Can. J. Math., Vol. XXXV, No. 3, 1983, pp. 509-515

POLYNOMIALS WITH COEFFICIENTS FROM A DIVISION RING

UNA BRAY AND GEORGE WHAPLES

1. Introduction. Let R be any division ring and let

$$(1)\quad f(X) = a_0 + a_1X + a_2X^2 + \ldots + a_nX^n, \quad (a_i \in R, 1 \leqq i \leqq n)$$

be a polynomial, in the indeterminate X, with coefficients in R. Note that the powers of X are always to the right of the coefficients. We denote the set of all such polynomials by $R[X]$.

B. Beck [**3**] proved the following theorem for the generalized quaternion division algebra; i.e., any division ring of dimension 4 over its center:

THEOREM 1. *If $f(X)$ is of degree n then $f(X)$ has either infinitely many or at most n zeros in R.*

Under a reasonable definition of multiplicity Beck also proved:

THEOREM 2. *Let $(c_1, c_2, \ldots, c_n)$ be a set of pairwise non-conjugate elements of R, and $(m_1, \ldots, m_N)$ positive integers such that $\sum m_i = n =$* deg *$f(x)$.*

(A) *If the m_i are all equal to 1, there is a unique $f(X)$ of degree n and leading coefficient 1 with $c_1, \ldots, c_n$ as its only zeros.*

(B) *If one of the m_i is greater than 1, then there are infinitely many $f(X)$ of degree $\sum_{i=1}^{n} m_i$, leading coefficient 1, c_i a zero of multiplicity m_i and no other zeros.*

In this paper we present elementary proofs that Theorems 1 and 2A hold for every division ring R and Theorem 2B holds for every division algebra.

2. Polynomials. Let R denote any division ring and k its center. We make $R[X]$ into a ring by defining addition in the usual way and multiplication, which we shall denote by "$\circ$", also in the usual way:

$$(2)\quad f(X) \circ g(X) = h(X) = \sum_{\nu=0}^{n} a_\nu g(X)X^\nu.$$

Received March 17, 1982 and in revised form August 18, 1982.

Obviously this product does not in general correspond to the product of polynomial functions but we do have:

PROPOSITION 1. *Let* $f(X)$ *and* $g(X)$ *be in* $R[X]$. (a) *Every zero of* $g(X)$ *is also a zero of* $f(X) \circ g(X)$. (b) *Every zero of* $f(X) \circ g(X)$ *which is not a zero of* $g(X)$ *is a conjugate of a zero of* $f(X)$. (c) *If* $c \in R$ *then* c *is a zero of* $g(X)$ $\Leftrightarrow g(X) = q(X) \circ (X-c)$ *for some* $q(X) \in R[X]$.

Proof. To prove (a) let c be a zero of $g(X)$ and substitute c for X in (2). For (b), let $h(X) = f(X) \circ g(X)$ and suppose $g(c) \neq 0$; then by (2) we have

$$h(c) = f(g(c)cg(c)^{-1})g(c)$$

which cannot be zero unless c is a conjugate of a zero of $f(X)$.

To prove (c) note that by the ordinary division algorithm (taking care to keep products of coefficients in the correct order) we can find an identity

$$g(X) = q(X) \circ (x - c) + b \quad \text{with } b \in R.$$

By (a), $g(c) = b$.

PROPOSITION 2. *Let* $f(X)$ *be given by* (1), *let* $c \in R$, *and let* S *denote the set of all* $y \in R$ *with* $f(ycy^{-1}) = 0$. *Then* S *equals the set of all nonzero* $y \in R$ *with*

$$(3) \quad \sum_{\nu=0}^{n} a_\nu y c^\nu = 0.$$

Proof. The left side of (3) equals $f(ycy^{-1}) \circ y$.

Now the left side of (3) defines a k-linear homogeneous function

$$y \to l(y) = \sum_{\nu} a_\nu y c^\nu$$

of $R \to R$ where k is the center of R. This is an analytic linear function $R \to R$ of the type studied in [**2**]. The set of solutions of (3) forms a vector space over k. If R should be finite dimensional over k then the set of solutions of (3) can be computed by solving a system of linear equations over k. Hence, when R is a division algebra, we can by a programmable computation find for each $c \in R$ the set of all $y \in R$ with $f(ycy^{-1}) = 0$.

PROPOSITION 3. *If* $f(X) \in R[X]$ *and* $f(X)$ *has two different conjugates of an element* $c \in R$ *as zeros then* $f(X)$ *has infinitely many conjugates of* c *as zeros.*

https://doi.org/10.4153/CJM-1983-028-1 Published online by Cambridge University Press

Proof. We may assume that $f(X)$ is given by (1), that $f(c) = 0$, and that (3) has a solution b_1 with $b_1 c b_1^{-1} \neq c$.

