Contemporary Mathematics

Wedderburn Polynomials over Division Rings, I

T. Y. Lam and André Leroy

Abstract. A Wedderburn polynomial over a division ring K is a minimal polynomial of an algebraic subset of K. Such a polynomial is always a product of linear factors over K, although not every product of linear polynomials is a Wedderburn polynomial. In this paper, we establish various properties and characterizations of Wedderburn polynomials over K, and show that these polynomials form a complete modular lattice that is dual to the lattice of full algebraic subsets of K. Throughout the paper, we work in the general setting of an Ore skew polynomial ring $K[t, S, D]$, where S is an endomorphism of K and D is an S-derivation on K.

§1. Introduction

The main purpose of this paper is to develop the theory of a class of polynomials over a division ring K, which we call *Wedderburn polynomials* (or simply W-polynomials). Roughly speaking, a W-polynomial over K is one which has "enough zeros" in K. (For a more precise definition, see (3.1).) In the case when K is a field, W-polynomials are simply those of the form $(t-a_1)\cdots(t-a_n)$, where $a_1, \dots, a_n$ are *distinct* elements of K. In the general case of a division ring K, a W-polynomial still has the form $(t-a_1)\cdots(t-a_n)$, although the a_i's need no longer be distinct. And even if the a_i's are distinct, $(t-a_1)\cdots(t-a_n)$ need not be a W-polynomial. The recognition of a W-polynomial turns out to be a very interesting problem over a division ring K.

The early work of Wedderburn [We] (ca. 1921) showed that, if $a \in K$ is an algebraic element over the center F of K, then the minimal polynomial of a over F (in the usual sense) is a W-polynomial in $K[t]$ (and in particular splits completely over K). This classical result of Wedderburn has led to much research on $K[t]$, and has found important applications to the study of subgroups and quotient groups of the multiplicative group K^*, central simple algebras of low degrees and crossed product algebras, PI-theory, Vandermonde and Wronskian matrices, Hilbert 90 theorems, polynomial dependence, and the theory of ordered division rings, etc.

1991 *Mathematics Subject Classification.* Primary 16D40, 16E20, 16L30; Secondary 16D70, 16E10, 16G30.

The research reported in this paper was partially supported by a grant from NSA (T.Y.L.) and a faculty grant from the Université d'Artois at Lens (A.L.).

For representative literature along these lines, the reader may consult [Al], [HR], [Ja_3], [La_1], [LL_1]-[LL_4], [Ro_1]-[Ro_3], [RS_1], [RS_2], [Se], and [Tr].

Our definition of W-polynomials was directly inspired by the afore-mentioned work of Wedderburn, although our W-polynomials will have coefficients in K, instead of in F. These W-polynomials are rather rich in structure, and seem to be quite basic in working with the polynomial theory over K. Some examples and a few characterizations of W-polynomials are given in §3. In §4, we introduce two of the main technical tools for analyzing W-polynomials; these are (essentially) self-maps from K to K, called respectively the *Φ-transform* and the *λ-transform.* These transforms are then used in §5 to derive results on factors and products of W-polynomials, and on idealizers of certain left ideals in the (Ore) polynomial ring. The notion of W-polynomials in the quadratic case turns out to be closely related to the solution of certain "metro-equations" in division rings; some applications in this direction are presented in §6. In §7, we establish a basic Rank Theorem (7.3), which relates the ranks of the union and the intersection of two algebraic sets in the spirit of the dimension equation in the theory of finite-dimensional vector spaces. We then show in §8 that the set of W-polynomials over K (suitably augmented if necessary) has the natural structure of a complete modular lattice, and that furthermore, this lattice is dual to the lattice of sets of roots of polynomials over K. The paper concludes with two sections on questions, examples, and applications.

It is relevant to point out that our W-polynomials are a special case of the class of "completely reducible" polynomials introduced by Ore in his seminal paper [Or]. While Ore's completely reducible polynomials are lcm's (least common multiples) of irreducible polynomials (over K), our W-polynomials are lcm's of *linear* polynomials. The use of linear polynomials enables us to relate the W-polynomials readily to their root sets, and thereby get the lattice duality mentioned in the last paragraph. Retrospectively, we find it rather surprising that this viewpoint was not exploited by Ore.

Following Ore [Or], we work in the setting of skew polynomials (rather than just ordinary polynomials) over the division ring K. This added degree of generality is definitely worthwhile considering that skew polynomials have become increasingly important with the growing interests in quantized structures and noncommutative geometry. The basic mechanism of skew polynomials is recalled in §2, where we also set up the terminology and general framework for the paper. As a matter of fact, once the general mechanism of skew polynomials is set in place, the work of developing the theory of W-polynomials is no more complicated in the skew case than in the ordinary case. Therefore, although skew polynomials may appear difficult to some, to try to avoid them in this paper would be an unnecessary sacrifice of generality. In giving basic examples for the theory, however, we will not hesitate to go back to the case of ordinary polynomials, where the indeterminate commutes with all coefficients.

As will be clear to the reader, this work builds on [La_1] and [LL_1], the basic parts of which thus constitute the main prerequisites for reading this paper. A certain technical result is also needed from [LL_4]. Otherwise, we have tried to make the results here as self-contained as possible. Further results on Wedderburn polynomials will be presented in a second part of this paper, [LL_5].

§2. Recapitulation

To work with skew polynomials, we start with a triple (K, S, D), where K is a division ring, S is a ring endomorphism of K, and D is an S-derivation on K. (The latter means that D is an additive endomorphism of K such that $D(ab) = S(a)D(b) + D(a)b$, $\forall a, b \in K$.) In this general setting, we can form $K[t, S, D]$, the *Ore skew polynomial ring* consisting of ("left") polynomials of the form $\sum b_i t^i$ $(b_i \in K)$ which are added in the usual way and multiplied according to the rule

$$tb = S(b)t + D(b) \quad (\forall\, b \in K). \tag{2.1}$$

In case $D = 0$ (resp. $S = I$), we'll write $K[t, S]$ (resp. $K[t, D]$) for the skew polynomial ring. Of course, when $(S, D) = (I, 0)$ (we'll refer to this as the "classical case"), $K[t, S, D]$ boils down to the usual polynomial ring $K[t]$ with a *central* indeterminate t. *Throughout this paper, we'll write* $R := K[t, S, D]$. It is easy to check that R admits an euclidean algorithm for right division, so R is a principal left ideal domain.

In working with R, it is important to be able to "evaluate" a polynomial $f(t) = \sum b_i t^i$ at any scalar $a \in K$, that is, to define $f(a)$. Following our earlier work [LL$_1$], we take $f(a)$ to be $\sum b_i N_i(a)$, where the "ith power function" N_i is defined inductively by

$$N_0(a) = 1, \quad \text{and} \quad N_i(a) = S(N_{i-1}(a))a + D(N_{i-1}(a)) \quad \forall\, a \in K. \tag{2.2}$$

That this gives the "right" definition of $f(a)$ is seen from the validity of the *Remainder Theorem* [LL$_1$: (2.4)]: there is a unique $q \in R$ such that

$$f(t) = q(t)(t - a) + f(a). \tag{2.2$'$}$$

From this, it follows immediately that $f(a) = 0$ iff $t - a$ is a right factor of $f(t)$. This important fact will be used freely throughout the paper. In case $f(a) = 0$, we say that a is a (right) root, or (right) zero, of f. (Throughout this paper, the word "root" will always mean right root.)

Readers of our earlier papers have sometimes been deterred by the apparently tricky definition of evaluation in (2.2). For these readers, the following remarks should bring some solace. First, it is entirely possible to take $f(a)$ $(\forall a \in K)$ as *defined* by the equation (2.2)$'$. Once this is done, it is not difficult to check that the second formula in (2.2) is simply dictated upon us. Or. from a module-theoretic viewpoint, if we identify the cyclic R-module $R/R(t - a)$ with K, then the action of $f(t)$ on the cyclic generator 1 is simply given by $f(t) \cdot 1 = f(a)$. Lastly, in case $D = 0$, the definition (2.2) simplifies down to $N_i(a) = S^{i-1}(a) \cdots S(a)a$ $(\forall i)$, which is a familiar expression in many ways, e.g. from the norm formula for cyclic Galois extensions. For more detailed explanations of these viewpoints, see [La$_1$] and [LL$_1$, LL$_2$].

Another remarkable fact about evaluating skew polynomials at scalars is the "Product Formula" [LL$_1$: (2.7)] for evaluating $f = gh$ at any $a \in K$:

$$(gh)(a) = \begin{cases} 0 & \text{if } h(a) = 0, \\ g(a^{h(a)})h(a) & \text{if } h(a) \neq 0. \end{cases} \tag{2.3}$$

Here, for any $c \in K^*$, a^c denotes $S(c)ac^{-1} + D(c)c^{-1}$, which is called the (S, D)-*conjugate* of a (by c). With this general conjugation notation, it is easy to verify by a direct calculation that

$$(a^c)^d = a^{dc} \quad \text{for any } c,\, d \in K^*. \tag{2.4}$$

However, we must caution the reader that, in general, $(ab)^c$ need *not* be equal to $a^c b^c$. Also, in using the expression a^c, we have to constantly keep in mind that this is the (S, D)-conjugacy notation, not to be confused with the usual exponentiation (meaningful in the division ring K when the exponent is an integer). For instance, the "usual" exponentiation a^{-1} would mean the inverse of a, while the (S, D)-conjugate of a by -1 is $S(-1)a(-1)^{-1} + D(-1)(-1)^{-1}$, which is just a! (More generally, it is useful to note that $a^{-c} = a^c$.)

In the following, we shall write

$$\Delta^{S,D}(a) := \{\, a^c : \; c \in K^* \,\}; \tag{2.5}$$

this is called the (S, D)-*conjugacy class* of a. All such classes form a partition of K. For instance, $\Delta^{S,D}(0)$ is the set of all logarithmic derivatives $\{D(c)c^{-1} : \; c \in K^*\}$. And, in the classical case, $\Delta^{I,0}(a)$ is just the "usual" conjugacy class

$$\Delta(a) = \{cac^{-1} : \; c \in K^*\}.$$

A routine extension of the Gordon-Motzkin Theorem (using the Product Formula) shows the following (cf. [La_1: Thm. 4] and [La_2: (16.4)]):

Proposition 2.6. (1) *If $f \in R$ has degree n, then f can have roots in at most n (S, D)-conjugacy classes of K.* (2) *If $f(t) = (t - a_1) \cdots (t - a_n)$, then each root of f in K is (S, D)-conjugate to some a_i.*

Next, we introduce two basic notations for this paper. For $g \in R$, let

$$V(g) := \{a \in K : \; g(a) = 0\}, \tag{2.7}$$

and for any subset $\Delta \subseteq K$, let $I(\Delta)$ be the left ideal

$$\{g \in R : \; g(\Delta) = 0\}. \tag{2.8}$$

We'll say that the set Δ is *algebraic* (or, more precisely, (S, D)-algebraic) if $I(\Delta) \neq \{0\}$. In this case, the monic generator of $I(\Delta)$ is called the *minimal polynomial* of Δ; we denote it by f_Δ. The degree of f_Δ is called the *rank* of the algebraic set[1] Δ; we denote it by $\text{rk}(\Delta)$. According to the Remainder Theorem, f_Δ is just the (monic) "llcm" (least left common multiple) of the linear polynomials $\{t - d : \; d \in \Delta\}$. As in [$\text{La}_1$: Lemma 5], it is easy to see that f_Δ has always the form $(t - a_1) \cdots (t - a_n)$ where each a_i is (S, D)-conjugate to some element of Δ.

Of course, all of the above was inspired in part by classical algebraic geometry. Going a little further, we get a theory of *polynomial dependence* (or P-dependence for short) for the elements of K. By definition, an element b is P-dependent on an algebraic set Δ if $g(b) = 0$ for every $g \in I(\Delta)$. We see easily that the set of elements P-dependent on Δ is precisely $V(f_\Delta)$, which we shall henceforth call the "P-closure" of Δ and denote by $\overline{\Delta}$. As in [La_1], we can also define *P-independence*

[1] For technical reasons, it is convenient to define the rank of a *non-algebraic* set too: it is simply taken to be the symbol ∞. The "minimal polynomial" for a non-algebraic set is taken to be 0.

and the notion of a *P-basis* for an algebraic set Δ in a natural manner. The cardinality of a P-basis for Δ is just $\mathrm{rk}(\Delta)$. If $\{b_1, \ldots, b_r\}$ is a P-basis of Δ, then f_Δ is in fact the llcm of the linear polynomials $\{t - b_i : 1 \le i \le r\}$. We refer the reader to [La_1: §4] (see also [Tr]) for the rudiments of the theory of P-dependence. Although this theory was developed in [La_1] in the case $D = 0$, it holds word-for-word also in the (S, D)-case.)

§3. Wedderburn Polynomials: Examples and Characterizations

We now come to two of the central themes of the paper.

Definition 3.1. An algebraic set $\Delta \subseteq K$ is said to be *full* if $\Delta = \overline{\Delta}\,(= V(f_\Delta))$. A (monic) polynomial $f \in R$ is said to be a *Wedderburn polynomial* (or simply a W-polynomial) if $f = f_{V(f)}$.

From (3.1), it is easy to see that an algebraic set Δ is full iff $\Delta = V(f)$ for some nonzero $f \in R$, and a polynomial $f \in R$ is a W-polynomial iff $f = f_\Delta$ for some algebraic set Δ. Let us first give some examples of full algebraic sets.

Examples 3.2.

(1) The empty set $\emptyset$ is a full algebraic set, with minimal polynomial 1 (and rank 0). In particular, 1 is a W-polynomial.

(2) Any singleton $\{a\}$ is also always a full algebraic set, with minimal polynomial $t - a$. Thus, any monic linear polynomial is a W-polynomial.

(3) Consider a doubleton set $\Delta = \{a, b\}$. By the product Formula (2.3), it is easy to see that Δ has minimal polynomial

$$f_\Delta(t) = \big(t - b^{b-a}\big)(t - a).$$

Thus, any quadratic of this form is a W-polynomial. Notice that, by symmetry, we have automatically

$$(3.3) \qquad \big(t - b^{b-a}\big)(t - a) = \big(t - a^{a-b}\big)(t - b).$$

However, a doubleton set may not be full, as the example of $\{i, j\}$ over the quaternions shows. (The P-closure of $\{i, j\}$ is the set of all quaternions of square -1.)

(4) *If an (S, D)-conjugacy class $\Delta := \Delta^{S,D}(a)$ happens to be algebraic, then Δ is full.* In fact, if $f := f_\Delta$, then as we have noted in §2, there is a splitting $f(t) = (t - a_1) \cdots (t - a_n)$, where $a_i \in \Delta$. By (2.6), we have $V(f) \subseteq \Delta$, and so $\Delta = V(f)$ is full.

(5) For any algebraic set Δ, the P-closure $\overline{\Delta}$ is the smallest full algebraic set containing Δ, and $\overline{\overline{\Delta}} = \overline{\Delta}$. If $\Delta \subseteq \Delta^{S,D}(a)$, then $\overline{\Delta} \subseteq \Delta^{S,D}(a)$ as well: this follows from (2.6).

For later reference, we state two more convenient characterizations of W-polynomials. The proofs are easy, and can be found in [LL_4: (2.7)].

