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

$$(1) \quad \chi \alpha = \gamma \chi + \beta$$

have a solution $\chi$ 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 $\lambda$ a commutative indeterminate over $R$. Using the notation of Ore,\(^2\) a polynomial $a(\lambda) \in R[\lambda]$ of degree $n$,

$$(2) \quad a(\lambda) = a_n \lambda^n + a_{n-1} \lambda^{n-1} + \cdots + a_0,$$

will be called reduced if $a_n = 1$. The unique reduced polynomial $m(\lambda) \in F[\lambda]$ of minimum degree for which $m(a) = 0$ 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 $\alpha$ will be denoted by $R_\alpha$.

The substitution of an element of $R$ for $\lambda$ in the polynomial (1) is not well defined, as $\lambda$ 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 $\beta$ has been substituted for $\lambda$ on the right in (2), so that

$$(3) \quad a^r(\beta) = a_n \beta^n + a_{n-1} \beta^{n-1} + \cdots + a_0.$$

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 $a^r(\beta) = 0$, $\beta$ 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, $\beta$ 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

---

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

That is, the remainder on dividing a polynomial \( b(\lambda) \) on the right by \( (\lambda - \alpha) \) is \( b^*(\alpha) \). 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\(^4\) \( R \). It is frequently proven for special cases.

**Lemma A.** If \( c(\lambda) = a(\lambda)b(\lambda) \), then \( b^*(\alpha) = 0 \) implies \( c^*(\alpha) = 0 \).

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

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

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

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

**Lemma B.** If \( c(\lambda) = a(\lambda)b(\lambda) \), then \( (\lambda - \alpha)^{\prime}c(\lambda) \) if and only if \( (\lambda - \alpha)^{\prime}a(\lambda)b^*(\alpha) \).

From (4), \( c(\lambda) = a(\lambda)g(\lambda)(\lambda - \alpha) + a(\lambda)b^*(\alpha) \), 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 \( \tau = b^*(\alpha) \neq 0 \), then \( a^*(\tau + \tau^{-1}) = 0 \) if and only if \( c^*(\alpha) = 0 \).

This result was obtained by Wedderburn,\(^4\) and later by Richardson\(^6\) and, in a more general form, by Ore.\(^2\) Another result of Wedderburn's\(^4\) is the following lemma.

**Lemma C.** If \( a^*(\tau \alpha \tau^{-1}) = 0 \) for all nonzero elements \( \tau \in R \), then \( m_a(\lambda) \rvert a(\lambda) \).

The following fundamental theorem was obtained by Wedderburn\(^4\) for division algebras and holds equally well for algebraic division rings.

**Lemma D.** If \( m_a(\lambda) \) is of degree \( n \), then there exist elements \( \alpha_1(=\alpha), \alpha_2, \ldots, \alpha_n \) in \( R \) such that

\(^4\) See C. C. MacDuffee, *Vectors and matrices* (Carus Mathematical Monographs, No. 7), Mathematical Association of America, 1943, Theorem 36.


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

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 \( \sigma_{11}, \sigma_{12}, \ldots, \sigma_{1n} \in R \) such that, if

\[ \sigma_{i+1} = \sigma_i \sigma_{i+1} \sigma_{i-1}^{-1}, \quad i = 1, 2, \ldots, n - 1, \]

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

\[ \alpha_i = \sigma_{i} \sigma_{i-1} \sigma_{i-1}^{-1}, \quad i = 2, 3, \ldots, n. \]

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

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

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

\[ \alpha_i(\lambda) = \alpha_{k+1}(\lambda - \alpha_k)(\lambda - \alpha_{k-1}) \cdots (\lambda - \alpha_i), \quad i = 1, 2, \ldots, k, \]

and

\[ b_i(\lambda) = (\lambda - \alpha_k)(\lambda - \alpha_{k-1}) \cdots (\lambda - \alpha_i), \quad i = 1, 2, \ldots, k. \]

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

It is apparent that \( m_\alpha(\tau \sigma^{-1}) = 0 \) 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 \( \beta \) is a transform of \( \alpha \).