Let $R^{(c)}$ denote the set of all $z \in R$ which commute with c. It is a sub-division-ring of R which contains the commutative subfield $k(c)$. If $l(y) = \sum a_\nu y c^\nu$ then

$$l(yz) = l(y)\cdot z \quad \text{for every } z \in R^{(c)}.$$

Note that l is a right $R^{(c)}$-linear map from R to R. The set of solutions of $l(y) = 0$ is a right vector space over $R^{(c)}$.

Let $b_0, \dots, b_n$ be a family of solutions of (3) which are linearly independent as elements of the right vector space over $R^{(c)}$. Under the assumption of the first paragraph 1 and b_1 are linearly independent because $b_1 \notin R^{(c)}$. (The case $n = 1$ suffices for this proof, but we shall need the more general case later.) Let Y be the set of all

$$\sum_{\nu=0}^{n} b_\nu z_\nu \neq 0, \quad z_\nu \in R^{(c)},$$

and S_Y the set of all ycy^{-1} with $y \in Y$. Then S_Y is a subset of the set of all conjugates of c which are zeros of $f(X)$.

If

$$y = \sum_{\nu=0}^{n} b_\nu z_\nu \quad \text{and} \quad y' = \sum_{\nu=0}^{n} b_\nu z'\nu$$

then

$$ycy^{-1} = y'cy'^{-1} \Leftrightarrow y' = yz \quad \text{for some } z \in R^{(c)};$$

and from this we can see that the elements of S_Y are in one-one correspondence with the set of equivalence classes of $n + 1$ tuplets $(z_0, z_1, \dots, z_n) \neq 0$ of elements of $R^{(c)}$ where we define $(z_0, \dots, z_n)$ and $(z'_0, \dots, z'_n)$ to be equivalent if and only if there is a z in $R^{(c)}$ with

$$(z_0', \dots, z'_n) = (z_0 z, z_1 z, \dots, z_n z).$$

So we may say: The elements of S_Y are in one-one correspondence with the elements of an n-dimensional right projective space over $R^{(c)}$ [**1**].

If $R^{(c)}$ has infinitely many elements and $n \geqq 1$ then such a projective space has infinitely many elements. Now the proof of Proposition 3 is reduced to proving:

PROPOSITION 4. *If R is any noncommutative division ring and $c \in R$ then $R^{(c)}$ has infinitely many elements.*

https://doi.org/10.4153/CJM-1983-028-1 Published online by Cambridge University Press

Proof. We use the following notation: If A and B are division rings with $B \subset A$ then $(A:B)$ denotes the dimension of A as a left vector space over B. If $A \supset B \supset C$, then $(A:C) = (A:B)(B:C)$ if $(A:B)$ and $(B:C)$ are finite.

Suppose $R^{(c)}$ is a finite set. Then k must be a finite field and $(R^{(c)}:k)$ finite; since $k(c)$ is a commutative field with $k(c) \subset R^{(c)}$, then $(k(c):k)$ is also finite. By Theorem 13, part 4, of [**2**], taking $A = k(c)$, we get

$$(R:R^{(c)}) = (k(c):k).$$

Therefore, $(R:k) = (R:R^{(c)})(R^{(c}:k)$ is finite, hence R has only finitely many elements and, by a theorem of Wedderburn, [**7**], R is commutative.

PROPOSITION 5. *Let $c_1, c_2, \dots, c_n$ be a set of pairwise non-conjugate elements of R. Then there is a unique monic $f(X) \in R[X]$ of degree n such that:*

(a) *$c_1, c_2, \dots, c_n$ are zeros of $f(X)$.*
(b) *Every zero of $f(X)$ is a conjugate of one of the c_i.*
(c) *If $h(X)$ has all the c_i as zeros then $h(X) = q(X) \circ f(X)$ for a $q(X) \in R[X]$.*

Proof. This is true for $n = 1$ by Proposition 1. We use induction, assuming $n > 1$ and that the proposition is true for all sets of fewer than n pairwise non-conjugate elements. In particular there is a unique monic $g(X)$ satisfying our conditions for the set $c_1, c_2, \dots, c_{n-1}$. Consider the polynomial

$$f(X) = (X - u) \circ g(X)$$

where u is an undetermined element of R. For each $c \in R$,

$$f(c) = g(c)c - ug(c)$$

so $f(c) = 0$ if and only if $g(c) = 0$ or $u = g(c)cg(c)^{-1}$. If we take

$$u = g(c_n)c_n g(c_n)^{-1},$$

then $f(X)$ has properties (a) and (b): Namely, if $f(c') = 0$ then either $g(c') = 0$ so that c' is a conjugate of one of the $c_1, \dots, c_{n-1}$ by our induction assumption, or $c' = g(c')^{-1}ug(c')$ is a conjugate of c_n.

