Mitt. Math. Ges. Hamburg **27** (2008), 223–234

Linear equations in quaternionic variables

By Drahoslava Janovská & Gerhard Opfer

Abstract

We study the quaternionic linear system which is composed out of terms of the form $l_n(x) := \sum_{p=1}^{n} a_p x b_p$ with quaternionic constants a_p, b_p and a variable number n of terms. In the first place we investigate one equation in one variable. If $n = 2$ the corresponding equation, which is normally called Sylvester's equation will be treated completely by using only quaternionic algebra. For larger n a transition to the isomorphic (4×4) real matrix case is investigated. Sufficient conditions for non singularity will be obtained by using results from fixed point theorems. Connections to the Kronecker product are presented. The general case of a linear quaternionic system is treated, where each unknown is contained in a sum of the form mentioned above. As a tool the so-called column operator and its properties are used. An analogue of the Kronecker product for quaternionic systems involving terms of the form $\mathbf{AXB}$ is given.

Keywords: Linear equations in quaternions, Sylvester's equation in quaternions, Kronecker product for quaternionic matrices

AMS Subject classification: 11R52, 12E15, 12Y05, 65F40

1 Introduction

Because of the non commutativity of the (skew) field of quaternions, which we will denote by $\mathbb{H}$, the simplest one dimensional linear functions $l_n : \mathbb{H} \to \mathbb{H}$ already have the form

$$l_n(x) := \sum_{p=1}^{n} a_p x b_p, \quad a_p, b_p, x \in \mathbb{H}, a_p b_p \neq 0,\ p = 1, 2, \ldots, n, \tag{1.1}$$

where n may be an arbitrary, positive integer. The additional conditions $a_p b_p \neq 0$ for all $1 \leq p \leq n$ ensure that the sum does not reduce to a sum with less terms.

The set of positive integers will be denoted by $\mathbb{N}$, the set of integers by $\mathbb{Z}$, and $\mathbb{R}, \mathbb{C}$ will denote the real and complex numbers, respectively. The set of matrices with m rows and n columns and with elements from $\mathbb{F}$ will be denoted by $\mathbb{F}^{m\times n}$, where $\mathbb{F}$ is either $\mathbb{R}$ or $\mathbb{C}$ or $\mathbb{H}$.

It should be noted, that l_n is a linear function over $\mathbb{R}$ but not over $\mathbb{H}$. As a function over $\mathbb{H}$ it is just additive, i. e., $l_n(x+y) = l_n(x)+l_n(y)$. We will call l_n *singular* if there is $x \neq 0$ with $l_n(x) = 0$. Or in other words, l_n is singular, if the homogeneous equation $l_n(x) = 0$ has a non zero solution. If l_n is non singular, it is clear that $l_n(x) = c$ has a unique solution for all $c \in \mathbb{H}$.

Throughout this paper we shall use the well known fact, that a linear mapping $l : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is expressible by a unique matrix $\mathbf{A} \in \mathbb{R}^{m\times n}$ in the sense that $l(\mathbf{x}) = \mathbf{A}\mathbf{x}, \mathbf{x} \in \mathbb{R}^n$, which is the ordinary matrix times vector product. See [5, HORN & JOHNSON, Section 0.2], [6, HORN & JOHNSON, Lemma 4.3.2]. In particular, the mapping l_n defined in (1.1) may be taken as a linear mapping $l_n : \mathbb{R}^4 \rightarrow \mathbb{R}^4$ over $\mathbb{R}$. We will come to the matrix representation later. A form very similar to that given in (1.1) is known in real matrix theory. Define the linear mapping $\mathbf{L} : \mathbb{R}^{k\times l} \rightarrow \mathbb{R}^{j\times m}$ by

$$\mathbf{L}(\mathbf{X}) := \underbrace{\mathbf{A}}_{j\times k}\underbrace{\mathbf{X}}_{k\times l}\underbrace{\mathbf{B}}_{l\times m}. \tag{1.2}$$

This map $\mathbf{L}$ may also be regarded as a linear map of $\mathbb{R}^{kl} \rightarrow \mathbb{R}^{jm}$ by putting $\mathbf{x} := \mathit{col}(\mathbf{X})$ where $\mathbf{x}$ consists of one column formed by all columns of $\mathbf{X}$ from the left to the right. We note, that the column operator *col* is linear over $\mathbb{R}$. According to the general theory there must be a matrix $\mathbf{P}(\mathbf{A},\mathbf{B}) \in \mathbb{R}^{jm\times kl}$ such that

$$\mathit{col}(\mathbf{L}(\mathbf{X})) = \mathbf{P}(\mathbf{A},\mathbf{B})\mathit{col}(\mathbf{X}).$$

The matrix $\mathbf{P}(\mathbf{A},\mathbf{B})$ is known in the literature as *Kronecker product* (see [6, HORN & JOHNSON, Section 4.3]) and it is defined by

$$\mathbf{P}(\mathbf{A},\mathbf{B}) := \begin{pmatrix} b_{11}\mathbf{A} & b_{21}\mathbf{A} & \cdots & b_{l1}\mathbf{A} \\ b_{12}\mathbf{A} & b_{22}\mathbf{A} & \cdots & b_{l2}\mathbf{A} \\ \vdots & \vdots & & \vdots \\ b_{1m}\mathbf{A} & b_{2m}\mathbf{A} & \cdots & b_{lm}\mathbf{A} \end{pmatrix} \in \mathbb{R}^{jm\times kl}. \tag{1.3}$$

This implies that a matrix equation of the form

$$\begin{aligned} \mathbf{L}_n(\mathbf{X}) &:= \sum_{p=1}^{n} \mathbf{A}_p\mathbf{X}\mathbf{B}_p = \mathbf{C}, \qquad (1.4) \\ \mathbf{A}_p &\in \mathbb{R}^{j\times k}, \mathbf{B}_p \in \mathbb{R}^{l\times m}, \mathbf{C} \in \mathbb{R}^{j\times m}, \mathbf{X} \in \mathbb{R}^{k\times l}, p = 1,2,\ldots,n, \end{aligned}$$

can be transformed into the ordinary matrix equation

$$\Big(\sum_{p=1}^{n} \mathbf{P}(\mathbf{A}_p,\mathbf{B}_p)\Big)\mathit{col}(\mathbf{X}) = \mathit{col}(\mathbf{C}). \tag{1.5}$$

This system has jm equations in kl unknowns. The pattern (1.4)⇒(1.5) gives some guideline for the quaternionic case.

2 Sylvester's equation in one quaternionic variable

Let us consider the equation $l_n(x) = c$ with given $c \in \mathbb{H}$, and where l_n is defined in (1.1). We will see that for $n \leq 2$ a solution can be found just by applying quaternionic algebra, but that for $n \geq 3$ this seems not to be possible. For $n = 1$ we have $l_1(x) := axb$ and it is clear that l_1 is singular if and only if $ab = 0$. For $n = 2$ we have

$$l_2(x) := \sum_{p=1}^{2} a_p x b_p = c, \quad a_1 b_1 \neq 0, a_2 b_2 \neq 0.$$

By premultiplication from the left by a_1^{-1} and by b_2^{-1} from the right the equation takes the form

$$x b_1 b_2^{-1} + a_1^{-1} a_2 x = a_1^{-1} c b_2^{-1}.$$

We simplify the notation to

$$Ax + xB = C; \quad A := a_1^{-1} a_2 \neq 0, B := b_1 b_2^{-1} \neq 0, C := a_1^{-1} c b_2^{-1}. \tag{2.1}$$

This equation is usually called *Sylvester's equation*. The notion of *equivalence* of two quaternions is very helpful and will be introduced in the next definition.

Definition 2.1 Two quaternions a, b are said to be *equivalent* if there is an invertible quaternion h such that

$$h^{-1} a h = b.$$

In this connection, algebraists (see [13, V. D. WAERDEN, 1960, p. 35]) usually use the term *conjugate*, which, however, for quaternions is not a good choice. By this definition, two real numbers are equivalent if and only if they coincide. A complex number z is equivalent to its complex conjugate $\overline{z}$. The general situation is governed by the following lemma.

Lemma 2.2 *Two quaternions a, b are equivalent if and only if*

$$|a| = |b| \text{ and } \Re a = \Re b,$$

where $\Re z$ refers to the real part, the first component of $z \in \mathbb{H}$.

Proof: [9, JANOVSKÁ & OPFER]. $\square$

Theorem 2.3 *Sylvester's equation $l_2(x) := Ax + xB = C; A, B, C, x \in \mathbb{H}, AB \neq 0$, is singular if and only if A and $-B$ are equivalent. If it is non singular, its solution is*

$$x = f_l^{-1}(C + A^{-1} C \overline{B}) = (C + \overline{A} C B^{-1}) f_r^{-1} \text{ where} \tag{2.2}$$

$$f_l := 2\Re B + A + |B|^2 A^{-1}; \quad f_r := 2\Re A + B + |A|^2 B^{-1}. \tag{2.3}$$

Proof: Let $A, -B$ be equivalent, i. e. there is an $h \neq 0$ such that $B = -h^{-1} A h$. Thus, $l_2(x) = Ax - x h^{-1} A h$ and $l_2(h) = 0$, hence l_2 is singular. Let l_2 be singular, i. e. there is an $x \neq 0$ such that $l_2(x) := Ax + xB = 0$. Multiplying from the left by x^{-1} shows that $A, -B$ are equivalent. For the remainder we show that $f_l = 0$ if and only if $A, -B$ are equivalent. Put $q^2 := |B|^2/|A|^2$ and $A := (A_1, A_2, A_3, A_4), B :=$

(B_1, B_2, B_3, B_4). Let $0 = f_l = 2\Re B + A + |B|^2 A^{-1} = 2\Re B + A + q^2\overline{A} = (2B_1+(1+q^2)A_1, (1-q^2)A_2, (1-q^2)A_3, (1-q^2)A_4)$. If $q^2 = 1$ we have $|B|^2 = |A|^2$ and $2B_1 + 2A_1 = 0$. According to Lemma 2.2, $A, -B$ are equivalent. Let $q^2 \neq 1$. Then, $2B_1 + (1 + q^2)A_1 = 0, A_2 = A_3 = A_4 = 0$. Thus, A is real and $A_1 \neq 0$. Then, $0 = 2B_1 + (1 + q^2)A_1 = 2B_1 + (1 + |B|^2/|A_1|^2)A_1$. Multiplying by A_1 we obtain $(A_1 + B_1)^2 + B_2^2 + B_3^3 + B_4^2 = 0$ and therefore, B is also real and $A_1 = -B_1$, implying $q^2 = B_1^2/A_1^2 = 1$, a contradiction. Let $A, -B$ be equivalent. Then, $f_l = -2\Re A + A + |A|^2\overline{A}/|A|^2 = -2\Re A + A + \overline{A} = 0$. A very similar proof works for f_r. The first solution formula follows from $A^{-1}l_2(x)\overline{B} + l_2(x) = f_l x$ and from $l_2(x) = C$. For the second solution formula we form $\overline{A}l_2(x)B^{-1} + l_2(x) = xf_r$. $\square$

In (2.2) we have two solution formulas for the same x, the solution of Sylvester's equation. Nevertheless, for numerical purposes we should make a difference. If $|A| > 0$ but close to zero, then the first formula involving A^{-1} should be avoided. More generally, we should compare $|A|$ and $|B|$. If $|A| \leq |B|$ we should use the second formula, otherwise the first one. This would allow the inclusion of the cases $A = 0$ or $B = 0$ (not both).

In a paper by [11, R. E. Johnson, 1944] we find a treatment of an equation of type $Ax + xB = C$ over an algebraic division ring. However, the solution formula (2.2) was not given. The same equation with quaternionic matrices was treated by [7, Huang, 1996]. We return to this paper a little later.

3 Linear equations of general type in one quaternionic variable

We will treat equation (1.1) without a restriction on n, the number of terms. We introduce two mappings $\imath_1, \imath_2 : \mathbb{H} \to \mathbb{R}^{4\times 4}$. The first one, $\imath_1$, is the well known isomorphism between $\mathbb{H}$ and $\mathbb{R}^{4\times 4}$ and the other one is obtained by introducing a new multiplication in $\mathbb{H}$ namely $a \star b = ba$, where ba is the standard quaternion product. These mappings are

$$\imath_1(a) := \begin{pmatrix} a_1 & -a_2 & -a_3 & -a_4 \\ a_2 & a_1 & -a_4 & a_3 \\ a_3 & a_4 & a_1 & -a_2 \\ a_4 & -a_3 & a_2 & a_1 \end{pmatrix} \in \mathbb{R}^{4\times 4}, \tag{3.1}$$

$$\imath_2(a) := \begin{pmatrix} a_1 & -a_2 & -a_3 & -a_4 \\ a_2 & a_1 & a_4 & -a_3 \\ a_3 & -a_4 & a_1 & a_2 \\ a_4 & a_3 & -a_2 & a_1 \end{pmatrix} \in \mathbb{R}^{4\times 4}. \tag{3.2}$$

Let us denote the set of matrices of the form $\imath_1(a)$ by $\mathbb{H}_\mathbb{R}$ and the set of matrices of the form $\imath_2(a)$ by $\mathbb{H}_\mathbb{P}$. Both sets constitute skew fields. The elements of $\mathbb{H}_\mathbb{P}$ will be called *pseudo quaternions*. The essential properties of $\imath_1$ and $\imath_2$ are given in the next lemma.

Lemma 3.1 *Let $a = (a_1, a_2, a_3, a_4) \in \mathbb{H}$ and $col(a) := (a_1, a_2, a_3, a_4)^{\mathrm{T}}$, the column vector consisting of the four components of a. For $a, b, c \in \mathbb{H}$ we have*

$$\begin{aligned} \imath_2(ab) &= \imath_2(b)\imath_2(a), &(3.3)\\ \imath_1(a)\imath_2(b) &= \imath_2(b)\imath_1(a), &(3.4)\\ col(ab) &= \imath_1(a)col(b) = \imath_2(b)col(a), &(3.5)\\ col(abc) &= \imath_1(a)\imath_1(b)col(c) = \imath_2(c)\imath_2(b)col(a), &(3.6)\\ col(abc) &= \imath_1(a)\imath_2(c)col(b). &(3.7)\end{aligned}$$

Proof: A comparison of both sides of (3.3) and (3.4) yields the desired result. For (3.5) to (3.7) see [4, GÜRLEBECK & SPRÖSSIG, Lemma 1.23, p. 6] and [2, ARAMANOVITSCH, Appendix A No. 8., p. 1252]. □

Since [2, ARAMANOVITSCH, 1995] gives the formulas (3.5) to (3.7) without any quotation and explanation, we assume that these formulas belong to a class of formulas which are already known for a longer time.

Formula (3.7) is now the key for transforming any equation of the form $l_n(x) = c$ into a standard matrix equation.

Theorem 3.2 *Let $l_n(x) := \sum_{p=1}^{n} a_p x b_p$ with quaternionic entries be given. Then $l_n(x) = c$ is equivalent to the (4×4)-matrix equation*

$$\Big(\sum_{p=1}^{n} \imath_1(a_p)\imath_2(b_p)\Big)col(x) = col(c). \tag{3.8}$$

Proof: Follows immediately from formula (3.7). □

Corollary 3.3 *The linear function l_n defined in Theorem 3.2 is singular if and only if* $\det\big(\sum_{p=1}^{n} \imath_1(a_p)\imath_2(b_p)\big) = 0$.

Proof: Follows from the isomorphic representation (3.8) in Theorem 3.2. □

What happens, if we apply the Kronecker product to $l(x) := axb = c$ in the isomorphic matrix form

$$\imath_1(a)\imath_1(x)\imath_1(b) = \imath_1(c).$$

Applying (1.5) we obtain

$$\mathbf{P}(\imath_1(a), \imath_1(b))col(\imath_1(x)) = col(\imath_1(c)), \tag{3.9}$$

and, using (1.3), we obtain

$$\mathbf{P}(\imath_1(a), \imath_1(b)) = \begin{pmatrix} b_1\imath_1(a) & b_2\imath_1(a) & b_3\imath_1(a) & b_4\imath_1(a) \\ -b_2\imath_1(a) & b_1\imath_1(a) & b_4\imath_1(a) & -b_3\imath_1(a) \\ -b_3\imath_1(a) & -b_4\imath_1(a) & b_1\imath_1(a) & b_2\imath_1(a) \\ -b_4\imath_1(a) & b_3\imath_1(a) & -b_2\imath_1(a) & b_1\imath_1(a) \end{pmatrix} \in \mathbb{R}^{16\times 16}.$$

Since the second, third, and fourth four components of $col(\imath_1(x))$ are essentially permutations of the first four components of $\mathbf{x} := (x_1, x_2, x_3, x_4)^{\mathrm{T}}$, the first four rows of the Kronecker product can be (essentially) expressed by column permutations of the matrix $\imath_1(a)$. More precisely, define $\mathbf{I}_1 = \mathrm{diag}(1,1,1,1)$, the identity matrix and

$$\mathbf{I}_2 = \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{pmatrix}; \mathbf{I}_3 = \begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}; \mathbf{I}_4 = \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}.$$

Then, the first four rows of $\mathbf{P}(\imath_1(a), \imath_1(b))col(\imath_1(x))$ take the form

$$\big(b_1\imath_1(a)\mathbf{I}_1 + b_2\imath_1(a)\mathbf{I}_2 + b_3\imath_1(a)\mathbf{I}_3 + b_4\imath_1(a)\mathbf{I}_4\big)\mathbf{x}. \tag{3.10}$$

Theorem 3.4 *Let $a, b \in \mathbb{H}$ be given. Then $l_1(x) := axb \in \mathbb{H}$ and $\imath_1(l_1(x)) = \imath_1(a)\imath_1(x)\imath_1(b) \in \mathbb{R}^{4\times4}$ are isomorphix expressions in $\mathbb{H}, \mathbb{R}^{4\times4}$, respectively. Let $\mathbf{x} = (x_1, x_2, x_3, x_4)^{\mathrm{T}}$. Then, the first four rows of the corresponding equation (3.9) defined by the Kronecker product have the explicit form given in (3.10) and this form coincides with $\imath_1(a)\imath_2(b)\mathbf{x}$ according to (3.8).*

Proof: According to Theorem 3.2 one has to show that $b_1\imath_1(a)\mathbf{I}_1 + b_2\imath_1(a)\mathbf{I}_2 + b_3\imath_1(a)\mathbf{I}_3 + b_4\imath_1(a)\mathbf{I}_4 = \imath_1(a)\imath_2(b)$. This is a little tedious and left out here. $\square$

For Sylvester's equation we were able to characterize the singular cases. This seems to be impossible for the general case. However, by applying some results from fixed point theorems, it is possible to find a sufficient condition for non singularity.

Theorem 3.5 *Let $l_n := \sum_{p=1}^{n} a_p x b_p$ with $n \geq 3, a_p, b_p \in \mathbb{H}$ and $a_p b_p \neq 0$ for all $p = 1, 2, \ldots, n$. If there is a $1 \leq p_0 \leq n$ such that*

$$\kappa := \frac{\sum_{p\neq p_0}^{n} |a_p||b_p|}{|a_{p_0}||b_{p_0}|} < 1, \tag{3.11}$$

then, l_n is non singular.

Proof: Choose $1 \leq p_0 \leq n$. We multiply $l_n(x) = c$ from the left by $a_{p_0}^{-1}$ and from the right by $b_{p_0}^{-1}$. Then we obtain $x + \sum_{p\neq p_0}^{n} a'_p x b'_p = c'$ where $a'_p = a_{p_0}^{-1} a_p, b'_p = b_p b_{p_0}^{-1}, p \neq p_0, c' = a_{p_0}^{-1} c b_{p_0}^{-1}$. Put $x = c' - \sum_{p\neq p_0}^{n} a'_p x b'_p := f(x)$. This is a linear fixed point equation and

$$f(x) - f(y) = \sum_{p\neq p_0}^{n} a'_p (y - x) b'_p = a_{p_0}^{-1} \big(\sum_{p\neq p_0}^{n} a_p (y - x) b_p \big) b_{p_0}^{-1}.$$

Going to absolute values and applying the triangle inequality yields

$$|f(x) - f(y)| \leq \kappa |x - y|,$$

where κ is defined in (3.11). If one of the posible κs is less than one, Banach's fixed point theorem [1] gives the desired result. $\square$

Corollary 3.6 *Let $l_n := \sum_{p=1}^{n} a_p x b_p$ with $n \geq 3, a_p, b_p \in \mathbb{H}$ and $a_p b_p \neq 0$ for all $p = 1, 2, \ldots, n$. If there is a constant $k > 0$ and an index $1 \leq p_0 \leq n$ such that $|a_p||b_p| \leq k$ for all $p \neq p_0$ and $|a_{p_0}||b_{p_0}| \geq k^2$, then, l_n is non singular if $k > n-1$. For $k := n-1$ this is not necessarily true.*

Proof: Under the stated conditions we have

$$\kappa \leq \frac{(n-1)k}{k^2} = \frac{(n-1)}{k} < 1.$$

The result follows from Theorem 3.5. In order to prove the last part, let $n = 3$, and $a_1 := (1,1,1,1), b_1 = 1, a_2 := (1,1,1,-1), b_2 = (-1,1,1,1), a_3 = 1$, $b_3 = (1,1,-1,-1)$ and choose $p_0 = 2$. We have $|a_1||b_1| = k = n-1 = 2, |a_2||b_2| = 4 = k^2, |a_3||b_3| = 2 = k$. But, the linear function $l_3(x) := a_1 x + a_2 x b_2 + x b_3$ is singular. In order to find that out apply Corollary 3.3 and compute

$$\det\big(\imath_1(a_1) + \imath_1(a_2)\imath_2(b_2) + \imath_2(b_3)\big).$$

The second column of the involved matrix is zero, thus, $\det = 0$. $\square$

We derived some sufficient conditions for non singularity by applying Banach's fixed point theorem. However, this does not imply that we advocate the use of fixed point iterations. The linear function l_n is treated in more detail for the case $n = 3$ in [10].

4 Linear systems in quaternionic variables

Let $\mathbf{A}$ be a quadratic matrix with quaternionic entries. Then, the system $\mathbf{A}\mathbf{x} = \mathbf{b}$ and the eigenvalue problem $\mathbf{A}\mathbf{x} = \mathbf{x}\lambda$ (observe, the position of λ) have very much in common with their real and complex relatives. See [14, ZHANG, 1997]. In particular, the quadratic system could be solved by elimination, or in other words, there is an $\mathbf{LU}$ decomposition (under certain conditions). However, determinants do not exist which in particular implies that there are no characteristic equations. See [3, FAN, 2003].

Consider the simple example

$$\begin{aligned} xa + by &= f, \\ cx + dy &= g. \end{aligned} \tag{4.1}$$

This system already creates problems if we would use only quaternionic algebra. See also [8] for systems of this type. Instead we apply the column operator *col* to the system (4.1) and use its properties collected in Lemma 3.1. We obtain

$$\begin{aligned} col(xa) + col(by) = \imath_2(a)col(x) + \imath_1(b)col(y) &= col(f), \\ col(cx) + col(dy) = \imath_1(c)col(x) + \imath_1(d)col(y) &= col(g). \end{aligned} \tag{4.2}$$

Define

$$\mathbf{A} := \begin{pmatrix} \imath_2(a) & \imath_1(b) \\ \imath_1(c) & \imath_1(d) \end{pmatrix}, \quad \mathbf{z} := \begin{pmatrix} col(x) \\ col(y) \end{pmatrix}, \mathbf{h} := \begin{pmatrix} col(f) \\ col(g) \end{pmatrix}.$$

Then (4.1) reads in real terms

$$\mathbf{A}\mathbf{z} = \mathbf{h}. \tag{4.3}$$

Example 4.1 For problem (4.1) choose $a := \mathbf{k}, b := \mathbf{j}, c := \mathbf{i}, d := 1 + \mathbf{k}, f := (-11, 11, 3, -5), g := (-5, 0, 9, 16)$, where $\mathbf{i}, \mathbf{j}, \mathbf{k}$ have the standard meaning. The right hand side is chosen such that $x = (1, 2, 3, 4), y = (5, 6, 7, 8)$ is the solution which will also be reproduced by solving (4.3).

The technique described for the (2×2) system can be applied in a very general situation. Define

$$l_{n_{jk}}^{(jk)}(x_k) := \sum_{p=1}^{n_{jk}} a_p^{(jk)} x_k b_p^{(jk)}; a_p^{(jk)}, b_p^{(jk)} \in \mathbb{H}\backslash\{0\}, 1 \le j, k \le n,\ n_{jk} \in \mathbb{N} \tag{4.4}$$

and consider the quaternionic $n \times n$ system

$$\sum_{k=1}^{n} l_{n_{jk}}^{(jk)}(x_k) = c^{(j)}, \quad j = 1, 2, \ldots, n. \tag{4.5}$$

Theorem 4.2 *Let $\imath_1, \imath_2$ be defined as in (1.2), (1.3), respectively. The quaternionic $(n \times n)$ system (4.5) is equivalent to the real $(4n \times 4n)$ system*

$$\mathbf{A}\mathbf{z} = \mathbf{c}, \tag{4.6}$$

where

$$\mathbf{A} := (\, \mathbf{a}_{jk} \,), j, k = 1, 2, \ldots, n, \quad \mathbf{z} := \begin{pmatrix} col(x_1) \\ col(x_2) \\ \vdots \\ col(x_n) \end{pmatrix}, \mathbf{c} := \begin{pmatrix} col(c^{(1)}) \\ col(c^{(2)}) \\ \vdots \\ col(c^{(n)}) \end{pmatrix},$$

and where

$$\mathbf{a}_{jk} := \sum_{p=1}^{n_{jk}} \imath_1(a_p^{(jk)}) \imath_2(b_p^{(jk)}) \in \mathbb{R}^{4 \times 4}.$$

Proof: Apply the column operator *col* to (4.5) and use Theorem 3.2. □

5 The Kronecker product for quaternionic systems composed out of $\mathbf{AXB}$

Let $\mathbf{L}(\mathbf{X}) := \mathbf{AXB}$ with the same dimensions as defined in (1.2) where, however, all matrices should have quaternions as entries and the dimensions will be set in capital letters. Thus, $\mathbf{L} : \mathbb{H}^{K\times L} \to \mathbb{H}^{J\times M}$. Again, $\mathbf{L}$ may be regarded as a linear real mapping $\mathbf{L} : \mathbb{R}^{4KL} \to \mathbb{R}^{4JM}$. According to the general theory there must be a real matrix $\mathbf{\Pi}(\mathbf{A},\mathbf{B}) \in \mathbb{R}^{4JM\times 4KL}$ such that

$$col(\mathbf{L}(\mathbf{X})) = \mathbf{\Pi}(\mathbf{A},\mathbf{B})col(\mathbf{X}). \tag{5.1}$$

We have to extend the definition of the column operator col to quaternionic matrices $\mathbf{X} \in \mathbb{H}^{K\times L}$ as follows: First replace $\mathbf{X}$ by a quaternionic column vector of size $KL \times 1$ by putting all columns of $\mathbf{X}$ into one column going from left to right. Then, replace all quaternions by a real 4×1 column. Eventually, $col(\mathbf{X}) \in \mathbb{R}^{4KL\times 1}$. The Kronecker product $\mathbf{P}(\mathbf{A},\mathbf{B})$ introduced in (1.3) does not coincide with $\mathbf{\Pi}(\mathbf{A},\mathbf{B})$. In order to see this let $J = K = L = M = 1$ and, correspondingly,

$$\mathbf{L}(x) := axb, a, b, x \in \mathbb{H}.$$

Here we have $\mathbf{P}(a,b) = ab$ but $\mathbf{P}(a,b)x \neq \mathbf{L}(x) = axb$. Applying formula (3.7) of Lemma 3.1 yields the correct formula $\mathbf{\Pi}(a,b) = \imath_2(a)\imath_1(b)$. The question is, how $\mathbf{\Pi}(\mathbf{A},\mathbf{B})$ has to be defined in general. The already mentioned paper by [7, Huang] does not give information in this direction. We have to extend the definitions of $\imath_1, \imath_2$ given in (3.1), (3.2) to matrices. Let $\mathbf{A} \in \mathbb{H}^{J\times K}$ and $\mathbf{A} = (a_{j,k}), j = 1,2,\ldots,J$, $k = 1,2,\ldots,K$. Then, put

$$\imath_1(\mathbf{A}) = (\imath_1(a_{j,k})) \in \mathbb{R}^{4J\times 4K}, \quad \imath_2(\mathbf{A}) = (\imath_2(a_{j,k})) \in \mathbb{R}^{4J\times 4K}. \tag{5.2}$$

In the first place we present an analogue of Lemma 3.1 for matrices. For this purpose we introduce *block transposition* for real matrices $\mathbf{A} \in \mathbb{R}^{4J\times 4K}$ denoted by $\mathbf{A}^{\mathrm{B}}$ where four by four blocks remain together. Let $\mathbf{A} = (a_{jk})$ consist of the (4×4) blocks

$$b_{jk} := a_{(4(j-1)+1:4(j-1)+4,4(k-1)+1:4(k-1)+4)}, j = 1,2,\ldots,J; k = 1,2,\ldots,K.$$

Then, we define

$$\mathbf{A}^{\mathrm{B}} := \begin{pmatrix} b_{11} & b_{21} & \ldots & b_{J1} \\ b_{12} & b_{22} & \ldots & b_{J2} \\ \vdots & \vdots & \ldots & \vdots \\ b_{1K} & b_{2K} & \ldots & b_{JK} \end{pmatrix} \in \mathbb{R}^{4K\times 4J}. \tag{5.3}$$

It should be noted that the ordinary transposition and the block transposition do *not* have the property $(\mathbf{AB})^{\mathrm{T}} = \mathbf{B}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}$, $(\mathbf{AB})^{\mathrm{B}} = \mathbf{B}^{\mathrm{B}}\mathbf{A}^{\mathrm{B}}$. As an example take two arbitrary (4×4) matrices $\mathbf{A},\mathbf{B}$. Then the block transposition does not change the matrix, and $(\mathbf{AB})^{\mathrm{B}} = \mathbf{AB}$, and $\mathbf{B}^{\mathrm{B}}\mathbf{A}^{\mathrm{B}} = \mathbf{BA}$ which is in general different from $\mathbf{AB}$. On the other hand, $\mathbf{A} = (\mathbf{A}^{\mathrm{T}})^{\mathrm{T}}$, $\mathbf{A} = (\mathbf{A}^{\mathrm{B}})^{\mathrm{B}}$ are valid.

Lemma 5.1 *Let $\mathbf{A} \in \mathbb{H}^{J\times K}, \mathbf{B} \in \mathbb{H}^{K\times L}, \mathbf{C} \in \mathbb{H}^{L\times M}$. For integers m, n we use the notation $m : n := \{m, m+1, \ldots, n\}$. Then, for $j = 1, 2, \ldots, J$, $k = 1, 2, \ldots, K, l = 1, 2, \ldots, L, m = 1, 2, \ldots, M$ we have*

$$\begin{aligned}
\imath_1(\mathbf{AB}) &= \imath_1(\mathbf{A})\imath_1(\mathbf{B}), &(5.4)\\
\imath_2(\mathbf{AB}) &= \big(\imath_2(\mathbf{B}^{\mathrm{T}})\imath_2(\mathbf{A}^{\mathrm{T}})\big)^{\mathrm{B}}, &(5.5)\\
\imath_1(\mathbf{A})\imath_2(\mathbf{B}) &= \big(\imath_2(\mathbf{B}^{\mathrm{T}})\imath_1(\mathbf{A}^{\mathrm{T}})\big)^{\mathrm{B}}, &(5.6)\\
col(\mathbf{AB})_{4J(l-1)+1:4Jl} &= \imath_1(\mathbf{A})col(\mathbf{B})_{4K(l-1)+1:4Kl}, &(5.7)\\
col(\mathbf{AB})_{[j;L]} &= \imath_2(\mathbf{B}^{\mathrm{T}})col(\mathbf{A}^{\mathrm{T}})_{4K(j-1)+1:4Kj}, &(5.8)\\
[j;L] &:= j, j+J, j+2J, \ldots, j+(L-1)J, &(5.9)\\
col(\mathbf{ABC})_{4J(m-1)+1:4Jm} &= \imath_1(\mathbf{AB})col(\mathbf{C})_{4L(m-1)+1:4Lm}, &(5.10)\\
col(\mathbf{ABC})_{[j;M]} &= \big(\imath_2(\mathbf{BC})\big)^{\mathrm{B}}col(\mathbf{A}^{\mathrm{T}})_{4K(j-1)+1:4Kj}. &(5.11)
\end{aligned}$$

The abbreviation $[j;L]$ in (5.8) means that the numbers of the elements of $col(\mathbf{AB})_{[j;L]}$ do not follow the same scheme as on the right side, where the numbers are running consecutively. It means, that the first four computed components are $col(\mathbf{AB})_{4(j-1)+1:4(j-1)+4}$, the second four computed components are actually $col(\mathbf{AB})_{4(j-1+J)+1:4(j-1+J)+4}$, and the last four computed components are $col(\mathbf{AB})_{4(j-1+(L-1)J)+1:4(j-1+(L-1)J)+4}$.

