Hamburger Beiträge
zur Angewandten Mathematik

On one linear equation in one quaternionic unknown

Dedicated to Bernd Fischer
on the occasion of his 50th birthday

Drahoslava Janovská and Gerhard Opfer

The paper has been submitted to Linear Algebra Appl.
The final form may differ from this preprint.

Nr. 2007-14
September 2007

ON ONE LINEAR EQUATION IN ONE QUATERNIONIC UNKNOWN

DRAHOSLAVA JANOVSKÁ* AND GERHARD OPFER†

Dedicated to Bernd Fischer on the occasion of his 50th birthday

Abstract. We study quaternionic linear equations of type $\lambda_m(x) := \sum_{j=1}^m b_j x c_j = e$ with quaternionic constants b_j, c_j, e and arbitrary positive integer m. For $m = 2$ the resulting equation is called *Sylvester's equation*. For this case a complete solution (solution formula, determination of null space) will be given. For the general case we show that the solution can be found by a corresponding matrix equation of a particular simple form. This matrix form is connected with the centralizers of a quaternion and of its isomorphic image in $\mathbb{R}^{4\times 4}$. We present a complete determination of these centralizers. However, the mentioned matrix form does not inlude a detection of the singular cases. The determination of singular cases is to some extent possible by applying Banach's fixed point theorem from which we are able to deduce several sufficient conditions for non singular cases. We end the paper with a conjecture on the form of the inverse of a linear mapping and show that interpolation problems and recovery problems have in general no solution.

Key words. One linear equation in quaternions, Sylvester's equation in quaternions, Centralizers of quaternions.

AMS subject classifications. 11R52, 12E15, 12Y05, 65F40

1. Introduction. Linear mappings in one real or one complex variable are not of much interest from an algebraic point of view. The situation changes if we go to quaternionic linear mappings. In this area non trivial non singular linear mappings exist and we have to cope with the difficulty that it is usually impossible to apriorily distinguish between singular and non singular mappings.

Linear quaternionic mappings may be understood as the simplest form of systems of linear equations in quaternions. Such systems were investigated already by [11, ORE, 1931] with references back to papers of the 19th century. However, all treated equations are of the form $a_{11}x_1 + a_{12}x_2 + \cdots$, one coefficient on the left side of the unknowns. In non-commutative algebra also other forms exist, e. g. $ax + xd$ or $ax + bxc + xd$, and more generally $\sum_{j=1}^m b_j x c_j$. A first approach of linear systems including this general type of equation was made by the authors, [8]. However, a thorough investigation of one linear equation in one quaternionic variable is still missing. There is one exception, a paper by [10, R. E. JOHNSON, 1944] in which an equation of type $ax + xd = e$ was investigated over an algebraic division ring. The same topic with quaternionic matrices was treated by [5, HUANG, 1996].

2. Quaternionic linear mappings in one variable. We denote the (skew) field of quaternions by $\mathbb{H}$ and the field of real, complex numbers by $\mathbb{R}, \mathbb{C}$, respectively. The zero element of all three fields will be denoted by 0 and the multiplicative unit by 1. In $\mathbb{H}$ we will also use the notations $\mathbf{i} := (0,1,0,0)$, $\mathbf{j} := (0,0,1,0)$, $\mathbf{k} := (0,0,0,1)$. And $\mathbf{i}$ will be used in $\mathbb{C}$ with the ordinary meaning. We shall also use

$$\mathscr{N}_1 := 1; \quad \mathscr{N}_2 := \mathbf{i}; \quad \mathscr{N}_3 := \mathbf{j}; \quad \mathscr{N}_4 := \mathbf{k}. \tag{2.1}$$

The objective of this paper is to study *quaternionic linear mappings* $\lambda : \mathbb{H} \to \mathbb{H}$ over $\mathbb{R}$ which are defined by the property

$$\lambda(\gamma x + \delta y) = \gamma\lambda(x) + \delta\lambda(y) \text{ for all } x, y \in \mathbb{H} \text{ and all } \gamma, \delta \in \mathbb{R}. \tag{2.2}$$

All linear mappings have the property that $\lambda(0) = 0$. This follows from (2.2) by putting $x = y, \gamma = 1, \delta = -1$.

Let $\lambda(x_0) = 0$ for some specific $x_0 \in \mathbb{H}$. Then, we will call x_0 a *solution of the homogeneous equation*. Since $x_0 = 0$ is always a solution of the homogeneous equation, we call $x_0 = 0$ the *trivial solution* of the homogeneous equation. Let $e \in \mathbb{H}\backslash\{0\}$ be a given quaternion. The equation $\lambda(x) = e$ will be called *inhomogeneous* equation. The linearity has the following consequence: Let $\lambda(x_0) = 0$ and $\lambda(x_1) = e$. Then, $\lambda(\gamma x_0 + x_1) = e$ for all $\gamma \in \mathbb{R}$. For this reason, it is important to study the *null space* (or *kernel*) of λ defined by

$$\mathcal{N} := \{x \in \mathbb{H} : \lambda(x) = 0\}. \tag{2.3}$$

The null space $\mathcal{N}$ is a linear subspace of $\mathbb{H}$, regarded as a space over $\mathbb{R}$. Thus, the dimension of that space, regarded as a subspace of $\mathbb{R}^4$, may vary from zero to four. Linear mappings $\lambda : \mathbb{R} \to \mathbb{R}$ and also $\lambda : \mathbb{C} \to \mathbb{C}$ are easy to describe and $\lambda(x_0) = 0$ either implies that x_0 is the trivial solution or λ is the trivial mapping $\lambda(x) = 0$ for all x. Thus, linear mappings in one variable on $\mathbb{R}$ or on $\mathbb{C}$ are not of much interest from the algebraic point of view.

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

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

DEFINITION 2.1. A quaternionic linear mapping, $\lambda : \mathbb{H} \to \mathbb{H}$, will be called *non singular* if $\lambda(x) = e$ has a unique solution for all choices of $e \in \mathbb{H}$. The mapping λ will be called *singular* if it is not non singular.

LEMMA 2.2. *The linear mapping λ is singular if and only if the homogeneous equation $\lambda(x) = 0$ has non trivial solutions.*

Proof: Follows directly from the definition, since $x = 0$ is always a solution of the homogeneous equation. □

If we leave the trivial mapping $\lambda(x) = 0$ for all x aside, it is not at all obvious that singular linear mappings exist, in particular, if we compare with the non trivial linear mappings $\mathbb{R} \to \mathbb{R}$ and $\mathbb{C} \to \mathbb{C}$, which are always non singular. Let $\lambda, \mu : \mathbb{H} \to \mathbb{H}$ be two linear mappings. We can define the *composition* $\lambda \circ \mu$ by

$$(\lambda \circ \mu)(x) := \lambda(\mu(x)).$$

A composition of two linear mappings is again a linear mapping.

LEMMA 2.3. *Let $\lambda, \mu : \mathbb{H} \to \mathbb{H}$ be two linear mappings. (i) