POLYNOMIALS WITH A COMMON COMPOSITE

ROBERT M. BEALS, JOSEPH L. WETHERELL, AND MICHAEL E. ZIEVE

ABSTRACT. Let f and g be nonconstant polynomials over a field K. In this paper we study the pairs (f, g) for which the intersection $K[f] \cap K[g]$ is larger than K. We describe all such pairs in case K has characteristic zero, as a consequence of classical results due to Ritt. For fields K of positive characteristic we present various results, examples, and algorithms.

1. INTRODUCTION

Let f_1 and f_2 be nonconstant polynomials over a field K of characteristic $p \geq 0$. In this paper we examine whether f_1 and f_2 have a *common composite*, i.e., whether there are nonconstant $u, v \in K[x]$ such that $u(f_1(x)) = v(f_2(x))$. Any such polynomial $u(f_1(x))$ is a common composite.

It turns out that there are very precise results about common composites whose degree is not divisible by p. Namely, if f_1 and f_2 have such a common composite, then they have a common composite of degree $\operatorname{lcm}(\deg(f_1), \deg(f_2))$. Also, under the same hypotheses, there are $g_1, g_2, r \in K[x]$ with $\deg(r) = \gcd(\deg(f_1), \deg(f_2))$ such that $f_1 = g_1 \circ r$ and $f_2 = g_2 \circ r$. Further, in Theorem 5.1 we describe all possibilities for g_1 and g_2.

For common composites of degree divisible by p, the situation is much more complicated. In particular, we present counterexamples to the gcd and lcm results, as well as a sequence of examples of pairs of bounded-degree polynomials whose least-degree common composites have degrees growing without bound. As a substitute, we give an algorithm which quickly determines whether there is a common composite of degree less than any fixed bound. Further, we prove the following result describing necessary and sufficient criteria for two polynomials to have a common composite. In this result, $m_i(a)$ denotes the multiplicity of $x = a$ as a root of $f_i(x) - f_i(a)$, and $\overline{K}$ denotes an algebraic closure of K.

Theorem 1.1. *Polynomials $f_1, f_2 \in K[x] \setminus K$ have a common composite if and only if there is a nonempty finite subset A of $\overline{K}$ which admits a function $\ell : A \to \mathbb{Z}$ such that*

- *for $a \in A$ and $i \in \{1, 2\}$, $\ell(a)/m_i(a)$ is a positive integer; and*

Date: February 28, 2008.

We thank the anonymous referee for several valuable suggestions.

- *for $i \in \{1,2\}$, $a \in A$, and $b \in \overline{K}$, if $f_i(a) = f_i(b)$ then $b \in A$ and $\ell(a)/m_i(a) = \ell(b)/m_i(b)$.*

We suspect that 'most' pairs of polynomials (f_1, f_2) have no common composite. In characteristic zero, this follows from Ritt's 1922 results [10]; see also Theorem 5.1. However, in positive characteristic it is difficult to produce a pair of polynomials which one can prove do not have a common composite; in fact, it has even been conjectured that no such polynomials exist over finite fields [8]. This conjecture was disproved in the 1970's via clever examples in [2], [3] and [1]. However, the arguments in those papers seem to apply only to very carefully chosen polynomials. We give some general methods for proving two polynomials have no common composite. A special case of our results is as follows:

Theorem 1.2. *Suppose $f_1, f_2 \in K[x] \setminus K[x^p]$ and $\alpha, \beta \in \overline{K}$ satisfy $f_i(\alpha) = f_i(\beta)$ for both $i = 1$ and $i = 2$. Then f_1 and f_2 have no common composite if either of the following hold:*

- *$m_1(\alpha)m_2(\beta) \neq m_1(\beta)m_2(\alpha)$; or*
- *$f_1'(\alpha)f_2'(\beta) \neq f_1'(\beta)f_2'(\alpha)$, and $[K(\alpha) : K]$ is divisible by a prime greater than $\max(\deg(f_1), \deg(f_2))$.*

For instance, one can check that the first condition implies that $x^2 + x$ and $x^3 + x^2$ have no common composite over $\mathbb{F}_2$, and the second condition implies that $x^4 + x^3$ and $x^6 + x^2 + x$ have no common composite over $\mathbb{F}_2$. In fact, we expect that our most general version of the second condition will apply to 'most' pairs of polynomials over a finite field.

The existence of a common composite can be reformulated in several different ways. It is clearly equivalent to saying the intersection of the polynomial rings $K[f_1]$ and $K[f_2]$ is strictly bigger than K. We show moreover that it is also equivalent to saying the intersection $K(f_1) \cap K(f_2)$ is strictly bigger than K, i.e., it is equivalent to f_1 and f_2 having a common rational function composite.

Finally, we mention an application of the results in this paper. Suppose f_1 and f_2 have no common composite, and assume further that f_1 and f_2 are not both functions of any polynomial of degree more than 1. Then the x-resultant of $f_1(x) - u$ and $f_2(x) - v$ is an irreducible polynomial in $K[u, v]$ which is not a factor of any nonzero 'variables separated' polynomial $r(u) - s(v)$ with $r, s \in K[x]$. We know no other way to produce such irreducibles.

Various authors have considered common composites from different perspectives, using methods involving Riemann surfaces, power series, curves and differentials, and group theory, among others. Of special importance is Schinzel's book [12], which contains beautiful proofs using (in most cases) only basic properties of polynomials. In the first few sections of this paper, we include new proofs of some known results. Also, we include multiple proofs of some results, and numerous examples illustrating the different types of phenomena that can occur. We hope that this will lead to future work providing more insight into the mysteries surrounding common composites of polynomials.

We now describe the organization of this paper. In the next section we show that if two polynomials have a common composite over an extension of K, then they have a common composite over K. Then in Section 3 we show that existence of a common rational function composite implies existence of a common polynomial composite. In Section 4 we give results and examples addressing the degrees of common composites. In the next section we use these results to describe all polynomials which have a common composite of degree not divisible by $\operatorname{char}(K)$. In Section 6 we give an algorithm which quickly determines whether there is a common composite of degree less than some bound. In the final three sections we give criteria for existence or nonexistence of common composites, and in particular we prove generalizations of Theorems 1.1 and 1.2.

2. Reduction to the Case of Algebraically Closed Fields

Theorem 2.1. *If $f_1, f_2 \in K[x]$ have a common composite over the algebraic closure $\overline{K}$ of K, then they have a common composite over K. Moreover, the minimal degree of any common composite over K equals the minimal degree of any common composite over $\overline{K}$.*

Proof. Let n be the minimal degree of any common composite over $\overline{K}$. Then there are polynomials g_1, g_2, h in $\overline{K}[x]$ with

$$h = g_1 \circ f_1 = g_2 \circ f_2 \tag{1}$$

such that h has degree n. Let d_i be the degree of g_i. Equation (1) expresses a $\overline{K}$-linear dependence of the polynomials

$$1, f_1, f_1^2, \dots, f_1^{d_1}, f_2, f_2^2, \dots, f_2^{d_2}.$$

Letting V be the K-vector space spanned by these polynomials, we see that the $\overline{K}$-vector space $\overline{K} \otimes_K V$ has the same dimension as V. Thus the polynomials are linearly independent over $\overline{K}$ if and only if they are linearly dependent over K. □

Corollary 2.2. *If $f_1, f_2 \in K[x] \setminus K[x^p]$ have a common composite, then they have a common composite which is not in $K[x^p]$.*

Proof. Let $h \in K[x]$ be a minimal degree common composite of f_1 and f_2, and assume $h \in K[x^p]$. Write $h = g_1 \circ f_1 = g_2 \circ f_2$ with $g_1, g_2 \in K[x]$. Since an element of $K[x]$ lies in $K[x^p]$ if and only if its derivative is zero, our hypotheses imply $g_i \in K[x^p]$. Thus $g_i = \widehat{g}_i(x)^p$ for some $\widehat{g}_i \in \overline{K}[x]$, so $h = x^p \circ \widehat{g}_1 \circ f_1 = x^p \circ \widehat{g}_2 \circ f_2$, whence $\widehat{g}_1 \circ f_1 = \widehat{g}_2 \circ f_2$. In particular, f_1 and f_2 have a common composite in $\overline{K}[x]$ of degree less than $\deg(h)$, which contradicts Theorem 2.1. □

Remark. Theorem 2.1 was first proved by McConnell [8] in case K is infinite, and by Bremner and Morton [3] in general. Corollary 2.2 is due to Alexandru and Popescu [1].