Proposition 3.4. *For a monic $f \in R$ of degree n, the following are equivalent*:

(1) *f is a W-polynomial*;

(2) $\mathrm{rk}(V(f)) = n$ (“*f has enough zeros*”);

(3) *For any $p \in R$, $V(f) \subseteq V(p) \Longrightarrow p \in R \cdot f$.*

Several other characterizations of W-polynomials will be given later in [LL_5]. Here, we give a list of nontrivial examples (and one non-example) of Wedderburn polynomials.

Examples 3.5.

(1) *A monic quadratic polynomial $f \in R$ is a W-polynomial iff* $\operatorname{Card} V(f) \geq 2$. Indeed, if f is a W-polynomial, then $V(f)$ has rank 2 by (3.4), so it has at least two elements. Conversely, if $V(f)$ has at least two elements, clearly no linear polynomial can vanish on $V(f)$. Therefore, f must be the minimal polynomial of $V(f)$, so f is a W-polynomial. Note that the criterion $\operatorname{Card} V(f) \geq 2$ above for f to be a W-polynomial can also be expressed by saying that f *has at least two different factorizations into a product of monic linear factors.* (Here, "different" is taken in the absolute sense.)

For instance, over the real quaternions with $(S, D) = (I, 0)$, the polynomial

$$(\dagger) \qquad f(t) = t^2 - (i + j)t - k = (t - j)(t - i)$$

has a unique root $\{i\}$ (see [La_3: Ex. 16.3, p. 181]), and hence only one factorization (into monic linear factors) as above. Thus, f is not a W-polynomial. The polynomial $g(t) = t^2 + 1$ has infinitely many roots (namely, all conjugates of i), so g is a W-polynomial, with infinitely many factorizations. Finally, the polynomial $h(t) = t^2 - it + (k + 1)$ has the factorizations

$$h(t) = [t - (i - j)]\,(t - j) = (t + j)\,[t - (i + j)],$$

so h is a W-polynomial. In fact, one can show that $V(h) = \{j,\, i + j\}$, so the above are the *only* factorizations of h into monic linear factors. (For a more general perspective on this, see (6.3) and (6.4).)

(2) If K is a field and $(S, D) = (I, 0)$, the algebraic sets are precisely the finite subsets of K. From this, it follows that the W-polynomials are the polynomials of the form $(t - a_1) \cdots (t - a_n)$, where the a_i's are distinct elements in K. These are precisely the separable, completely split polynomials over K.

(3) In general, if $f(t) \in R$ is a W-polynomial with a splitting $(t - a_1) \cdots (t - a_n)$, the a_i's need not be distinct. We shall give two such examples here. For the first one, let K be a division ring of characteristic 2 in which there exist elements $a \neq b$ with $a^2 = b^2$. Then, for $(S, D) = (I, 0)$,

$$f(t) := (t - a)^2 = t^2 - a^2$$

has both root a and root b, so f is a W-polynomial over K. For the second example (in arbitrary characteristic), let k be any field, and let $K = k(x)$ be equipped with $S = I$ and the usual derivation $D = \frac{d}{dx}$. By (2.1), $N_2(b) = b^2 + D(b)$. Therefore, for $b = x^{-1}$, we have $N_2(b) = 0$. Thus, the polynomial $t^2 \in K[t, D]$ vanishes on x^{-1} as well as on 0. By (1) above, t^2 is a W-polynomial. (For yet another example, see (6.10)(3) below.)

(4) If a_i's are distinct elements in K, $(t - a_1) \cdots (t - a_n)$ need not be a W-polynomial: see the example ($\dagger$) in (1) above.

(5) Let $F = Z(K)$ (the center of K). *If $f(t) \in F[t]$ is an irreducible polynomial with a root $a \in K$, then f is a W-polynomial over K (with respect to $(S, D) =$*

$(I,0)$). In fact, f is the minimal polynomial of a over F (in the usual field-theoretic sense), so by Wedderburn's Theorem in [We], the usual conjugacy class $\Delta = \Delta(a)$ is algebraic with $f_\Delta = f(t)$ (and by (3.3)(4), $V(f) = \Delta$). Therefore, $f \in F[t]$ is a W-polynomial over K. In fact, the following proposition shows that *all* W-polynomials over K with coefficients in F "essentially" arise in this way.

Proposition 3.6. *Let $g(t)$ be a polynomial in $F[t]$, and let $(S,D) = (I,0)$ on K. Then $g(t)$ is a W-polynomial over K iff $g = f_1 \cdots f_r$ where each f_i is the minimal polynomial of some a_i over F and $a_1, \ldots, a_r$ are pairwise non-conjugate in K.*

Proof. First suppose $g(t)$ is a W-polynomial over K. If $a \in K$ is a root of g, then so is any conjugate of a. Therefore, $V(g)$ is the union of a finite number of distinct conjugacy classes, say $\Delta(a_1), \ldots, \Delta(a_r)$. Let $f_i := f_{\Delta(a_i)}$, which by (5) above is an irreducible polynomial in $F[t]$. Then $g \in K[t] \cdot f_i$ for each i, and so $g \in F[t] \cdot f_i$. From this, we see that $g \in F[t] \cdot f_1 \cdots f_r$. On the other hand, $f_1 \cdots f_r$ clearly vanishes on

$$\Delta(a_1) \cup \cdots \cup \Delta(a_r) = V(g),$$

which has minimal polynomial g. Therefore, $f_1 \cdots f_r \in K[t] \cdot g$, so we have $g = f_1 \cdots f_r$. Conversely, suppose g has the form $f_1 \cdots f_r$ described in the Proposition. By Wedderburn's Theorem, $\mathrm{rk}(\Delta(a_i)) = \deg(f_i)$, so by [La$_1$: Thm. 22], $\Delta := \Delta(a_1) \cup \cdots \cup \Delta(a_r)$ has rank $\sum_i \deg(f_i) = \deg\, g$. Since g vanishes on Δ, it follows that $g = f_\Delta$, so g is a W-polynomial over K. $\square$

§4. The Union Theorem, the Φ-Transform, and the λ-Transform

In this section, we shall obtain some preliminary results on the ranks of algebraic subsets of K, and set up two basic transformations called the Φ-transform and the λ-transform. All of these will be presented in the general (S,D)-setting, which turns out not to require any additional effort. We begin with the following observation on the degrees of the left least common multiple (llcm) and the right greatest common divisor (rgcd) of two given polynomials.

Degree Equation 4.1. *For any two nonzero polynomials $f, h \in R$, let $p = \mathrm{rgcd}(f, h)$, and $q = \mathrm{llcm}(f, h)$. Then*

$$\deg(f) + \deg(h) = \deg(p) + \deg(q).$$

Proof. This result is part of the folklore of the subject; see [Or: Ch. 1, (24)]. However, the proof given by Ore in this reference was quite indirect. For the convenience of the reader, we include a "modern" proof here. By the definitions of llcm and rgcd, we have $Rf \cap Rh = Rq$, and $Rf + Rh = Rp$. Thus, Noether's Isomorphism Theorem gives an R-module isomorphism $Rp/Rf \cong Rh/Rq$. Evaluating the left K-dimensions of both sides gives the desired formula. $\square$

Next, we observe the following special property for polynomials f which factor completely in $R = K[t,S,D]$.

Proposition 4.2. *Let $f, h \in R \setminus \{0\}$, and assume that f is a product of linear factors. Then $V(f) \cap V(h) = \emptyset$ iff $Rf + Rh = R$. In this case,*

$$\deg\bigl(\mathrm{llcm}(f, h)\bigr) = \deg(f) + \deg(h).$$

Proof. First assume $Rf + Rh = R$. Then $rf + sh = 1$ for suitable $r, s \in R$. If there exists $a \in V(f) \cap V(h)$, plugging a into the equation $rf + sh = 1$ would yield a contradiction, so we must have $V(f) \cap V(h) = \emptyset$. On the other hand, if $Rf + Rh \neq R$, then $Rf + Rh = Rp$, where $p := \mathrm{rgcd}(f, h)$ is *non-constant*. Write $f = f_1 p$ and $h = h_1 p$, where $f_1, h_1 \in R$. Since f is a product of linear factors, so is its factor p. (This follows from the fact that, if f is factored in any way into a product of irreducible factors, the degrees of these irreducible factors are uniquely determined; see, e.g. [Or: Ch.2, Th.1].) Thus, there exists $a \in V(p)$, and the equations $f = f_1 p$ and $h = h_1 p$ show that $a \in V(f) \cap V(h)$. Hence $V(f) \cap V(h) \neq \emptyset$. The last part of the Proposition now follows from (4.1). □

The Union Theorem 4.3. *Let Δ and Γ be algebraic sets in K, with minimal polynomials $f, h \in R$, of degrees n and m respectively. Then*

(1) *the minimal polynomial for $\Delta \cup \Gamma$ is $\mathrm{llcm}(f, h)$, and we have $\mathrm{rk}(\Delta \cup \Gamma) \leq n + m$.*

(2) *If $\overline{\Delta} \cap \overline{\Gamma} = \emptyset$, then equality holds in (1), and, if B_1, B_2 are respectively P-bases for Δ and Γ, then $B_1 \cup B_2$ is a P-basis for $\Delta \cup \Gamma$.*

Proof. A nonzero polynomial $r(t) \in R$ vanishes on $\Delta \cup \Gamma$ iff it is right divisible by f as well as by h. Therefore, the monic $r(t)$ of the smallest degree is exactly $q := \mathrm{llcm}(f, h)$. In particular, by (4.1), $\mathrm{rk}(\Delta \cup \Gamma) = \deg(q) \leq n + m$. This proves (1). To prove (2), assume that $\overline{\Delta} \cap \overline{\Gamma} = \emptyset$. This amounts to $V(f) \cap V(h) = \emptyset$. Since f is indeed a product of linear factors, (4.2) implies that

$$\deg(q) = \deg(f) + \deg(h) = n + m,$$

and the rest follows. □

A further refinement of (4.3) will appear later in the Rank Theorem (7.3). The following useful special case of (4.3) is essentially the "Union Theorem 22" in [La$_1$].

Corollary 4.4. (1) *Let Δ, Γ be algebraic sets such that no element of Δ is (S, D)-conjugate to an element of Γ. Then $\mathrm{rk}\,(\Delta \cup \Gamma) = \mathrm{rk}\,(\Delta) + \mathrm{rk}\,(\Gamma)$. If B_1, B_2 are respectively P-bases for Δ and Γ, then $B_1 \cup B_2$ is a P-basis for $\Delta \cup \Gamma$, and $\overline{\Delta \cup \Gamma} = \overline{\Delta} \cup \overline{\Gamma}$.*

(2) *If Δ_i $(1 \leq i \leq r)$ are algebraic sets contained in different (S, D)-conjugacy classes of K, then*

$$\mathrm{rk}\,\Big(\bigcup_{i=1}^{r} \Delta_i\Big) = \sum_{i=1}^{r} \mathrm{rk}\,(\Delta_i).$$

A P-basis for $\bigcup_{i=1}^{r} \Delta_i$ is given by a union of any P-bases for the Δ_i's, and we have the P-closure formula $\overline{\cup_i \Delta_i} = \bigcup_i \overline{\Delta_i}$.

Proof. By (2.6)(2), we have

$$\overline{\Delta} \subseteq \{x : \ x \text{ is } (S, D)\text{-conjugate to an element of } \Delta\},$$

and similarly for $\overline{\Gamma}$. By assumption, therefore, $\overline{\Delta} \cap \overline{\Gamma} = \emptyset$, so Theorem 4.3 applies to give the statement on rank and P-basis in (1). For the equation on P-closures in (1), it suffices to prove that $\overline{\Delta \cup \Gamma} \subseteq \overline{\Delta} \cup \overline{\Gamma}$. Consider any element $a \in K$ that is P-dependent on $\Delta \cup \Gamma$. Let $A = \Delta^{S,D}(a)$ (the (S,D)-conjugacy class of a, as defined in (2.5)). By the Excision Theorem in [La$_1$], a is already P-dependent on

$$A \cap (\Delta \cup \Gamma) = (A \cap \Delta) \cup (A \cap \Gamma).$$

Now, by the hypothesis on Δ and Γ again, one of the intersections $A \cap \Delta$ and $A \cap \Gamma$ must be empty. Say $A \cap \Gamma = \emptyset$. Then a is P-dependent on $A \cap \Delta$, and hence on Δ. This shows that $a \in \overline{\Delta}$, which completes the proof of (1). From this, (2) follows easily by induction. □

To get more refined results on the ranks of algebraic sets, we shall need some information on a certain "Φ-transform", which maps algebraic sets to algebraic sets in K. Let us now introduce the method of this Φ-transform. A few applications of this method will be given in (4.9) and (4.13); more applications of the Φ-transform will be given in the next section.

In the Product Formula (2.3) for evaluating gh at a, we first encountered the expression $a^{h(a)}$ (in case $h(a) \neq 0$). This led us to the following useful definition.

Definition 4.5. For $h \in R = K[t,S,D]$, we define the "Φ-transform" (associated to h)

$$\Phi_h : K \setminus V(h) \longrightarrow K$$

by $\Phi_h(x) = x^{h(x)}$, whenever $h(x) \neq 0$. (We do not attempt to define Φ_h on $V(h)$.) Note that Φ_h always preserves the (S,D)-conjugacy class of its argument x.

Examples 4.6.

(1) For a field K with $(S,D) = (I,0)$, the transform Φ_h (for any h) is the inclusion map $K \setminus V(h) \longrightarrow K$.

(2) Consider the case when h is a nonzero constant (polynomial) $c \in K^*$. Here, $\Phi_c(x) = x^c$ for all $x \in K$, so Φ_c is defined on all of K and is exactly (S,D)-conjugation by the element c. (In particular, Φ_1 is just the identity map on K.) In view of this example, we can think of the Φ-transform as a kind of generalization of (S,D)-conjugation.

(3) Suppose D is the inner S-derivation defined by $D(x) = ax - S(x)a$, where a is a fixed element of K. Then, for $h(t) = t - a$, we have, for any $x \neq a$:

$$\begin{aligned} \Phi_h(x) &= x^{x-a} = S(x-a)x(x-a)^{-1} + D(x-a)(x-a)^{-1} \\ &= [\,S(x-a)x + a(x-a) - S(x-a)a\,]\,(x-a)^{-1} \\ &= S(x-a) + a \\ &= S(x) + \big(a - S(a)\big). \end{aligned}$$

Thus, $\Phi_h : K \setminus \{a\} \longrightarrow K$ is just the map S followed by a translation by the constant $a - S(a)$. In particular, if $a = 0$ (for which $D = 0$), Φ_t is just the map S on K^*.

(4) Suppose an (S,D)-conjugacy class $\Delta^{S,D}(a)$ is algebraic of rank 2. Then its minimal polynomial f has the form $(t-b)(t-a)$ for some $b \in \Delta^{S,D}(a)$. Take $h(t) = t - a$. For any $c \in \Delta^{S,D}(a) \setminus \{a\}$, the fact that $f(c) = 0$ implies (by the Product Formula) that $\Phi_h(c) = b$. Therefore, the transform Φ_h restricted to $\Delta^{S,D}(a) \setminus \{a\}$ is the constant map taking everything to b.