Proof: Combine matrix multiplication rules with the results of Lemma 3.1. □

Finally, we come to the still missing Kronecker product.

Theorem 5.2 *Let $\mathbf{A} \in \mathbb{H}^{J\times K}, \mathbf{B} \in \mathbb{H}^{K\times L}, \mathbf{C} \in \mathbb{H}^{L\times M}$ and $\mathbf{b} := col(\mathbf{B}) \in \mathbb{R}^{4KL\times 1}$. The elements of $\mathbf{A}$ are denoted by $A_{jk}, j = 1, 2, \ldots, J; k = 1, 2, \ldots, K$ and correspondingly for the other matrices. Define*

$$\begin{aligned}
\mathbf{I}_{jm}(kl) &:= \imath_1(A_{jk})\imath_2(C_{lm}) \in \mathbb{R}^{4\times 4}, \quad j = 1, 2, \ldots, J, etc. &(5.12)\\
\mathbf{I}_{jm} &:= (\mathbf{I}_{jm}(1), \mathbf{I}_{jm}(2), \ldots, \mathbf{I}_{jm}(KL)) \in \mathbb{R}^{4\times 4KL}. &(5.13)
\end{aligned}$$

We have enumerated $\mathbf{I}_{jm}(kl)$ columnwise with respect to k, l in the form $\mathbf{I}_{jm}(1)$, $\mathbf{I}_{jm}(2), \ldots, \mathbf{I}_{jm}(KL)$. Thus, the order is $(1,1), (2,1), \ldots, (K,1); (1,2), (2,2), \ldots, (K,2); \ldots, (1,L), (2,L), \ldots, (K,L)$. In the same way we enumerate $\mathbf{I}_p$. Then,

$$col(\mathbf{ABC})_{4(p-1)+1:4(p-1)+4} = \mathbf{I}_p\mathbf{b}, \quad p = 1, 2, \ldots, JM,$$

and the Kronecker product is

$$\Pi(\mathbf{A}, \mathbf{C}) = \begin{pmatrix} \mathbf{I}_1 \\ \mathbf{I}_2 \\ \vdots \\ \mathbf{I}_{JM} \end{pmatrix} \in \mathbb{R}^{4JM\times 4KL}. \tag{5.14}$$

Proof: Use the matrix multiplication rules for $\mathbf{ABC}$ and apply (3.7). □

It is clear that also systems of type $\sum_{p=1}^{P} \mathbf{A}_p\mathbf{X}\mathbf{B}_p = \mathbf{C}$ could be treated in the same fashion.

Example 5.3 *We solve the* (2×2) *system* $\mathbf{AXB}=\mathbf{C}$ *for the following data:*

$$\mathbf{A} := \begin{pmatrix} \begin{pmatrix} 0\\ 2\\ 2\\ 0 \end{pmatrix} & \begin{pmatrix} 4\\ 5\\ -1\\ -5 \end{pmatrix} \\ \begin{pmatrix} 0\\ 2\\ 2\\ -1 \end{pmatrix} & \begin{pmatrix} -3\\ 3\\ -3\\ 2 \end{pmatrix} \end{pmatrix}, \mathbf{B} := \begin{pmatrix} \begin{pmatrix} 0\\ 4\\ -5\\ -4 \end{pmatrix} & \begin{pmatrix} -2\\ 2\\ 1\\ -4 \end{pmatrix} \\ \begin{pmatrix} -3\\ -5\\ 2\\ -1 \end{pmatrix} & \begin{pmatrix} 4\\ 3\\ -2\\ 3 \end{pmatrix} \end{pmatrix},$$

$$\mathbf{C} := \begin{pmatrix} \begin{pmatrix} 80\\ -51\\ 146\\ -187 \end{pmatrix} & \begin{pmatrix} -178\\ 77\\ -12\\ 29 \end{pmatrix} \\ \begin{pmatrix} 32\\ 152\\ 68\\ -20 \end{pmatrix} & \begin{pmatrix} -40\\ -65\\ 28\\ 89 \end{pmatrix} \end{pmatrix}, \mathbf{X} := \begin{pmatrix} \begin{pmatrix} 1\\ 1\\ 1\\ 1 \end{pmatrix} & \begin{pmatrix} 1\\ 2\\ 1\\ 2 \end{pmatrix} \\ \begin{pmatrix} 2\\ 1\\ 2\\ 1 \end{pmatrix} & \begin{pmatrix} 2\\ 2\\ 2\\ 2 \end{pmatrix} \end{pmatrix}.$$