3. Rational Composites and Polynomial Composites

Theorem 3.1. *If $f_1, f_2 \in K[x]$ satisfy $K(f_1) \cap K(f_2) \neq K$, then f_1 and f_2 have a common composite, and moreover any minimal-degree common composite h satisfies $K(f_1) \cap K(f_2) = K(h)$ and $K[f_1] \cap K[f_2] = K[h]$.*

Proof. We use Lüroth's theorem [12, Thm. 2], which asserts that any subfield of $K(x)$ which properly contains K must have the form $K(s)$. Thus, $K(f_1) \cap K(f_2) = K(\widehat{h})$ for some $\widehat{h} \in K(x)$. Write $\widehat{h} = g_1 \circ f_1 = g_2 \circ f_2$ with $g_1, g_2 \in K(x)$, and write $g_i = a_i/b_i$ with $a_i, b_i \in K[x]$ and $\gcd(a_i, b_i) = 1$. By inverting $\widehat{h}, g_1, g_2$ if necessary, we may assume $\deg(a_1) \geq \deg(b_1)$. Then

$$a_1(f_1(x)) \cdot b_2(f_2(x)) = a_2(f_2(x)) \cdot b_1(f_1(x)).$$

In particular, $a_1(f_1(x))$ must divide the right hand side. Since $\gcd(a_1, b_1) = 1$, some $K[x]$-linear combination of a_1 and b_1 equals 1; substituting $f_1(x)$ for x in this expression, it follows that 1 is a $K[x]$-linear combination of $a_1 \circ f_1$ and $b_1 \circ f_1$, so $\gcd(a_1 \circ f_1, b_1 \circ f_1) = 1$. Thus, $a_1(f_1(x))$ divides $a_2(f_2(x))$. By symmetry, they must divide each other, so there is a constant c such that

$$a_1(f_1(x)) = c \cdot a_2(f_2(x)).$$

In particular, $h_0 := a_1 \circ f_1$ is in $K(f_1) \cap K(f_2) = K(\widehat{h})$. But $\deg(h_0) = \deg(\widehat{h}) = [K(x) : K(\widehat{h})]$, so in fact $K(h_0) = K(\widehat{h})$.

Now let s be any common composite of f_1 and f_2. Then $s \in K(f_1) \cap K(f_2) = K(h_0)$, so $s = r \circ h_0$ with $r \in K(x)$. It follows as above that $r \in K[x]$: write $r = a/b$ with $a, b \in K[x]$ and $\gcd(a, b) = 1$, so $\gcd(a \circ h, b \circ h) = 1$, and since $b \circ h$ divides $a \circ h$ we must have $\deg(b \circ h) = 0$, whence b is constant. Thus $K[f_1] \cap K[f_2] = K[h_0]$. In particular, the minimal-degree common composites of f_1 and f_2 are precisely the polynomials $\ell \circ h_0$ with $\ell \in K[x]$ of degree one. The result follows. □

We can use this result to sharpen the conclusion of Theorem 2.1:

Corollary 3.2. *If $h \in \overline{K}[x]$ is a minimal-degree common composite of f_1 and f_2 over $\overline{K}$, then $\ell \circ h \in K[x]$ for some degree-one $\ell \in \overline{K}[x]$. In particular, if h is monic and has no constant term then $h \in K[x]$.*

Remark. The anonymous referee informed us that, with some effort, one can prove Corollary 3.2 via the linear algebra approach used to prove Theorem 2.1.

Another consequence of Theorem 3.1 is that the study of common composites can be reduced to the case where both polynomials have nonzero derivative:

Corollary 3.3. *For any $f_1, f_2 \in K[x]$, write $f_i = \widehat{f_i} \circ x^{p^{n_i}}$ with $n_i \geq 0$ and $\widehat{f_i} \in K[x] \setminus K[x^p]$, and suppose $n_1 \geq n_2$. For any perfect field L between K and $\overline{K}$, there exists $\widetilde{f_1} \in L[x] \setminus L[x^p]$ such that $\widehat{f_1} \circ x^{p^{n_1-n_2}} = x^{p^{n_1-n_2}} \circ \widetilde{f_1}$. Then f_1 and f_2 have a common composite over K if and only if $\widetilde{f_1}$ and $\widehat{f_2}$ have a common composite over L. Moreover, the common composites of f_1 and f_2 over L are precisely the polynomials of the form $x^{p^{n_1-n_2}} \circ h \circ x^{p^{n_2}}$, where h varies over the common composites of $\widetilde{f_1}$ and $\widehat{f_2}$ over L.*

Proof. From the equation defining $\widetilde{f_1}$, we see that $\widetilde{f_1}$ is gotten from $\widehat{f_1}$ by replacing each coefficient by its $p^{n_1-n_2}$-th root. Thus $\widetilde{f_1} \notin \overline{K}[x^p]$. Next, since $f_1 = x^{p^{n_1-n_2}} \circ \widetilde{f_1} \circ x^{p^{n_2}}$ and $f_2 = \widehat{f_2} \circ x^{p^{n_2}}$, the common composites of f_1 and f_2 over L are gotten by substituting $x^{p^{n_2}}$ into the common composites (over L) of $\widehat{f_2}$ and $\overline{f}_1 := x^{p^{n_1-n_2}} \circ \widetilde{f_1}$. Write $q := p^{n_1-n_2}$. If $\widetilde{f_1}$ and $\widehat{f_2}$ have a common composite, then its q-th power is a composite of $\overline{f}_1$; thus $\overline{f}_1$ and $\widehat{f_2}$ have a common composite if and only if $\widetilde{f_1}$ and $\widehat{f_2}$ do. Hence f_1 and f_2 have a common composite over L if and only if $\widetilde{f_1}$ and $\widehat{f_2}$ do, and by Theorem 2.1 the former condition is equivalent to f_1 and f_2 having a common composite over K. So suppose $\widetilde{f_1}$ and $\widehat{f_2}$ have a common composite (over L), and let $\widehat{h}$ be a common composite of minimal degree. By Theorem 3.1, the common composites of $\widetilde{f_1}$ and $\widehat{f_2}$ are precisely the polynomials $\psi \circ \widehat{h}$ with $\psi \in L[x]$. Corollary 2.2 implies that $\widehat{h} \notin L[x^p]$. Thus, $\psi \circ \widehat{h}$ is in $L[x^q]$ if and only if $\psi \in L[x^q]$, or equivalently $\psi \circ \widehat{h} \in L[\overline{f}_1]$. Hence the common composites of $\overline{f}_1$ and $\widehat{f_2}$ are the polynomials $\varphi \circ x^q \circ \widehat{h}$ with $\varphi \in L[x]$. Since L is perfect, the set of q-th powers in $L[x]$ equals $L[x^q]$, and the result follows. □

Remark. The first two parts of Theorem 3.1 were proved by Noether [9] in the case of characteristic zero, and by McConnell [8] in general. The third part of Theorem 3.1 was proved by Schinzel [12, Lemma 1, p. 18].

We now give another proof of Theorem 3.1 with a different flavor.

Second proof of Theorem 3.1. First assume $K(x)$ is a separable extension of $K(f_1) \cap K(f_2)$. By Lüroth's theorem, $K(f_1) \cap K(f_2) = K(h)$ for some $h \in K(x)$. By making a linear fractional change to h, we may assume that the infinite place of $K(x)$ lies over the infinite place of $K(h)$. Let N be the Galois closure of $K(x)/K(h)$, and let G, H, A, B be the subgroups of $\mathrm{Gal}(N/K(h))$ fixing $K(h)$, $K(x)$, $K(f_1)$, and $K(f_2)$. Let I be the inertia group in $N/K(h)$ of a place lying over the infinite place of $K(h)$. Then the corresponding inertia groups in $N/K(f_1)$ and $N/K(f_2)$ are $I \cap A$ and $I \cap B$. Since f_1 and f_2 are polynomials, we have $A = H(I \cap A)$ and $B = H(I \cap B)$. For any subgroup C of G, write C_I for $I \cap C$.

Thus $H\langle A_I, B_I\rangle = \langle A_I, B_I\rangle H$, so $H\langle A_I, B_I\rangle$ is a group and thus equals $\langle A, B\rangle = G$. Hence $HI = G$, so the infinite place of $K(x)$ is the unique place of $K(x)$ lying over the infinite place of $K(h)$, whence $h \in K[x]$. Moreover, since the infinite place of $K(f_i)$ is the unique place of $K(f_i)$ lying over the infinite place of $K(h)$, it follows that h is a common composite of f_1 and f_2.

Now assume $K(x)$ is an inseparable extension of $K(f_1) \cap K(f_2)$. By Lüroth's theorem, $\overline{K}(f_1) \cap \overline{K}(f_2) = \overline{K}(h)$ for some $h \in \overline{K}(x)$. Write $f_1 = x^{p^A} \circ \widehat{f}_1$ and $f_2 = x^{p^B} \circ \widehat{f}_2$ where $\widehat{f}_i \in \overline{K}[x] \setminus \overline{K}[x^p]$. Then there are $g_i \in \overline{K}(x)$ with $h = g_i \circ \widehat{f}_i$. Write $g_1 = x^{p^C} \circ \widehat{g}_1$ and $g_2 = x^{p^D} \circ \widehat{g}_2$ with $\widehat{g}_i \in \overline{K}(x) \setminus \overline{K}(x^p)$. Then $\overline{K}(x)/\overline{K}(\widehat{g}_i \circ \widehat{f}_i)$ is separable but $\overline{K}(\widehat{g}_i \circ \widehat{f}_i)/\overline{K}(h)$ is purely inseparable, so the latter extension is the maximal purely inseparable subextension of $\overline{K}(x)/\overline{K}(h)$; in particular, $\overline{K}(\widehat{g}_1 \circ \widehat{f}_1) = \overline{K}(\widehat{g}_2 \circ \widehat{f}_2)$. Thus $\overline{K}(x)$ is a separable extension of $\overline{K}(\widehat{f}_1) \cap \overline{K}(\widehat{f}_2)$, so the result proved in the previous paragraphs implies that $\overline{K}(\widehat{f}_1) \cap \overline{K}(\widehat{f}_2) = \overline{K}(r)$ for some $r \in \overline{K}[x]$ which is a common composite of $\widehat{f}_1$ and $\widehat{f}_2$. It follows easily that $\overline{K}(f_1) \cap \overline{K}(f_2) = \overline{K}(r^{p^{\max(A,B)}})$. Now Theorem 2.1 implies that f_1 and f_2 have a common composite $\widehat{r} \in K[x]$ with $\deg(\widehat{r}) = \deg(r^{p^{\max(A,B)}})$, and since $[K(x) : K(f_1) \cap K(f_2)] \geq [\overline{K}(x) : \overline{K}(f_1) \cap \overline{K}(f_2)]$, it follows that $K(f_1) \cap K(f_2) = K(\widehat{r})$ as desired.