For the applications we have in mind for the Φ-transform in §5, we shall need the next three propositions. The first one is a useful composition result for the Φ-transform.

Proposition 4.7. *Let $h(t) = p(t)q(t) \in R$, and $A := K \backslash V(h)$. Then $\Phi_h = \Phi_p \circ \Phi_q$ on A. In particular,* $\mathrm{im}(\Phi_h) \subseteq \mathrm{im}(\Phi_p)$.

Proof. For $a \in A$, we have $h(a) \neq 0$, so (2.3) gives $q(a) \neq 0$ and $p\big(a^{q(a)}\big) \neq 0$. Thus, Φ_q is defined on A, and Φ_p is defined on $\Phi_q(A)$. Our job is to prove the commutativity of the following diagram:

$$\begin{array}{ccc} A & \xrightarrow{\Phi_q} & \Phi_q(A) \\ & \Phi_h \searrow & \downarrow \Phi_p \\ & & \Phi_p\big(\Phi_q(A)\big). \end{array}$$

This is checked by the following calculation using (2.3) and (2.4):

$$\begin{aligned} \Phi_h(a) &= a^{h(a)} = a^{p(a^{q(a)})\,q(a)} \\ &= \big(a^{q(a)}\big)^{p(a^{q(a)})} = \Phi_p\left(a^{q(a)}\right) \\ &= \Phi_p\big(\Phi_q(a)\big) = \big(\Phi_p \circ \Phi_q\big)(a), \end{aligned}$$

which is valid for any $a \in A$. From this calculation, it follows immediately that $\mathrm{im}(\Phi_h) \subseteq \mathrm{im}(\Phi_p)$. □

Remark. The referee pointed out that a somewhat more conceptual proof of (4.7) is possible. One notes that, for $x \notin V(h)$, the element $\Phi_h(x)$ is uniquely characterized by the equation $\mathrm{llcm}\,(h(t),\, t-x) = (t - \Phi_h(x))h(t)$. The formula in (4.7) then easily follows upon computing $\mathrm{llcm}\,(p(t)q(t),\, t-x) = \mathrm{llcm}\,\{p(t)q(t),\, \mathrm{llcm}(q(t),\, t-x)\}$.

With the notation of the Φ-transform, we can rephrase the second case of the Product Formula (2.3) as follows. If $f = gh \in R$ and $a \in K$, then

$$f(a) = g\big(\Phi_h(a)\big)\; h(a) \quad \text{if } h(a) \neq 0. \tag{4.8}$$

Thus, for any $a \notin V(h)$, we have $a \in V(f)$ iff $\Phi_h(a) \in V(g)$. This observation leads easily to the following *explicit* way for constructing the minimal polynomial of a union of two algebraic sets in terms of the Φ-transform (cf. (4.3)(1)).

Proposition 4.9. *Let Γ be an algebraic set in K, with $h := f_\Gamma$. Then for any algebraic set Δ, $f_{\Delta\cup\Gamma} = f_{\Phi_h(\Delta\backslash\overline{\Gamma})}\, f_\Gamma$. (Recall that the P-closure $\overline{\Gamma}$ of Γ is simply given by $V(h)$, so Φ_h is defined on $\Delta \setminus \overline{\Gamma}$.)*

Proof. We do know, from (4.3), that $\Delta \cup \Gamma$ is algebraic. To find $f_{\Delta\cup\Gamma}$, we look for the monic polynomial f of the least degree that vanishes on $\Delta \cup \Gamma$. Since $f(\Gamma) = 0$, f has the form gh for some monic g. To make sure that $f(\Delta \setminus \Gamma) = 0$ too, we need to have $g\big(\Phi_h(\Delta \setminus \overline{\Gamma})\big) = 0$, by (4.8). The monic g of the least degree satisfying this is $f_{\Phi_h(\Delta\backslash\overline{\Gamma})}$. □

For use in later sections, we shall recall another transform, called the λ-transform, which we have introduced earlier in [LL$_3$].

Definition 4.10. For $h \in R$ and $b \in K$, we define the λ-transform $\lambda_{h,b} : K \longrightarrow K$ by taking

$$\lambda_{h,b}(d) = \begin{cases} 0 & \text{if } \ d = 0, \\ h(b^d)\, d & \text{if } \ d \neq 0. \end{cases} \tag{4.11}$$

The (S, D)-*centralizer* of b is defined to be the set

$$C^{S,D}(b) := \{0\} \cup \{c \in K^* : \ b^c = b\}.$$

This (S, D)-centralizer is easily seen to be a division subring of K. As noted in [LL$_3$], $\lambda_{h,b}$ is an endomorphism of K as a right vector space over $C := C^{S,D}(b)$. In fact, for any $d \in K^*$ and $c \in C^*$, we have (by (2.4)):

$$\lambda_{h,b}(dc) = h\big(b^{dc}\big)\, dc = h\big((b^c)^d\big)\, dc = h(b^d)\, dc = \lambda_{h,b}(d)\, c.$$

The additivity of $\lambda_{h,b}$ can be deduced from the calculation in the proof of (3.16) in [LL$_1$]. Alternatively, it can also be checked very quickly as follows. According to the Product Formula (in the case when the second factor is a constant polynomial), we have the relation

$$\lambda_{h,b}(d) = h(b^d)\, d = (h \cdot d)(b) \quad (\text{for any } \ d \in K^*). \tag{4.12}$$

Since $(h \cdot d)(b)$ is clearly additive in d, the desired conclusion follows. Incidentally, the formula (4.12) also provides a nice example for the λ-transform: taking $b = 0$ and $h = t^n$, we see that $\lambda_{t^n,0}$ is just the operator D^n, since (4.12) implies

$$\lambda_{t^n,0}(d) = (t^n \cdot d)(0) = (\text{const. term of } t^n \cdot d) = D^n(d) \quad (\forall\, d \in K).$$

It should come as no surprise to the reader that the λ-transform $\lambda_{h,b}$ is closely related to the Φ-transform Φ_h. In fact, in a manner of speaking, working with the λ-transform is equivalent to working with the Φ-transform. The following result summarizes the exact relationship between these two transforms, and records some of their key properties.

Proposition 4.13. (1) *For any* $d \in K^*$, *we have* $\Phi_h(b^d) = b^{\lambda_{h,b}(d)}$.

(2) *For* $d, e \in K^*$, $\Phi_h(b^d) = \Phi_h(b^e)$ *iff* $\lambda_{h,b}(d) \in \lambda_{h,b}(e) \cdot C^{S,D}(b)$.

(3) *For any* $d \in K^*$, $b^d \in \operatorname{im}(\Phi_h)$ *iff* $d \in \operatorname{im}(\lambda_{h,b})$.

(4) *If* h *has no zeros on* $\Delta := \Delta^{S,D}(b)$, *then* $\lambda_{h,b} : K \longrightarrow K$ *and* $\Phi_h : \Delta \longrightarrow \Delta$ *are both injective maps.*

(5) (*"Closure Property" of* $\operatorname{im}(\Phi_h)$.) *If* Δ *is any algebraic set contained in* $\operatorname{im}(\Phi_h)$, *then* $\overline{\Delta} \subseteq \operatorname{im}(\Phi_h)$.

Proof. To simplify the notation, let us write λ for $\lambda_{h,b}$ below (with h and b fixed).

(1) For any $d \in K^*$, the conjugation rule (2.4) gives

$$\Phi_h(b^d) = \big(b^d\big)^{h(b^d)} = b^{h(b^d)d} = b^{\lambda(d)}.$$

(2) Assume first that $\lambda(d) = \lambda(e) \cdot c$, for some $c \in C^{S,D}(b)$. Using (2.4) again, we get

$$b^{\lambda(d)} = b^{\lambda(e) \cdot c} = \big(b^c\big)^{\lambda(e)} = b^{\lambda(e)},$$

so by (1) we get $\Phi_h(b^d) = \Phi_h(b^e)$. Conversely, suppose $\Phi_h(b^d) = \Phi_h(b^e)$. By (1) again, we have $b^{\lambda(d)} = b^{\lambda(e)}$. Thus, $\lambda(d) = \lambda(e)\, c$ for some $c \in C^{S,D}(b)$.

(3) If $d = \lambda(d')$ for some d', then by (1):

$$b^d = b^{\lambda(d')} = \Phi_h\big(b^{d'}\big) \in \operatorname{im}(\Phi_h).$$

Conversely, suppose $b^d = \Phi_h(a)$ for some a. Since Φ_h preserves (S,D)-conjugacy classes, we have $a = b^e$ for some $e \in K^*$. Then $b^d = \Phi_h(b^e) = b^{\lambda(e)}$, so for some $c \in C^{S,D}(b)$, we have

$$d = \lambda(e)\, c = \lambda(ec) \in \operatorname{im}(\lambda).$$

(4) Assume that h has no zeros on Δ. Then

$$d \neq 0 \implies h(b^d) \neq 0 \implies \lambda(d) = h(b^d)\, d \neq 0.$$

Therefore, $\ker(\lambda) = 0$, so $\lambda : K \longrightarrow K$ is injective. Next, suppose $\Phi_h(b^d) = \Phi_h(b^e)$. By (2) above, we have $\lambda(d) = \lambda(e)\, c = \lambda(ec)$ for some $c \in C^{S,D_i}(b_i)$. The injectivity of λ implies that $d = ec$, and so $b^d = b^{ec} = (b^c)^e = b^e$.

(5) Since Δ intersects only finitely many (S,D)-conjugacy classes, we can write it as a disjoint union $\Delta_1 \cup \cdots \cup \Delta_n$, where the Δ_i 's lie in different classes. By (4.4)(2), we have $\overline{\Delta} = \bigcup_{i=1}^n \overline{\Delta_i}$. Thus, it is enough to handle the case when Δ lies in a *single* class $\Delta^{S,D}(b)$. Here, the quickest way to prove (5) is to use some of the results from [LL$_2$]. Write

$$\Delta = b^Y := \{b^y :\ y \in Y\},$$

where Y is some subset of K^*. By (3) above, $b^Y \subseteq \operatorname{im}(\Phi_h)$ implies that $Y \subseteq \operatorname{im}(\lambda)$. Writing $C := C^{S,D}(b)$ and $Y \cdot C$ for the (right) linear C-span of the set Y, we have $Y \cdot C \subseteq \operatorname{im}(\lambda)$, since $\operatorname{im}(\lambda)$ is a right C-space. By (3) again, we have therefore $b^{Y \cdot C} \subseteq \operatorname{im}(\Phi_h)$. Now by Th. 4.5 in [LL$_2$], $b^{Y\cdot C}$ is exactly the P-closure of b^Y. Hence we have $\overline{\Delta} = b^{Y\cdot C} \subseteq \operatorname{im}(\Phi_h)$. □

Remark 4.14. To see what the "closure property" (in (5)) means in a special case, take $h(t) = t$ and $D = 0$. In this case, by (4.6)(3), Φ_t is the map S on K^*. Hence $\operatorname{im}(\Phi_t) = S(K^*)$. The closure property tells us that *if an element $b \in K$ is P-dependent on a set $S(a_1), \ldots, S(a_n)$ for some a_i 's in K^*, then $b = S(a)$ for some $a \in K^*$.* This seems to be a somewhat nontrivial statement.

§5. Factors and Products of W-Polynomials

In this section, we shall study the Wedderburn polynomials *as a whole* in a fixed Ore skew polynomial ring $R = K[t,S,D]$. For the rest of the paper, let us write $\mathcal{W}(= \mathcal{W}(K,S,D))$ for the set of all W-polynomials in R. Our formation of the set $\mathcal{W}$ is, in part, motivated by the classical work of Oystein Ore. In [Or], Ore defined a *completely reducible polynomial* to be the llcm (least left common multiple) of a finite number of irreducible polynomials in R. Since linear polynomials are obviously irreducible, our W-polynomials are a special case of Ore's completely reducible polynomials.

In retrospect, it may seem a bit surprising that Ore himself did not study the class of W-polynomials (as a subclass of his completely reducible polynomials). We believe the reason may very well have been that Ore was not aware of the possibility of a theory of evaluation of skew polynomials at constants. Without such a theory, the interpretation of the llcm of linear polynomials $\{t - a_i\}$ as the minimal polynomial of the set $\{a_i\}$ is lacking, and as a result, such llcm's may not have invited particular attention. But, again retrospectively, since *linear* polynomials are a very special kind of irreducible polynomials, one should have expected *their* llcm's (the W-polynomials) to have a much richer structure than the llcm's of irreducible polynomials (Ore's completely reducible polynomials).

The main goal of this section is to establish some basic results on the factors and products of W-polynomials. We have to clarify what exactly is meant by the word "factor" in this paper. Throughout the sequel, we'll say that a polynomial p is a **factor** of another polynomial f if $f = f_1 p f_2$ for some polynomials $f_1, f_2 \in R$. Right and left factors of f have their usual meanings, and these are, of course, particular kinds of factors in our sense. The following result, which essentially goes back to Ore, gives an interesting description of W-polynomials in terms of its factors, and more specifically, its *quadratic* factors.

Factor Theorem 5.1. *For any monic $f \in R$, the following are equivalent*:

(1) *f is a W-polynomial*;

(2) *f splits completely,*[2] *and every monic factor of f is a W-polynomial*;

(3) *f splits completely, and every monic quadratic factor of f is a W-polynomial.*

Proof. In §2, we have already observed that any W-polynomial splits completely. Thus, (1) $\Longrightarrow$ (2) follows from [LL$_4$: (5.9)]. (2) $\Longrightarrow$ (3) being trivial, it only remains for us to prove (3) $\Longrightarrow$ (1). Assume (3), and write $f = g(t-a)$, where g is monic (as f is). Since g has the same properties as f, we may assume (by induction on $n := \deg(f)$) that $g \in \mathcal{W}$. Take a P-basis $\{d_1, \dots, d_{n-1}\}$ for $V(g)$, and write $g = g_i(t-d_i)$ for $1 \le i \le n-1$. Then $f = g_i(t-d_i)(t-a)$, so by assumption $(t-d_i)(t-a) \in \mathcal{W}$. If c_i is a root of $(t-d_i)(t-a)$ other than a, we have $d_i = c_i^{c_i - a}$ (for $1 \le i \le n-1$) by (4.8). Applying (4.9) for $\Gamma = \{a\}$ and $\Delta := \{c_1, \dots, c_{n-1}\}$, we see that the minimal polynomial of $\Delta \cup \Gamma$ is given by

$$f_{\Phi_{t-a}(\Delta)} \cdot (t-a) = f_{\{d_1,\dots,d_{n-1}\}} \cdot (t-a) = g(t-a) = f.$$

(Here, $f_{\{d_1,\dots,d_{n-1}\}} = g$ since $g \in \mathcal{W}$.) From this, we see that $f \in \mathcal{W}$. $\square$

Remarks 5.2.

(1) The result (5.1) is essentially a specialization of Theorem 3 in Chapter II of [Or] to W-polynomials. We presented here a treatment of (5.1) for two reasons. First, Ore's proof for his Theorem 3 has not been re-examined in the literature for quite some time, and is likely to be difficult for a modern reader to follow. In fact, we ourselves were not able to fill in some of the omitted steps in Ore's proof. Thus, it seems that an alternative treatment is desirable. Second, Ore's Theorem 3 was proved for the more general class of completely reducible polynomials. Since W-polynomials are so special (and also so nice!) in nature, it would seem reasonable to give a direct proof of (5.1) in our context without taking a detour into Ore's theory of completely reducible polynomials. For further generalizations of Ore's result, see, e.g. [Co$_2$: III.6.11].