Now, let $f(X)$ denote any monic polynomial of degree n with $c_1, \dots, c_n$ as zeros. (We have proven there is at least one such polynomial.) Let $h(X) \in R[X]$; then

$$h(X) = q(X) \circ f(X) + r(X) \quad \text{where } r(X) = 0,$$

or

$$r(X) = cr'(X) \quad \text{with } c \neq 0$$

and $r'(X)$ is either $= 1$ or is a monic polynomial of degree $m < n$. If $h(X)$ has each of $c_1, \ldots, c_n$ as zeros then so does $r'(X)$. Hence, $r'(X)$ cannot be a nonzero constant. If it has degree m it follows from our induction assumption that, for each subset of m elements contained in $\{c_1, c_2, \ldots, c_n\}$, $r'(X)$ must be the degree m polynomial associated with that subset by our proposition. However, condition (b) leads to a contradiction because there is more than one such subset. Therefore, $r'(X) = 0$ and our $f(X)$ satisfies condition (c). From this uniqueness follows.

3. Proof of theorem 1. Let $h(X)$ have a finite number, n, of zeros. By Proposition 3 this set of zeros is pairwise non-conjugate. By Proposition 5, $h(X) = q(X) \circ f(X)$ for a certain $f(X)$ of degree n. This means degree $h(X) \geqq n$.

PROPOSITION 6. *Suppose $f(X)$ has zeros in n distinct conjugacy classes and has two (hence, by Proposition 3, infinitely many) zeros in one of these classes. Then degree $f(X) > n$.*

Proof. Suppose $f(X)$ has $c_1', c_1, c_2, \ldots, c_n$ as zeros where c_1' and c_1 lie in the same conjugacy class and $c_1, c_2, \ldots, c_n$ in different ones. We can construct

$$g(X) = (X - u) \circ (X - c_1)$$

with c_1' and c_1 as zeros; by Proposition 1, all its zeros lie in the conjugacy class of c_1. By the division algorithm and the proof of Proposition 6, we get

$$f(X) = q(X) \circ g(X) + r(X) \quad \text{with degree } r(X) \leqq 1.$$

Since both c_1 and c_1' are zeros of $f(X)$ and $g(X)$ we see $r(X) = 0$. The set of conjugacy classes of zeros of $f(X)$ is contained in the union of the set of conjugacy classes of zeros of $q(X)$ and $g(X)$. The former set of conjugacy classes contains at least $n - 1$ elements, which implies degree $q(X) \geqq n - 1$ by Proposition 5, and degree $f(X) \geqq n + 1$.

4. Proof of theorem 2. Theorem 2A follows at once from Propositions 5 and 6.

For Theorem 2B we need a definition of multiplicity. From Theorem 5, page 34 of [**8**] it follows that if $f(X) \in R[X]$ has two factorizations into irreducible factors then these factors can be placed into a one-one correspondence so that corresponding factors are similar; from the definition of similarity given in [**6**], this implies for each $c \in R$ that the

https://doi.org/10.4153/CJM-1983-028-1 Published online by Cambridge University Press

number of factors $(X - u)$ with u conjugate to c is the same for both factorizations.

Definition. We say $c \in R$ is a zero of $f(X)$ of *multiplicity* m if $f(c) = 0$ and in every factorization of $f(X)$ into monic irreducible factors exactly m of the factors are of the form $(X - c')$ with c' conjugate to c.

To prove Theorem 2B we use induction. Suppose $c_1, c_2, \ldots, c_r$ is a pairwise non-conjugate set of elements of R. From Theorem 2B we already know that we can find a unique $g(X)$ of degree r which has the c_i as its only zeros, each with multiplicity 1. We use induction: Assume c_1 is not in the center and we have a polynomial of degree

$$m = \sum_{\rho=1}^{r} m_\rho$$

which has each c_i as a zero of multiplicity $m_i \geqq 1$, and has no other zeros. We want to construct infinitely many $h(X)$ with degree $m + 1$, the same set of zeros, and c_1 of multiplicity $m_1 + 1$. Suppose $h(X) = (X - u) \circ g(X)$. If X is any element of R with $g(X) \neq 0$, it follows from (2) that

$$h(X) = (g(X)Xg(X)^{-1} - u)g(X).$$

Let $X = yc_1y^{-1}$ be a conjugate of c_1 different from c_1; then:

$$(4) \quad h(X) = (g(yc_1y^{-1})yc_1y^{-1}(g(yc_1y^{-1}))^{-1} - u)g(yc_1y^{-1}).$$