We have shown that $K(f_1) \cap K(f_2) = K(h)$ where $h \in K[x]$ is a common composite of f_1 and f_2. For any common composite $\widehat{h}$ of f_1 and f_2, we have $K(\widehat{h}) \subseteq K(h)$, and moreover the infinite place of $K(h)$ is the unique place of $K(h)$ lying over the infinite place of $K(\widehat{h})$; thus $\widehat{h} = r(h)$ for some $r \in K[x]$. □

This second proof generalizes at once to intersections of higher-genus function fields:

Proposition 3.4. *Let F be a finite extension of $K(x)$, and let F_1 and F_2 be subfields of F which contain K. Suppose F is a finite separable extension of $F_0 := F_1 \cap F_2$. If a place P of F is totally ramified in both F/F_1 and F/F_2, then P is totally ramified in F/F_0.*

In Section 7 (following Theorem 7.4) we give a third proof of Theorem 3.1, which is a different type of constructive proof.

4. Degree Constraints

In this section we examine the possible degrees of common composites of f_1 and f_2. By Theorem 3.1, the set of degrees of common composites equals the set of multiples of some integer n, so it suffices to analyze n, which is the minimal degree of any common composite. Clearly any common composite has degree divisible by $\operatorname{lcm}(\deg(f_1), \deg(f_2))$. Conversely, in characteristic zero we now show that if there is a common composite then there is one of this minimal degree. More generally this holds if there is a common composite of degree not divisible by $p := \operatorname{char}(K)$:

Theorem 4.1. *If $f_1, f_2 \in K[x]$ have a common composite, then they have a common composite of degree* $\operatorname{lcm}(\deg(f_1), \deg(f_2))p^s$ *for some $s \geq 0$. (Here we use the convention $0^0 = 1$.)*

Proof. First assume $f_1, f_2 \notin K[x^p]$. Let $h(x)$ be a common composite of minimal degree. Corollary 2.2 and Theorem 3.1 imply that $K(x)/K(h(x))$

is separable and $K(h) = K(f_1) \cap K(f_2)$. Let L be the Galois closure of $K(x)/K(h(x))$, and let G, A, B, H be the subgroups of $\mathrm{Gal}(L/K(h(x)))$ fixing $h(x)$, $f_1(x)$, $f_2(x)$, and x. Then $G = \langle A, B\rangle$. Let P be a place of L lying over the infinite place of $K(h(x))$, and let I be the inertia group of P in $L/K(h(x))$. Since h is a polynomial, $G = HI$.

For any group C with $H \le C \le G$, let $C_I := C \cap I$. Clearly C contains HC_I, and since $G = HI$ we have $C = HC_I$. Moreover, $[C : H] = [C_I : H_I]$. Since $H\langle A_I, B_I\rangle = \langle A_I, B_I\rangle H$, the set $H\langle A_I, B_I\rangle$ is a group and thus equals $\langle A, B\rangle = G$. Hence $I = G_I = \langle A_I, B_I\rangle$. Recall the structure of inertia groups (cf., e.g., [13, Cor. 4 to Prop. 7, § IV.2]): I is the semidirect product $V \rtimes D$ where V is a normal p-subgroup and D is cyclic of order not divisible by p. Since $I = \langle VA_I, VB_I\rangle$, we have $I/VH_I = \langle VA_I/VH_I, VB_I/VH_I\rangle$, and since these are cyclic groups we see that $[I : VH_I]$ is the least common multiple of $[VA_I : VH_I]$ and $[VB_I : VH_I]$. Finally, $\deg(h) = [I : H_I]$ and $\deg(f_1) = [A_I : H_I]$, so $[V : VH_I]$ and $[VA_I : VH_I]$ are the maximal divisors of $\deg(h)$ and $\deg(f_1)$ which are not divisible by p, whence $\deg(h) = \mathrm{lcm}(\deg(f_1), \deg(f_2))p^t$ with $t \ge 0$.

Now for arbitrary $f_1, f_2 \in K[x]$ having a common composite, Corollary 3.3 implies that the minimal degree of any common composite of f_1 and f_2 over $\overline{K}$ is a power of p times the minimal degree of any common composite of two related polynomials $\widetilde{f_1}, \widehat{f_2} \in \overline{K}[x] \setminus \overline{K}[x^p]$, where both $\deg(f_1)/\deg(\widetilde{f_1})$ and $\deg(f_2)/\deg(\widehat{f_2})$ are powers of p. Since $\widetilde{f_1}, \widehat{f_2} \notin \overline{K}[x^p]$, it follows from above that the minimal degree of any common composite of $\widetilde{f_1}$ and $\widehat{f_2}$ is $\mathrm{lcm}(\deg(\widetilde{f_1}), \deg(\widehat{f_2}))p^s$ for some $s \ge 0$, so the minimal degree of any common composite of f_1 and f_2 over $\overline{K}$ is $\mathrm{lcm}(\deg(f_1), \deg(f_2))p^t$ with $t \ge 0$. The result now follows from Theorem 2.1. □

Remark. Theorem 4.1 was proved by Engstrom [5] in the case of characteristic zero, and his proof extends at once to the case where f_1 and f_2 have a common composite of degree not divisible by p (cf. [12, Thm. 5]). This elegant proof is completely different from ours (for instance it depends on nothing beyond the division algorithm in $K[x]$), and it would be interesting to try to extend Engstrom's argument to prove our full result. In case f_1 and f_2 have a common composite of degree not divisible by p, our proof is essentially a modernized account of an argument due to Ritt [10]; an alternate treatment of Ritt's proof in this case, using fields and power series instead of groups and inertia groups, is in [8]. The basic ideas in [8] can be discerned by scrutinizing the proof of [6, Thm. 3.6], though significant effort is required since the latter proof contains errors in nearly every line. An incorrect generalization of Theorem 4.1 is given as [1, Thm. 2.1]; specifically, they assert that the result for degrees also holds for the ramification indices under any prescribed place of $K(x)$. A counterexample is $f_1 = x^2$ and $f_2 = x^3 - x$ over $K = \mathbb{C}$ at the place $x = 1$, since $x = 1$ is unramified in $K(x)/K(f_1)$ and $K(x)/K(f_2)$ but ramifies in $K(x)/(K(f_1) \cap K(f_2)) = K(x)/K((x^3 - x)^2)$.

The mistake in the proof of [1, Thm. 2.1] is the assertion that the completions of $K(f_1)$ and $K(f_2)$ (at places under the prescribed place of $K(x)$) intersect in the completion of $K(f_1) \cap K(f_2)$, which is not generally true.

It is not possible to remove the power of p from the conclusion of Theorem 4.1. For example, there are polynomials over $\mathbb{F}_2$ of degrees 11 and 13 whose least-degree common composite has degree $143 \cdot 2^{60}$, and there are polynomials over $\mathbb{F}_2$ of degrees 1447 and 1451 whose least-degree common composite has degree $1447 \cdot 1451 \cdot 2^{1048350}$. These are special cases of the following result.

Proposition 4.2. *Suppose $p := \operatorname{char}(K)$ is nonzero. If $p \nmid n$ then a minimal-degree common composite of x^n and $x^{p^r} - x$ is $(x^{p^{rd}} - x)^n$, where d is the multiplicative order of $p^r \bmod n$. If $p \nmid nm$ and $n, m > 1$ then a minimal-degree common composite of x^n and $(x-1)^m$ is $(x^{p^d} - x)^{\operatorname{lcm}(m,n)}$, where d is the multiplicative order of $p \bmod \operatorname{lcm}(m,n)$.*

Proof. Let $\overline{K}$ be an algebraic closure of K, and let ζ be a primitive n-th root of unity in $\overline{K}$. Then $\overline{K}(x)/\overline{K}(x^n)$ is Galois with group generated by $\sigma : x \mapsto \zeta x$, and $\overline{K}(x)/\overline{K}(x^{p^r} - x)$ is Galois with group H consisting of the various maps $x \mapsto x + \alpha$ with $\alpha \in \mathbb{F}_{p^r}$. The subgroup G of $\operatorname{Aut}_{\overline{K}} \overline{K}(x)$ generated by σ and H consists of the maps $x \mapsto \mu x + \nu$ where $\mu \in \langle \zeta \rangle$ and $\nu \in \mathbb{F}_{p^r}(\zeta)$. Here $\#G = np^{rd}$, where $d := [\mathbb{F}_{p^r}(\zeta) : \mathbb{F}_{p^r}]$ is the multiplicative order of p^r mod n. The group G fixes $h(x) := (x^{p^{rd}} - x)^n$, so since $\deg h = \#G$ we see that $\overline{K}(h)$ is the subfield of $\overline{K}(x)$ fixed by G, whence $\overline{K}(h)$ is the intersection of $\overline{K}(x^n)$ and $\overline{K}(x^{p^r} - x)$. This shows that h is a minimal-degree common composite of x^n and $x^{p^r} - x$ over $\overline{K}$. Since $h \in K[x]$, it is also a minimal-degree common composite over K.