[2] By this, we mean that f can be written as a product of linear polynomials in R.

(2) One may wonder if, in the statement of (5.1)(3), the word "factor" can be replaced by "right factor". The following example shows that this is *not* the case. Let $R = K[t, S]$ where $K = \mathbb{Q}((x))$ and S is the $\mathbb{Q}$-endomorphism of K defined by $S(x) = x^2$. Clearly,

$$h(t) := t\,(t - x) = \left(t - x^2\right) t \in \mathcal{W}.$$

It is easy to see that $(t - x)\,t \notin \mathcal{W}$, and hence $f(t) := (t - x)\,t\,(t - x) \notin \mathcal{W}$ by (5.1). But the reader can check that the only monic quadratic *right* factor of $f(t)$ is $h(t)$, and hence "all" such right factors of f are W-polynomials.

(3) Note that (3) $\Rightarrow$ (1) in (5.1) *does not* mean that $f(t) = (t - a_1) \cdots (t - a_n) \in \mathcal{W}$ if $(t - a_i)(t - a_{i+1}) \in \mathcal{W}$ for every $i < n$. (An obvious counterexample is given by $t(t-1)t$ over a field K.) In other words, the "factor" condition in (5.1)(3) must be imposed on *every* monic quadratic polynomial q such that $f = f_1 q f_2$ for some $f_1, f_2 \in R$.

(4) For any W-polynomial $f(t) = (t - a_n) \cdots (t - a_1)$, let $f_i(t) = (t - a_i) \cdots (t - a_1)$ $(i \leq n)$. By (5.1)(2), $f_i \in \mathcal{W}$, so we have $\operatorname{rk} V(f_i) = i$ (for each i). In particular, we have strict inclusions

$$\{a_1\} = V(f_1) \subsetneq V(f_2) \subsetneq \cdots \subsetneq V(f_n).$$

This generalizes a result of Haile and Rowen [HR: Prop. 1.1] in several ways. First our result holds in the (S, D) setting, and for general W-polynomials (instead of minimal polynomials of algebraic elements over the center). Second, the above shows that not only $V(f_i) \subsetneq V(f_{i+1})$, but actually $\operatorname{rk} V(f_i) < \operatorname{rk} V(f_{i+1})$.

Our next goal is to obtain some necessary and sufficient conditions for a product of two (monic) polynomials to be a W-polynomial (Theorem 5.6). In preparation for this, we first prove the following key result concerning the "llcm" of two polynomials, one of which is Wedderburn: this is an interesting application of the "closure property" in (4.13)(5).

Proposition 5.3. *Let $\ell = \operatorname{llcm}(f,\, h)$, where $h \in R$ is monic and $f \in \mathcal{W}$, and let $\ell = pf = gh$, where $p,\, g \in R$ are monic. Then $g \in \mathcal{W}$, and $V(g) \subseteq \operatorname{im}(\Phi_h)$.*

Note that the first conclusion here is a generalization of the fact that a left (monic) factor of a Wedderburn polynomial is Wedderburn, by considering the special case where $f = gh$. The second conclusion in this special case gives a necessary condition for gh to be Wedderburn, which will turn out to be sufficient as well, if $g,\, h \in \mathcal{W}$.

Proof of (5.3). Let $\Pi := \Phi_h\big(V(f) \setminus V(h)\big) \subseteq \operatorname{im}(\Phi_h)$. For $a \in V(f) \setminus V(h)$, we have by (4.8):

$$0 = \ell(a) = g(\Phi_h(a))\, h(a) \implies g(\Phi_h(a)) = 0,$$

so $g(\Pi) = 0$. Let g_0 be the minimal polynomial of Π. Reversing the argument above, we see that $g_0 h$ vanishes on $V(f) \setminus V(h)$, and hence on $V(f)$. Since $f \in \mathcal{W}$, we have $g_0 h \in Rf$. Thus, $g_0 h$ is a common left multiple of f and h. Since $\deg(g_0) \leq \deg(g)$, we must have $g = g_0 \in \mathcal{W}$. Finally, by the closure property (4.13)(5), $V(g) = V(g_0) = \overline{\Pi} \subseteq \operatorname{im}(\Phi_h)$, as desired. $\square$

A second result we need for the proof of (5.6) is a certain characterization of $\operatorname{im}(\Phi_h)$ for polynomials $h \in \mathcal{W}$. This depends rather heavily on some results in [LL$_4$]. Specifically, we'll need from that paper the symmetry theorem on W-polynomials [LL$_4$: (4.5)], which states that a monic polynomial h belongs to $\mathcal{W}$ iff $hR = \bigcap_j (t-b_j)R$ for some set of elements $\{b_j\}$ in K. Following [LL$_4$], we write

$$V'(h) := \{b \in K : \ h \in (t-b)R\}; \tag{5.4}$$

this is the set of "left roots" of h. If $h \in \mathcal{W}$ and $\deg(h) = r$, [LL$_4$: (4.5)] also implies that one can write $hR = \bigcap_{j=1}^r (t-b_j)R$ for suitable $b_1, \dots, b_r \in K$. In analogy with the case of right roots, we shall call any such set $\{b_1, \dots, b_r\}$ a *P-basis* for the left root set $V'(h)$.

Proposition 5.5. *Let $h \in \mathcal{W}$, and let $\{b_1, \dots, b_r\}$ be any P-basis of $V'(h)$. Then* $\operatorname{im}(\Phi_h) = \bigcap_j \operatorname{im}(\Phi_{t-b_j})$.

Proof. The inclusion "$\subseteq$" follows from the last statement of (4.7). For the reverse inclusion, let $x \in \bigcap_j \operatorname{im}(\Phi_{t-b_j})$, say $x = c_j^{c_j - b_j}$, where $c_j \neq b_j$. Then the minimal polynomial f_j for $\{b_j, c_j\}$ has the form $(t-x)(t-b_j)$ as well as the form $(t-y_j)(t-c_j)$, where $y_j := b_j^{b_j - c_j}$ (see (3.3)). According to [LL$_4$: (4.4)], the left ideal representation $Rf_j = R(t-b_j) \cap R(t-c_j)$ leads to a right ideal representation $f_j R = (t-x)R \cap (t-y_j)R$. Therefore,

$$\begin{aligned} (t-x)hR &= (t-x) \cdot \textstyle\bigcap_j (t-b_j)R \\ &= \textstyle\bigcap_j (t-x)(t-b_j)R \\ &= \textstyle\bigcap_j \left[(t-x)R \cap (t-y_j)R \right] \\ &= (t-x)R \cap \textstyle\bigcap_j (t-y_j)R. \end{aligned}$$

By [LL$_4$: (4.5)], this implies that $(t-x)h \in \mathcal{W}$. It then follows from (5.3) that $x \in \operatorname{im}(\Phi_h)$. $\square$

Having proved (5.3) and (5.5), we can now formulate various criteria for a product of two W-polynomials to be a W-polynomial.

Theorem 5.6. *For $f := gh \in R$ where g, h are monic, the following are equivalent:*

(1) $f \in \mathcal{W}$.

(2) g, $h \in \mathcal{W}$, *and* $V(g) \subseteq \operatorname{im}(\Phi_h)$.

(3) g, $h \in \mathcal{W}$, *and some P-basis B of $V(g)$ is contained in* $\operatorname{im}(\Phi_h)$.

(4) g, $h \in \mathcal{W}$, *and $(t-a)(t-b) \in \mathcal{W}$ for every $a \in V(g)$ and $b \in V'(h)$.*

(5) g, $h \in \mathcal{W}$, *and, for some P-basis $\{a_i\}$ of $V(g)$ and some P-basis $\{b_j\}$ of $V'(h)$, we have $(t-a_i)(t-b_j) \in \mathcal{W}$ for all i, j.*

In case $f \in \mathcal{W}$, a P-basis for $V(f)$ is given by $A \cup C$ where C is a P-basis for $V(h)$ and A is any subset of $K \setminus V(h)$ that is mapped bijectively by Φ_h to a P-basis for $V(g)$.

Proof. (1) $\Longrightarrow$ (2) follows from the Factor Theorem and (5.3), and (2) $\Longrightarrow$ (3) is trivial. (We have (3) $\Longrightarrow$ (2) too, by the "closure property" of $\operatorname{im}(\Phi_h)$. But we can get by below without using this.)

(3) $\Longrightarrow$ (1). For the P-basis B for $V(g)$ given in (3), take any set $A \subseteq K \setminus V(h)$ that Φ_h maps bijectively onto B. Since (by assumption) $g,\, h \in \mathcal{W}$, we have $f_B = g$, and $f_{V(h)} = h$. By (4.9),

$$f_{A\cup V(h)} = f_{\Phi_h(A\setminus V(h))}\, f_{V(h)} = f_B\, f_{V(h)} = gh.$$

This shows that $gh \in \mathcal{W}$, proving (1). Now take a P-basis C for $V(h)$. Then,

$$f_{A\cup C} = f_{A\cup V(h)} = gh = f.$$

Since $|A \cup C| = |A| + |C| = \deg(f)$, $A \cup C$ is necessarily a P-basis for $V(f)$. This proves the claim in the last paragraph of the theorem.

Next, (1) $\Longrightarrow$ (4) follows from the Factor Theorem, and (4) $\Longrightarrow$ (5) is clear. So we can complete the proof of (5.6) with the following last step.

(5) $\Longrightarrow$ (3). Since $(t - a_j)(t - b_j) \in \mathcal{W}$, it has a root $c_j \neq b_j$, and hence $a_i = c_j^{c_j - b_j} \in \mathrm{im}(\Phi_{t-b_j})$ for all i, j. Thus, $a_i \in \mathrm{im}(\Phi_h)$ by (5.5). Since this holds for all i, we have (3). $\square$

Remarks 5.7. (A) The advantage of the criterion (5) is that it reduces the checking of $f \in \mathcal{W}$ to verifying that a certain *finite* set of quadratic polynomials are W-polynomials. In §6, we'll see that quadratic W-polynomials are detected by the solvability of certain (S, D)-metro equations. In view of this result (see (6.6)), (5.6) has the effect of reducing the testing of $f = gh \in \mathcal{W}$ to the solvability of a *finite number* of (S, D)-metro equations.

(B) Note that the criterion $V(g) \subseteq \mathrm{im}(\Phi_h)$ in (2) above has a very clear meaning in the classical case when K is a field and $(S, D) = (I, 0)$. Here, we have $g(t) = (t - b_1)\cdots(t - b_r)$ where the b_i 's are distinct, and $h(t) = (t - a_1)\cdots(t - a_s)$ where the a_j 's are distinct. Since (by (4.6)(1)) $\mathrm{im}(\Phi_h) = K \setminus \{a_1, \ldots, a_s\}$, the condition $V(g) \subseteq \mathrm{im}(\Phi_h)$ amounts to $\{b_1, \ldots, b_r\}$ and $\{a_1, \ldots, a_s\}$ being disjoint, which is, of course, the expected criterion for gh to be again a W-polynomial.

There is another major characterization for $gh \in \mathcal{W}$ that is not yet covered in Theorem 5.6. This characterization involves idealizers of left (principal) ideals in the ring R. Recall that, for $g \in R$, the *idealizer* of the left ideal $Rg \subseteq R$ is defined to be

$$\mathbb{I}_R(Rg) = \{k \in R : \; gk \in Rg\}, \tag{5.8}$$

which is just the largest subring of R in which Rg is an ideal. Repeating a part of the proof of (5.3) (with $f,\, h$ there replaced by $g,\, k$), we obtain easily the following characterization of the idealizer $\mathbb{I}_R(Rg)$ via the Φ-transform, (The "closure property" argument in the proof of (5.3) is not needed for this.)

Proposition 5.9. *For any polynomials $g',\, k \in R$ and any W-polynomial g, we have*

$$g'k \in Rg \iff \Phi_k(V(g) \setminus V(k)) \subseteq V(g').$$

In particular, $k \in \mathbb{I}_R(Rg)$ iff $\Phi_k(V(g) \setminus V(k)) \subseteq V(g)$.

Using this result, we can now give some other criteria for a product of two W-polynomials $g,\, h$ to be a W-polynomial, in terms of the solvability of the equation $ug + hv = 1$, and in terms of the idealizer $\mathbb{I}_R(Rg)$ of Rg.

Theorem 5.10. *For $g,\, h \in \mathcal{W}$, the following are equivalent:*

(1) $gh \in \mathcal{W}$.
(2) $1 \in Rg + hR$.
(3) $\mathbb{I}_R(Rg) \subseteq Rg + hR$.
(4) *For every* $k \in R$ *such that* $\Phi_k\big(V(g)\setminus V(k)\big) \subseteq V(g)$, *we have* $k \in Rg+hR$.
(5) *There exists* $k \in R$ *such that* $V(g) \subseteq \Phi_k\big(V(g)\setminus V(k)\big)$ *and* $k \in Rg+hR$.

Proof. The equivalence of (3) and (4) follows from (5.9). In the following, we shall prove the equivalence of (2), (3), and then prove the equivalence of (1), (2) and (5).
(3) $\Longrightarrow$ (2) is trivial, since $1 \in \mathbb{I}_R(Rg)$.
(2) $\Longrightarrow$ (3) Write $1 = ug + hv$, where $u,\, v \in R$. For any $k \in \mathbb{I}_R(Rg)$, we have $gk \in Rg$, and so

$$k = 1 \cdot k \in (Rg + hR) \cdot k \subseteq R \cdot gk + hR \subseteq Rg + hR.$$

(2) $\Longrightarrow$ (5) is trivial, since (2) implies that (5) holds for $k = 1$. (Recall that Φ_1 is the identity map on K.)
(5) $\Longrightarrow$ (1). Let k be as in (5), and write $k = ug + hv$ where $u,\, v \in R$. To prove (1), it suffices (according to (5.6)) to verify that $V(g) \subseteq \mathrm{im}(\Phi_h)$. For any $a \in V(g)$, write $a = \Phi_k(b)$ for a suitable $b \in V(g) \setminus V(k)$. Since

$$0 \neq k(b) = (ug)(b) + (hv)(b) = (hv)(b),$$

we have

$$a = \Phi_k(b) = b^{k(b)} = b^{(hv)(b)} = \Phi_{hv}(b) = \Phi_h(\Phi_v(b)),$$

where the last equality follows from (4.7). Therefore, $a \in \mathrm{im}(\Phi_h)$, as desired.
(1) $\Longrightarrow$ (2). This is the hardest (and perhaps the most interesting) implication of all. To begin its proof, assume that $gh \in \mathcal{W}$. Fix a P-basis A for $V(h)$, and extend it to a P-basis $A \cup B$ for $V(gh)$. Let $g' = f_B$ (the minimal polynomial of B). Then $gh = h'g'$ for some $h' \in R$. Since $gh \in \mathcal{W}$, it is the minimal polynomial of $A \cup B$, and thus $gh = \mathrm{llcm}\{g',\, h\}$. Now $A \cup B$ is P-independent, so we have $V(g') \cap V(h) = \emptyset$. By (4.2), this implies that there exist $u,\, v \in R$ such that $ug' + vh = 1$. *We claim that*

$$\{p \in R :\ ph \in Rg'\} = Rg. \tag{5.11}$$

The inclusion "$\supseteq$" is clear from $gh = h'g'$, so we only need to prove "$\subseteq$". Let $p \in R$ be such that $ph \in Rg'$. Then ph is right divisible by both h and by g', and hence by $\mathrm{llcm}\{h,\, g'\} = gh$. Writing $ph = qgh$ (for a suitable $q \in R$), we see by cancellation of h that $p = qg \in Rg$, thus proving (5.11). Left-multiplying $1 = ug' + vh$ by h, we get $hug' + hvh = h$, and so $(hv - 1)h \in Rg'$. By (5.11), we have then $hv - 1 \in Rg$, from which we get $1 \in Rg + hR$, as desired. ☐

Remarks 5.12. (1) For readers who are familiar with P. M. Cohn's book [Co$_2$], it is relevant to point out that, in a left principal ideal domain, the condition $1 \in Rg + hR$ is equivalent to the existence of what Cohn called a "comaximal relation" $gh = h'g'$ in R (see [Co$_2$: p. 28, p. 171]). Theorem 5.10 above is partly inspired by Cohn's result.

