Math. J. Okayama Univ. **41** (1999), 45–62

UNIVERSAL FACTORIZATION EQUALITIES FOR QUATERNION MATRICES AND THEIR APPLICATIONS

YONGGE TIAN

ABSTRACT. We present in this paper two types of universal factorization equalities for real quaternions, as well as for matrices of real quaternions. These universal factorization equalities can serve as a valuable tool for developing matrix analysis over the real quaternion algebra.

1. INTRODUCTION

Let $a = a_0 + a_1 i + a_2 j + a_3 k$ be an element over the real quaternion algebra $\mathbb{H}$, where a_0—a_3 are four real numbers, $i^2 = j^2 = k^2 = -1$ and $ijk = -1$. As two fundamental facts, it is best known in algebra theory that a has two complex and real matrix representations as follows

$$\psi(a) := \begin{pmatrix} a_0 + a_1 i & -(a_2 + a_3 i) \\ a_2 - a_3 i & a_0 - a_1 i \end{pmatrix},$$

$$\text{and} \quad \phi(a) := \begin{pmatrix} a_0 & -a_1 & -a_2 & -a_3 \\ a_1 & a_0 & -a_3 & a_2 \\ a_2 & a_3 & a_0 & -a_1 \\ a_3 & -a_2 & a_1 & a_0 \end{pmatrix}, \tag{1.1}$$

and $\mathbb{H}$ is algebraically isomorphic to the matrix algebras composed by all $\psi(a)$ and $\phi(a)$, respectively. In this article we shall reveal two new facts on the relationship between a quaternion and their two matrix representations, which can be stated that for any $a \in \mathbb{H}$, there are two independent unitary matrices P and Q over $\mathbb{H}$ such that a satisfies the following two similarity factorizations

$$P\text{diag}(\, a, \ a\,)P^* = \psi(a), \qquad Q\text{diag}(\, a, \ a, \ a, \ a\,)Q^* = \phi(a). \tag{1.2}$$

1991 *Mathematics Subject Classification.* 1523, 15A33.

Key words and phrases. Quaternions, universal factorizations, matrix representations, block matrices, linear equations, matrix equations.

where P and Q have no relations with a. Based on these two equalities, we shall derive a variety of new properties on quaternions, and matrices of quaternions.

Throughout, $\mathbb{R}$ and $\mathbb{C}$ stand for the real and the complex number fields, respectively. For any $a = a_0 + a_1 i + a_2 j + a_3 k \in \mathbb{H}$, we used the notations, $\mathrm{Re}\, a = a_0$; $\mathrm{Im}\, a = a_1 i + a_2 j + a_3 k$; $\bar{a} = a_0 - a_1 i - a_2 j - a_3 k$; $|a| = \sqrt{a\bar{a}} = (a_0^2 + a_1^2 + a_2^2 + a_3^2)^{1/2}$, $\overrightarrow{a} = (\, a_0,\ a_1,\ a_2,\ a_3\,)^T$. For any quaternion matrix $A = (a_{st}) \in \mathbb{H}^{m\times n}$, $A^* = (\overline{A})^T = (\overline{a_{ts}})$ stands for the *conjugate transpose* of A. A square quaternion matrix A of size n is said to be *unitary* if $AA^* = A^*A = I_n$, the identity matrix.

2. Basic results

In this section, we first establish two universal similarity factorization equalities for quaternions, and then present various operation properties derived from them for quaternions.

Theorem 2.1. *Let $a = a_0 + a_1 i + a_2 j + a_3 k \in \mathbb{H}$ be given. Then aI_2 satisfies the universal similarity factorization equality*

$$P\begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix}P^* = \begin{pmatrix} a_0 + a_1 i & -(a_2 + a_3 i) \\ a_2 - a_3 i & a_0 - a_1 i \end{pmatrix} = \psi(a) \in \mathbb{C}^{2\times 2}, \tag{2.1}$$

where

$$P = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -i \\ -j & k \end{pmatrix} \tag{2.2}$$

is a unitary matrix over $\mathbb{H}$.

Proof. Note that $a - iai = 2a_0 + 2a_1 i$, $a + iai = 2a_2 j + 2a_3 k$. We have

$$\begin{aligned} P(aI_2)P^* &= \frac{1}{2}\begin{pmatrix} a - iai & aj + iak \\ -ja + kai & -(jaj + kak) \end{pmatrix} \\ &= \frac{1}{2}\begin{pmatrix} a - iai & (a + iai)j \\ -j(a + iai) & -j(a - iai)j \end{pmatrix} \\ &= \begin{pmatrix} a_0 + a_1 i & (a_2 j + a_3 k)j \\ -j(a_2 j + a_3 k) & -j(a_0 + a_1 i)j \end{pmatrix} \\ &= \begin{pmatrix} a_0 + a_1 i & -(a_2 + a_3 i) \\ a_2 - a_3 i & a_0 - a_1 i \end{pmatrix}, \end{aligned}$$

establishing (2.1). □

Theorem 2.2. *Let $a = a_0 + a_1 i + a_2 j + a_3 k \in \mathbb{H}$ be given. Then aI_4 satisfies the following universal similarity factorization*

$$Q\begin{pmatrix} a & & & \\ & a & & \\ & & a & \\ & & & a \end{pmatrix}Q^* = \begin{pmatrix} a_0 & -a_1 & -a_2 & -a_3 \\ a_1 & a_0 & -a_3 & a_2 \\ a_2 & a_3 & a_0 & -a_1 \\ a_3 & -a_2 & a_1 & a_0 \end{pmatrix} = \phi(a) \in \mathbb{R}^{4\times 4}, \tag{2.3}$$

where the matrix Q has the following independent expression

$$Q = Q^* = \frac{1}{2}\begin{pmatrix} 1 & i & j & k \\ -i & 1 & k & -j \\ -j & -k & 1 & i \\ -k & j & -i & 1 \end{pmatrix}, \tag{2.4}$$

which is a unitary matrix over $\mathbb{H}$.

Proof. First we build a unitary matrix $U = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -j \\ -i & -k \end{pmatrix}$ over $\mathbb{H}$, and then calculate

$$\begin{aligned} U\begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix}U^* &= \frac{1}{2}\begin{pmatrix} a - jaj & (a + jaj)i \\ -i(a + jaj) & -i(a - jaj)i \end{pmatrix} \\ &= \begin{pmatrix} a_0 + a_2 j & -a_1 + a_3 j \\ a_1 + a_3 j & a_0 - a_2 j \end{pmatrix} \\ &= \begin{pmatrix} a_0 & -a_1 \\ a_1 & a_0 \end{pmatrix} + \begin{pmatrix} a_2 & a_3 \\ a_3 & -a_2 \end{pmatrix} j \\ &= A_0 + A_1 j := A, \end{aligned}$$

where A_0 and A_1 are two real matrices. Next we build another block unitary matrix $V = \frac{1}{\sqrt{2}}\begin{pmatrix} I_2 & -I_2 i \\ -I_2 j & I_2 k \end{pmatrix}$ over $\mathbb{H}$, and then calculate

$$\begin{aligned} V\begin{pmatrix} A & 0 \\ 0 & A \end{pmatrix}V^* &= \frac{1}{2}\begin{pmatrix} A - iAi & (A + iAi)j \\ -j(A + iAi) & -j(A - iAi)j \end{pmatrix} \\ &= \begin{pmatrix} A_0 & -A_1 \\ A_1 & A_0 \end{pmatrix} = \phi(a). \end{aligned}$$

Substituting $A = U\mathrm{diag}(a,\ a)U^*$ in its left-hand side gives

$$V\mathrm{diag}(U,\ U)\mathrm{diag}(a,\ a,\ a,\ a)\mathrm{diag}(U^*,\ U^*)V^* = \phi(a). \tag{2.5}$$

Set

$$Z = V\mathrm{diag}(\,U,\ U\,) = \frac{1}{\sqrt{2}}\begin{pmatrix} U & -iU \\ -jU & kU \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1 & -j & -i & k \\ -i & -k & -1 & -j \\ -j & -1 & k & i \\ -k & i & -j & 1 \end{pmatrix}.$$

Then (2.5) is $Z\mathrm{diag}(\,a,\ a,\ a,\ a\,)Z^* = \phi(a)$. Next set

$$J = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Then it is easy to see that $J\mathrm{diag}(\,a,\ a,\ a,\ a\,)J = \mathrm{diag}(\,a,\ a,\ a,\ a\,)$. Thus we have $ZJ\mathrm{diag}(\,a,\ a,\ a,\ a\,)JZ^* = \phi(a)$. Finally, let us set $Q = ZJ$. Then we have (2.3) and (2.4). □

Some direct consequences can directly be derived from (2.1) and (2.3), most of them are well known.

Corollary 2.3. *Let $a = a_0 + a_1 i + a_2 j + a_3 k \in \mathbb{H}$ be given. Then a and its two matrix representations $\psi(a)$ and $\phi(a)$ satisfy*

$$a = \frac{1}{2}E_2\psi(a)E_2^*, \quad \text{and} \quad a = \frac{1}{4}E_4\phi(a)E_4^*,$$

where $E_2 = (\,1,\ j\,)$ and $E_4 = (\,1,\ i,\ j,\ k\,)$.

Corollary 2.4. *Let $a,\ b \in \mathbb{H}$, and $\lambda \in \mathbb{R}$. Then*

(a) $a = b \iff \psi(a) = \psi(b) \iff \phi(a) = \phi(b)$.
(b) $\psi(a+b) = \psi(a) + \psi(b), \qquad \phi(a+b) = \phi(a) + \phi(b)$.
(c) $\psi(ab) = \psi(a)\psi(b), \qquad \phi(ab) = \phi(a)\phi(b)$.
(d) $\psi(\lambda a) = \psi(a\lambda) = \lambda\psi(a), \quad \phi(\lambda a) = \phi(a\lambda) = \lambda\phi(a)$.
(e) $\psi(1) = I_2, \quad \phi(1) = I_4$.
(f) $\psi(\overline{a}) = \psi^*(a), \qquad \phi(\overline{a}) = \phi^T(a)$.
(g) $\psi(a^{-1}) = \psi^{-1}(a), \qquad \phi(a^{-1}) = \phi^{-1}(a), \qquad a \neq 0$.
(h) $\det[\psi(a)] = |a|^2, \qquad \det[\phi(a)] = |a|^4$.

We can also introduce from (2.3) another real matrix representation of a as follows.

$$\tau(a) := L\phi^T(a)L = \begin{pmatrix} a_0 & -a_1 & -a_2 & -a_3 \\ a_1 & a_0 & a_3 & -a_2 \\ a_2 & -a_3 & a_0 & a_1 \\ a_3 & a_2 & -a_1 & a_0 \end{pmatrix}, \tag{2.6}$$

where $L = \mathrm{diag}(\,1,\ -1,\ -1,\ -1\,)$. Some basic operation properties on $\tau(a)$ are

$$\tau(a+b) = \tau(a) + \tau(b), \qquad \tau(ab) = \tau(b)\tau(a), \qquad \tau(\overline{a}) = \tau^T(a). \tag{2.7}$$

Combining the two real matrix representations of quaternions with their real vector representations, we find the following important result.

Theorem 2.5. *Let* $a,\ b,\ x \in \mathbb{H}$. *Then*

$$\overrightarrow{ax} = \phi(a)\overrightarrow{x}, \qquad \overrightarrow{xb} = \tau(b)\overrightarrow{x}, \tag{2.8}$$

$$\overrightarrow{axb} = \phi(a)\tau(b)\overrightarrow{x} = \tau(b)\phi(a)\overrightarrow{x}, \tag{2.9}$$

and

$$\phi(a)\tau(b) = \tau(b)\phi(a). \tag{2.10}$$

Proof. It is easy to see from (2.3) and (2.6) that for all $a \in \mathbb{H}$, its real vector representation $\overrightarrow{a}$ can be expressed as

$$\overrightarrow{a} = \phi(a)\alpha_4^T, \qquad \overrightarrow{a} = \tau(a)\alpha_4^T, \qquad \alpha_4 = (\,1,\ 0,\ 0,\ 0\,).$$

Thus by Corollary 2.4(c) and (2.7), we get

$$\begin{aligned}
\overrightarrow{ax} &= \phi(ax)\alpha_4^T = \phi(a)\phi(x)\alpha_4^T = \phi(a)\overrightarrow{x},\\
\overrightarrow{xb} &= \tau(xb)\alpha_4^T = \tau(b)\tau(x)\alpha_4^T = \tau(b)\overrightarrow{x},\\
\overrightarrow{axb} &= \overrightarrow{a(xb)} = \phi(a)\overrightarrow{(xb)} = \phi(a)\tau(b)\overrightarrow{x},\\
\overrightarrow{axb} &= \overrightarrow{(ax)b} = \tau(b)\overrightarrow{(ax)} = \tau(b)\phi(a)\overrightarrow{x}.
\end{aligned}$$

These four equalities are exactly the results in (2.8) and (2.9). Note that $\overrightarrow{x}$ in (2.9) is, in fact, an arbitrary real vector when x runs over $\mathbb{H}$. Thus (2.10) follows. □

The three formulas in (2.8) and (2.9) can be applied to deal with various linear equations over $\mathbb{H}$, which we will consider in the next section.

3. How to solve linear equations over $\mathbb{H}$

Since $\mathbb{H}$ is a noncommutative algebra, the general system of linear equations over $\mathbb{H}$ should have the following two-sided form

$$\left\{\begin{array}{ccccccccc}
a_{11}x_1b_{11} & + & a_{12}x_2b_{12} & + & \cdots & + & a_{1n}x_nb_{1n} & = & c_1,\\
a_{21}x_1b_{21} & + & a_{22}x_2b_{22} & + & \cdots & + & a_{2n}x_nb_{2n} & = & c_2,\\
\cdots & & \cdots & & \cdots & & \cdots & & \cdots,\\
a_{m1}x_1b_{m1} & + & a_{m2}x_2b_{m2} & + & \cdots & + & a_{mn}x_nb_{mn} & = & c_m,
\end{array}\right. \tag{3.1}$$

where a_{st}, b_{st} and $c_s(1 \le s \le m,\ 1 \le t \le n)$ are given, and $x_t(1 \le t \le n)$ is an unknown quaternion. This kind of system over a division algebra was first examined in [14] through a specifically defined determinant. Now applying the formula in (2.9) to the two sides of each equation in (3.1), we obtain a new system as follows

$$(3.2)\qquad \begin{cases} \phi(a_{11})\tau(b_{11})\overrightarrow{x_1} & + & \cdots & + & \phi(a_{1n})\tau(b_{1n})\overrightarrow{x_n} & = & \overrightarrow{c_1} \\ \vdots & & & & \vdots & & \vdots \\ \phi(a_{m1})\tau(b_{m1})\overrightarrow{x_1} & + & \cdots & + & \phi(a_{mn})\tau(b_{mn})\overrightarrow{x_n} & = & \overrightarrow{c_m}, \end{cases}$$

or simply

$$(3.3)\qquad \begin{pmatrix} \phi(a_{11})\tau(b_{11}) & \cdots & \phi(a_{1n})\tau(b_{1n}) \\ \vdots & & \vdots \\ \phi(a_{m1})\tau(b_{m1}) & \cdots & \phi(a_{mn})\tau(b_{mn}) \end{pmatrix} \begin{pmatrix} \overrightarrow{x_1} \\ \vdots \\ \overrightarrow{x_n} \end{pmatrix} = \begin{pmatrix} \overrightarrow{c_1} \\ \vdots \\ \overrightarrow{c_m} \end{pmatrix},$$

which shows that all systems of linear equations over $\mathbb{H}$ can be solved by transforming them into conventional systems of linear equations over $\mathbb{R}$.

As a special case, we next consider the following well-known linear equation

$$(3.4)\qquad ax - xb = c$$

over $\mathbb{H}$, which was examined in [6, 10, 18, 21], but they did not get a complete resolution to this fundamental equation.

According to (2.8), the equation (3.4) is equivalent to

$$(3.5)\qquad [\,\phi(a) - \tau(b)\,]\overrightarrow{x} = \overrightarrow{c},$$

which is a simple system of linear equations over $\mathbb{R}$. In order to symbolically solve it, we need to examine some operation properties on the matrix $\phi(a) - \tau(b)$.

Lemma 3.1. *Let* $a = a_0 + a_1 i + a_2 j + a_3 k$, $b = b_0 + b_1 i + b_2 j + b_3 k \in \mathbb{H}$ *be given, and denote* $\theta(a,\ b) := \phi(a) - \tau(b)$. *Then*

(a) *The determinant of* $\theta(a,\ b)$ *is*

$$\begin{aligned} |\theta(a,\ b)| &= [\,s^2 + (\,|\mathrm{Im}\,a| - |\mathrm{Im}\,b|\,)^2\,][\,s^2 + (\,|\mathrm{Im}\,a| + |\mathrm{Im}\,b|\,)^2\,] \\ &= s^4 + 2s^2[\,(|\mathrm{Im}\,a|^2 + |\mathrm{Im}\,b|^2)\,] + (\,|\mathrm{Im}\,a|^2 - |\mathrm{Im}\,b|^2\,)^2, \end{aligned}$$

where $s = a_0 - b_0$.

(b) $\theta(a,\ b)$ *is a normal matrix, and has four eigenvalues as follows*

$$\lambda = (a_0 - b_0) \pm i|\,|\mathrm{Im}\,a| \pm |\mathrm{Im}\,b|\,|.$$

(c) *If* $a_0 \neq b_0$, *or* $|\mathrm{Im}\,a| \neq |\mathrm{Im}\,b|$, *then* $\theta(a,\ b)$ *is nonsingular and its inverse can be expressed as*

$$\begin{aligned} \theta^{-1}(a,\ b) &= \phi^{-1}(\,a^2 - 2b_0 a + |b|^2\,)[\,\phi(a) - \tau(\overline{b})\,] \\ &= \phi^{-1}[\,2(a_0 - b_0)a + |b|^2 - |a|^2\,][\,\phi(a) - \tau(\overline{b})\,], \end{aligned}$$

and

$$\begin{aligned}\theta^{-1}(a,\ b) &= \tau^{-1}(\, b^2 - 2a_0 b + |a|^2\,)[\,\phi(\overline{a}) - \tau(b)\,]\\ &= \tau^{-1}[\,2(b_0 - a_0)b + |a|^2 - |b|^2\,][\,\phi(\overline{a}) - \tau(b)\,].\end{aligned}$$

(d) *If* $a_0 = b_0$ *and* $|\mathrm{Im}\, a| = |\mathrm{Im}\, b|$, *then* $\theta(a,\ b)$ *is singular and has a generalized inverse as follows*

$$\theta^{-}(a,\ b) = \frac{-1}{4|\mathrm{Im}\, a|^2}\theta(a,\ b) = \frac{1}{4|\mathrm{Im}\, a|^2}[\,\tau(\mathrm{Im}\, b) - \phi(\mathrm{Im}\, a)\,].$$

Proof. It is a known result (see [3] and [21]) that for all $a,\ b \in \mathbb{H}$, there are nonzero $p,\ q \in \mathbb{H}$ such that

$$a = p(\,a_0 + |\mathrm{Im}\, a|i\,)p^{-1} = p\hat{a}p^{-1}, \quad \text{and} \quad b = q(\,b_0 + |\mathrm{Im}\, b|i\,)q^{-1} = q\hat{b}q^{-1}.$$

Now applying Corollary 2.4(c) and (2.7) to both of them we obtain

$$\phi(a) = \phi(p)\phi(\hat{a})\phi(p^{-1}), \qquad \tau(b) = \tau(q^{-1})\tau(\hat{b})\tau(q).$$

Thus from Corollary 2.5 we can derive

$$\begin{aligned}|\theta(a,\ b)| &= |\,\phi(p)\phi(\hat{a})\phi(p^{-1}) - \tau(q^{-1})\tau(\hat{b})\tau(q)\,|\\ &= |\phi(p)|\,|\,\phi(\hat{a}) - \phi(p^{-1})\tau(q^{-1})\tau(\hat{b})\tau(q)\phi(p)\,|\,|\phi(p^{-1})|\\ &= |\,\phi(\hat{a}) - \tau(q^{-1})\tau(\hat{b})\tau(q)\,|\\ &= |\tau(q^{-1})|\,|\tau(q)\phi(\hat{a})\tau(q^{-1}) - \tau(\hat{b})\,|\,|\tau(q)| = |\,\phi(\hat{a}) - \tau(\hat{b})\,|,\end{aligned}$$

in which case, substituting $\hat{a} = a_0 + |\mathrm{Im}\, a|i$ and $\hat{b} = b_0 + |\mathrm{Im}\, b|i$ into it may produce Part (a). From the structure of $\theta(a,\ b)$ and $|\theta(a,\ b)|$, we also know that

$$|\,\lambda I_4 - \theta(a,\ b)\,| = [\,(\lambda - s\,)^2 + (\,|\mathrm{Im}\, a| - |\mathrm{Im}\, b|\,)^2][(\lambda - s\,)^2 + (\,|\mathrm{Im}\, a| + |\mathrm{Im}\, b|\,)^2].$$

Thus we have Part (b). The normality of $\theta(a,\ b)$ can be seen from the following equality

$$\begin{aligned}\theta(a,\ b) + \theta^T(a,\ b) &= \phi(a) - \tau(b) + \phi^T(a) - \tau^T(b)\\ &= \phi(a) - \tau(b) + \phi(\overline{a}) - \tau(\overline{b})\\ &= \phi(a + \overline{a}) - \tau(b + \overline{b}) = 2(a_0 - b_0)I_4.\end{aligned}$$

The results in Part (c) come from the following two equalities

$$\begin{aligned}[\,\phi(a) - \tau(\overline{b})\,][\,\phi(a) - \tau(b)\,] &= \phi(a^2) - 2b_0\phi(a) + |b|^2 I_4\\ &= \phi(\,a^2 - 2b_0 a + |b|^2\,),\\ [\,\phi(\overline{a}) - \tau(b)\,][\,\phi(a) - \tau(b)\,] &= \tau(b^2) - 2a_0\tau(b) + |a|^2 I_4\\ &= \tau(\,b^2 - 2a_0 b + |a|^2\,).\end{aligned}$$

Finally, under the conditions that $a_0 = b_0$ and $|\mathrm{Im}\, a| = |\mathrm{Im}\, b|$, it is easily seen that

$$\theta(a,\ b) = \phi(a) - \tau(b) = \phi(\mathrm{Im}\, a) - \tau(\mathrm{Im}\, b).$$

From it and a simple fact $(\mathrm{Im}\, a)^2 = (\mathrm{Im}\, b)^2 = -|\mathrm{Im}\, a|^2$, we can easily deduce the following equality $\theta^3(a,\ b) = -4|\mathrm{Im}\, a|^2\theta(a,\ b)$. So we have Part (d). □

Based on Lemma 3.1, we have the following several results.

Theorem 3.2. *Let $a \in \mathbb{H}$ and $a \notin \mathbb{R}$. Then the general solution of the equation*

$$ax = xa \tag{3.6}$$

is

$$x = p - \frac{1}{|\mathrm{Im}\, a|^2}(\mathrm{Im}\, a)p(\mathrm{Im}\, a), \tag{3.7}$$

where $p \in \mathbb{H}$ is arbitrary, or equivalently,

$$x = \lambda_0 + \lambda_1 a, \tag{3.8}$$

where $\lambda_1,\ \lambda_2 \in \mathbb{R}$ are arbitrary.

Proof. According to (2.8), the equation (3.6) is equivalent to $[\,\phi(a) - \tau(a)\,]\overrightarrow{x} = \theta(a,\ a)\overrightarrow{x} = 0$, and the general solution of this equation can be expressed as

$$\overrightarrow{x} = 2[I_4 - \theta^-(a,\ a)\theta(a,\ a)]\overrightarrow{p},$$

where $\overrightarrow{p}$ is an arbitrary real vector. Now substituting Lemma 3.1(d) in it, we get

$$\begin{aligned}\overrightarrow{x} &= 2\left[\, I_4 + \frac{1}{4|\mathrm{Im}\, a|^2}\theta^2(a,\ a)\right]\overrightarrow{p} \\ &= 2\left[\, I_4 - \frac{1}{4|\mathrm{Im}\, a|^2}(\, 2|\mathrm{Im}\, a|^2 I_4 + 2\phi(\mathrm{Im}\, a)\tau(\mathrm{Im}\, a)\,)\right]\overrightarrow{p} \\ &= \left[\, I_4 - \frac{1}{|\mathrm{Im}\, a|^2}\phi(\mathrm{Im}\, a)\tau(\mathrm{Im}\, a)\right]\overrightarrow{p}.\end{aligned}$$

Returning it to quaternion form by (2.8) and (2.9), we have (3.7). Next let $p = (\mathrm{Im}\, a)q$ in (3.7), where $q \in \mathbb{H}$ is arbitrary. Then (3.7) becomes

$$\begin{aligned} x = (\mathrm{Im}\, a)q + q(\mathrm{Im}\, a) &= 2(\mathrm{Re}\, q)(\mathrm{Im}\, a) + (\mathrm{Im}\, a)(\mathrm{Im}\, q) + (\mathrm{Im}\, q)(\mathrm{Im}\, a) \\ &= t_0 + t_1(\mathrm{Im}\, a),\end{aligned}$$

where $t_0,\ t_1 \in \mathbb{R}$, which is equivalent to (3.8). □

Theorem 3.3. *Let $a,\ b \in \mathbb{H}$ be given. Then*

(a) [3] *The linear equation*

$$ax = xb \tag{3.9}$$

has a nonzero solution, i.e., a and b are similar, if and only if

$$\mathrm{Re}\, a = \mathrm{Re}\, b, \quad \textit{and} \quad |\mathrm{Im}\, a| = |\mathrm{Im}\, b|. \tag{3.10}$$

(b) *In that case, the general solution of* (3.9) *is*

$$x = p - \frac{1}{|\mathrm{Im}\, a||\mathrm{Im}\, b|}(\mathrm{Im}\, a)p(\mathrm{Im}\, b), \tag{3.11}$$

where $p \in \mathbb{H}$ is arbitrary; in particular, if $b \neq \bar{a}$, i.e., $\mathrm{Im}\, a + \mathrm{Im}\, b \neq 0$, *then the general solution of* (3.9) *can be written as*

$$x = \lambda_1(\mathrm{Im}\, a + \mathrm{Im}\, b) + \lambda_2[\,|\mathrm{Im}\, a|\,|\mathrm{Im}\, b| - (\mathrm{Im}\, a)(\mathrm{Im}\, b)\,], \tag{3.12}$$

where $\lambda_1,\ \lambda_2 \in \mathbb{R}$ are arbitrary.

Proof. According (2.8), the equation (3.9) is equivalent to

$$[\,\phi(a) - \tau(b)\,]\vec{x} = \theta(a,\, b)\vec{x} = 0, \tag{3.13}$$

and this equation has a nonzero solution if and only if $|\theta(a,\, b)| = 0$, which is equivalent, by Lemma 3.1(a), to (3.10). In that case, the general solution of this equation can be expressed as

$$\vec{x} = 2[\,I_4 - \theta^-(a,\, b)\theta(a,\, b)\,]\vec{p},$$

where $\vec{p}$ is an arbitrary real vector. Now substituting $\theta^-(a,\, b)$ in Lemma 3.1(d) in it, we get

$$\begin{aligned} \vec{x} &= 2\left[\, I_4 + \frac{1}{4|\mathrm{Im}\, a|^2}\theta^2(a,\, b)\right]\vec{p} \\ &= 2\left[\, I_4 - \frac{1}{4|\mathrm{Im}\, a|^2}(2|\mathrm{Im}\, a|^2 I_4 + 2\phi(\mathrm{Im}\, a)\tau(\mathrm{Im}\, b))\right]\vec{p} \\ &= \left[\, I_4 - \frac{1}{|\mathrm{Im}\, a|^2}\phi(\mathrm{Im}\, a)\tau(\mathrm{Im}\, b)\right]\vec{p}. \end{aligned}$$

Returning it to quaternion form by (2.8) and (2.9), we have (3.11). If $b \neq \bar{a}$ in (3.9), then we set $p = (\mathrm{Im}\, a)$ and $p = |\mathrm{Im}\, a|\,|\mathrm{Im}\, b|$ in (3.11), respectively, and (3.11) becomes

$$x_1 = \mathrm{Im}\, a + \mathrm{Im}\, b, \qquad x_2 = |\mathrm{Im}\, a|\,|\mathrm{Im}\, b| - (\mathrm{Im}\, a)(\mathrm{Im}\, b).$$

Thus (3.12) is also a solution to (3.9) under (3.10). The independence of x_1 and x_2 can be seen from two simple facts that $\mathrm{Re}\, x_1 = 0$ and $\mathrm{Re}\, x_2 \neq 0$. Therefore (3.12) is exactly the general solution to (3.9), since the rank of $\theta(a,\, b)$ is two under (3.10). □

A direct consequence of Theorem 3.2 is given below.