Now let η be a primitive m-th root of unity in $\overline{K}$. Then the extension $\overline{K}(x)/\overline{K}((x-1)^m)$ is Galois with group generated by $\gamma : x \mapsto 1 + \eta(x-1)$. Let H be the subgroup of $\operatorname{Aut}_{\overline{K}} \overline{K}(x)$ generated by σ and γ. Then H contains the commutator $\gamma^{-1}\sigma^{-1}\gamma\sigma : x \mapsto x - (\eta - 1)(\zeta - 1)$. One easily checks that H consists of the maps $x \mapsto \mu x + \nu$ where $\mu \in \langle \zeta, \eta \rangle$ and $\nu \in \mathbb{F}_p(\zeta, \eta)$. Moreover, H fixes $j(x) := (x^{p^d} - x)^{\operatorname{lcm}(m,n)}$, where d is the multiplicative order of $p \bmod \operatorname{lcm}(m,n)$. Since $\deg(j) = \#H$, it follows as above that j is a minimal-degree common composite of x^n and $(x-1)^m$ over $\overline{K}$, and hence over K. □

Remark. The second part of Proposition 4.2 was first proved by Bremner and Morton [3].

In case $\operatorname{char}(K) = 0$, Theorem 4.1 says that if $f_1, f_2 \in K[x]$ have a common composite then they have one of degree $\operatorname{lcm}(\deg(f_1), \deg(f_2))$. Proposition 4.2 shows that this is no longer true when $\operatorname{char}(K) > 0$. Specifically, for any prime p, the least-degree common composite of x^2 and $x^2 - x$ over $\mathbb{F}_p$ is $(x^p - x)^2$. Thus, two degree-2 polynomials can have lowest-degree common composite of arbitrarily large degree. We make the following definition to give a framework for recovering some analogue of the characteristic zero result, by restricting to polynomials over a fixed field, or polynomials over fields of a fixed characteristic.

Definition. Given integers $n_1, n_2 > 1$ and a field K, let $N(n_1, n_2, K)$ be the supremum of the integers $r(f_1, f_2, K)$, where

- $f_1, f_2 \in K[x]$ have a common composite and satisfy $\deg(f_1) = n_1$ and $\deg(f_2) = n_2$, and
- $r(f_1, f_2, K)$ is the lowest degree of any common composite of f_1 and f_2.

For any prime number p (and for $p = 0$), let $N(n_1, n_2, p)$ be the supremum of the values $N(n_1, n_2, K)$, where K varies over all fields of characteristic p.

Theorem 4.1 implies that $N(n_1, n_2, 0) = \operatorname{lcm}(n_1, n_2)$, and more generally that $r(f_1, f_2, K) = \operatorname{lcm}(n_1, n_2) \operatorname{char}(K)^s$. However, Proposition 4.2 shows that in positive characteristic there are examples with arbitrarily large s. But our examples have $n_1 + n_2 \to \infty$, and we do not know whether one can bound s in terms of n_1, n_2 and K, or even just in terms of n_1 and n_2. In fact, every example we know (when $p = \operatorname{char}(K) > 0$) satisfies $s \le \operatorname{lcm}(n_1, n_2)$. We now prove that, when $n_1 = n_2 = 2$, we can actually take $s \le 1$:

Proposition 4.3. *Any two degree-2 polynomials over a field of characteristic $p > 0$ have a common composite of degree $2p$.*

Proof. Let K be a field of characteristic p, and let f_1 and f_2 be degree-2 polynomials in $K[x]$. If $K(x)/K(f_1)$ is not separable, then $p = 2$ and $f_1 = ax^2 + b$, so f_2^2 is a common composite of f_1 and f_2 of degree $2p$. Henceforth assume $K(x)/K(f_1)$ and $K(x)/K(f_2)$ are separable. Thus these extensions are Galois. Moreover, writing $f_1 = ax^2 - bx + c$, the Galois group of $K(x)/K(f_1)$ is generated by $x \mapsto b/a - x$. Thus, $\operatorname{Gal}(K(x)/K(f_1))$ and $\operatorname{Gal}(K(x)/K(f_2))$ are generated by $\sigma_1 : x \mapsto \alpha_1 - x$ and $\sigma_2 : x \mapsto \alpha_2 - x$, for some $\alpha_1, \alpha_2 \in K$. Now, $K(f_1) \cap K(f_2)$ is the subfield of $K(x)$ fixed by $H := \langle \sigma_1, \sigma_2 \rangle$. Since σ_1 and σ_2 have order 2, they generate a dihedral group of order twice the order of the composite map $\sigma_1\sigma_2 : x \mapsto (\alpha_1 - x) \circ (\alpha_2 - x) = \alpha_1 - \alpha_2 + x$. Since the latter map has order 1 or p, it follows that $\#H \mid 2p$. Now the result follows from Theorem 3.1. □

The anonymous referee suggested the following alternate proof:

Second proof of Proposition 4.3. Let K be a field of characteristic p, and let f_1 and f_2 be degree-2 polynomials in $K[x]$. By replacing f_i with $\ell_i \circ f_i$ for a suitable degree-1 polyomial $\ell_i \in K[x]$, we may assume $f_1 = x^2 + ax$ and $f_2 = x^2 + bx$. If $a = b$ then f_1 is already a common composite; hence we assume $a \ne b$. Then f_1 and f_2 have a common composite of degree at most $2n$ if and only if the polynomials $1, f_1, f_2, f_1^2, f_2^2, \dots, f_1^n, f_2^n$ are linearly dependent. These polynomials span the same space as the polynomials $1, f_1 - f_2, f_2, f_1^2 - f_2^2, f_2^2, \dots, f_1^n - f_2^n, f_2^n$. Since the leading term of f_2^i is x^{2i}, and the leading term of $f_1^i - f_2^i$ is $i(a-b)x^{2i-1}$, the matrix of coefficients of these polynomials is triangular, and it has no zero entries on the main diagonal if and only if $n < p$. Thus f_1 and f_2 have a common composite of degree $2p$, and no common composite of lower degree. □

It would be interesting to determine further values of $N(n_1, n_2, p)$, or even to determine whether these values are finite. One can attempt to produce infinite values of $N(n_1, n_2, p)$ by modifying the proof of Theorem 4.1. Below is a group-theoretic example satisfying the conditions used in that proof, such that H_I, A_I, B_I, I have orders $1, 2, 3, 2 \cdot 3^{2n+1}$ respectively, where n can be any positive integer. If this group-theoretic setup could be realized by polynomials f_1 and f_2 in characteristic 3, then there would be polynomials of degrees 2 and 3 whose lowest-degree common composite has degree $2 \cdot 3^{2n+1}$.

Example 4.4. *Let I be the group generated by a, b, c subject to the relations $b^{3^n} = c^{3^n} = a^6 = 1$, $bc = cb$, $a^{-1}ba = c^{-1}$, $a^{-1}ca = bc$. One can check that $I = VC$ where $V = \langle b, c, a^2 \rangle$ has order 3^{2n+1} and $C = \langle a^3 \rangle$ has order 2. Now $\langle a^3 b, a^2 b \rangle$ contains a and b and thus contains $c = a^{-1}b^{-1}a$. Hence $I = \langle A_I, B_I \rangle$ where $A_I := \langle a^3 b \rangle$ and $B_I := \langle a^2 b \rangle$. Finally, one can check that $\#A_I = 2$ and $\#B_I = 3$.*

In a subsequent paper we will show that the above configuration does not happen, and in fact we will compute $N(2, 3, p)$. However, our proof uses a different framework, and does not give a simple explanation why the above configuration doesn't occur. It would be interesting to know a general constraint on the Galois groups associated to a polynomial which would preclude this setup from being realizable.

Theorem 4.1 is a 'least common multiple' result. We now give a companion 'greatest common divisor' result.

Theorem 4.5. *If $f_1, f_2 \in K[x]$ have a common composite of degree not divisible by p, then there are $g_1, g_2, r \in K[x]$ with $\deg(r) = \gcd(\deg(f_1), \deg(f_2))$ such that $f_1 = g_1 \circ r$ and $f_2 = g_2 \circ r$.*

Proof. We use the notation from the proof of Theorem 4.1. Thus $H(A_I \cap B_I) = (A_I \cap B_I)H$, so $H(A_I \cap B_I)$ is a group, and equals $A \cap B$. By Lüroth's theorem, the subfield of L fixed by $H(A_I \cap B_I)$ has the form $K(r(x))$ for some rational function $r(x)$. By making a linear fractional change to $r(x)$ if necessary, we may assume that the infinite place of $K(r(x))$ lies under the infinite place of $K(x)$. Since the latter place is totally ramified in $K(x)/K(r(x))$, it follows that $r(x)$ is a polynomial. Moreover, the infinite place of $K(r(x))$ is the unique place lying over the infinite place of $K(f(x))$, so $f_1 = g_1 \circ r$ for some polynomial g_1, and likewise $f_2 = g_2 \circ r$.

