \mathrm{S}_{p^m} \times \mathrm{S}_{p^m} \hookrightarrow \mathrm{S}_n$.

VI. Conclusion of the proof of Theorem 4. It remains to show that G has no fixed F-points in $\overline{X}$. A fixed F-point for G in $\overline{X}$ corresponds to a 2-dimensional G-invariant subspace L of F^n such that $D \subset L \subset X$. It will thus suffice to prove the following claim: no G-invariant 2-dimensional subspace of F^n is contained in X.

Since F is an algebraically closed field of characteristic 2 and G is an abelian p-group, where $p \neq 2$, the G-representation on F^n is completely reducible. More precisely, F^n decomposes as $F_{\mathrm{reg}}^{p^m}[1] \oplus F_{\mathrm{reg}}^{p^m}[2]$, the direct sum of the regular representations of the two factors of $G = (\mathbb{Z}/p\mathbb{Z})^m \times (\mathbb{Z}/p\mathbb{Z})^m$. Each $F_{\mathrm{reg}}^{p^m}[i]$ further decomposes as the direct sum of p^m one-dimensional invariant spaces

$$V_\chi[i] := \mathrm{Span}_F(\chi(g_1), \dots, \chi(g_{p^m}))\,,$$

where $\chi \colon (\mathbb{Z}/p\mathbb{Z})^m \to F^*$ is a multiplicative character. Thus $F^n = F_{\mathrm{reg}}^{p^m}[1] \oplus F_{\mathrm{reg}}^{p^m}[2]$ is the direct sum of the two-dimensional subspace

$$(F^n)^G = V_0[1] \oplus V_0[2] = \{(\underbrace{a, \dots, a}_{p^m \text{ times}}, \underbrace{b, \dots, b}_{p^m \text{ times}}) \mid a, b \in F\}\,,$$

where 0 denotes the trivial character of $(\mathbb{Z}/p\mathbb{Z})^m$, and $2p^m - 2$ distinct non-trivial 1-dimensional representations $V_\chi[i]$, where $i = 1, 2$, and χ ranges over the non-trivial characters $(\mathbb{Z}/p\mathbb{Z})^m \to F^*$. Note that $\chi(g)^p = \chi(g^p) = 1$ for any character $\chi \colon (\mathbb{Z}/p\mathbb{Z})^m \to F^*$, and thus

$$\chi_1(g_1)^p + \cdots + \chi(g_{p^m})^p = \underbrace{1 + \cdots + 1}_{p^m \text{ times}} = p^m = 1 \text{ in } F.$$

(Recall that $\mathrm{char}(F) = 2$ and p is odd.) Since one of the defining equations (6) for X is $x_1^p + \cdots + x_n^p = 0$, we conclude that none of the $2p^m$ G-invariant 1-dimensional subspaces $V_\chi[i]$ is contained in X, and the claim follows. $\square$

VII. Proof of Theorem 5. Let L_0 be the 5-dimensional K-linear subspace of L given by $\mathrm{tr}(y) = 0$. Let Y be the cubic threefold in $\mathbb{P}^4_K = \mathbb{P}(L_0)$ given by $\sigma_3(y) = 0$ (or equivalently,

$\mathrm{tr}(y^3) = 0)$. It is easy to see that Y is a cone, with vertex $1 \in L_0$, over a cubic surface $\overline{Y}$ in $\mathbb{P}^3_K := \mathbb{P}(L_0/K)$, defined over K. Note that $\overline{Y}$ is a K-form of the variety $\overline{X}$ we considered in the proof of Theorem 4. Applying the Jacobian criterion to the defining equations (6) of $\overline{X}$, (with $p = 3$ and $n = 6$), we see that $\overline{X}$ is a smooth surface, and hence, so is $\overline{Y}$. Either condition (a) or (b) implies that there exists a $y \in L - K$ such that $\mathrm{tr}(y) = \mathrm{tr}(y^3) = 0$. Equivalently, $\overline{Y}(K) \neq \emptyset$. It remains to show that we can choose a *generator* $y \in L$ with $\mathrm{tr}(y) = \mathrm{tr}(y^3) = 0$ or equivalently, that $\overline{Y}$ has a rational point away from of the "diagonal" hyperplanes $x_i = x_j$ in $\mathbb{P}^3$, $1 \leqslant i < j \leqslant 6$. (Note that the individual diagonal hyperplanes are defined over $\overline{K}$, but their union is defined over K.)

Suppose K is an infinite field. Since $\overline{Y}(K) \neq \emptyset$, $\overline{Y}$ is unirational; see [Ko]. Hence, K-points are dense in $\overline{Y}$, so that one (and in fact, infinitely many) of them lie away from the diagonal hyperplanes. Thus we may assume without loss of generality that $K = \mathbb{F}_q$ is a finite field of order q, where q is a power of 2, and $L = \mathbb{F}_{q^6}$. (Note that in this case both conditions (a) and (b) are satisfied.) If $y \in L$ is not a generator, it will lie in $\mathbb{F}_{q^2}$ or $\mathbb{F}_{q^3}$. Clearly $\mathrm{tr}(y) \neq 0$ for any $y \in \mathbb{F}_{q^2} - \mathbb{F}_q$ and $\mathrm{tr}(y) = \mathrm{tr}(y^3) = 0$ for any $y \in \mathbb{F}_{q^3}$. Thus a non-generator $y \in L$ satisfies $\mathrm{tr}(y) = \mathrm{tr}(y^3) = 0$ if and only if $y \in \mathbb{F}_{q^3}$. In geometric language, elements of $\mathbb{F}_{q^3}$ are the K-points of a line in $\overline{Y}$, defined over $K = \mathbb{F}_q$. We will denote this line by Z. It suffices to show that $\overline{Y}$ contains a K-point away from Z.

By [Ma, Corollary 27.1.1], $|\overline{Y}(K)| \geq q^2 - 7q + 1$. On the other hand, since $Z \simeq \mathbb{P}^1$ over K, $|Z(K)| = q + 1$. Thus for $q > 8$, $\overline{Y}$ has a K-point away from Z. In the remaining cases, where $q = 2$, 4 and 8, we will exhibit an explicit irreducible polynomial over $\mathbb{F}_q$ of the form (2):

$t^6 + t + 1$ is irreducible over $\mathbb{F}_2$ see [Ch, p. 199],

$t^6 + t^2 + t + \alpha$ is irreducible, over $\mathbb{F}_4$, where $\alpha \in \mathbb{F}_4 - F_2$, and

$t^6 + t + \beta$ is irreducible over $\mathbb{F}_8$, for some $\beta \in \mathbb{F}_8 - \mathbb{F}_2$; see [GT, Table 5]. □

VIII. Concluding remarks. (1) Theorem 1 extends a 1861 result of C. Hermite [He], which asserts that every separable extensions L/K of degree 5 has a generator $y \in L$ with $\sigma_1(y) = \sigma_3(y) = 0$. Surprisingly, Hermite's theorem is valid in any characteristic; see [Kr, Main Theorem (a)] or [Co, Theorem 2.2].

(2) It is natural to ask if results analogous to Theorem 1 are true for separable field extensions L/K of degree n, other than 5 and 6: does L/K always have a generator $y \in L$ with $\sigma_1(y) = \sigma_3(y) = 0$? If n can be written in the form $3_1^m + 3^{m_2}$ for some integers $m_1 > m_2 \geqslant 0$, then the answer is "no" in any characteristic (other than 3); see [Re, Theorem 1.3(c)], [RY$_2$, Theorem 1.8]. Some partial results for other n can be found in [Co, §4].

(3) Using the Going Up Theorem for G-fixed points [RY$_1$, Proposition A.4], our proof of Theorem 4 can be modified, to yield the following stronger statement. Suppose K'/K_n is a finite field extension of degree prime to p. Set $L' := L_n \otimes_{K_n} K'$. Then there is no $y \in L' - K'$ such that $\mathrm{tr}(y) = \mathrm{tr}(y^2) = \cdots = \mathrm{tr}(y^p) = 0$. In particular, under the assumptions of Theorem 3, there is no $y \in L' - K'$ with $\sigma_1(y) = \sigma_3(y) = 0$ for any finite field extension K'/K_n of degree prime to 3.

(4) Our argument shows that the G-action on $\overline{X}$ is not versal in the sense of [Se$_2$, Section I.5] or [DR]. Otherwise $\overline{X}$ would have a G-fixed point; see [DR, Remark 2.7]. Moreover, in view of Remark (3) above, the G-action on $\overline{X}$ is not even p-versal. Since $G \subset \mathrm{A}_n \subset \mathrm{S}_n$, the same is true of the A_n- and S_n-actions on $\overline{X}$. This answers a question raised by J.-P. Serre in a letter to the author in 2005.

(5) Theorem 3 corrects an inaccuracy in the statement of Joubert's theorem in [Re, Theorem 1.1], where the assumption that $\mathrm{char}(K) \neq 2$ was inadvertently left out.

(6) In the case where $K = \mathbb{F}_q$ is a finite field, Theorem 5 was communicated to the author by F. Voloch, along with an alternative proof, which is reproduced below with his permission.

"As in your comment after Theorem 4, it is enough to find y in $\mathbb{F}_{q^6}$, not in a smaller field, with $\mathrm{tr}(y) = \mathrm{tr}(y^3) = 0$, where the trace is to $\mathbb{F}_q$. These conditions are equivalent to the existence of $x, z \in \mathbb{F}_{q^6}$ with $y = x^q - x$, $y^3 = z^q - z$, so $z^q - z = (x^q - x)^3$. Letting $u = z + x^3$, we get an affine plane curve $u^q - u = x^{2q+1} + x^{q+2}$ over $\mathbb{F}_{q^6}$ (here q is a power of 2). It is a general fact that any affine plane curve of the form $u^q - u = f(x)$, where $f(x)$ polynomial of degree d prime to q, has genus $(q-1)(d-1)/2$, and its smooth projective model has exactly one point at infinity. In particular, our curve has genus $q(q-1)$. By the Weil bound the number of points on the smooth projective model of this curve is at least $q^6 + 1 - 2q(q-1)q^3$. There is one point at infinity and at most q^5 points with $y = x^q - x \in \mathbb{F}_{q^3}$; these are the bad points. If $q > 2$, our curve has a good point, one that gives rise to a generator of $\mathbb{F}_{q^6}$ over $\mathbb{F}_q$, because $q^6 + 1 - 2q(q-1)q^3 > 1 + q^5$ for any $q > 2$. For $q = 2$, I can exhibit an explicit 'Joubert polynomial', as in your formula (2). In fact, there are exactly two irreducible Joubert polynomials over $\mathbb{F}_2$, $t^6 + t + 1$ and $t^6 + t^4 + t^2 + t + 1$."

Acknowledgement. The author is grateful to M. Florence, H. Kraft, D. Lorenzini, J.-P. Serre, F. Voloch, and the anonymous referee for helpful comments.

References

[Ch] R. Church, *Tables of irreducible polynomials for the first four prime moduli*, Ann. of Math. (2) **36** (1935), 198–209. MR1503219

[Co] D. F. Coray, *Cubic hypersurfaces and a result of Hermite*, Duke Math. J. **54** (1987), 657–670. MR0899410

[DR] A. Duncan, Z. Reichstein, *Versality of algebraic group actions and rational points on twisted varieties*, *J. Alg. Geom.*, to appear, arXiv:1109.6093

[GT] D. H. Green and I. S. Taylor, *Irreducible polynomials over composite Galois fields and their applications in coding techniques*, Proc. Inst. Elec. Engrs. **121** (1974), 935–939. MR0434611

[He] C. Hermite, *Sur l'invariant du dix-huitième ordre des formes du cinquième degré*, J. Crelle **59** (1861), 304-305.

[Jo] P. Joubert, *Sur l'equation du sixième degré*, C-R. Acad. Sc. Paris **64** (1867), 1025-1029.

[Ko] J. Kollár, *Unirationality of cubic hypersurfaces*, J. Inst. Math. Jussieu **1** (2002), 467–476. MR1956057

[Kr] H. Kraft, *A result of Hermite and equations of degree 5 and 6*, J. Algebra **297** (2006), 234–253. MR2206857

[Ma] Yu. I. Manin, *Cubic forms. Algebra, geometry, arithmetic.* Translated from the Russian by M. Hazewinkel. Second edition. North-Holland Mathematical Library, 4. North-Holland Publishing Co., Amsterdam, 1986. MR0833513

[Re] Z. Reichstein, *On a theorem of Hermite and Joubert*, Canad. J. Math. **51** (1999), 69–95. MR1692919

[RY$_1$] Z. Reichstein and B. Youssin, *Essential dimensions of algebraic groups and a resolution theorem for G-varieties*, with an appendix by J. Kollár and E. Szabó, Canad. J. Math. **52** (2000), 1018–1056. MR1782331

[RY$_2$] Z. Reichstein and B. Youssin, *Conditions satisfied by characteristic polynomials in fields and division algebras*, J. Pure Appl. Algebra **166** (2002), 165–189. MR1868544

[Se$_1$] J.-P. Serre, *Galois cohomology*, translated from the French by Patrick Ion and revised by the author, Springer, Berlin, 1997. MR1466966

[Se$_2$] J.-P. Serre, Cohomological invariants, Witt invariants, and trace forms, notes by Skip Garibaldi, in *Cohomological invariants in Galois cohomology*, 1–100, Univ. Lecture Ser., 28, Amer. Math. Soc., Providence, RI. MR1999384

Department of Mathematics, University of British Columbia, Vancouver, CANADA
E-mail address: `reichst` *at* `math.ubc.ca`, *web page*: `www.math.ubc.ca/~reichst`