Then, the Kronecker product $\Pi(\mathbf{A},\mathbf{B})$ *is*

$$\begin{pmatrix}
2 & -8 & 8 & -18 & -45 & -37 & 20 & -5 & 6 & 4 & 8 & 14 & 10 & 46 & 19 & -6 \\
-8 & -18 & 2 & 8 & -5 & 5 & 13 & 60 & -8 & 14 & 6 & -4 & -24 & 16 & -34 & -25 \\
8 & 2 & 18 & 8 & -20 & 45 & 35 & -13 & -4 & 6 & -14 & 8 & 41 & 4 & -30 & 4 \\
-18 & 8 & 8 & -2 & -37 & 20 & -45 & 5 & 14 & 8 & -4 & -6 & 16 & -15 & 14 & -44 \\
-2 & -13 & 4 & -18 & -19 & 34 & 5 & -15 & 5 & 6 & 13 & 11 & 32 & -7 & -10 & -6 \\
-13 & -14 & 2 & 12 & 10 & -5 & 39 & -11 & -6 & 15 & 3 & -9 & 5 & 16 & -12 & 28 \\
4 & 2 & 22 & 3 & 35 & 15 & -11 & -14 & 1 & 9 & -13 & 10 & -4 & -30 & -2 & 17 \\
-18 & 12 & 3 & -6 & 9 & 19 & 10 & 35 & 17 & 3 & -2 & -7 & -12 & -2 & -31 & -10 \\
-6 & -4 & 12 & -2 & -37 & 11 & 4 & 13 & -2 & -2 & -14 & -10 & 14 & -45 & -18 & 1 \\
-12 & -2 & -6 & 4 & 7 & 1 & -29 & 28 & 14 & -10 & -2 & 2 & 19 & -12 & 45 & 4 \\
4 & -6 & 2 & 12 & 16 & 23 & 23 & 19 & 2 & -2 & 10 & -14 & -42 & -19 & 14 & -15 \\
-2 & 12 & 4 & 6 & 1 & 32 & -17 & -19 & -10 & -14 & 2 & 2 & -15 & -4 & 1 & 48 \\
-10 & -3 & 10 & -4 & 11 & 22 & 13 & 1 & 1 & -4 & -17 & -6 & -33 & -8 & 3 & 4 \\
-11 & 2 & -8 & 6 & -2 & -11 & 19 & 17 & 12 & -13 & 2 & 5 & -2 & -9 & -2 & -33 \\
2 & -4 & 6 & 13 & 19 & -13 & 7 & -14 & -1 & -6 & 7 & -16 & -9 & 32 & -3 & -8 \\
0 & 14 & 5 & 2 & 17 & -1 & -14 & 17 & -14 & -11 & 0 & 5 & 2 & 3 & 34 & -3
\end{pmatrix}$$

and, by applying MATLAB, the numerical solution of $\Pi(\mathbf{A},\mathbf{B})col(\mathbf{X}) = col(\mathbf{C})$ *is* $col(\mathbf{X})$*, where* $\mathbf{X}$ *is given above.*

References

[1] S. BANACH, Théorie des opérateurs linéaires, Warschau, 1932.

[2] L. I. ARAMANOVITSCH, *Quaternion Non-linear Filter for Estimation of Rotating body Attitude*, Mathematical Meth. in the Appl. Sciences, **18** (1995), 1239-1255.

[3] J. FAN, *Determinants and multiplicative functionals on quaternionic matrices*, Linear Algebra Appl. **369** (2003), 193–201.

[4] K. GÜRLEBECK & W. SPRÖSSIG, Quaternionic and Clifford Calculus for Physicists and Engineers, Wiley, Chichester, 1997, 371 p.

[5] R. A. HORN & C. R. JOHNSON, Matrix Analysis, Cambridge University Press, Cambridge, 1992, 561p.

[6] R. A. HORN & C. R. JOHNSON, Topics in Matrix Analysis, Cambridge University Press, Cambridge, 1991, 607p.

[7] L. HUANG, *The matrix equation $AXB - GXD = E$ over the quaternion field*, Linear Algebra Appl. **234** (1996), 197–208.

[8] D. JANOVSKÁ & G. OPFER, *Linear equations in quaternions*, Numerical Mathematics and Advanced Applications, Proceedings of ENUMATH 2005, A. B. Castro, D. Gómez, P. Quintela, P. Saldago (eds), Springer, Berlin, Heidelberg, New York,2006, 945-953, ISBN 3-540-34287-7.

[9] D. JANOVSKÁ & G. OPFER, *Computing quaternionic roots by Newton's method*, Electron. Trans. Numer. Anal., **26** (2007), 82–102.
http://etna.math.kent.edu/vol.26.2007/pp82-102.dir/pp82-102.html

[10] D. JANOVSKÁ & G. OPFER, *On one linear equation in one quaternionic unknown*, Hamburger Beiträge zur Angewandten Mathematik, Nr. 2007-14, September 2007, 34 p. Dedicated to Bernd Fischer on the occasion of his 50th birthday.
http://www.math.uni-hamburg.de/research/papers/hbam/hbam2007-14.pdf.

[11] R. E. JOHNSON, *On the equation $\chi\alpha = \gamma\chi + \beta$ over an algebraic division ring*. Bull. Amer. Math. Soc., **50** (1944), 202–207.

[12] O. ORE, *Linear equations in non-commutative fields*, Ann. of Math., 2nd series, **32** (1931), 463–477.

[13] B. L. VAN DER WAERDEN, Algebra, 5. Aufl., Springer, Berlin, Göttingen, Heidelberg, 1960, 292 p.

[14] F. ZHANG, *Quaternions and matrices of quaternions*, Linear Algebra Appl. **251** (1997), 21–57.

Eingegangen am 15. 8. 2008

Drahoslava Janovská
Institute of Chemical Technology
Prague, Department of Mathematics
Technická 5
166 28 Prague 6, Czech Republic
janovskd@vscht.cz

Gerhard Opfer
University of Hamburg
Faculty for Mathematics, Informatics, and Natural Sciences [MIN], Area OA
Bundestraße 55, 20146 Hamburg, Germany
opfer@math.uni-hamburg.de