It remains only to determine the degree of r, which equals $[H(A_I \cap B_I) : H] = \#(A_I \cap B_I)$. Since A_I and B_I are subgroups of the cyclic group I, we have $\#(A_I \cap B_I) = \gcd(\#A_I, \#B_I)$. Thus the degree of r is $\gcd([A : H], [B : H]) = \gcd(\deg(f_1), \deg(f_2))$. □

Remark. Theorem 4.5 was proved by Engstrom [5] in the case of characteristic zero, and his proof extends at once to the general case (cf. [12, Thm. 5]).

The situation is the same as for Theorem 4.1: Engstrom's argument (as simplified by Schinzel) uses just polynomials, and no Galois theory. Our Galois-theoretic proof is a modernized version of an argument of Ritt's [10], and a complicated field-theoretic version of Ritt's argument (with numerous errors) is in [6].

Note that the hypothesis on the degree in Theorem 4.5 is necessary—for instance, if $f_1 = x^2 + ax$ and $f_2 = x^2 + bx$ with $a \neq b$, then certainly there is no r satisfying the conclusion of Theorem 4.5, but Proposition 4.3 says that f_1 and f_2 have a common composite over any field of positive characteristic.

Theorems 4.1 and 4.5 show that the existence of a common composite of degree not divisible by p is a very unusual occurrence. For instance, if polynomials f_1 and f_2 of the same degree have such a common composite, then $f_1 = \ell \circ f_2$ for some degree-1 polynomial ℓ.

5. The Tame Case

In this section we describe all pairs of polynomials $f_1, f_2 \in K[x]$ which have a common composite of degree not divisible by $\operatorname{char}(K)$. The statement of the result involves the Dickson polynomials, which are defined as follows. For any $\alpha \in K$ and $n > 0$, define $D_n(x, \alpha) \in K[x]$ by

$$D_n(x, \alpha) = \sum_{i=0}^{\lfloor n/2 \rfloor} \frac{n}{n-i} \binom{n-i}{i} (-\alpha)^i x^{n-2i}.$$

The key property of Dickson polynomials is that $D_n(x + \alpha/x, \alpha) = x^n + (\alpha/x)^n$.

Theorem 5.1. *Suppose $f_1, f_2 \in K[x]$ satisfy $\deg(f_1) \geq \deg(f_2) > 1$ and $\operatorname{char}(K) \nmid \deg(f_1)\deg(f_2)$. Then f_1 and f_2 have a common composite of degree not divisible by $\operatorname{char}(K)$ if and only if there are degree-1 polynomials $\ell_1, \ell_2 \in K[x]$ and a polynomial $h(x) \in K[x]$ of degree $\gcd(\deg(f_1), \deg(f_2))$ such that either*

1. *$f_1 = \ell_1 \circ x^r P(x^n) \circ h(x)$ and $f_2 = \ell_2 \circ x^n \circ h(x)$, where $r, n > 0$ and $P \in K[x]$; or*
2. *$f_1 = \ell_1 \circ D_m(x, \alpha) \circ h(x)$ and $f_2 = \ell_2 \circ D_n(x, \alpha) \circ h(x)$, where $\alpha \in K$ and $m, n > 0$.*

Proof. Suppose f_1 and f_2 have a common composite of degree not divisible by $p := \operatorname{char}(K)$. By Theorem 4.5, there are $g_1, g_2, h \in K[x]$ such that $f_i = g_i \circ h$ and $\deg(h) = \gcd(\deg(f_1), \deg(f_2))$. By Theorem 4.1, g_1 and g_2 have a common composite of degree $\operatorname{lcm}(\deg(g_1), \deg(g_2))$. Now the result follows from Theorem 5.2 below. □

In the following result, if ℓ is a degree-1 polynomial over a field K, we write $\ell^{\langle -1 \rangle}$ to denote the functional inverse of ℓ; thus $\ell^{\langle -1 \rangle}$ is the unique degree-1 polynomial over K for which $\ell^{\langle -1 \rangle}(\ell(x)) = x$, or equivalently $\ell(\ell^{\langle -1 \rangle}(x)) = x$.

Theorem 5.2 (Zannier)**.** *Suppose $a, b, c, d \in K[x]$ satisfy $\deg(a) = \deg(d) = m > 1$ and $\deg(b) = \deg(c) = n > 1$, where $\gcd(m, n) = 1$ and $m > n$ and $a'c' \neq 0$. Then $a(b) = c(d)$ holds if and only if there are degree-1 polynomials $\ell_1, \ell_2, \ell_3, \ell_4 \in K[x]$ such that either*

1. *$\ell_1 \circ a \circ \ell_3^{\langle -1 \rangle} = x^r P(x)^n$ and $\ell_3 \circ b \circ \ell_2 = x^n$ and $\ell_1 \circ c \circ \ell_4^{\langle -1 \rangle} = x^n$ and $\ell_4 \circ d \circ \ell_2 = x^r P(x^n)$, where $P \in K[x]$ and $r = m - n \deg(P) > 0$; or*
2. *$\ell_1 \circ a \circ \ell_3^{\langle -1 \rangle} = D_m(x, \alpha^n)$ and $\ell_3 \circ b \circ \ell_2 = D_n(x, \alpha)$ and $\ell_1 \circ c \circ \ell_4^{\langle -1 \rangle} = D_n(x, \alpha^m)$ and $\ell_4 \circ d \circ \ell_2 = D_m(x, \alpha)$, where $\alpha \in K$.*

Remark. Theorem 5.2 was proved by Zannier [15]; an alternate exposition of his proof is in [12, Thm. 8]. Previously special cases had been proved by Ritt [10], Levi [7], Dorey and Whaples [4], Schinzel [11], and Tortrat [14].

Theorem 5.1 shows in a strong sense that, when $\text{char}(K) = 0$, very few pairs of polynomials (f_1, f_2) have a common composite. We suspect that the same qualitative behavior occurs in positive characteristic, but it is difficult to prove significant results in this direction.

6. Fiber-finding

In this section we give two algorithms which produce either a common composite or a proof that there is no such of degree less than a prescribed bound. The idea of the first algorithm is simple: if f_1 and f_2 have a common composite h, then any $\alpha, \beta \in \overline{K}$ with $f_1(\alpha) = f_1(\beta)$ also satisfy $h(\alpha) = h(\beta)$. Thus, starting with some $\alpha \in \overline{K}$, we compute all $\beta \in \overline{K}$ with $f_1(\alpha) = f_1(\beta)$. Then for each β we compute all $\gamma \in \overline{K}$ with $f_2(\beta) = f_2(\gamma)$. Note that $h(\gamma) = h(\beta) = h(\alpha)$. Continuing this process, we find more and more elements of $\overline{K}$ which have the same h-value. This gives a lower bound on the degree of h; conversely, we show in Section 7 that, if this process produces only finitely many elements of $\overline{K}$, then we can determine whether f_1 and f_2 have a common composite.

In the second algorithm we work with polynomials over K rather than elements of $\overline{K}$. In this case it is convenient to assume the f_i are monic. Suppose we have a nonconstant $r \in K[x]$ which divides our hypothesized (minimal degree) common composite h. Let m be the minimal polynomial of f_i mod r, i.e., m is the minimal degree monic polynomial in $K[x]$ such that $m \circ f_i$ is divisible by r. Then $m \circ f_i$ divides h. By iterating this process, we can quickly build up large-degree factors of h. We can start this process with $r_0 = x$. After one step, we have $r_1 = (x - f_1(0)) \circ f_1 = f_1 - f_1(0)$. The polynomials r_j alternate between composites of f_1 and composites of f_2. Therefore, if this process ever stabilizes (by giving $r_j = r_{j+1}$ for some $j > 0$), then the final r_j is a minimal degree common composite of f_1 and f_2.

Example 6.1. *Let $f_1 = x^2$ and $f_2 = x^3 + x^2 + x$, where $\text{char}(K) = 3$. Then we start with $r_1 := f_1 = x^2$. The minimal polynomial of f_2 mod r_1 is x^2, so we put $r_2 := x^2 \circ f_2 = x^6 - x^5 - x^3 + x^2$. The minimal polynomial of f_1 mod*

r_2 is $x^5-x^4-x^2+x$, so we put $r_3 := (x^5-x^4-x^2+x)\circ f_1 = x^{10}-x^8-x^4+x^2$. The minimal polynomial of f_2 mod r_3 is $m := x^6+x^5+x^3+x^2$, so we put $r_4 := m \circ f_2 = x^{18}-x^{14}-x^6+x^2$. Finally, the minimal polynomial of f_1 mod r_4 is $x^9-x^7-x^3+x$, and $r_4 = (x^9-x^7-x^3+x)\circ f_1$, so r_4 is a minimal-degree common composite.