Corollary 3.4. *Let $a \in \mathbb{H}$ be given with $a \notin \mathbb{C}$. Then the equation*

$$ax = x(\,\mathrm{Re}\, a + |\mathrm{Im}\, a| i\,) \tag{3.14}$$

always has a nonzero solution, namely, $a \sim \mathrm{Re}\, a + |\mathrm{Im}\, a| i$, and its general solution is

$$x = \lambda_1[\,|\mathrm{Im}\, a| i + \mathrm{Im}\, a\,] + \lambda_2[\,|\mathrm{Im}\, a| - (\mathrm{Im}\, a) i\,], \tag{3.15}$$

where $\lambda_1, \ \lambda_2 \in \mathbb{R}$ are arbitrary.

The equality (3.14) can easily help to find powers and n-th roots of quaternions, this work was previously examined in [2, 13, 16].

Now for the nonhomogeneous linear equation (3.4), we have the following two general results, which give a complete resolution for (3.4).

Theorem 3.5. *Let $a, \ b \in \mathbb{H}$ be given with $a \sim b$. Then the equation in (3.4) has a solution if and only if*

$$ac = c\bar{b}, \tag{3.16}$$

in which case, the general solution of (3.4) can be written as

$$x = \frac{1}{4|\mathrm{Im}\, a|^2}(cb - ac) + p - \frac{1}{|\mathrm{Im}\, a|^2}(\mathrm{Im}\, a)p(\mathrm{Im}\, b), \tag{3.17}$$

where $p \in \mathbb{H}$ is arbitrary.

Proof. According to (2.8), (3.4) can equivalently be written as

$$[\,\phi(a) - \tau(b)\,]\overrightarrow{x} = \theta(a, \ b)\overrightarrow{x} = \overrightarrow{c}. \tag{3.18}$$

This equation is solvable if and only if

$$\theta(a, \ b)\theta^{-}(a, \ b)\overrightarrow{c} = \overrightarrow{c},$$

which is equivalent to

$$\phi(\mathrm{Im}\, a)\tau(\mathrm{Im}\, b)\overrightarrow{c} = |\mathrm{Im}\, a|^2\overrightarrow{c}.$$

Returning it to quaternion form by (2.8) and (2.9) produces

$$(\mathrm{Im}\, a)c(\mathrm{Im}\, b) = |\mathrm{Im}\, a|^2 c,$$

which is equivalent to $c(\mathrm{Im}\, b) = -(\mathrm{Im}\, a)c$ and then (3.16). In that case, the general solution of (3.18) can be expressed as

$$\overrightarrow{x} = \theta^{-}(a, \ b)\overrightarrow{c} + 2[I_4 - \theta^{-}(a, \ b)\theta(a, \ b)]\overrightarrow{p},$$

where $\overrightarrow{p}$ is an arbitrary real vector. Returning it to quaternion form, we find (3.17). □

Theorem 3.6. *Let a, $b \in \mathbb{H}$ be given with a and b not similar, that is,* $\mathrm{Re}\, a \neq \mathrm{Re}\, b$ *or* $|\mathrm{Im}\, a| \neq |\mathrm{Im}\, b|$. *Then* (3.4) *has a unique solution*

$$x = (\, 2sa + |b|^2 - |a|^2\,)^{-1}(ac - c\bar{b}) = (\, cb - \bar{a}c\,)(\, 2sb + |b|^2 - |a|^2\,)^{-1}, \tag{3.19}$$

where $s = \mathrm{Re}\, a - \mathrm{Re}\, b$.

Proof. Under the assumption of this theorem, $\theta(a,\ b) = \phi(a) - \tau(b)$ is nonsingular by Lemma 3.1(c). Hence (3.5) has a unique solution as follows

$$\vec{x} = \theta^{-1}(a,\ b)\vec{c} = \phi^{-1}(\, 2sa + |b|^2 - |a|^2\,)(\, \phi(a) - \tau(\bar{b})\,)\vec{c},$$

$$\vec{x} = \theta^{-1}(a,\ b)\vec{c} = \tau^{-1}(\, -2sb + |a|^2 - |b|^2\,)(\, \phi(\bar{a}) - \tau(b)\,)\vec{c}.$$

Returning them to quaternion expressions yields (3.19). □

4. Two universal factorization equalities for quaternion matrices

The two universal similarity factorization equalities in Section 2 can also be extended to all quaternion matrices. Next are the corresponding results.

Theorem 4.1. *Let* $A = A_0 + A_1 i + A_2 j + A_3 k \in \mathbb{H}^{m\times n}$ *be given, where* A_0—$A_3 \in \mathbb{R}^{m\times n}$. *Then* $\mathrm{diag}(\,A,\ A\,)$ *satisfies the following universal factorization equality*

$$P_{2m}\begin{pmatrix} A & 0 \\ 0 & A \end{pmatrix}P_{2n}^{*} = \begin{pmatrix} A_0 + A_1 i & -(A_2 + A_3 i) \\ A_2 - A_3 i & A_0 - A_1 i \end{pmatrix} := \Psi(A) \in \mathbb{C}^{2m\times 2n}, \tag{4.1}$$

where P_{2m} *and* P_{2n}^{*} *are the following two unitary matrices over* $\mathbb{H}$

$$P_{2m} = \frac{1}{\sqrt{2}}\begin{pmatrix} I_m & -iI_m \\ -jI_m & kI_m \end{pmatrix}, \qquad P_{2n}^{*} = \frac{1}{\sqrt{2}}\begin{pmatrix} I_n & jI_n \\ iI_n & -kI_n \end{pmatrix}. \tag{4.2}$$

In particular, if $m = n$, *then* (4.1) *becomes a universal similarity factorization equality over* $\mathbb{H}$.

Proof. Calculating the matrix product in the left-hand side of (4.1) yields

$$\begin{aligned} P_{2m}\begin{pmatrix} A & 0 \\ 0 & A \end{pmatrix}P_{2n}^{*} &= \frac{1}{2}\begin{pmatrix} I_m & -iI_m \\ -jI_m & kI_m \end{pmatrix}\begin{pmatrix} A & 0 \\ 0 & A \end{pmatrix}\begin{pmatrix} I_n & jI_n \\ iI_n & -kI_n \end{pmatrix} \\ &= \frac{1}{2}\begin{pmatrix} A - iAi & Aj + iAk \\ -jA + kAi & -(jAj + kAk) \end{pmatrix} \\ &= \frac{1}{2}\begin{pmatrix} A - iAi & (A + iAi)j \\ -j(A + iAi) & -j(A - iAi)j \end{pmatrix}. \end{aligned}$$

It is easy to verify that $A - iAi = 2A_0 + 2A_1 i$ and $A + iAi = 2A_2 j + 2A_3 k$. Substituting both of them in the right-hand side of the above equality, we obtain (4.1). □

Theorem 4.2. *Let* $A = A_0 + A_1 i + A_2 j + A_3 k \in \mathbb{H}^{m\times n}$ *be given, where* A_0—$A_3 \in \mathbb{R}^{m\times n}$. *Then* $\mathrm{diag}(\,A,\ A,\ A,\ A\,)$ *satisfies the following universal factorization equality*

$$(4.3)\qquad Q_{4m}\begin{pmatrix} A & & & \\ & A & & \\ & & A & \\ & & & A \end{pmatrix} Q_{4n}^* = \begin{pmatrix} A_0 & -A_1 & -A_2 & -A_3 \\ A_1 & A_0 & -A_3 & A_2 \\ A_2 & A_3 & A_0 & -A_1 \\ A_3 & -A_2 & A_1 & A_0 \end{pmatrix} := \Phi(A) \in \mathbb{R}^{4m\times 4n},$$

where Q_{4t} *is the following unitary matrix over* $\mathbb{H}$

$$(4.4)\qquad Q_{4t} = Q_{4t}^* = \frac{1}{2}\begin{pmatrix} I_t & iI_t & jI_t & kI_t \\ -iI_t & I_t & kI_t & -jI_t \\ -jI_t & -kI_t & I_t & iI_t \\ -kI_t & jI_t & -iI_t & I_t \end{pmatrix}, \qquad t = m,\ n.$$

In particular, if $m = n$, *then* (4.3) *becomes a universal similarity equality over* $\mathbb{H}$.

