ON THE EQUATION $\chi\alpha=\gamma\chi+\beta$ OVER AN ALGEBRAIC DIVISION RING

R. E. JOHNSON

1. Introduction and notation. The main purpose of this paper is to give necessary and sufficient conditions in order that the equation

$$\chi\alpha = \gamma\chi + \beta \tag{1}$$

have a solution χ over an algebraic division ring. In case a solution exists, it is given explicitly if it is unique; otherwise, a method of obtaining one of the solutions is given. The application of the results to a quaternion algebra is discussed in the final section.

Let R be a division ring algebraic over its separable[1] center F, and λ a commutative indeterminate over R. Using the notation of Ore,[2] a polynomial $a(\lambda)\in R[\lambda]$ of degree n,

$$a(\lambda) = \alpha_n\lambda^n + \alpha_{n-1}\lambda^{n-1} + \cdots + \alpha_0, \tag{2}$$

will be called reduced if $\alpha_n=1$. The unique reduced polynomial $m(\lambda)\in F[\lambda]$ of minimum degree for which $m(\alpha)=0$ will be labelled $m_\alpha(\lambda)$. It is apparent that $m_\alpha(\lambda)$ is irreducible over $F[\lambda]$. The ring of all elements of R which commute with α will be denoted by R_α.

The substitution of an element of R for λ in the polynomial (1) is not well defined, as λ commutes with elements of R, whereas the elements of R do not all commute among themselves. However, unilateral substitution is well defined. We shall use the symbol $a^r(\beta)$ to mean that β has been substituted for λ on the right in (2), so that

$$a^r(\beta) = \alpha_n\beta^n + \alpha_{n-1}\beta^{n-1} + \cdots + \alpha_0. \tag{3}$$

Left substitution is defined similarly—as there is a complete duality between left and right substitution in our case, we shall discuss right substitution only. If $a^r(\beta)=0$, β is called a right root of $a(\lambda)$. The notation $a(\lambda)\,|^r b(\lambda)$ is used to mean that $a(\lambda)$ is a right factor of $b(\lambda)$. As is well known, β is a right root of $a(\lambda)$ if and only if $(\lambda-\beta)\,|^r a(\lambda)$.

2. Preliminary lemmas. A division algorithm exists over $R[\lambda]$. The particular case of interest here is given by

Presented to the Society, November 27, 1943; received by the editors September 27, 1943.

[1] That is, no irreducible polynomial in $F[\lambda]$ has a multiple root in R.

[2] O. Ore, *Theory of noncommutative polynomials*, Ann. of Math. vol. 34 (1933) pp. 481–508.

$$b(\lambda) = q(\lambda)(\lambda - \alpha) + b^r(\alpha). \tag{4}$$

That is, the remainder on dividing a polynomial $b(\lambda)$ on the right by $(\lambda-\alpha)$ is $b^r(\alpha)$. For any two elements $a(\lambda)$, $b(\lambda)$ of $R[\lambda]$, there exists a unique reduced greatest common right divisor and a unique reduced least common left multiple.

The following lemma is true for any ring[3] R. It is frequently proven for special cases.

LEMMA A. *If $c(\lambda)=a(\lambda)b(\lambda)$, than $b^r(\alpha)=0$ implies $c^r(\alpha)=0$.*

To prove this for a general ring, let $a(\lambda)=\sum_{i=0}^{n}\alpha_i\lambda^i$, $b(\lambda)=\sum_{i=0}^{m}\beta_i\lambda^i$: then

$$c(\lambda) = \sum_{i=0}^{n}\alpha_i\left(\sum_{j=0}^{m}\beta_j\lambda^j\right)\lambda^i. \tag{5}$$

From this form, it is apparent that $c^r(\alpha)=0$ if $b^r(\alpha)=0$.

In any polynomial ring which possesses a division algorithm, the following lemma holds.

LEMMA B. *If $c(\lambda)=a(\lambda)b(\lambda)$, then $(\lambda-\alpha)\,|^r c(\lambda)$ if and only if $(\lambda-\alpha)\,|^r a(\lambda)b^r(\alpha)$.*

From (4), $c(\lambda)=a(\lambda)q(\lambda)(\lambda-\alpha)+a(\lambda)b^r(\alpha)$, and the lemma follows. Over a division ring, this lemma can be put in the following form.

LEMMA B′. *If $c(\lambda)=a(\lambda)b(\lambda)$ and $\tau=b^r(\alpha)\neq 0$, then $a^r(\tau+\tau^{-1})=0$ if and only if $c^r(\alpha)=0$.*

This result was obtained by Wedderburn,[4] and later by Richardson[5] and, in a more general form, by Ore.[2] Another result of Wedderburn's[4] is the following lemma.

LEMMA C. *If $a^r(\tau\alpha\tau^{-1})=0$ for all nonzero elements $\tau\in R$, then $m_\alpha(\lambda)\,|\,a(\lambda)$.*

The following fundamental theorem was obtained by Wedderburn[4] for division algebras and holds equally well for algebraic division rings.

LEMMA D. *If $m_\alpha(\lambda)$ is of degree n, then there exist elements $\alpha_1(=\alpha)$, $\alpha_2, \cdots, \alpha_n$ in R such that*

[3] See C. C. MacDuffee, *Vectors and matrices* (Carus Mathematical Monographs, No. 7), Mathematical Association of America, 1943, Theorem 36.

[4] J. H. M. Wedderburn, *On division algebras*, Trans. Amer. Math. Soc. vol. 22 (1921) pp. 130–131.

[5] A. R. Richardson, *Equations over a division algebra*, Messenger of Mathematics vol. 57 (1928) pp. 1–6.

$$(6) \qquad m_\alpha(\lambda) = (\lambda - \alpha_n)(\lambda - \alpha_{n-1}) \cdots (\lambda - \alpha_1).$$

A particular factorization of $m_\alpha(\lambda)$ is needed in the proof of Theorem 2. To obtain this, we establish the following lemma.

LEMMA D′. *There exist elements* $\sigma_{11}, \sigma_{12}, \cdots, \sigma_{1n} \in R$ *such that, if*

$$(7) \qquad \sigma_{i+1\,j} = \sigma_{ij}\alpha - \sigma_{i\,i-1}\alpha\sigma_{i\,i-1}^{-1}\sigma_{ij}, \qquad i, j = 1, 2, \cdots, n-1,$$

where $\sigma_{i0} = 1$ *for all* i, *then* $m_\alpha(\lambda)$ *has the factorization* (6) *for*

$$(8) \qquad \alpha_i = \sigma_{i\,i-1}\alpha\sigma_{i\,i-1}^{-1}, \qquad i = 2, 3, \cdots, n.$$

That this is true can be seen inductively. Assume that $\sigma_{11}, \sigma_{12}, \sigma_{13}, \cdots, \sigma_{1\,k-1}$ exist, $k < n$, so that

$$m_\alpha(\lambda) = a_{k+1}(\lambda)(\lambda - \alpha_k)(\lambda - \alpha_{k-1}) \cdots (\lambda - \alpha_1),$$

where $\alpha_1, \alpha_2, \cdots, \alpha_k$ are given by (7) and (8). Let

$$a_i(\lambda) = a_{k+1}(\lambda)(\lambda - \alpha_k)(\lambda - \alpha_{k-1}) \cdots (\lambda - \alpha_i), \quad i = 1, 2, \cdots, k,$$

and

$$b_i(\lambda) = (\lambda - \alpha_k)(\lambda - \alpha_{k-1}) \cdots (\lambda - \alpha_i), \qquad i = 1, 2, \cdots, k.$$

From Lemma C, there must exist an element $\sigma_{1k} \in R$ such that $b_1^r(\sigma_{1k}\alpha\sigma_{1k}^{-1}) \neq 0$. Then $\sigma_{1k}\alpha\sigma_{1k}^{-1} - \alpha_1 \neq 0$, and by Lemma B′, $a_2^r[(\sigma_{1k}\alpha\sigma_{1k}^{-1} - \alpha_1)\sigma_{1k}\alpha\sigma_{1k}^{-1}(\sigma_{1k}\alpha\sigma_{1k}^{-1} - \alpha_1)^{-1}] = 0$ or $a_2^r(\sigma_{2k}\alpha\sigma_{2k}^{-1}) = 0$. Now, as $b_1(\lambda) = b_2(\lambda)(\lambda - \alpha_1)$, and $b_1^r(\sigma_{1k}\alpha\sigma_{1k}^{-1}) \neq 0$, $b_2^r(\sigma_{2k}\alpha\sigma_{2k}^{-1}) \neq 0$ from Lemma B′. Thus $\sigma_{2k}\alpha\sigma_{2k}^{-1} - \alpha_2 \neq 0$, so that $a_3^r[(\sigma_{2k}\alpha\sigma_{2k}^{-1} - \alpha_2)\sigma_{2k}\alpha\sigma_{2k}^{-1}(\sigma_{2k}\alpha\sigma_{2k}^{-1} - \alpha_2)^{-1}] = 0$ or $a_3^r(\sigma_{3k}\alpha\sigma_{3k}^{-1}) = 0$. By induction, $a_i^r(\sigma_{ik}\alpha\sigma_{ik}^{-1}) = 0$, $i = 1, 2, \cdots, k+1$, so that we can select $\alpha_{k+1} = \sigma_{k+1\,k}\alpha\sigma_{k+1\,k}^{-1}$.

It is apparent that $m_\alpha(\tau\alpha\tau^{-1}) = 0$ for all nonzero $\tau \in R$. That all roots of $m_\alpha(\lambda)$ are of this form is given by the following lemma.

LEMMA E. *If* $m_\alpha(\beta) = 0$, *then* β *is a transform of* α.

To prove this, let $m_\alpha(\lambda) = (\lambda - \beta)a(\lambda)$. From Lemma C, there must exist an element $\tau \in R$ such that $a^r(\tau\alpha\tau^{-1}) \neq 0$. Thus, in view of Lemma B′, $\beta = \sigma\alpha\sigma^{-1}$, where $\sigma = a^r(\tau\alpha\tau^{-1})$.

3. **Principal theorems.** If either α or γ is in F, equation (1) becomes trivial. Therefore we shall assume that both α and γ are not in F. Define $\nu_0 = \beta$, and, in general,

$$\nu_i = \gamma^i\beta + \gamma^{i-1}\beta\alpha + \cdots + \gamma\beta\alpha^{i-1} + \beta\alpha^i, \qquad i = 1, 2, \cdots.$$

Then, if $m(\lambda) = \sum_{i=0}^n \mu_i\lambda^i$ is any polynomial in $F[\lambda]$, any χ which is a solution of (1) is also a solution[6] of

$$\chi\alpha m(\alpha) = \gamma m(\gamma)\chi + \sum_{i=0}^{n} \mu_i\nu_i. \tag{9}$$

The discussion of (1) is divided quite naturally into two cases. The first case, which is the easier of the two, is for γ not a transform of α. The second case is for γ a transform of α.

Case 1. As α and γ are not transforms of each other, $m_\alpha(\gamma)\neq 0$ in view of Lemma E. Thus, if we let $m(\lambda)$ of (9) be $m_\alpha(\lambda)$, we obtain

$$\chi = -[m_\alpha(\gamma)]^{-1}\gamma^{-1}\left(\sum_{i=0}^{n} \mu_i\nu_i\right) \tag{10}$$

as the unique solution of (9). A substitution of this value of χ in (1) shows that it is also a solution of (1). As any solution of (1) is also a solution of (9), (10) gives the unique solution of (1). We have thus established the following theorem:

THEOREM 1. *If α and γ are not transforms of each other, then*

$$\chi\alpha = \gamma\chi + \beta$$

has a unique solution. If γ is not zero, this solution is given by (10).

Case 2. The remaining considerations are for $\gamma=\tau\alpha\tau^{-1}$. It is apparent that the methods of Case 1 now fail, as $m_\alpha(\gamma)=0$. Thus a new approach must now be made.

Equation (1) can now be put in the form

$$\chi\alpha = \tau\alpha\tau^{-1}\chi + \beta.$$

This equation has a solution if and only if the equation

$$\tau^{-1}\chi\alpha = \alpha\tau^{-1}\chi + \tau^{-1}\beta$$

has a solution. Therefore we need only consider an equation of the form

$$\chi\alpha = \alpha\chi + \beta. \tag{11}$$

The existence of solutions of this equation is given by the following theorem.

THEOREM 2. *Let α be an element of R not in F with minimum polynomial $m_\alpha(\lambda)=a(\lambda)(\lambda-\alpha)$ and β be a nonzero element of R. Then the equation*

[6] See M. H. Ingraham and H. C. Trimble, *On the matric equation $TA=BT+C$*, Amer. J. Math. vol. 63 (1941) p. 13.

$$\chi\alpha = \alpha\chi + \beta$$

has a solution χ *in* R *if and only if* $a^r(\beta\alpha\beta^{-1}) = 0$.

PROOF. We shall first assume that there exists an element $\chi \in R$ such that (11) is satisfied. Then, as $m_\alpha(\chi\alpha\chi^{-1}) = 0$ and $\chi\alpha \neq \alpha\chi$, we have by Lemma B′ that $a^r([\chi\alpha - \alpha\chi]\alpha[\chi\alpha - \alpha\chi]^{-1}) = 0$. Thus $a^r(\beta\alpha\beta^{-1}) = 0$, and the first part of the theorem is established.

On the other hand, suppose that $a^r(\beta\alpha\beta^{-1}) = 0$. We shall now use the particular factorization of $m_\alpha(\lambda)$ given in Lemma D′. Let the polynomials $b_{ij}(\lambda)$ be defined by

$$b_{ij}(\lambda) = (\lambda - \alpha_i)(\lambda - \alpha_{i-1}) \cdots (\lambda - \alpha_j), \qquad i \geqq j = 1, 2, \cdots, n.$$

Also, let $\beta_1 = \beta$, and recursively,

$$\beta_i = \beta_{i-1}\alpha - \alpha_i\beta_{i-1}, \qquad i = 2, 3, \cdots, n. \tag{12}$$

There must exist an integer k such that $b^r_{k2}(\beta\alpha\beta^{-1}) \neq 0$, $b^r_{k+1\,2}(\beta\alpha\beta^{-1}) = 0$. As in the proof of Lemma D′, the successive application of Lemma B′ yields

$$b^r_{k+1\ i+1}(\beta_i\alpha\beta_i^{-1}) = 0, \qquad i = 1, 2, \cdots, k.$$

The last application gives $b^r_{k+1\ k+1}(\beta_k\alpha\beta_k^{-1}) = 0$, so that $\alpha_{k+1} = \beta_k\alpha\beta_k^{-1}$. From (8), $\alpha_{k+1} = \sigma_{k+1\ k}\alpha\sigma^{-1}_{k+1\ k}$: thus there must exist an element $\delta_k \in R_\alpha$ such that $\beta_k = \sigma_{k+1\ k}\delta_k$. Now let us assume that there exist elements $\delta_j \in R_\alpha$ and an integer m such that

$$\beta_i = \sum_{j=i}^{k} \sigma_{i+1\ j}\delta_j, \qquad i = m, m+1, \cdots, k.$$

Then it follows from (7), (8), and (12) that

$$\beta_{m-1}\alpha - \alpha_m\beta_{m-1} = \sum_{j=m}^{k} (\sigma_{mj}\alpha - \alpha_m\sigma_{mj})\delta_j,$$

so that

$$\left(\beta_{m-1} - \sum_{j=m}^{k} \sigma_{mj}\delta_j\right)\alpha = \alpha_m\left(\beta_{m-1} - \sum_{j=m}^{k} \sigma_{mj}\delta_j\right).$$

As $\alpha_m = \sigma_{m\ m-1}\alpha\sigma^{-1}_{m\ m-1}$, there must exist an element $\delta_{m-1} \in R_\alpha$ such that

$$\beta_{m-1} = \sum_{j=m-1}^{k} \sigma_{mj}\delta_j.$$

By induction,

$$\beta = \sum_{j=1}^{k} \sigma_{2j}\delta_j.$$

From (7),

$$\beta = \sum_{j=1}^{k} \sigma_{1j}\delta_j\alpha - \alpha\sum_{j=1}^{k}\sigma_{1j}\delta_j$$

and thus, for any $\delta \in R_\alpha$,

$$\chi = \sum_{j=1}^{k} \sigma_{1j}\delta_j + \delta \tag{13}$$

is a solution of (11).

4. Special considerations. As a special case of Theorem 2, consider R as the ring of quaternions over a formally real field F, $R=F(1, i, j, k)$. If we let $\bar{\alpha}$ denote the conjugate of α, α not in F, then

$$m_\alpha(\lambda) = (\lambda - \bar{\alpha})(\lambda - \alpha).$$

Thus $a(\lambda)=(\lambda-\bar{\alpha})$, and Theorem 2 can be written in the following form.

COROLLARY 1. *If R is a quaternion algebra over a formally real field F and α is an element of R not in F, then*

$$\chi\alpha = \alpha\chi + \beta$$

has a solution if and only if

$$\beta\alpha = \bar{\alpha}\beta. \tag{14}$$

Having obtained one solution of (11) from (13), say χ_1, then all solutions are given by $\chi_1+\delta$, $\delta\in R_\alpha$. It is observed that (11) cannot have a solution if $\beta\in R_\alpha$—as $a^r(\beta\alpha\beta^{-1})=a^r(\alpha)$ in this case, and $a^r(\alpha)$ cannot be zero due to the separability of F. However, it is not true that (11) always has a solution if β is not in R_α. A simple example to show this is as follows: let R be the ring of quaternions over a formally real field F. For $\beta=i+j$ and $\alpha=i$, (14) is not satisfied, and thus (11) can have no solution.

WASHINGTON, D. C.