The above algorithms are actually two incarnations of the same idea. In the first algorithm we explore the fiber $\{\zeta : h(\zeta) = h(\alpha)\}$. Letting Z be the set of ζ's seen up to a given step, we can put $r := \prod_{\zeta\in Z}(x-\zeta)$. We know $r(x)$ divides $h(x)-h(\alpha)$; we may assume $h(\alpha)=0$, so r divides h. Suppose the next step involves equating values of f_i, and let $\widehat{Z}$ be the next set of ζ's. Let $v(x)$ be obtained by eliminating z from the system:

$$\begin{aligned} r(z) &= 0 \\ f_i(z) &= f_i(x) \end{aligned}$$

(i.e., $v(x)$ generates the intersection of the ideal $(r(z), f_i(z)-f_i(x))$ with $K[x]$). Then every root of v lies in $\widehat{Z}$, and every element of $\widehat{Z}$ is a root of v. It turns out that $v = m \circ f_i$, where m is the minimal polynomial of f_i mod r. Thus, both our algorithms produce the same set of elements of $\overline{K}$ at each step; the main difference between them is that the second algorithm keeps track of multiplicities, while the first does not.

Here is an example where the second algorithm can be used to prove that two polynomials have no common composite.

Example 6.2. *Let $f_1 = x^2-x$ and $f_2 = x^3-x^2$. We start with $r_1 := f_1$. Inductively, we show that $r_{2j+1} = f_1^{2^j}$ and $r_{2j+2} = f_2^{2^j}$. Indeed, if $r_{2j+1} = f_1^{2^j}$ then its roots $x=0$ and $x=1$ each have multiplicity 2^j; since $x=0$ and $x=1$ are roots of f_2 of multiplicities 2 and 1, it follows that $r_{2j+2} = f_2^{2^j}$. Thus the roots of r_{2j+2} are again $x=0$ and $x=1$, this time with multiplicities 2^{j+1} and 2^j; since $x=0$ and $x=1$ are simple roots of f_1, it follows that $r_{2j+3} = f_1^{2^{j+1}}$. Since the degrees of the r_j grow without bound, f_1 and f_2 have no common composite.*

If we apply the first algorithm with $\alpha = 0$ to the polynomials in the above example, we quickly find a stable set $Z = \{0,1\}$. This example is better understood in the context of the next two sections: the first algorithm terminates with $Z = \{0,1\}$ because that set is *compatible* (see Section 7). The second algorithm fails to terminate because the set Z is *inconsistent* (see Example 8.1).

While we suspect that the above example illustrates a rare situation, it is worth modifying the second algorithm so that, if the set Z of roots of r stabilizes, we check Z for consistency.

7. Compatible Consistent Sets

Let f_1 and f_2 be nonconstant polynomials over K. If f_1 and f_2 have a common composite h then, for any $\alpha \in \overline{K}$, the h-fiber $\{\beta \in \overline{K} : h(\beta) = h(\alpha)\}$

is a finite subset of $\overline{K}$ which is simultaneously a union of f_1-fibers and a union of f_2-fibers. We generalize this to arbitrary f_1 and f_2 (which might not have a common composite) as follows:

Definition 7.1. A nonempty finite subset of $\overline{K}$ is *compatible* if it is simultaneously a union of f_1-fibers and a union of f_2-fibers.

We will show that, if there is a compatible set, then there is a common composite precisely when a certain easily checkable condition is met. To motivate this extra condition, assume again that f_1 and f_2 have a common composite h. For each $a \in \overline{K}$, let $\ell(a)$ be the ramification index of $x = a$ in the extension $\overline{K}(x)/\overline{K}(h(x))$; in other words, $\ell(a)$ is the multiplicity of $x = a$ as a root of $h(x) - h(a)$. Likewise, let $m_i(a)$ be the ramification index of $x = a$ in the extension $\overline{K}(x)/\overline{K}(f_i(x))$. Then $m_i(a)$ divides $\ell(a)$, and moreover if $a, b \in \overline{K}$ satisfy $f_i(a) = f_i(b)$ for some i then the ramification index of $f_i(x) = f_i(a)$ in $\overline{K}(f_i(x))/\overline{K}(h(x))$ is

$$\frac{\ell(a)}{m_i(a)} = \frac{\ell(b)}{m_i(b)}.$$

In general, when f_1 and f_2 are not assumed to have a common composite, we make the following definition. Again, $m_i(a)$ is the ramification index of $x = a$ in the extension $\overline{K}(x)/\overline{K}(f_i(x))$.

Definition 7.2. A subset $A \subseteq \overline{K}$ is *consistent* if there is a function ℓ on A such that

(1) for each $a \in A$, $\ell(a)$ is a positive integer multiple of both $m_1(a)$ and $m_2(a)$; and
(2) for $a, b \in A$ and $i \in \{1, 2\}$, if $f_i(a) = f_i(b)$ then $\ell(a)/m_i(a) = \ell(b)/m_i(b)$.

The above discussion implies

Proposition 7.3. *If f_1 and f_2 have a common composite of degree n, then every element of $\overline{K}$ is contained in a compatible set of size at most n, and every subset of $\overline{K}$ is consistent via the labeling defined by the ramification index in $\overline{K}(x)/(\overline{K}(f_1) \cap \overline{K}(f_2))$.*

We now prove a converse result, which implies Theorem 1.1:

Theorem 7.4. *If there is a compatible consistent set A, then f_1 and f_2 have a common composite over K. Explicitly, if $\ell : A \to \mathbb{Z}$ is a consistent labeling on A, then $h := \prod_{a \in A}(x - a)^{\ell(a)}$ is a common composite over $\overline{K}$.*

Proof. We may assume f_1 and f_2 are monic. Let $B_1 = \{f_1(a) : a \in A\}$, and for each $b \in B_1$ pick an element $a_b \in A$ with $f_1(a_b) = b$. Let $A_1 = \{a_b : b \in$

$B_1\}$. Now we compute

$$\begin{aligned}\prod_{a\in A}(x-a)^{\ell(a)} &= \prod_{\widehat{a}\in A_1}\prod_{\substack{a\in A\\ f_1(a)=f_1(\widehat{a})}}(x-a)^{\ell(a)}\\ &= \prod_{\widehat{a}\in A_1}\Big(\prod_{\substack{a\in A\\ f_1(a)=f_1(\widehat{a})}}(x-a)^{m_1(a)}\Big)^{\ell(\widehat{a})/m_1(\widehat{a})}\\ &= \prod_{\widehat{a}\in A_1}(f_1(x)-f_1(\widehat{a}))^{\ell(\widehat{a})/m_1(\widehat{a})}\\ &= \Big(\prod_{\widehat{a}\in A_1}(x-f_1(\widehat{a}))^{\ell(\widehat{a})/m_1(\widehat{a})}\Big)\circ f_1(x),\end{aligned}$$

where the two middle equalities hold because A is consistent and compatible, respectively. Thus, the polynomial $h := \prod_{a\in A}(x-a)^{\ell(a)}$ is a composite of f_1 over $\overline{K}$; but likewise it is a composite of f_2 over $\overline{K}$, so it is a common composite over $\overline{K}$. It follows by Theorem 2.1 that f_1 and f_2 have a common composite over K. □

This result has several consequences. For one thing, it gives yet another proof of the first part of Theorem 3.1, namely that $K(f_1)\cap K(f_2)\neq K$ implies f_1 and f_2 have a common composite: for in this case $K(f_1)\cap K(f_2) = K(h)\neq K$, and the proof of Proposition 7.3 shows there are compatible consistent subsets of $\overline{K}$. More importantly, in Theorem 7.4 we exhibited a specific common composite h over $\overline{K}$. The shape of this polynomial h enables us to control the ramification in a minimal-degree common composite in terms of the ramification in f_1 and f_2; in a subsequent paper we will show how this can be used to prove that two polynomials have no common composite.

Corollary 7.5. *If $f_1, f_2\in K[x]$ have a common composite, then the ramification index of $x=a$ in $\overline{K}(x)/(\overline{K}(f_1)\cap\overline{K}(f_2))$ is a divisor of $\ell(a)$, for any consistent labeling ℓ on any compatible set containing a.*

Another consequence of Theorem 7.4 is a description of the minimal compatible sets, in case there is a common composite. We need a lemma before stating the result:

Lemma 7.6. *If f_1 and f_2 have a common composite, and $A\subset\overline{K}$ is a minimal compatible set, then there is a consistent labeling $\ell_0 : A\to\mathbb{Z}$ such that every consistent labeling $\ell : A\to\mathbb{Z}$ has the form $\ell = n\ell_0$ with n a positive integer.*

Proof. Pick some $a\in A$ and some consistent labeling $\ell : A\to\mathbb{Z}$. Since A is a minimal compatible set, for any $b\in A$ there is a finite sequence $a_1,\dots,a_r$ of elements of A, where $a=a_1$ and $b=a_r$, such that (for each j) a_j and a_{j+1} have the same image under either f_1 or f_2. If $f_i(a_j)=f_i(a_{j+1})$ then $\ell(a_j)/m_i(a_j)=\ell(a_{j+1})/m_i(a_{j+1})$, so $\ell(a_{j+1})=\ell(a_j)m_i(a_{j+1})/m_i(a_j)$.