(2) In the special case where $g(t) = t - b$, there is an alternative proof for (1) $\Longrightarrow$ (2) in (5.10) that yields an "explicit" expression $u_1 g + hv_1 = 1$. In fact,

if $(t-b)h \in \mathcal{W}$, we know from (5.3) that $b \in \mathrm{im}(\Phi_h)$, and therefore, by (4.13)(3), $1 = \lambda_{h,b}(a) = h(b^a)a$ for some $a \in K^*$. By the Product Formula, we can rewrite this as $(h \cdot a)(b) = 1$. Then by the Remainder Theorem (applied to $h \cdot a$ "divided by" $t-b$), we have $h \cdot a = q(t)(t-b) + 1$ for some $q \in R$. Therefore, we have a solution for $u_1 g + hv_1 = 1$ with $u_1 := -q$ of degree one less than $\deg(h)$, and with $v_1 := a \in K$.

(3) In the general case, the referee pointed out that, starting from *any* equation $ug + hv = 1$ $(u, v \in R)$, one can always derive a new equation of the same type with the degree bounds $\deg(v) < \deg(g)$ and $\deg(u) < \deg(h)$ (assuming that $\deg(g), \deg(h) > 0$). In fact, using the division algorithm to write $v = wg + v_1$ with $\deg(v_1) < \deg(g)$, we get

$$1 = ug + hv = ug + h(wg + v_1) = (u + hw)g + hv_1.$$

Setting $u_1 = u + hw \in R$, we have

$$\deg(u_1 g) = \deg(hv_1) = \deg(h) + \deg(v_1) < \deg(h) + \deg(g).$$

Therefore, $\deg(u_1) < \deg(h)$, along with $\deg(v_1) < \deg(g)$.

Combining (5.10) with (5.6), we reach the following curious conclusion.

Corollary 5.13. *For $g, h \in \mathcal{W}$, let $\{a_i\}$ and $\{b_j\}$ be as in* (5.6)(5). *Then we can solve the equation $ug + hv = 1$ iff, for each i, j, we can solve the equation $1 = u_{ij}(t - b_j) + (t - a_i)v_{ij}$.*

§6. Algebraic Conjugacy Classes and (S, D)-Metro Equations

We begin this section by giving some applications of the results in §5 to the theory of algebraic (S, D)-conjugacy classes initiated in [LL$_2$]. Such algebraic conjugacy classes and their minimal polynomials have many special properties, as the following two results show.

Theorem 6.1. *Suppose $\Delta := \Delta^{S,D}(b)$ is an algebraic set, with minimal polynomial f_Δ. Then:*

(1) *A monic polynomial $h \in R$ is a right factor of f_Δ iff $h \in \mathcal{W}$ and $h(t) = (t - b_r) \cdots (t - b_1)$ for some $b_1, \ldots, b_r \in \Delta$.*

(2) *Suppose each of $h_1, \ldots, h_r$ is a monic right factor of f_Δ. Then $h_r \cdots h_1$ is a right factor of f_Δ iff it is a W-polynomial.* (*For instance, for $d \in K^*$, $(t - b^d)h_1$ is a right factor of f_Δ iff $d \in \mathrm{im}(\lambda_{h_1,b})$.*)

Proof. (1) First assume $f_\Delta \in R \cdot h$. By by the Factor Theorem, $h \in \mathcal{W}$. As a W-polynomial, h has a splitting $(t - b_r) \cdots (t - b_1)$ where each b_i is (S, D)-conjugate to some element of $V(h)$. Since $V(h) \subseteq V(f_\Delta) = \Delta$, and Δ is closed under (S, D)-conjugation, we have each $b_i \in \Delta$. Conversely, if $h \in \mathcal{W}$ and

$$h(t) = (t - b_r) \cdots (t - b_1) \quad \text{with } b_i \in \Delta,$$

then by (2.6), $V(h) \subseteq \Delta = V(f_\Delta)$. Therefore, it follows from (3.4) that h is a right factor of f_Δ.

(2) The "only if" part follows from the Factor Theorem. For the converse, assume that $h_r \cdots h_1 \in \mathcal{W}$. By (1), each h_i is a product of linear factors of the form $t - d$ where $d \in \Delta$. Then, $h_r \cdots h_1$ has the same property. Therefore, by (1) again,

$h_r \cdots h_1$ is a right factor of f_Δ. The statement in parentheses follows from (5.6) and (4.13)(3), by letting $r = 2$ and $h_2 = t - b^d$. □

Note that the "if" parts of the Theorem remain true with the adjective "right" replaced by "left" everywhere. In fact, by [LL$_2$: (5.2)], f_Δ is a left invariant polynomial in the sense that $f_\Delta R \subseteq R f_\Delta$.[3] For such a polynomial, any right factor is automatically a left factor.

Theorem 6.2. *Let $\Delta_i = \Delta^{S,D}(b_i)$ $(1 \le i \le r)$ be different algebraic conjugacy classes of K, and let $h(t) \in K[t, S, D]$ be a polynomial with no zeros on $\bigcup_i \Delta_i$. Then*

(a) *For each i, $\lambda_{h,b_i} : K \to K$ is a bijection, and $\Phi_h : \Delta_i \to \Delta_i$ is a bijection.*

(b) *Assume that h above is a W-polynomial. Then, for any W-polynomial $g(t)$ with $V(g) \subseteq \bigcup_i \Delta_i$, we have $f := gh \in \mathcal{W}$, and $V(f) = V(h) \cup \Phi_h^{-1}(V(g))$.*

Proof. (a) Since Δ_i is an *algebraic* (S, D)-conjugacy class, K is finite-dimensional as a right vector space over the division ring $C_i := C^{S,D}(b_i)$, by [LL$_2$: Th. (5.10)]. Now the lack of zeros of h on Δ_i means that the right C_i-linear map $\lambda_{h,b_i} : K \to K$ has a zero kernel. Therefore, λ_{h,b_i} must be an isomorphism. From this, it follows from (3) and (4) of (4.13) that $\Phi_h : \Delta_i \longrightarrow \Delta_i$ is a bijection.
(b) Let h and g be as in (b). Then $V(g) \subseteq \bigcup_i \Delta_i$ implies that $V(g) \subseteq \operatorname{im}(\Phi_h)$, since Φ_h is a bijection from $\bigcup_i \Delta_i$ to itself. Thus, by (5.6), $f = gh \in \mathcal{W}$, and the Product Formula implies that $V(f) = V(h) \cup \Phi_h^{-1}(V(g))$. □

Theorem 6.2 has some interesting consequences, which we shall now explore.

Corollary 6.3. *Let $c \in K$ be such that the class $\Delta := \Delta^{S,D}(c)$ is (S, D)-algebraic. If $t - c$ is a factor of $f(t) \in K[t, S, D]$ (in the sense of §5), then f has a root in Δ. In particular, if a polynomial has c as a left root, then it has a (right) root that is (S, D)-conjugate to c.*

Proof. Write $f(t) = h'(t)(t - c)h(t)$, where h, h' are suitable polynomials. We may assume h has no root in Δ, for otherwise we are done already. By (6.2)(a), Φ_h is then a bijection from Δ to itself. Therefore, $c = \Phi_h(c')$ for some $c' \in \Delta$. Evaluating $f_0(t) := (t-c)h(t)$ on c' by the Product Formula, we see that $f_0(c') = \big(\Phi_h(c') - c\big)h(c') = 0$. Thus, c' is a root of f_0, and hence of f. □

Remark. In (6.3), the assumption that the class of c is (S, D)-algebraic turns out to be essential, even in the classical case when $(S, D) = (I, 0)$. An example to this effect, with $\deg(f) = 2$, will be constructed in §10 below.

Corollary 6.4. *Let $\{\Delta_i = \Delta^{S,D}(b_i)\}$ be n different algebraic conjugacy classes of K. For any $b \notin \bigcup_i \Delta_i$, $f(t) = (t - b_n) \cdots (t - b_1)(t - b)$ is a W-polynomial, with $V(f) = \{a_1, \dots, a_n, b\}$ where $a_i \in \Delta_i$ for each i.*

Proof. The conclusion is clear if $n = 0$, and follows easily from (6.2) by induction on n. In the inductive step, we take $h(t)$ to be

$$(t - b_{n-1}) \cdots (t - b_1)(t - b),$$

[3] In [LL$_2$], we have called f_Δ right invariant instead. The referee pointed out that f_Δ should be called *left invariant* in accordance with the usage in [Co$_2$: p. 203].

and $g(t)$ to be $t - b_n$. (Alternatively, we could also have applied (6.3).) $\square$

Remark. In the case when $(S, D) = (I, 0)$ and K is algebraic over its center F, all conjugacy classes of K are algebraic, so the above corollary implies that, whenever $b_1, \ldots, b_n$ are pairwise non-conjugate elements in K, the zeros of the polynomial $(t - b_n) \cdots (t - b_1)$ are $\{a_1, \ldots, a_n\}$, where each a_i is conjugate to b_i (and $a_1 = b_1$). This result has been proved independently by Lok Sun Siu, in the case where $\dim_F K < \infty$. [As Siu has pointed out, this result may be viewed as the converse to the result that, if $b_1, \ldots, b_n$ are pairwise non-conjugate in K, there is a unique monic polynomial of degree n that vanishes on $\{b_1, \ldots, b_n\}$ (namely, the minimal polynomial of this set).] This result is, however, not true in general for *centrally infinite* division algebras, as we will see in an example in §10.

Next we come to the topic of metro equations. In the theory of division rings, the study of the equation $ax - xb = d$ has had a long history, going back to the work of R. E. Johnson [Jo] and N. Jacobson [Ja$_2$] in the 1940s. By a slight abuse of the terminology of P. M. Cohn ([Co$_2$], [Co$_3$]), we shall call $ax - xb = d$ a "metro equation" over K. (For an account on the origin of this terminology, see [Co$_3$: p.418].) It turns out that the notion of metro equations bears a close relationship to that of Wedderburn polynomials. In the following, we'll try to explain this interesting relationship. In the process of doing so, we actually obtain an extension of the metro equation notion to the general (S, D)-setting, which did not seem to have been introduced before.

For $a, b, d \in K$, let us call

$$ax - S(x)b - D(x) = d \tag{6.5}$$

the (S, D)-*metro equation* (associated with a, b, d). (Of course, when $(S, D) = (I, 0)$, (6.5) boils down to the ordinary metro equation $ax - xb = d$.) In the case $d = 0$, (6.5) has an obvious solution $x = 0$, so in the following, we'll assume $d \neq 0$ whenever (6.5) is considered. The following result gives the precise relationship between (6.5) and *quadratic* Wedderburn polynomials, in the general (S, D)-setting. (We continue to write $\mathcal{W}$ for the set of W-polynomials in $R = K[t, S, D]$.)

Theorem 6.6. *For any a, $b \in K$ and $d \in K^*$, the following are equivalent:*

(1) *The (S, D)-metro equation $ax - S(x)b - D(x) = d$ has a solution in K;*

(2) *The equation $b^x = a - dx^{-1}$ has a solution $x \in K^*$;*

(3) $(t - b^d)(t - a) \in \mathcal{W}$.

(4) $1 \in R \cdot (t - b^d) + (t - a) \cdot R$.

In fact, $x \in K^*$ is a root for the equation in (2) (or the equation in (1)) iff b^x is a root of $(t - b^d)(t - a)$ different from a.

Proof. In view of the definition of (S, D)-conjugation, the equation in (2) amounts to

$$S(x)bx^{-1} + D(x)x^{-1} = a - dx^{-1}.$$

Right multiplying this by x and transposing, we obtain the (S, D)-metro equation in (1). This shows that (1) $\Longleftrightarrow$ (2). (Note that, since $d \in K^*$, any solution x for (1) is necessarily nonzero.)

Next, we prove (2) $\iff$ (3). For $h(t) := t - a$, we may rewrite the equation in (2) in the form

$$-d = (b^x - a)x = h(b^x)x = \lambda_{h,b}(x).$$

Thus, (2) amounts to $d \in \operatorname{im}(\lambda_{h,b})$, or equivalently, $b^d \in \operatorname{im}(\Phi_h)$ (by (4.13)(3)). By (1) $\iff$ (3) in (5.6), this last condition is equivalent to $(t - b^d)h(t) \in \mathcal{W}$.

Finally, (3) $\iff$ (4) follows from (5.10) since $t-a$ and $t-b^d$ are W-polynomials. The proof for the last statement in the Theorem can be easily extracted from the arguments above. $\square$

Having nailed down the basic connection between W-polynomials and the (S, D)-metro equations, it is now a simple matter to apply the result (6.1) to get useful information on such metro equations. The two corollaries below are extensions of classical results to the general (S, D)-setting; see the explanations after the proof of (6.8).

Corollary 6.7. *Let $\Delta := \Delta^{S,D}(b)$ be an algebraic (S, D)-conjugacy class, and let $a \in K \setminus \Delta$, and $d \in K^*$. Then $(t - b^d)(t - a) \in \mathcal{W}$, and the (S, D)-metro equation (6.5) has a unique solution in K.*

Proof. This follows by taking $p(t) = t - a$ in (6.1)(3), and then applying (6.6). Since the map $\lambda_{p,b}$ is *bijective* in this case, the proof of (6.6) shows that the solution for the (S, D)-metro equation exists and is unique (in K). $\square$

We can also deduce easily some criteria for the (S, D)-metro equation (6.5) to be solvable, in the case when a and b lie in a single (S, D)-conjugacy class that is algebraic.

Corollary 6.8. *Let $\Delta := \Delta^{S,D}(a)$ be an algebraic (S, D)-conjugacy class, with minimal polynomial $f(t) \in R$. Write $f(t) = g(t)(t - a)$, and let $b \in \Delta$, and $d \in K^*$. Then the following are equivalent:*

(1) *The (S, D)-metro equation $ax - S(x)b - D(x) = d$ has a solution in K;*

(2) $f(t) \in R \cdot (t - b^d)(t - a)$;

(3) $(t - b^d)(t - a) \in \mathcal{W}$.

(4) $g(b^d) = 0$.

Proof. (1) $\iff$ (3) is directly from (6.6). By cancellation, (2) amounts to $g(t) \in R \cdot (t - b^d)$, which, in turn, amounts to $g(b^d) = 0$ (by the Remainder Theorem). Therefore, we have (2) $\iff$ (4).

(2) $\Longrightarrow$ (1). By the remarks made before (3.2), $f = f_\Delta \in \mathcal{W}$. Therefore, by the Factor Theorem 5.1, (2) implies that $(t - b^d)(t - a) \in \mathcal{W}$. Now (1) follows from (6.6).

(1) $\Longrightarrow$ (2). Reversing the argument, (1) implies that $q(t) := (t - b^d)(t - a) \in \mathcal{W}$, by (6.6). Since $a,\, b^d \in \Delta$, (6.1)(2) yields $f(t) \in R \cdot q(t)$. $\square$

In the classical case where $(S, D) = (I, 0)$, (6.7) and (6.8) are well known, and can be found in Theorem 8.5.4 of [Co$_3$]. In this case, (6.8) was first proved by P. M. Cohn in [Co$_1$], and the special case when $b = a$ goes back to R. E. Johnson [Jo: Thm. 2]. But even in this classical case, our proofs differ substantially from those of Cohn and Johnson.

The following special case of (6.8) will perhaps help us better appreciate its meaning.

Corollary 6.9. *Suppose* $\Delta := \Delta^{S,D}(a)$ *is algebraic of rank* 2, *with minimal polynomial* $f(t) = (t-e)(t-a)$, *and let* $b \in \Delta$, $d \in K^*$. *Then* $ax - S(x)b - D(x) = d$ *is solvable in* K *iff* $b^d = e$.

For instance, when $(S, D) = (I, 0)$ and K is the division ring of the real quaternions, this Corollary says that, if $a, b \notin \mathbb{R}$ are conjugate quaternions, then $ax - xb = d$ has a solution iff $db = \overline{a}\, d$ where $\overline{a}$ is the quaternionic conjugate of a. In the case when $a = b$, this was noted by R. E. Johnson in 1944 (see [Jo: Cor. 1]).

By specializing the parameters a, b, d and varying the choices of S and D in (6.6), we get many nice examples of quadratic Wedderburn polynomials. Let us record some explicit ones.

Examples 6.10.

(1) $t(t-a) \in K[t, S]$ *is a W-polynomial iff* $a \neq 0$. (This follows from (6.6) by setting $D = 0$, $b = 0$, and $d = 1$.)

(2) *For* $d \in K^*$, $\big(t - D(d)d^{-1}\big)t \in K[t, S, D]$ *is a W-polynomial iff* $d \in \mathrm{Im}(D)$. (This follows from (6.6) by setting $a = b = 0$.)

(3) Let K be the division hull of the Weyl algebra $\mathbb{Q}\langle u, v\rangle$ with the relation $uv - vu = 1$, and take $(S, D) = (I, 0)$. *Then the quadratic polynomial* $f(t) = (t-u)^2 \in K[t]$ *is a W-polynomial.* (This follows from (6.6) by setting $a = b = u$ and $d = 1$. More directly, it also follows by checking that $f(t)$ vanishes on both u and $u - v^{-1}$.) In fact, it can be shown that *all powers* $(t-u)^n \in K[t]$ are W-polynomials; the proof of this will be given in [LL_5]. This is noteworthy since $(t-u)^n$ (for $n \geq 2$) is *not* a W-polynomial over $\mathbb{Q}(u)$, but "becomes" a W-polynomial when we pass from the field $\mathbb{Q}(u)$ to the division ring K. On the other hand, it is easy to see that, if $(K_1, S_1, D_1) \subseteq (K_2, S_2, D_2)$ (in the sense that S_2, D_2 restrict to S_1, D_1), any W-polynomial in $K_1[t, S_1, D_1]$ remains a W-polynomial in $K_2[t, S_2, D_2]$.

§7. The Rank Theorem

This section will be devoted to some further applications of the Factor Theorem (5.1), particularly to questions on the union and intersection of algebraic sets and their ranks. The principal result here is the Rank Theorem 7.3. The proof of this requires the basic proposition below, which will also turn out to be crucial for the applications to modular lattices we have in mind for §8.

Proposition 7.1. *The intersection of any nonempty family of full algebraic sets* $\{\Delta_j : j \in J\}$ *is also a full algebraic set, with minimal polynomial given by* $\mathrm{rgcd}\{f_{\Delta_j} : j \in J\}$.

Proof. Let $f_j = f_{\Delta_j}$, for every $j \in J$. Then $V(f_j) = \Delta_j$ since Δ_j is full. Let $x \in K$ be any element that is P-dependent on $\Delta := \bigcap_{j \in J} \Delta_j$. Then x is P-dependent on each Δ_j and hence $f_j(x) = 0$. Therefore, $x \in \bigcap_j V(f_j) = \Delta$. This shows that Δ is a *full* algebraic set. Let $p := f_\Delta$. Of course, p is a right common divisor of the f_j's. To see that it is the *greatest* right common divisor, consider *any* $g \in R$ that right divides all f_j. By the Factor Theorem, g *is a W-polynomial.* On the other hand,

$$V(g) \subseteq \bigcap_j V(f_j) = \bigcap_j \Delta_j = \Delta = V(p),$$

since Δ is full. Therefore, (3.4) implies that g is a right divisor of p. This shows that $p = \mathrm{rgcd}\{f_{\Delta_j} : j \in J\}$. □

Remark 7.2. The fact that $\bigcap_j \Delta_j$ has minimal polynomial $\mathrm{rgcd}\{f_{\Delta_j} : j \in J\}$ is generally not true if the algebraic sets Δ_j are not all *full*. For instance, in the division ring K of real quaternions, $\Delta = \{i\}$ is full and $\Gamma = \{j, k\}$ is not full. The rgcd of $f_\Delta = t - i$ and $f_\Gamma = t^2 + 1$ is $t - i$. But $\Delta \cap \Gamma = \emptyset$ has minimal polynomial 1.

With the aid of (the finite case of) (7.1), we can also translate our earlier Degree Equation (4.1) into the following, which provides an ultimate refinement to (4.3).

The Rank Theorem 7.3. *For any two algebraic sets* Δ *and* Γ, *we have*

$$\mathrm{rk}(\Delta) + \mathrm{rk}(\Gamma) = \mathrm{rk}(\Delta \cup \Gamma) + \mathrm{rk}(\overline{\Delta} \cap \overline{\Gamma}). \tag{7.4}$$

In particular, $\mathrm{rk}(\Delta \cup \Gamma) = \mathrm{rk}(\Delta) + \mathrm{rk}(\Gamma)$ *iff* $\overline{\Delta} \cap \overline{\Gamma} = \emptyset$.

Proof. Let $f := f_\Delta = f_{\overline{\Delta}}$, $h := f_\Gamma = f_{\overline{\Gamma}}$, and $p := \mathrm{rgcd}(f, h)$, $q := \mathrm{llcm}(f, h)$. Then $q = f_{\Delta \cup \Gamma}$ by (4.3), and $p = f_{\overline{\Delta} \cap \overline{\Gamma}}$ by (7.1). Therefore, the Degree Equation

$$\deg(f) + \deg(h) = \deg(p) + \deg(q)$$

in (4.1) transcribes into (7.4). The last statement in the Theorem follows immediately from this equation. □

Remark 7.5. The above Rank Theorem may be viewed as an analogue of the well-known dimension equation for the sum and intersection of two finite-dimensional subspaces in a given vector space. This analogy, however, may belie the depth of (7.4). As a matter of fact, the usual approach to the dimension equation for vector subspaces *does not* seem to work for the proof of (7.3). After observing that $\mathrm{rk}(\Delta \cup \Gamma) = \mathrm{rk}(\overline{\Delta} \cup \overline{\Gamma})$ (which is quite easy to prove), we need only prove (7.3) in the case when Δ and Γ are both full. Following the "usual" proof, we would start with a P-basis $\{c_1, \ldots, c_r\}$ for $\Delta \cap \Gamma$, and complete this to a P-basis $\{c_1, \ldots, c_r, a_{r+1}, \ldots, a_n\}$ for Δ, and to a P-basis $\{c_1, \ldots, c_r, b_{r+1}, \ldots, b_m\}$ for Γ, where $n = \mathrm{rk}(\Delta)$, and $m = \mathrm{rk}(\Gamma)$. It is easy to see that the union $\Delta \cup \Gamma$ is P-dependent on the set

$$C := \{c_1, \ldots, c_r, a_{r+1}, \ldots, a_n, b_{r+1}, \ldots, b_m\}, \tag{7.6}$$

and hence we get

$$\mathrm{rk}(\Delta \cup \Gamma) \le |C| \le n + m - r = \mathrm{rk}(\Delta) + \mathrm{rk}(\Gamma) - \mathrm{rk}(\Delta \cap \Gamma).$$

To see that equality holds here would require proving that the set C in (7.6) is *P-independent.* This fact does not seem to be easily checkable from the definition of (and known facts about) P-independence, although, of course, we do know it to be true once we have proved (7.4).

§8. A Lattice Duality

Results such as (4.3) and (7.1) in the previous sections lead us quickly to the construction of several lattices, as follows. For the first one, consider the poset $\mathcal{F} = \mathcal{F}(K, S, D)$ of all *full* algebraic sets in K (with respect to (S, D)), where the partial ordering is given by inclusion:

$$(8.1) \qquad \Delta \leq \Gamma \iff \Delta \subseteq \Gamma \quad (\text{for } \Delta,\, \Gamma \in \mathcal{F}\,).$$

This poset is a *lattice*, with $\Delta \wedge \Gamma$ given by $\Delta \cap \Gamma$ (which lies in $\mathcal{F}$ by (7.1)), and with $\Delta \vee \Gamma$ given by $\overline{\Delta \cup \Gamma}$. (The union $\Delta \cup \Gamma$ is algebraic, but may not be *full*, as the case $|\,\Delta\,| = |\,\Gamma\,| = 1$ already shows.) Note that the lattice $\mathcal{F}$ has a smallest element, given by the empty set $\emptyset$ (see (3.2)(1)).

In the majority of cases, $\mathcal{F}$ will not have a largest element. (In fact, $\mathcal{F}$ has a largest element iff K itself is an algebraic set; e.g. in the case when K is a finite field.) Technically, however, it would be convenient to have a largest element. We can achieve this by introducing a lattice $\mathcal{F}^*$, which is defined to be just $\mathcal{F}$ if K happens to be (S, D)-algebraic, and $\mathcal{F}$ adjoined with one point K otherwise (this point being larger than all other points). We shall call $\mathcal{F}^*$ the *augmented lattice* of full algebraic sets.

To get a lattice in the context of (skew) polynomials, we consider the set $\mathcal{W} = \mathcal{W}(K, S, D)$ of all W-polynomials, partially ordered by:

$$(8.2) \qquad f \leq h \iff Rf \subseteq Rh \iff h \text{ is a right divisor of } f \quad (\text{for } f,\, h \in \mathcal{W}).$$

The poset $\mathcal{W}$ is again a lattice: for $f,\, h \in \mathcal{W}$ as above, $f \vee h$ is given by $\operatorname{rgcd}(f,\, h)$ (this being a W-polynomial by the Factor Theorem), and $f \wedge h$ is given by $\operatorname{llcm}(f,\, h)$ (this being a W-polynomial since it is the minimal polynomial of $V(f) \cup V(h)$ by (4.3)). The lattice $\mathcal{W}$ has a largest element, given by $1 \in \mathcal{W}$, and it will have a smallest element iff K happens to be (S, D)-algebraic. (The smallest element in the latter case is the minimal polynomial of K itself. For instance, if $K = \mathbb{F}_q$ and $(S, D) = (I, 0)$, then this smallest element is $t^q - t$.) In analogy with the case of full algebraic sets, we can introduce an *augmented W-polynomial lattice* $\mathcal{W}^*$, which is defined to be $\mathcal{W}$ if K is (S, D)-algebraic, and $\mathcal{W}$ adjoined with the polynomial 0 otherwise (this being smaller than all other W-polynomials).

Remark 8.3. Of course, there are two other lattices lurking in the background of the ones we have introduced. If we write $\mathcal{L} = \mathcal{L}_R$ for the set of *all* (principal) left ideals in R, then $\mathcal{L}$ is a (well-known) lattice under the usual partial ordering given by inclusion. For convenience, we shall "identify" a monic polynomial f with the principal left ideal $R \cdot f$ it generates. Then, by (4.3) and (7.1), $\mathcal{W}$ and $\mathcal{W}^*$ are *sublattices* of $\mathcal{L}$. Similarly, we can look at the lattice $\mathcal{A}$ of all algebraic subsets of K, with the point K adjoined if necessary. Here, the partial ordering is again given by inclusion, and $\Delta \vee \Gamma$ is simply given by $\Delta \cup \Gamma$ (without taking the closure). We have the inclusions $\mathcal{F} \subseteq \mathcal{F}^* \subseteq \mathcal{A}$, although here $\mathcal{F}$ and $\mathcal{F}^*$ are no longer sublattices of $\mathcal{A}$. If we define mappings

$$\sigma : \mathcal{L} \longrightarrow \mathcal{A} \quad \text{and} \quad \tau : \mathcal{A} \longrightarrow \mathcal{L}$$

by taking zero sets and taking vanishing polynomials, we get a *Galois connection* between $\mathcal{L}$ and $\mathcal{A}$ (in the sense of [St: p.xx]). The posets $\mathcal{W}^*$ and $\mathcal{F}^*$ are precisely the sets of "closed points" under this Galois connection. The fact that these two posets are anti-isomorphic under the maps σ and τ is a general conclusion deducible from the basic theory of Galois connections.

Theorem 8.4. *For a fixed triple (K, S, D), we have the following*:

(1) *$\mathcal{F}^*$ and $\mathcal{W}^*$ are both complete modular lattices. The maps $\Delta \mapsto f_\Delta$ and $f \mapsto V(f)$ (extended in the obvious way) define mutually inverse lattice dualities between $\mathcal{F}^*$* and *$\mathcal{W}^*$.*

(2) *The map* "rk" *(extended in the obvious way) is the "dimension function" on the modular lattice $\mathcal{F}^*$ in the sense of lattice theory,*[4] *and the degree map* "deg" *is the "dual dimension function" on the modular lattice $\mathcal{W}^*$.*

(3) *The (nontrivial) minimal elements (the so-called atoms) of the lattice $\mathcal{F}^*$ are the singleton subsets of K, and the (proper) maximal elements of the lattice $\mathcal{W}$ are the monic linear polynomials in R.*

(4) *A subset $A \subseteq K$ is P-independent in our sense iff the singleton (full algebraic) sets $\{a\}$ $(a \in A)$ are independent in the lattice $\mathcal{F}^*$ in the sense of lattice theory, as expounded, e.g. in* [CD: p. 46].