Let $g(X) = \sum b_\nu X^\nu$; then

$$g(yc_1y^{-1}) = l(y)y^{-1} \quad \text{where } l(y) = \sum b_\nu y c_1{}^\nu;$$

and (4) becomes:

$$h(X) = (l(y)y^{-1}yc_1y^{-1}(l(y)y^{-1})^{-1} - u)g(X),$$

i.e.,

$$(5) \quad h(yc_1y^{-1}) = (l(y)c_1(l(y))^{-1} - u)g(yc_1y^{-1}).$$

Now we claim: If $(k(c_1):k)$ is finite there is a conjugate u of c_1 which is not of the form $l(y)c_1(l(y))^{-1}$ for any $y \notin R(c_1)$.

To see this note first that, from Proposition 2 and our assumptions about zeros of $g(X)$, it follows that $l(y) = 0$ if and only if $y \in R^{(c_1)}$. As in the proof of Proposition 3, l is an $R^{(c_1)}$-linear map $R \to R$; by Proposition 2,

$$l(y) = 0 \Leftrightarrow g(yc_1y^{-1}) = 0 \Leftrightarrow y \in R^{(c_1)}.$$

https://doi.org/10.4153/CJM-1983-028-1 Published online by Cambridge University Press

Suppose now that $k(c_1)$ is of finite dimension d over the center k. Then as in the proof of Proposition 4, $(R:R^{(c)}) = d$ also. So l is an $R^{(c_1)}$-linear map $R \to R$ with a kernal of dimension 1 over $R^{(c_1)}$. Its image is of dimension $d - 1$ over $R^{(c_1)}$. ($d \geqq 2$ because we assumed c_1 is not in the center.)

By the argument used in the proof of Proposition 3, the set of all conjugates of c_1 in r is in one-one correspondence with the set of elements of a $(d - 1)$ dimensional right projective space over $R^{(c_1)}$ while under the same correspondence the conjugates of the special form

$$l(y)c_1(l(y))^{-1} \quad \text{with } l(y) \neq 0$$

correspond to a $(d - 2)$ dimensional space. Hence, there are infinitely many choices for u, conjugate to c_1, such that $h(X)$ has no zeros except $c_1, \ldots, c_r$ and the multiplicity of c_i is $m_i + 1$. This proves Theorem 2B in case R is an algebra.

We have as yet found no proof or counterexample of Theorem 2B for the general case.

References

1. E. Artin, *Geometric algebra* (Wiley Interscience Series, 1957).
2. E. Artin and G. Whaples, *The theory of simple rings*, American Journal of Math. *65* (1943), 87-107.
3. B. Beck, *Sur les équations polynomiales dans les quaternions*, L'Enseignment Math. *25*.
4. B. Gordon and T. S. Motzkin, *On the zeros of polynomials over division rings*, Trans. Amer. Math. Soc. *116* (1965), 218-226, correction ibid. *122* (1966), 547.
5. I. N. Herstein, *Conjugates in division rings*, Proc. Amer. Math. Soc. *7* (1956), 1021-1022.
6. M. H. Ingraham and M. C. Wolf, *Relative linear sets and similarity of matrices whose elements belong to a division algebra*, Trans. Amer. Math. Soc. *42* (1934), 16-31.
7. N. Jacobson, *The structure of rings* (American Math. Society, Colloquium Publ. *37*, 1956).
8. ——— *The theory of rings*, 5th edition, Math. Survey Series No. 2 (Amer. Soc. Series, 1978).
9. I. Niven, *Equations in quaternions*, American Math. Monthly *48* (1941), 654-661.
10. ——— *The roots of quaternions*, American Math. Monthly *49* (1942), 386-388.
11. O. Ore, *The theory of non-commutative polynomials*, Annals of Math. II *34* (1933), 480-508.
12. J. Sylvester, *Mathematical papers*, Volume IV (Cambridge University Press, 1912).
13. L. Wolf, *Simularity of matrices in which the elements are real quaternions*, Bulletin of the Amer. Math. Soc. *42* (1936), 737-743.

Smith College,
Northampton, Massachusetts;
University of Massachusetts,
Amherst, Massachusetts