Thus, we can express $\ell(b)$ as $\ell(a)$ times a rational number whose numerator and denominator are products of values of m_1 and m_2. It follows that any other compatible labeling must be a rational number times ℓ. Conversely, a rational multiple of ℓ is a consistent labeling if and only $\ell(b)/m_i(b) \in \mathbb{Z}$ for every $b \in A$ and $i \in \{1, 2\}$. The result follows. □

Corollary 7.7. *Suppose f_1 and f_2 have a common composite, and let h be a common composite of minimal degree. Then the minimal compatible sets $A \subset \overline{K}$ are precisely the sets $\{b \in \overline{K} : h(b) = h(a)\}$ with $a \in \overline{K}$. Moreover, if ℓ_0 is the minimal consistent labeling on A, then $\ell_0(a)$ is the multiplicity of $x = a$ as a root of $h(x) - h(a)$, and furthermore $\sum_{a \in A} \ell_0(a) = \deg(h)$. Finally, writing $\widehat{h} := \prod_{a \in A}(x - a)^{\ell_0(a)}$, we have $\widehat{h}(x) - \widehat{h}(0) \in K[x]$, and there is a degree-one $\mu \in K[x]$ such that $\widehat{h}(x) = \widehat{h}(0) + \mu(h(x))$.*

Proof. For any $a \in \overline{K}$, let $\ell(a)$ denote the ramification index of $x = a$ in $\overline{K}(x)/\overline{K}(h(x))$. The fiber $S = \{b \in \overline{K} : h(b) = h(a)\}$ is compatible, and ℓ is a consistent labeling on S. Note that $\sum_{b \in S} \ell(b) = \deg(h)$. Let A be a minimal compatible set contained in S. Then Theorem 7.4 implies that $\widehat{h} := \prod_{b \in A}(x - b)^{\ell(b)}$ is a common composite over $\overline{K}$. By minimality of $\deg(h)$, we must have $\deg(h) \leq \deg(\widehat{h})$, so $A = S$ and $\deg(h) = \deg(\widehat{h})$. Likewise, ℓ must be the minimal consistent labeling on A, since otherwise using a smaller labeling in Theorem 7.4 would produce a common composite of degree lower than $\deg(h)$. Now $\widehat{h}(x) = \widehat{h}(0) + \mu(h(x))$ for some degree-one $\mu \in \overline{K}[x]$. Since $\widehat{h}$ is monic and $h \in K[x]$, the leading coefficient of μ must be in K. Since the constant terms of both h and $(\widehat{h}(x) - \widehat{h}(0))$ are in K, we have $\mu(0) \in K$. This completes the proof. □

8. Inconsistent Sets

In this section we give examples of $f_1, f_2 \in K[x]$ for which there is an inconsistent subset of $\overline{K}$. By Proposition 7.3, this implies there is no common composite. We begin by reworking Example 6.2.

Example 8.1. *Consider $f_1 = x^2 - x$ and $f_2 = x^3 - x^2$ over any field K. We claim that $\{0, 1\}$ is inconsistent. For, suppose there were a function ℓ on $\{0, 1\}$ satisfying the properties of Definition 7.2. Since $f_1(0) = f_1(1)$ and $f_2(0) = f_2(1)$, we would have*

$$\frac{\ell(0)}{m_1(0)} = \frac{\ell(1)}{m_1(1)} \qquad \textit{and} \qquad \frac{\ell(0)}{m_2(0)} = \frac{\ell(1)}{m_2(1)},$$

so

$$\frac{m_1(0)}{m_1(1)} = \frac{\ell(0)}{\ell(1)} = \frac{m_2(0)}{m_2(1)}.$$

But $m_1(0) = m_1(1) = m_2(1) = 1$ and $m_2(0) = 2$, contradiction.

In the above example the set $\{0, 1\}$ is compatible, but this property is not used in proving there is no common composite. (By contrast, we crucially used this property when we treated these polynomials in Example 6.2.)

It is not difficult to construct similar examples involving noncompatible inconsistent sets—for instance, one could replace f_1 by $(x^2-x)(x^2-x-1)$.

Our next example involves a larger inconsistent set.

Example 8.2. *Consider $f_1 = x^3 + x + 1$ and $f_2 = x^4 + x + 1$ in $\mathbb{F}_3[x]$. We claim that $A := \{0, -1, i, i-1\}$ is inconsistent. For, suppose there is a consistent labeling ℓ on A. Since $f_1(i) = f_1(0) = 1$ and $m_1(i) = m_1(0) = 1$, we have $\ell(i) = \ell(0)$. Since $f_2(0) = f_2(-1) = 1$ and $m_2(0) = 1$ and $m_2(-1) = 3$, we have $\ell(-1) = 3\ell(0)$. Since $f_1(-1) = f_1(i-1) = -1$ and $m_1(-1) = m_1(i-1) = 1$, we have $\ell(i-1) = \ell(-1)$. Since $f_2(i-1) = f_2(i) = i-1$ and $m_2(i-1) = m_2(i) = 1$, we have $\ell(i) = \ell(i-1)$. Thus*

$$\ell(i) = \ell(i-1) = \ell(-1) = 3\ell(0) = 3\ell(i),$$

contradicting the fact that $\ell(i)$ is nonzero.

These two examples generalize as follows:

Theorem 8.3. *Suppose $c_1, \ldots, c_{2d} \in \overline{K}$ satisfy $f_1(c_i) = f_1(c_{i+1})$ for odd i and $f_2(c_i) = f_2(c_{i+1})$ for even i (where $c_{2d+1} := c_1$). If f_1 and f_2 have a common composite then*

$$1 = \prod_{i=1}^{d} \frac{m_1(c_{2i-1})}{m_2(c_{2i-1})} \frac{m_2(c_{2i})}{m_1(c_{2i})}. \tag{2}$$

Proof. Suppose f_1 and f_2 have a common composite, and let ℓ be a consistent labeling on $\overline{K}$. Then $\ell(c_i)/\ell(c_{i+1})$ equals $m_1(c_i)/m_1(c_{i+1})$ if i odd, and equals $m_2(c_i)/m_2(c_{i+1})$ otherwise. The desired formula follows by computing the product of all $2d$ terms $\ell(c_i)/\ell(c_{i+1})$. □

We do not know how often one can satisfy the criteria of this Proposition. Namely, if one begins with a value c_1 such that $m_1(c_1) > 1$ (i.e., $f_1'(c_1) = 0$), then how likely is it that there exist $c_2, \ldots, c_{2d}$ such that $f_1(c_i) = f_1(c_{i+1})$ for odd i and $f_2(c_i) = f_2(c_{i+1})$ for even i? If such c_i do exist, one would expect that 'usually' Equation (2) is not satisfied. However, we suspect that it is rare for such c_i to exist.

As an extreme example in this direction, we note that there are polynomials f_i for which $\overline{K}$ is consistent, even though the f_i have no common composite:

Example 8.4. *Let $f_1 = x^2$ and $f_2 = (x-1)^2$ be polynomials over $\mathbb{Q}$. Then $m_i(\alpha) = 1$ for all $\alpha \in \overline{\mathbb{Q}}$ and $i \in \{1, 2\}$, except that $m_1(0) = 2$ and $m_2(1) = 2$. Thus, the constant function $\ell = 2$ is a consistent labeling on $\overline{\mathbb{Q}}$. However, any compatible subset S of $\overline{\mathbb{Q}}$ would have to be closed under the map $x \mapsto -x$ (since $-x$ and x are in the same fiber of f_1), and likewise S would be closed under $x \mapsto 2-x$. But then S would be closed under the composite map $x \mapsto 2+x$, contradicting finiteness of S. Hence there is no compatible subset of $\overline{\mathbb{Q}}$, so f_1 and f_2 have no common composite.*

9. Derivatives

In the previous section we gave a method which, in certain special cases, enables one to prove that two polynomials f_1 and f_2 have no common composite. In this section we give a more robust method for this.

Proposition 9.1. *Suppose that $f_1, f_2 \in K[x]$ have a common composite h, and suppose $\alpha, \beta \in \overline{K}$ satisfy $h'(\alpha)h'(\beta) \neq 0$ and $f_i(\alpha) = f_i(\beta)$ for both $i = 1$ and $i = 2$. Then $f_1'(\alpha)f_2'(\beta) = f_1'(\beta)f_2'(\alpha)$.*

Proof. Writing $h = F_i \circ f_i$ with $F_i \in K[x]$, we have

$$\begin{aligned} h'(\alpha) &= F_i'(f_i(\alpha)) \cdot f_i'(\alpha) \\ h'(\beta) &= F_i'(f_i(\beta)) \cdot f_i'(\beta) = F_i'(f_i(\alpha)) \cdot f_i'(\beta). \end{aligned}$$

Since $h'(\beta) \neq 0$, this implies

$$\frac{h'(\alpha)}{h'(\beta)} = \frac{f_i'(\alpha)}{f_i'(\beta)}.$$

Since the left side of this equation does not depend on i, the result follows. □

Example 9.2. *Consider $f_1 = x^3$ and $f_2 = x^2 + x$ over $K = \mathbb{F}_2$. Letting ω be a primitive cube root of unity in $\overline{K}$, we see that $f_i(\omega^j) = 1$ for each $i, j \in \{1, 2\}$. Since $f_1'(\omega)f_2'(\omega^2) \neq f_1'(\omega^2)f_2'(\omega)$, Proposition 9.1 implies that every common composite h of f_1 and f_2 must satisfy $h'(\omega)h'(\omega^2) = 0$. In this instance, we know by Proposition 4.2 that f_1 and f_2 have a common composite, and that a minimal-degree common composite is $\widehat{h} := (x^4 + x)^3$. And indeed, $\widehat{h}'(\omega) = \widehat{h}'(\omega^2) = 0$.*

This example illustrates how to use Proposition 9.1 to prove a property of common composites, assuming such composites exist. We now build this into a criterion enabling us to prove nonexistence of a common composite in some cases.