(5) *For $f \leq g \in \mathcal{W}$, the "interval" $[f, g]$ in the lattice $\mathcal{W}$ is isomorphic to the lattice of all R-submodules of Rg/Rf.*

Proof. First let us clarify the phrase "extended in the obvious way" (used twice above, in (1) and in (2)). When K itself is not (S, D)-algebraic (which is the majority of cases), we have the adjoined points K in $\mathcal{F}^*$ and 0 in $\mathcal{W}^*$. We simply make these correspond to each other. This is reasonable, since $V(0)$ is indeed K, and the "minimal polynomial" of K can only be taken to be 0 if K is not algebraic. (See Footnote (1).) To define "rk" and "deg" on the adjoined points, of course we use the usual conventions: $\deg(0) = -\infty$, and $\mathrm{rk}(K) = \infty$ if K is not (S, D)-algebraic.

Certainly the maps set up in the Theorem are inverses of one another, on $\mathcal{F}^*$ and on $\mathcal{W}^*$. If $f \leq h$ in $\mathcal{W}^*$ (or even in $\mathcal{L}$ as in (8.3)), then $f \in Rh$, and so $V(f) \geq V(h)$ in $\mathcal{F}^*$. On the other hand, if $\Delta \leq \Gamma$ in $\mathcal{F}^*$ (or even in $\mathcal{A}$ as in (8.3)), then $f_\Gamma \in R \cdot f_\Delta$, and so $f_\Delta \geq f_\Gamma$ in $\mathcal{W}^*$. This checks, in particular, that our maps define poset dualities (and hence lattice dualities) between $\mathcal{F}^*$ and $\mathcal{W}^*$.

In view of this lattice duality, it is sufficient to show that $\mathcal{W}^*$ is a complete modular lattice. Now it is well-known that $\mathcal{L}$ (defined in (8.3)) is a complete modular lattice. To prove the same for $\mathcal{W}^*$, it is convenient to view $\mathcal{W}^*$ as a sublattice of $\mathcal{L}$. Consider any subset T of $\mathcal{W}^*$. By the Factor Theorem (5.1), it is clear that the join $\bigvee T \in \mathcal{L}$ is actually in $\mathcal{W}^*$. As for the meet $\bigwedge T \in \mathcal{L}$, write $T = \{Rf_i : i \in I\}$, where (as we may assume) each $f_i \in \mathcal{W}$. Then $\bigcap_i Rf_i = Rf$ for some monic $f \in R$. It is easy to see that f is the minimal polynomial of $\bigcup_i V(f_i)$, and therefore $\bigwedge T = Rf \in \mathcal{W}^*$. (It is possible that $\bigcup_i V(f_i)$ is no

[4]Recall from Crawley and Dilworth [CD: p. 27] that, in a (semi)modular lattice L, the *dimension* of an element $x \in L$ is the integer n such that the interval $[0, x] \subseteq L$ has a maximal chain of length $n + 1$. The referee pointed out, however, that the term "dimension function" has been used in different senses by other authors, and that, in the context here, "height function" may be a more suggestive term to use.

longer (S, D)-algebraic. In this case, we simply have $f = 0 \in \mathcal{W}^*$. Otherwise, $f \in \mathcal{W}$.) What the above remarks showed is that $\mathcal{W}^*$ is a *complete sublattice* of the complete lattice $\mathcal{L}$, in the sense of lattice theory (see, e.g. [CD: p. 11]). From this observation, it follows right away that $\mathcal{W}^*$ is also a complete modular lattice. This establishes (1).

The statement (2) of the Theorem concerning "rk" and "deg" is now clear from the definition of "dimension functions" recalled in Footnote (4). (3) follows from this (and is easy to see in any case without (2)).

For (4), consider the set of singletons $\{\{a\} : a \in A\}$ in the complete lattice $\mathcal{F}^*$. Such a set is *independent* in the lattice sense if

$$(*) \qquad \{a\} \wedge \bigvee \{\{b\} : a \neq b \in A\} = \emptyset \in \mathcal{F}^* \quad (\forall\, a \in A).$$

Here, $\bigvee \{\{b\} : a \neq b \in A\}$ is given by the P-closure of $A \setminus \{a\}$ if this set is (S, D)-algebraic, and is given by K otherwise. Thus, the negation of the statement $(*)$ means that some $a \in A$ is P-dependent on its complement $A \setminus \{a\}$, and this means exactly that $A \subseteq K$ is not a P-independent set. This proves the assertion (4).

To prove (5), note that if $f \in \mathcal{W}$, any monic $h \in R$ with $Rh \supseteq Rf$ is also in $\mathcal{W}$ by the Right Factor Theorem. Therefore, given $f \leq g$ in $\mathcal{W}$, the interval $[f, g]$ in $\mathcal{W}$ is precisely isomorphic to the lattice of all submodules of the quotient R-module Rg/Rf. $\square$

In the proof above, the modularity of $\mathcal{F}^*$ was not proved directly, but was rather deduced from that of $\mathcal{W}^*$. It is, therefore, of interest to record the following statement, which essentially amounts to the modular law for $\mathcal{F}^*$.

Corollary 8.5. *Let* Γ, Π *and* Δ *be algebraic sets, where* Γ, Π *are full, and* $\Delta \subseteq \Gamma$. *If* $x \in \Gamma$ *is P-dependent on* $\Pi \cup \Delta$, *then it is already P-dependent on the smaller set* $(\Gamma \cap \Pi) \cup \Delta$.

Proof. The modular law in $\mathcal{F}^*$, applied to the full algebraic sets Γ, Π, and $\overline{\Delta}$, says that

$$(8.6) \qquad \Gamma \wedge (\Pi \vee \overline{\Delta}) = (\Gamma \wedge \Pi) \vee \overline{\Delta}.$$

Suppose $x \in \Gamma$ is P-dependent on $\Pi \cup \Delta$. Then clearly x belongs to the LHS of (8.6). According to this equation, x must belong to the RHS, which means that x is P-dependent on $(\Gamma \cap \Pi) \cup \overline{\Delta}$. But then x is already P-dependent on $(\Gamma \cap \Pi) \cup \Delta$, as desired. $\square$

Remark 8.7. If one of Γ, Π is not full, the conclusion in the Corollary may not hold. For instance, in the real quaternions, $\Delta = \{j\}$ and $\Pi = \{k\}$ are full, but $\Gamma = \{i, j\}$ is not full. Here, $x = i$ is P-dependent on $\Pi \cup \Delta = \{j, k\}$, but is obviously *not* P-dependent on $(\Gamma \cap \Pi) \cup \Delta = \emptyset \cup \Delta = \{j\}$.

§9. Questions and Examples

In this section, we shall pose, and answer, some natural questions concerning the behavior of W-polynomials and algebraic (S, D)-conjugacy classes.

The first question is prompted by the original form of Wedderburn's Theorem in [We]. If an element $a \in K$ is algebraic over $F = Z(K)$, with minimal polynomial

$f(t) \in F[t]$, and if $(t-a_1)\cdots(t-a_n)$ is any complete factorization of f in $K[t]$, Wedderburn observed that the product of the linear factors $t-a_i$ is unchanged if they are permuted cyclically. Now, the polynomial f can also be interpreted as the minimal polynomial of the algebraic set $\Delta(a)$ (the usual conjugacy class of a). Taking this point of view, we can in fact give the following generalization of Wedderburn's result.

Proposition 9.1. *Let $R = K[t, D]$ (with $S = I$), and let $\Delta^{I,D}(a)$ be an algebraic (I, D)-conjugacy class of K, with minimal polynomial $f(t) = (t-a_1)\cdots(t-a_n)$. Then f is in the center of R, and the product of the linear factors in this factorization is unchanged if they are permuted cyclically.*

Proof. In Lemma 5.2 of [LL$_2$], it is proved that f is a *left invariant* polynomial in the sense that $R \cdot f$ is an ideal (that is, $f \cdot R \subseteq R \cdot f$). Now we'll use an argument due to S. Amitsur. For any $a \in K$, we have $fa = a'f$ for some $a' \in K$. A comparison of the leading coefficients of both sides (under the assumption that $S = I$) shows that $a' = a$. Similarly, $f \cdot t = (bt + c) \cdot f$ for some $b, c \in K$, and a comparison of the coefficients of terms of degree $n+1$ and n shows that $b = 1$ and $c = 0$. This shows that $ft = tf$. Since R is generated as a ring by K and t, the above work shows that f is central. The last conclusion now follows easily, since if f is any central element in a domain, then $f = gh \Longrightarrow f = hg$. $\square$

The above extension of Wedderburn's result leads naturally to the following question in the general (S, D)-setting.

Question 9.2. *If $f \in R = K[t, S, D]$ is the minimal polynomial of an algebraic (S, D)-conjugacy class $\Delta^{S,D}(a)$, and $(t-a_1)\cdots(t-a_n)$ is a complete splitting of f in R, is the product of the linear factors unchanged if they are permuted cyclically?*

We shall show that the answer to Question (9.2) is "no", even in the case when $\Delta^{S,D}(a)$ has rank 2 and $D = 0$. To construct our counterexample, we begin with a division ring K with an automorphism S such that $S^2 = I_a$, where $S(a) = a$. (Here, I_a denotes the inner automorphism of K sending any $x \in K$ to axa^{-1}.) By Theorem 5.17 in [LL$_2$], the class $\Delta := \Delta^{S,0}(a)$ is algebraic of rank 2. Fix an element $b := a^c \neq a$ in Δ. Then, $\{a, b\}$ is a P-basis of Δ, so according to (3.3)(3), the minimal polynomial f_Δ is given by $(t - b^{b-a})(t-a)$. Now

$$b^{b-a} = (a^c)^{b-a} = a^{(b-a)c} = a^{bc-ac} = a^{S(c)a-ac}.$$

Let us write $d := S(c)a - ac$, so that $f_\Delta(t) = (t-a^d)(t-a)$. The following lemma gives a criterion for the two linear factors to be permutable.

Lemma 9.3. *In the above notations, $f_\Delta(t)$ is also given by $(t-a)(t-a^d)$ iff $a^d \in K^S$ (the fixed-point set of S).*

Proof. By direct expansion, we have

$$(9.4) \qquad \begin{aligned} (t-a^d)(t-a) &= t^2 - \big(a^d + S(a)\big)\, t + a^d a, \\ (t-a)(t-a^d) &= t^2 - \big(a + S(a^d)\big)\, t + a a^d. \end{aligned}$$

Since $S(a) = a$, these are equal iff $a^d \in K^S$ and $a^d a = a a^d$. Now the latter amounts to $a^d = a a^d a^{-1} = S^2(a^d)$, so it is already implied by the former.

Therefore, the criterion for the equality of the two polynomials in (9.4) is simply $a^d \in K^S$. □

To construct an explicit counterexample to (9.2) (in the rank two case and with $D = 0$), it thus suffices to produce a suitable pair (K, S) with $a, b, c, d \in K$ as above such that $a^d \notin K^S$. This can be done as follows. Start with a rational function field $k(c)$ where k is any field, and let σ be the k-endomorphism on $k(c)$ defined by $\sigma(c) = c^2$. We then construct the skew polynomial ring $k(c)[\,a; \sigma^2\,]$ (with the twist law $ac = \sigma^2(c)a = c^4 a$). This is a principal left ideal domain, so it has a classical left ring of quotients, $K := k(c)(a, \sigma^2)$, which is a division ring. Now define a k-endomorphism on K by the rules: $S(a) = a$, $S(c) = c^2$. (Note that the relation $ac = c^4 a$ is respected by S, since $S(a)S(c) = ac^2 = c^8 a$, while also $S(c^4)S(a) = c^8 a$.) We do have here $S^2 = I_a$, since

$$I_a(a) = a = S^2(a), \quad \text{and} \quad I_a(c) = aca^{-1} = c^4 = S^2(c).$$

(In particular, it follows that S is an automorphism of K.) Now, in the notation of (9.3), $d := S(c)a - ac = c^2 a - ac = (c^2 - c^4)a$, and so

$$\begin{aligned} a^d &= S(d)ad^{-1} \\ &= S(c^2 - c^4)S(a) \cdot a \cdot a^{-1}\, S(c^2 - c^4) \\ &= (c^4 - c^8)\, a\, (c^4 - c^8) \\ &= (c^4 - c^8)(c^8 - c^{16})\, a. \end{aligned}$$

This element is clearly *not* fixed by S, so we have constructed the needed counterexample.

Remark. In the above construction, we made heavy use of the fact that the automorphism S on K is *not* the identity. Indeed, in the case $S = I$, *if* $\Delta = \Delta^{I,D}(a)$ *has rank* 2 *and* $f_\Delta(t) = (t-b)(t-a)$ $(a,\, b \in K)$, *then* $f_\Delta(t) = (t-a)(t-b)$ by (9.1).

Since the counterexample produced above for Question 9.2 involved the use of $(S, D) \neq (I, 0)$, one might get the impression that things have gone awry as a result of the (S, D)-twist. To correct this impression, let us now go back to the untwisted case $(S, D) = (I, 0)$, and consider the following alternate question to (9.2) that is also prompted by Wedderburn's cyclic permutation result in [We]:

Question 9.5. *If* f *is a W-polynomial in* $R = K[t]$ *and* $(t - a_1) \cdots (t - a_n)$ *is a complete splitting of* f *over* K, *is the product of the linear factors unchanged if they are permuted cyclically?*

In the following, we shall show by constructing some explicit counterexamples that the answer to this question is also "no" in general. Again, it turns out that counterexamples can be found already in degree two. We begin our considerations here by a close examination of *cubic* minimal polynomials of elements over the center.

Proposition 9.6. *Suppose* $a \in K$ *is cubic over* $F = Z(K)$, *with minimal polynomial* $f(t) = (t - c)(t - b)(t - a)$, *where* $b,\, c \in K$ *are conjugate to* a. *Then* $(t - a)(t - b) \in \mathcal{W}$ *iff* $a,\, b,\, c$ *pairwise commute (in which case the splitting field of* f *over* F *is embeddable in* K).

Proof. By the Factor Theorem (5.1), $(t-b)(t-a) \in \mathcal{W}$. If $ab = ba$, then of course $(t-a)(t-b) \in \mathcal{W}$. Conversely, suppose $(t-a)(t-b) \in \mathcal{W}$. By (6.1)(1), we have $f(t) = (t-d)(t-a)(t-b)$ for some $d \in K$. Since f is central, the original factorization of f also gives $f(t) = (t-a)(t-c)(t-b)$. Thus,

$$(t-d)(t-a) = (t-a)(t-c).$$

This gives $d + a = a + c$, and $da = ac$. Therefore, $d = c$, and $ca = ac$. It follows that $F(a, c)$ is a field, which must contain b (since $cba \in F^*$.). We have thus proved that a, b, c pairwise commute, and that $F(a, c)$ is a splitting field of f that is embedded in K. ☐

Proposition 9.7. *Suppose K is a central F-division algebra of dimension* 9*, and suppose $F(a)/F$ is a non-Galois cubic extension contained in K. If $f(t) = (t-c)(t-b)(t-a) \in F[t]$ is the minimal polynomial of a over F as in* (9.6)*, then $(t-b)(t-a) \in \mathcal{W}$, but $(t-a)(t-b) \notin \mathcal{W}$.*

Proof. If $(t-a)(t-b) \in \mathcal{W}$, then, as in the proof of (9.6), $ac = ca$ and $F(a, c)$ is a splitting field for f. Since $\dim_F K = 9$, $F(a)$ is a *maximal* subfield of K. Thus, we must have $F(a) = F(a, c)$, so $F(a)/F$ is Galois, a contradiction. Thus, $(t-a)(t-b)$ *cannot* be a W-polynomial. ☐

It is now easy to produce an explicit example (in the classical case $(S, D) = (I, 0)$) where $(t-b)(t-a) \in \mathcal{W}$ but $(t-a)(t-b) \notin \mathcal{W}$. We can take, for instance, Dickson's 9-dimensional cyclic $\mathbb{Q}$-division algebra K generated by two elements x, v with the relations

(9.8) $$v^3 + v^2 - 2v - 1 = 0, \quad x^3 = 2, \quad \text{and} \quad xv = (v^2 - 2)x.$$

(See [La$_2$: p. 239].) Here, if we choose $a = x$, $F(a) = F(x)$ is a *non-Galois* cubic extension of $\mathbb{Q}$ contained in K. A straightforward calculation shows that the minimal polynomial of x over F has a Wedderburn splitting

(9.9) $$t^3 - 2 = \big[t + (v+1)x\big](t - vx)(t - x).$$

Thus, it follows from (9.7) that $(t - vx)(t - x)$ is a W-polynomial, while $(t-x)(t-vx)$ is not. On the other hand, if we choose $a = v$, then $F(a) = F(v)$ is a (Galois) cyclic extension, and the minimal polynomial of v over F has the splitting

(9.10) $$t^3 + t^2 - 2t - 1 = \big[t - (1 - v - v^2)\big]\big[t - (v^2 - 2)\big](t - v),$$

already in $F(v)[t]$. Here, of course, the product of any two of the three linear factors is a W-polynomial over $F(v)$ and over K.

It is worth pointing out that, in view of (6.6), an example where $(t-b)(t-a) \in \mathcal{W}$ but $(t-a)(t-b) \notin \mathcal{W}$ has also the following interpretation in terms of metro equations: *for such elements $a, b \in K$, the metro equation $ax - xb = 1$ has a solution in K, but $bx - xa = 1$ does not.*

§10. Left Root/Right Root Counterexample, and an Application

From (6.3), it follows that if an element $c \in K$ belongs to an *algebraic* (S, D)-conjugacy class, then whenever a polynomial has c as a left root, it has also a (right) root that is (S, D)-conjugate to c. We shall now show by an example that this statement is no longer true if $\Delta^{S,D}(c)$ is not assumed to be (S, D)-algebraic. In fact, our example is given in the simple ("untwisted") case when $(S, D) = (I, 0)$. We shall construct a division ring K with a *quadratic* polynomial $f(t) = (t-c)(t-b) \in K[t]$ such that the left root c has no conjugate in K that is a (right) root. Note that such a polynomial f will have the following properties: (1) b and c must be nonconjugate, (2) f has a unique root b (so it is not a Wedderburn polynomial), and (3) c is necessarily transcendental over the center of K, according to (6.3).

Throughout this section, K denotes a division ring (with $(S, D) = (I, 0)$), and $C_K(x)$ denotes the centralizer of an element $x \in K$. We begin our construction with the following observation on split quadratic polynomials.

Lemma 10.1. *Let $b, c \in K$, and $f(t) = (t-c)(t-b) \in K[t]$. Then $f(t)$ has a root conjugate to c iff there exists $r \in K^*$ such that $rc - br \in C_K(c)$.*

Proof. For any conjugate rcr^{-1} of c, we have

$$f(rcr^{-1}) = rc^2r^{-1} - (b+c)rcr^{-1} + cb = \left(rc^2 - (b+c)rc + cbr\right) r^{-1}.$$

It follows that $f(rcr^{-1}) = 0$ iff $(rc - br)c = c(rc - br)$, that is, iff $rc - br \in C_K(c)$. □

Now consider a twisted Laurent series division ring $K = k((x, \sigma))$, where k is a division ring, and σ is an automorphism on k. We shall write σ in the exponential form: $a \mapsto a^\sigma$, and let k_0 be the division ring of fixed points of σ. Fixing an element $y \in k^*$, we shall apply the lemma to the quadratic polynomial

$$f(t) = (t-x)(t-y^{-1}x) \in K[t] = k((x, \sigma))[t]. \tag{10.2}$$

Lemma 10.3. *$f(t)$ has a root in K conjugate to x iff $y \in k^*$ has the form $y = (a^\sigma + \varepsilon)\, a^{-1}$ for some $a \in k^*$ and some $\varepsilon \in k_0$.*

Proof. By a slight abuse of notation, let σ also denote the inner automorphism on K induced by x: $r^\sigma = xrx^{-1}$ for $r \in K$. (This will not cause any confusion since the new σ extends the given σ on k.) We look for $r \in K^*$ such that

$$rx - y^{-1}xr = (r - y^{-1}xrx^{-1})x \in C_K(x);$$

that is, $r - y^{-1}r^\sigma \in C_K(x)$. If $r = dx^n + ex^{n+1} + \cdots$ where $d, e, \ldots \in k$, $d \neq 0$, then

$$\begin{aligned} r - y^{-1}r^\sigma &= (dx^n + ex^{n+1} + \cdots) - y^{-1}(d^\sigma x^n + e^\sigma x^{n+1} + \cdots) \\ &= (d - y^{-1}d^\sigma)\, x^n + (e - y^{-1}e^\sigma)\, x^{n+1} + \cdots . \end{aligned}$$

If this belongs to $C_K(x)$, then $d - y^{-1}d^\sigma \in k_0$, that is, $y = d^\sigma(d-\varepsilon)^{-1}$ where $\varepsilon \in k_0 \setminus \{d\}$. Writing $a = d - \varepsilon \in k^*$, we have $y = (a^\sigma + \varepsilon)\, a^{-1}$. Conversely, if y has this form, then $d - y^{-1}d^\sigma \in k_0$ for some $d \in k^*$. Choosing $r = d$, we'll have $r - y^{-1}r^\sigma \in C_K(x)$ by (a special case of) the calculation above. □

Note that $f \in K[t]$ in (10.2) has a left root x. To construct such a polynomial for which no conjugate of x is a (right) root, we need only construct a pair (k, σ)

in which there is an element $y \in k^*$ *not* of the form $(a^\sigma + \varepsilon)\, a^{-1}$, where $a \in k^*$ and $\varepsilon \in k_0$. This can be accomplished as follows.

Lemma 10.4. *Let $k = \mathbb{R}(y)$ where y is an indeterminate, and let σ be the $\mathbb{R}$-automorphism on k defined by $\sigma(y) = 2y$. Then y is not of the form $(a^\sigma + \varepsilon)\, a^{-1}$, where $a \in k^*$ and $\varepsilon \in k_0$.*

Proof. Here, $k_0 = \mathbb{R}$. Assume $y = (a^\sigma + \varepsilon)\, a^{-1}$, or equivalently $ya - \varepsilon = a^\sigma$, where $a = h(y)/g(y)$, with $h,\, g \in \mathbb{R}[y] \setminus \{0\}$, and $\varepsilon \in k_0 = \mathbb{R}$. We may assume that h/g is reduced to lowest terms. Then we have

$$\frac{y\, h(y) - \varepsilon\, g(y)}{g(y)} = \frac{h(2y)}{g(2y)}. \tag{10.5}$$

The fraction on the right is reduced to lowest terms; hence so is the fraction on the left (since $\deg\left(g(y)\right) = \deg\left(g(2y)\right)$). This implies that $g(2y)$ is a scalar multiple of $g(y)$, which in turn implies that $g(y) = \alpha \cdot y^m$, where $\alpha \in \mathbb{R}^*$, and $m \geq 0$. If $m \geq 1$, the LHS of (10.5) is not in lowest terms, as both numerator and denominator have a factor of y. Hence $m = 0$, $g(y) = \alpha$, and $h(2y) = y\, h(y) - \varepsilon\, \alpha$. This is impossible, again by considering degrees in y. $\square$

In the above example, the division ring K in (10.2) is not of the centrally finite type. Indeed, if K is of the centrally finite type, we know that such an example is impossible. This observation leads to the following curious consequence of (6.3) and (10.3).

Proposition 10.6. *Let $(k,\, \sigma)$ be a centrally finite division ring equipped with an automorphism σ of finite inner order (i.e. a positive power of σ is an inner automorphism). Then any $y \in k$ has the form $(a^\sigma + \varepsilon)\, a^{-1}$ where $a \in k^*$ and $\varepsilon \in \{0,\, 1\}$.*

Proof. Under the given hypotheses, it is known that $K = k((x,\, \sigma))$ is also a centrally finite division ring (see, e.g. [Pi: p. 384-385]). Applying (6.3) to the quadratic polynomial $f(t)$ in (10.2) for any $y \in k^*$, we know that f has a root in K conjugate to x. Thus, by (10.3), y has the form $(a^\sigma + \varepsilon)\, a^{-1}$, where $a \in k^*$, and $\varepsilon \in k$ with $\sigma(\varepsilon) = \varepsilon$. If $\varepsilon = 0$, we get $y = a^\sigma\, a^{-1}$; otherwise,

$$y = (a^\sigma\, \varepsilon^{-1} + 1)\, \varepsilon\, a^{-1} = (b^\sigma + 1)\, b^{-1}, \quad \text{with} \quad b := a\, \varepsilon^{-1} \in k^*.$$

This proves the Proposition for $y \neq 0$. If $y = 0$, the Proposition holds by taking $\varepsilon = 1$ and $b = -1$. $\square$

This Proposition does not supersede the Hilbert 90 Theorem, but rather, supplements it. For instance, if $\sigma^n = I$ in (10.6), and $y \in k$ is such that $y^{\sigma^{n-1}} \cdots y^\sigma y$ is *not* equal to 1, (10.6) implies that y has the form $(a^\sigma + 1)\, a^{-1}$ for some $a \in k^*$. This works with only a centrally finite assumption on k, and with no assumptions on the restriction of σ to the center of k.

References

[Al] A. A. Albert: *On ordered algebras*, Bull. A.M.S. **45**(1940), 521-522.

[Co_1] P. M. Cohn: *The range of derivations on a skew field and the equation $ax - xb = c$*, J. Indian Math. Soc. **37**(1973), 1-9.

[Co_2] P. M. Cohn: *Free Rings and Their Relations*, 2nd Edition, London Math. Soc. Monograph No. 19, Academic Press, London/New York, 1985.

[Co_3] P. M. Cohn: *Skew Fields. Theory of General Division Rings*, Encyclopedia in Math., Vol. 57, Cambridge Univ. Press, Cambridge, 1995.

[CD] P. Crawley and R. P. Dilworth: *Algebraic Theory of Lattices*, Prentice Hall Inc., Englewood Cliffs, N.J., 1973.

[HR] D. E. Haile and L. H. Rowen: *Factorization of polynomials over division algebras*, Algebra Colloq. **2**(1995), 145-156.

[Ja_1] N. Jacobson: *The Theory of Rings*, Math. Surveys, No. 2, Amer. Math. Soc., Providence, R.I., 1943.

[Ja_2] N. Jacobson: *The equation $x' \equiv xd - dx = b$*, Bull. A.M.S. **50**(1944), 902-905.

[Ja_3] N. Jacobson: *Finite-Dimensional Division Algebras over Fields,* Springer-Verlag, Berlin-Heidelberg-New York, 1996.

[Jo] R. E. Johnson: *On the equation $\chi\alpha = \gamma\chi + \beta$ over an algebraic division ring*, Bull. A.M.S. **50**(1944), 202-207.

[La_1] T. Y. Lam: *A general theory of Vandermonde matrices*, Expositiones Mathematicae **4**(1986), 193-215.

[La_2] T. Y. Lam: *A First Course in Noncommutative Rings*, Graduate Texts in Math., Vol. **131**, Springer-Verlag, Berlin-Heidelberg-New York, 1991.

[La_3] T. Y. Lam: *Exercises in Classical Ring Theory*, Problem Books in Mathematics, Springer-Verlag, Berlin-Heidelberg-New York, 1995.

[LL_1] T. Y. Lam and A. Leroy: *Vandermonde and Wronskian matrices over division rings*, J. Algebra **119**(1988), 308-336.

[LL_2] T. Y. Lam and A. Leroy: *Algebraic conjugacy classes and skew polynomial rings*, in: "Perspectives in Ring Theory", (F. van Oystaeyen and L. Le Bruyn, eds.), Proceedings of the Antwerp Conference in Ring Theory, pp. 153-203, Kluwer Academic Publishers, Dordrecht/Boston/London, 1988.

[LL_3] T. Y. Lam and A. Leroy: *Hilbert 90 Theorems for division rings*, Trans. A.M.S. **345**(1994), 595-622.

[LL_4] T. Y. Lam and A. Leroy: *Principal one-sided ideals in Ore polynomial rings,* in *Algebra and Its Applications* (D.V. Huynh, S.K. Jain and S.R. López-Permouth, eds.), Contemp. Math. **259**, pp. 333-352, Amer. Math. Soc., Providence, R.I., 2000.

[LL_5] T. Y. Lam and A. Leroy: *Wedderburn polynomials over division rings,* II, in preparation.

[Or] O. Ore: *Theory of noncommutative polynomials*, Annals of Math. **34**(1933), 480-508.

[Pi] R. S. Pierce: *Associative Algebras,* Graduate Texts in Math. **88**, Springer-Verlag, Berlin-Heidelberg-New York, 1982.

[Ro_1] L. H. Rowen: *Polynomial Identities in Ring Theory*, Academic Press, London-Toronto-New York, 1980.

[Ro_2] L. H. Rowen: *Wedderburn's method and algebraic elements in simple artinian rings*, Contemp. Math. **124**(1991), 179-202.

[Ro_3] L. H. Rowen: *Polynomials over division rings, and their applications*, in "Ring Theory, Granville, Ohio, 1992" (S. K. Jain and S. T. Rizvi, eds.), pp. 287-301, World Scientific Publ. Co., Singapore-Hong Kong, 1993.

[RS_1] L. H. Rowen and Y. Segev: *The finite quotients of the multiplicative group of a division algebra of degree* 3 *are solvable*, Israel J. Math. **111**(1999), 373-380.

[RS_2] L. H. Rowen and Y. Segev: *The multiplicative group of a division algebra of degree* 5 *and Wedderburn's factorization theorem*, in *Algebra and Its Applications* (D.V. Huynh, S.K. Jain and S.R. López-Permouth, eds.), Contemp. Math. **259**, pp. 475-486. Amer. Math. Soc., Providence, R.I., 2000.

[Se] Y. Segev: *Some applications of Wedderburn's factorization theorem*, Bull. Austral. Math. Soc. **59**(1999), 105-110.

[St] B. Stenström: *Rings of Quotients,* Grundlehren d. Math. Wissenschaften **217**, Springer-Verlag, Berlin-Heidelberg-New York, 1975.

[Tr] J. Treur: *Separate zeros and Galois extensions of skew fields*, J. Algebra **120**(1989), 392-405.

[We] J. H. M. Wedderburn: *On division algebras*, Trans. A.M.S. **22**(1921), 129-135.

Department of Mathematics, University of California, Berkeley, CA 94720
E-mail address: lam@math.berkeley.edu

Department of Mathematics, Université d'Artois, 62307 Lens Cedex, France
E-mail address: leroy@euler.univ-artois.fr