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

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

\[ \nu_i = \gamma \beta + \gamma^{-1} \beta \alpha + \cdots + \gamma \beta \alpha^{i-1} + \beta \alpha^i, \quad i = 1, 2, \ldots. \]

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

\[ \chi m(\alpha) = \gamma m(\gamma) + \sum_{i=0}^{n} \mu \nu_i. \]

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

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

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

as the unique solution of (9). A substitution of this value of \( \chi \) 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 \( \alpha \) and \( \gamma \) are not transforms of each other, then

\[ \chi \alpha = \gamma \chi + \beta \]

has a unique solution. If \( \gamma \) 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

\[ \chi \alpha = \alpha \chi + \beta. \]

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

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

\[ \chi x = \alpha x + \beta \]

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

**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 \( \chi \neq \alpha \chi \), we have by Lemma B' that \( \alpha^*([\chi \alpha - \alpha \chi] \alpha [\chi \alpha - \alpha \chi]^{-1}) = 0 \). Thus \( \alpha^*(\beta_\alpha \beta^{-1}) = 0 \), and the first part of the theorem is established.

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

\[ b_{ij}(\lambda) = (\lambda - \alpha_i)(\lambda - \alpha_{i-1}) \cdots (\lambda - \alpha_j), \quad i \geq j = 1, 2, \ldots, n. \]

Also, let \( \beta_1 = \beta \), and recursively,

\[ \beta_i = \beta_{i-1} \alpha - \alpha_i \beta_{i-1}, \quad i = 2, 3, \ldots, n. \]

There must exist an integer \( k \) such that \( b_{ik}(\beta_\alpha \beta^{-1}) \neq 0 \) and \( b_{k+1}(\beta_\alpha \beta^{-1}) = 0 \). As in the proof of Lemma D', the successive application of Lemma B' yields

\[ b_{k+1+i+1}(\beta_\alpha \beta^{-1}) = 0, \quad i = 1, 2, \ldots, k. \]

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

\[ \beta_i = \sum_{j=i}^{k} \sigma_{i+1} \delta_i, \quad i = m, m + 1, \ldots, k. \]

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

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

so that

\[ (\beta_{m-1} - \sum_{j=m}^{k} \sigma_{m} \delta_j) \alpha = \alpha_m (\beta_{m-1} - \sum_{j=m}^{k} \sigma_{m} \delta_j). \]

As \( \alpha_m = \sigma_m \alpha_m^{-1} \alpha_m^{-1} \), there must exist an element \( \delta_m \in R_\alpha \) such that

\[ \beta_m = \sum_{j=m}^{k} \sigma_{m} \delta_j. \]

By induction,

\[ \beta = \sum_{j=1}^{k} \sigma_{j} \delta_j. \]
From (7),
\[ \beta = \sum_{j=1}^{k} \sigma_1 \delta_j \alpha - \sum_{j=1}^{k} \sigma_1 \delta_j \]
and thus, for any \( \delta \in R_a \),
\[ \chi = \sum_{j=1}^{k} \sigma_1 \delta_j + \delta \]
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 \( \alpha \) denote the conjugate of \( \alpha \), \( \alpha \) not in \( F \), then
\[ m_\alpha(\lambda) = (\lambda - \alpha)(\lambda - \alpha). \]
Thus \( a(\lambda) = (\lambda - \alpha) \), 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 \( \alpha \) is an element of \( R \) not in \( F \), then
\[ \chi \alpha = \alpha \chi + \beta \]
has a solution if and only if
\[ (14) \quad \beta \alpha = \alpha \beta. \]

Having obtained one solution of (11) from (13), say \( \chi_1 \), then all solutions are given by \( \chi_1 + \delta, \delta \in R_a \). It is observed that (11) cannot have a solution if \( \beta \in R_a \)—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 \( \beta \) is not in \( R_a \). 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.