Lemma 9.3. *Suppose $f_1, f_2 \in K[x] \setminus K[x^p]$ have a common composite, and let h be a minimal-degree common composite. For any $\alpha \in \overline{K}$ such that $[K(\alpha) : K]$ is divisible by a prime greater than $\max(\deg(f_1), \deg(f_2))$, we have $h'(\alpha) \neq 0$.*

Proof. By Proposition 7.3, there is a compatible consistent set $A \subset \overline{K}$ containing α. Assume A is the minimal such set; then A consists of all $\beta \in \overline{K}$ for which there is a finite sequence of elements of $\overline{K}$, starting with α and ending with β, such that consecutive members of the sequence have the same image under either f_1 or f_2. Our condition on the degrees implies that the large prime dividing $[K(\alpha) : K]$ also divides $[K(\gamma) : K]$ for each γ in the sequence, so this prime divides $[K(\beta) : K]$, whence $f_i'(\beta) \neq 0$. Thus $\ell = 1$ is the minimal consistent labeling on A, so Corollary 7.7 implies that $\widehat{h}(x) := \prod_{a \in A}(x - a)$ satisfies $\widehat{h}(x) - \widehat{h}(0) = \mu(h(x))$ for some degree-one $\mu \in K[x]$. In particular, since $\widehat{h}'(\alpha) \neq 0$, we must have $h'(\alpha) \neq 0$. □

Combining the previous two results gives our desired criterion:

Corollary 9.4. *Suppose* $f_1, f_2 \in K[x] \setminus K[x^p]$ *and* $\alpha, \beta \in \overline{K}$ *satisfy* $f_i(\alpha) = f_i(\beta)$ *for both* $i = 1$ *and* $i = 2$*, and also* $[K(\alpha) : K]$ *is divisible by a prime greater than* $\max(\deg(f_1), \deg(f_2))$*. If* $f_1'(\alpha)f_2'(\beta) \neq f_1'(\beta)f_2'(\alpha)$ *then* f_1 *and* f_2 *have no common composite.*

Example 9.5. *Consider* $f_1 = x^4+x^3$ *and* $f_2 = x^6+x^2+x$ *over* $\mathbb{F}_2$*. One can check that* $\psi(x) := x^{14}+x^{10}+x^9+x^8+x^7+x^6+x^4+x+1$ *is irreducible over* $\mathbb{F}_2$*. For any root* α *of* ψ*, let* $\beta = \alpha^{128}$*. Then* $f_i(\alpha) = f_i(\beta)$ *for each* i*, but* $f_1'(\alpha)f_2'(\beta) \neq f_1'(\beta)f_2'(\alpha)$*, so Corollary 9.4 implies the* f_i *have no common composite.*

Our proof of Corollary 9.4 generalizes at once to prove the following:

Theorem 9.6. *For* $f_1, f_2 \in K[x] \setminus K[x^p]$*, suppose* $c_1, \dots, c_{2d} \in \overline{K}$ *satisfy* $f_1(c_i) = f_1(c_{i+1})$ *for odd* i *and* $f_2(c_i) = f_2(c_{i+1})$ *for even* i *(where we define* $c_{2d+1} := c_1$*). Suppose further that* $[K(c_1) : K]$ *is divisible by a prime greater than* $\max(\deg(f_1), \deg(f_2))$*. If*

$$\prod_{i=1}^{d} \left(f_1'(c_{2i-1})f_2'(c_{2i})\right) \neq \prod_{i=1}^{d} \left(f_2'(c_{2i-1})f_1'(c_{2i})\right)$$

then f_1 *and* f_2 *have no common composite.*

One can check that there is no loss in only applying this result when the c_i are distinct.

Example 9.7. *Consider* $f_1 = x^2 + x$ *and* $f_2 = x^4 + x^3 + x$ *over* $\mathbb{F}_2$*. The two primitive cube roots of unity have the same image as one another under both* f_1 *and* f_2*, but they have degree* 2 *over* $\mathbb{F}_2$ *so the above result does not apply. For* $d < 5$*, this is the only choice of distinct* c_i*'s such that* $f_1(c_i) = f_1(c_{i+1})$ *for odd* i *and* $f_2(c_i) = f_2(c_{i+1})$ *for even* i*. But for* $d = 5$ *we can choose* $(c_1, \dots, c_{10}) := (w, w^{268}, w^4, w^{49}, w^{16}, w^{196}, w^{64}, w^{784}, w^{256}, w^{67})$ *where* $w^{10} + w^9 + w^4 + w^2 = 1$*. Since* $[\mathbb{F}_2(w) : \mathbb{F}_2] = 10$*, these* c_i *satisfy all the hypotheses of Theorem 9.6, so* f_1 *and* f_2 *have no common composite.*

We suspect that Theorem 9.6 applies to 'most' pairs of polynomials over a finite field. This intuition has been reinforced by various examples we have computed. Our intuition is based on the following reasoning: the c_i are defined by $2d$ equations in $2d$ variables, so 'at random' we expect to find solutions. Specifically, we can apply the fiber-finding algorithm to the indeterminate $\alpha = t$ in $K[t]$. This gives a polynomial $r_{2d} \in K[t, x]$ such that $r_{2d}(c_1, c_1) = 0$, narrowing the choices for c_1 to a finite set. It may happen that no such choice for c_1 leads to a solution for $c_2, \dots, c_{2d}$ with the c_i distinct, but this seems unlikely to happen except in unusual circumstances. Finally, as we vary d, it seems there should be some d for which a corresponding c_1 is defined over an extension of K of degree divisible by a large prime, and moreover 'at random' the products of derivatives expressed in Theorem 9.6 are almost certainly distinct.

Unfortunately, there are cases where two polynomials have no common composite, but this nonexistence cannot be proved with Theorem 9.6.

Example 9.8. *Consider $f_1 = x^2 + x$ and $f_2 = x^6 + x$ over $\mathbb{F}_2$. Since $f_i'(x) = 1$, there are no c_j's satisfying the hypotheses of Theorem 9.6.*

In a subsequent paper we will develop further methods for proving nonexistence of a common composite, and in particular we will show that the polynomials in the above example have no common composite.

References

[1] V. Alexandru and N. Popescu, *On subfields of $k(x)$*, Rend. Sem. Mat. Univ. Padova **75** (1986), 257–273.

[2] A. J. Berkson, *Polynomial subfields over perfect fields*, Nordisk Mat. Tidskr. **21** (1973), 29–30.

[3] A. Bremner and P. Morton, *Polynomial relations in characteristic p*, Quart. J. Math. Oxford (2) **29** (1978), 335–347.

[4] F. Dorey and G. Whaples, *Prime and composite polynomials*, J. Algebra **28** (1974), 88–101.

[5] H. T. Engstrom, *Polynomial substitutions*, Amer. J. Math. **63** (1941), 249–255.

[6] M. D. Fried and R. E. MacRae, *On the invariance of chains of fields*, Illinois J. Math. **13** (1969), 165–171.

[7] H. Levi, *Composite polynomials with coefficients in an arbitrary field of characteristic zero*, Amer. J. Math. **64** (1942), 389–400.

[8] A. McConnell, *Polynomial subfields of $k(x)$*, J. Reine Angew. Math. **266** (1974), 136–139.

[9] E. Noether, *Körper und Systeme rationaler Funktionen*, Math. Ann. **76** (1915), 161–196.

[10] J. F. Ritt, *Prime and composite polynomials*, Trans. Amer. Math. Soc. **23** (1922), 51–66.

[11] A. Schinzel, Selected Topics on Polynomials, University of Michigan Press, Ann Arbor, 1982.

[12] A. Schinzel, Polynomials with Special Regard to Reducibility, Cambridge University Press, Cambridge, 2000.

[13] J.-P. Serre, Local Fields, Springer-Verlag, New York, 1979.

[14] P. Tortrat, *Sur la composition des polynômes*, Colloq. Math. **55** (1988), 329–353.

[15] U. Zannier, *Ritt's second theorem in arbitrary characteristic*, J. Reine Angew. Math. **445** (1993), 175–203.

Center for Communications Research, 805 Bunn Drive, Princeton, NJ 08540-1966, USA.

E-mail address: `beals@idaccr.org`

Center for Communications Research, 4320 Westerra Court, San Diego, CA 92121-1967, USA.

E-mail address: `jlwether@alum.mit.edu`

Center for Communications Research, 805 Bunn Drive, Princeton, NJ 08540-1966, USA.

E-mail address: `zieve@math.rutgers.edu`

URL: `http://www.math.rutgers.edu/~zieve/`