The proof of (4.3) is much analogous to that Theorem 2.2. Hence we omit it here. Some direct consequences of the above two theorems are listed below.

Corollary 4.3. *Let* $A = A_0 + A_1 i + A_2 j + A_3 k \in \mathbb{H}^{m\times n}$ *be given. Then* A *and its two adjoint matrices* $\Psi(A)$ *and* $\Phi(A)$ *satisfy the following four equalities*

$$(4.5)\qquad A = \frac{1}{2}E_{2m}\Psi(A)E_{2n}^*, \qquad A = \frac{1}{4}E_{4m}\Phi(A)E_{4n}^*,$$

and

$$(4.6)\quad \Psi(A)E_{2n}^*E_{2n} = E_{2m}^*E_{2m}\Psi(A), \qquad \Phi(A)E_{4n}^*E_{4n} = E_{4m}^*E_{4m}\Phi(A),$$

where $E_{2t} = (\,I_t,\ jI_t\,)$ *and* $E_{4t} = (\,I_t,\ iI_t,\ jI_t,\ kI_t\,)$, $t = m,\ n$.

Proof. The two equalities in (4.1) and (4.3) can also be written as

$$\mathrm{diag}(\,A,\ A\,) = P_{2m}^*\Psi(A)P_{2n}, \qquad \mathrm{diag}(\,A,\ A,\ A,\ A\,) = Q_{4m}^*\Phi(A)Q_{4n}.$$

Thus A can be expressed as

$$A = (\,I_m,\ 0\,)P_{2m}^*\Psi(A)P_{2n}(\,I_n,\ 0\,)^T,$$
$$A = (\,I_m,\ 0,\ 0,\ 0\,)Q_{4m}^*\Phi(A)Q_{4n}(\,I_n,\ 0,\ 0,\ 0\,)^T.$$

Written in explicit forms, they are the desired results in (4.5). On the other hand, note that

$$E_{2m}^{*}A = \Psi(A)E_{2n}^{*}, \qquad AE_{2n} = E_{2m}\Psi(A).$$

Thus

$$(E_{2m}^{*}A)E_{2n} = \Psi(A)E_{2n}^{*}E_{2n}, \qquad E_{2m}^{*}(AE_{2n}) = E_{2m}^{*}E_{2m}\Psi(A).$$

Since $(E_{2m}^{*}A)E_{2n} = E_{2m}^{*}(AE_{2n})$, the first one in (4.6) holds. In the same manner, we can also show the second one in (4.6). □

The results of the following corollary are well-known or easy to prove.

Corollary 4.4. *Let A, $B \in \mathbb{H}^{m\times n}$, $C \in \mathbb{H}^{n\times p}$ and $\lambda \in \mathbb{R}$. Then*

(a) $A = B \iff \Psi(A) = \Psi(B) \iff \Phi(A) = \Phi(B)$.

(b) $\Psi(A + B) = \Psi(A) + \Psi(B)$, $\quad \Phi(A + B) = \Phi(A) + \Phi(B)$.

(c) $\Psi(AC) = \Psi(A)\Psi(C)$, $\quad \Phi(AC) = \Phi(A)\Phi(C)$.

(d) $\Psi(\lambda A) = \Psi(A\lambda) = \lambda\Psi(A)$, $\quad \Phi(\lambda A) = \Phi(A\lambda) = \lambda\Phi(A)$.

(e) $\Psi(A^{*}) = \Psi^{*}(A)$, $\quad \Phi(A^{*}) = \Phi^{T}(A)$.

(f) $\mathrm{rank}A = \frac{1}{2}\mathrm{rank}\Psi(A)$, $\quad \mathrm{rank}A = \frac{1}{4}\mathrm{rank}\Phi(A)$.

(g) *if A is a nonsingular matrix of size m, then*

$$\Psi(A^{-1}) = \Psi^{-1}(A), \qquad \Phi(A^{-1}) = \Phi^{-1}(A), A^{-1} = \frac{1}{4}E_{4m}\Phi^{-1}(A)E_{4m}^{*},$$

where E_{2m} and E_{4m} are as in (4.5).

(h) *The Moore-Penrose inverse of $A^{\dagger}$ of A, i.e., the unique solution X to the following four equations*

$$AXA = A, \qquad XAX = X, \qquad (AX)^{*} = AX, \qquad (XA)^{*} = XA,$$

satisfies

$$\Psi(A^{\dagger}) = \Psi^{\dagger}(A), \qquad \Phi(A^{\dagger}) = \Phi^{\dagger}(A),$$

$$A^{\dagger} = \frac{1}{2}E_{2n}\Psi^{\dagger}(A)E_{2m}^{*}, \qquad A^{\dagger} = \frac{1}{4}E_{4n}\Phi^{\dagger}(A)E_{4m}^{*},$$

where E_{2m}, E_{2n}, E_{4m} and E_{4n} are as in (4.5).

From the structure of two adjoint matrices of a quaternion matrix, we can also derive the following results.

Theorem 4.5. *Let $A = A_0 + A_1 i + A_2 j + A_3 k \in \mathbb{H}^{m\times n}$ be given. Then*

(a) *The complex adjoint matrix $\Psi(A)$ of A satisfies the following equality*

$$\overline{\Psi(A)} = K_{2m}\Psi(A)K_{2n}^{-1}, \tag{4.7}$$

where $K_{2t} = \begin{pmatrix} 0 & -I_t \\ I_t & 0 \end{pmatrix}$, $\quad t = m,\ n$.

(b) *The real adjoint matrix* $\Phi(A)$ *of* A *satisfies the following equalities*

(4.8)
$$\Phi(A) = R_{4m}\Phi(A)R_{4n}^{-1}, \quad \Phi(A) = S_{4m}\Phi(A)S_{4n}^{-1}, \quad \Phi(A) = T_{4m}\Phi(A)T_{4n}^{-1},$$

where

(4.9)
$$R_{4t} = \begin{pmatrix} 0 & -I_t & 0 & 0 \\ I_t & 0 & 0 & 0 \\ 0 & 0 & 0 & I_t \\ 0 & 0 & -I_t & 0 \end{pmatrix},$$
$$S_{4t} = \begin{pmatrix} 0 & 0 & -I_t & 0 \\ 0 & 0 & 0 & -I_t \\ I_t & 0 & 0 & 0 \\ 0 & I_t & 0 & 0 \end{pmatrix}, \quad t = m,\ n,$$

(4.10)
$$T_{4t} = \begin{pmatrix} 0 & 0 & 0 & -I_t \\ 0 & 0 & I_t & 0 \\ 0 & -I_t & 0 & 0 \\ I_t & 0 & 0 & 0 \end{pmatrix}, \quad t = m,\ n.$$

The correctness of (4.7) and (4.8) can directly be verified by multiplying the matrices in them.

The two equalities in (4.1) and (4.3) can directly be applied to deal with various matrix problems over quaternion algebra. In most cases, we can easily extend various known results in real and complex matrix theory to quaternion algebra.

We next just present a general method to solve linear matrix equations over $\mathbb{H}$. We illustrate this method by considering the followineg linear matrix equation

(4.11)
$$A_1XB_1 + \cdots + A_lXB_l = C.$$

Theorem 4.6. *Let* $A_s \in \mathbb{H}^{m\times n}$, $B_s \in \mathbb{H}^{p\times q}$ *and* $C \in \mathbb{H}^{m\times q}$ *be given,* $s = 1, 2, \cdots, l$. *Then the matrix equation* (4.11) *has a solution* $X \in \mathbb{H}^{n\times p}$, *if and only if the following complex matrix equation*

(4.12)
$$\Psi(A_1)Y\Psi(B_1) + \cdots + \Psi(A_l)Y\Psi(B_l) = \Psi(C)$$

has a solution $Y \in \mathbb{C}^{2n\times 2p}$, *in which case, if* Y *is a solution to* (4.12), *then the following quaternion matrix*

(4.13)
$$X = \frac{1}{4}(\,I_n,\ jI_n\,)(\,Y + K_{2n}^{-1}\overline{Y}K_{2p}\,)\begin{pmatrix} I_p \\ -jI_p \end{pmatrix}$$

is a solution to (4.11), *where* K_{2n} *and* K_{2p} *are defined in* (4.7).

Proof. Suppose first that (4.11) has a solution $X \in \mathbb{H}^{n\times p}$. By applying Corollary 4.4(a)—(c) to the both sides of (4.11), we obtain

$$\Psi(A_1)\Psi(X)\Psi(B_1) + \cdots + \Psi(A_l)\Psi(X)\Psi(B_l) = \Psi(C),$$

which shows that $Y = \Psi(X)$ is a solution to (4.12). Conversely, assume that (4.12) has a solution

$$Y = \begin{pmatrix} Y_1 & Y_2 \\ Y_3 & Y_4 \end{pmatrix} \in \mathbb{C}^{2n\times 2p}, \quad Y_1,\ Y_2,\ Y_3,\ Y_4 \in \mathbb{C}^{n\times p}.$$

Then taking the complex conjugate on both sides of (4.12), we have

$$\overline{\Psi(A_1)}\,\overline{Y}\,\overline{\Psi(B_1)} + \cdots + \overline{\Psi(A_l)}\,\overline{Y}\,\overline{\Psi(B_l)} = \overline{\Psi(C)}. \tag{4.14}$$

According to (4.7), the matrix equation (4.14) can also be written as

$$K_{2m}\Psi(A_1)K_{2n}^{-1}\overline{Y}K_{2p}\Psi(B_1)K_{2q}^{-1} + \cdots + K_{2m}\Psi(A_l)K_{2n}^{-1}\overline{Y}K_{2p}\Psi(B_l)K_{2q}^{-1} = K_{2m}\Psi(C)K_{2q}^{-1},$$

which can be simplified to

$$\Psi(A_1)(K_{2n}^{-1}\overline{Y}K_{2p})\Psi(B_1) + \cdots + \Psi(A_l)(K_{2n}^{-1}\overline{Y}K_{2p})\Psi(B_l) = \Psi(C).$$

This equality shows that $K_{2n}^{-1}\overline{Y}K_{2p}$ is also a solution to (4.12). Thus the sum of

$$\hat{Y} = \frac{1}{2}(Y + K_{2n}^{-1}\overline{Y}K_{2p})$$

is also a solution to (4.12). Written in a block matrix, this sum is

$$\hat{Y} = \frac{1}{2}\begin{pmatrix} Y_1 & Y_2 \\ Y_3 & Y_4 \end{pmatrix} + \frac{1}{2}\begin{pmatrix} \overline{Y_4} & -\overline{Y_3} \\ -\overline{Y_2} & \overline{Y_1} \end{pmatrix} = \begin{pmatrix} Z_1 & -Z_2 \\ \overline{Z_2} & \overline{Z_1} \end{pmatrix}, \tag{4.15}$$

where $Z_1 = \frac{1}{2}(Y_1 + \overline{Y_4})$ and $Z_2 = \frac{1}{2}(-Y_2 + \overline{Y_3})$. From it we construct a quaternion matrix as follows

$$X = Z_1 + Z_2 j = \frac{1}{2}(I_n,\ jI_n)\hat{Y}\begin{pmatrix} I_p \\ -jI_p \end{pmatrix}.$$

Then its complex adjoint matrix apparently is

$$\Psi(X) = \hat{Y} = \begin{pmatrix} Z_1 & -Z_2 \\ \overline{Z_2} & \overline{Z_1} \end{pmatrix}.$$

Since this $\Psi(X)$ is a solution to (4.12), the corresponding matrix X in (4.13) by Corollary 4.4(a)—(c) is, of course, a solution to (4.11). □

An analogous result is given below.

Theorem 4.7. *Let $A_s \in \mathbb{H}^{m\times n}$, $B_s \in \mathbb{H}^{p\times q}$ and $C \in \mathbb{H}^{m\times q}$ be given, $s = 1, 2, \cdots, l$. Then the matrix equation in* (4.11) *has a solution $X \in \mathbb{H}^{n\times p}$ if and only if the following real matrix equation*

$$\Phi(A_1)Y\Phi(B_1) + \cdots + \Phi(A_l)Y\Phi(B_l) = \Phi(C) \tag{4.16}$$

has a solution $Y \in \mathbb{R}^{4n\times 4p}$; in which case, if Y is a solution to (4.16), *then the following quaternion matrix derived from Y*

$$X = \frac{1}{16}E_{4n}(Y + R_{4n}^{-1}YR_{4p} + S_{4n}^{-1}YS_{4p} + T_{4n}^{-1}YT_{4p})E_{4p}^{*} \tag{4.17}$$

is a solution to (4.11), *where E_{4n} and E_{4p} are as in* (4.5), *and R_{4n}, R_{4p}, S_{4n}, S_{4p}, T_{4n} and T_{4p} are defined given in* (4.9) and (4.10).

Proof. We only show that if

$$Y = \begin{pmatrix} Y_{11} & Y_{12} & Y_{13} & Y_{14} \\ Y_{21} & Y_{22} & Y_{23} & Y_{24} \\ Y_{31} & Y_{32} & Y_{33} & Y_{34} \\ Y_{41} & Y_{42} & Y_{43} & Y_{44} \end{pmatrix}, \quad Y_{uv} \in \mathbb{R}^{n\times p}, \quad u, \ v = 1, \ 2, \ 3, \ 4 \tag{4.18}$$

is a solution to (4.16), then the quaternion matrix given in (4.17) is a solution to (4.11). In fact, according to (4.8),

$$\begin{aligned} \Phi(A_s) &= R_{4m}\Phi(A_s)R_{4n}^{-1}, & \Phi(B_s) &= R_{4p}\Phi(B_s)R_{4q}^{-1}, & \Phi(C) &= R_{4m}\Phi(C)R_{4q}^{-1}, \\ \Phi(A_s) &= S_{4m}\Phi(A_s)S_{4n}^{-1}, & \Phi(B_s) &= S_{4p}\Phi(B_s)S_{4q}^{-1}, & \Phi(C) &= S_{4m}\Phi(C)S_{4q}^{-1}, \\ \Phi(A_s) &= T_{4m}\Phi(A_s)T_{4n}^{-1}, & \Phi(B_s) &= T_{4p}\Phi(B_s)T_{4q}^{-1}, & \Phi(C) &= T_{4m}\Phi(C)T_{4q}^{-1}. \end{aligned}$$

Substituting them into (4.16), respectively, and simplifying the corresponding equation, we get three equations as follows

$$\begin{aligned} \Phi(A_1)(R_{4n}^{-1}YR_{4p})\Phi(B_1) + \cdots + \Phi(A_l)(R_{4n}^{-1}YR_{4p})\Phi(B_l) &= \Phi(C), \\ \Phi(A_1)(S_{4n}^{-1}YS_{4p})\Phi(B_1) + \cdots + \Phi(A_l)(S_{4n}^{-1}YS_{4p})\Phi(B_l) &= \Phi(C), \\ \Phi(A_1)(T_{4n}^{-1}YT_{4p})\Phi(B_1) + \cdots + \Phi(A_l)(T_{4n}^{-1}YT_{4p})\Phi(B_l) &= \Phi(C). \end{aligned}$$

These three equations show that if Y is a solution to (4.16), then $R_{4n}^{-1}YR_{4p}$, $S_{4n}^{-1}YS_{4p}$ and $T_{4n}^{-1}YT_{4p}$ are also solutions to (4.16). Thus the following sum

$$\hat{Y} = \frac{1}{4}(\ Y + R_{4n}^{-1}YR_{4p} + S_{4n}^{-1}YS_{4p} + T_{4n}^{-1}YT_{4p}\ ) \tag{4.19}$$

is also a solution to (4.16). Now substituting (4.18) in (4.20) and then simplifying the expression, we get

$$(4.20) \qquad \hat{Y} = \begin{pmatrix} Z_0 & -Z_1 & -Z_2 & -Z_3 \\ Z_1 & Z_0 & -Z_3 & Z_2 \\ Z_2 & Z_3 & Z_0 & -Z_1 \\ Z_3 & -Z_2 & Z_1 & Z_0 \end{pmatrix},$$

where

$$Z_0 = \frac{1}{4}(Y_{11} + Y_{22} + Y_{33} + Y_{44}), \quad Z_1 = \frac{1}{4}(Y_{21} - Y_{12} + Y_{43} - Y_{34}),$$

$$Z_2 = \frac{1}{4}(Y_{31} - Y_{42} - Y_{13} + Y_{24}), \quad Z_3 = \frac{1}{4}(Y_{41} + Y_{32} - Y_{23} - Y_{14}).$$

The process from (4.18) to (4.20) needs much space, so we have to omit it here. From (4.20) we construct a quaternion matrix as follows

$$X = Z_0 + Z_1 i + Z_2 j + Z_3 k = \frac{1}{4} E_{4n} \hat{Y} E_{4p}^*.$$

Then its real adjoint matrix apparently is $\Phi(X) = \hat{Y}$. Since the matrix $\Phi(X) = \hat{Y}$ is a solution to (4.16), the corresponding matrix X in (4.17) by Corollary 4.4(a)—(c) is, of course, a solution to (4.11). □

The above two theorems show that the solvability and solution of the matrix equation (4.11) can all be determined by its two complex and real adjoint matrix equations in (4.12) and (4.16).

The methods showing in Theorems 4.6 and 4.7 can also be applied to deal with any other linear matrix equations over $\mathbb{H}$.

Conclusions. In this paper, we have established a group of universal factorization equalities for quaternions, as well as for matrices of quaternions. These equalities clearly reveal the relationship between the real quaternion algebra and the complex matrix algebra, as well as, the real matrix algebra. Based on the results presented in the paper, it is expected that a perfect theory on matrix analysis over the real quaternion algebra can be routinely established. Furthermore, the methods and results developed in this paper can also extend to octonions and matrices of octionions. We shall present them in another paper.

The author wishes to thank Professors George P. H. Styan and Agnes M. Herzberg for their guidance and encouragement. The work was supported by the Natural Sciences and Engineering Research Council of Canada.

References

[1] H. G. Baker, How to solve the quaternion equation $AZ + ZC = E$, preprint, 1996.

[2] L. Brand, The roots of a quaternion, *Amer. Math. Monthly* 49(1942), 519–520.
[3] J. L. Brenner, Matrices of quaternions, *Pacific J. Math.* 1(1951), 329–335.
[4] P. M. Cohn, The range of the derivation and the equation $ax - xb = c$, *J. Indian Math. Soc.* 37(1973), 1–9.
[5] P. M. Cohn, Skew Field Constructions, Cambridge U. P., London, 1977.
[6] J. Groß, G. Trenkler and S. Troschke, Quaternions: Further contributions to a matrix oriented approach, preprint, 2000.
[7] L. Huang, The matrix equation $AXB - CXD = E$ over the quaternion field, *Linear Algebra Appl.* 234(1996), 197–208.
[8] H. C. Lee, Eigenvalues of canonical forms of matrices with quaternion coefficients, *Proc. Roy. Irish. Acad.Sect. A* 52(1949), 253–260.
[9] N. Jacobson, Basic Algebra I, Freeman, New York, 1974.
[10] R. E. Johnson, On the equation $\chi\alpha = \gamma\chi + \beta$ over an algebraic division ring, *Bull. Amer. Math. Soc.* 50(1944), 202–207.
[11] P. R. Michael, Quaternionic linear and quadratic equations, *J. Natur. Geom.* 11(1997), 101–106.
[12] I, Niven, Equations in quaternions, *Amer. Math. Monthly* 48(1941), 654–661.
[13] I, Niven, The roots of a quaternion, *Amer. Math. Monthly* 49(1942), 386–388.
[14] O. Ore, Linear equations in non-commutative fields, *Ann. Math.* 32(1931), 463–477.
[15] A. R. Richardson, Simultaneous linear equations over a division algebra, *Proc. London Math. Soc.* 28(1928), 395–420.
[16] H. E. Salzer, An elementary note on powers of quaternions, *Amer. Math. Monthly* 59(1952), 298–300.
[17] Y. Tian, Universal similarity factorization equalities over real Clifford algebras, *Adv. Appl. Clifford Agebras* 8(1998), 365–402.
[18] X, Wang, Quaternion matrix and the re-nonnegative definite solutions to the quaternion matrix inverse problem $AX = B$, *Math. J. Okayama Univ.* 39(1997), 61–69.
[19] N. A. Wiegmann, Some theorems on matrices with real quaternion elements, *Canad. J. Math.* 7(1955), 191–201.
[20] L. A. Wolf, Similarity of matrices in which the elements are real quaternions, *Bull. Amer. Math. Soc.* 42(1936), 737–743.
[21] F. Zhang, Quaternions and matrices of quaternions, *Linear Algebra Appl.* 251(1997), 21–57.

Yongge Tian
Department of Mathematics and Statistics
Queen's University
Kingston, Ontario, Canada K7L 3N6
e-mail address: e-mail:ytian@mast.queensu.ca

(Received May 8, 2000)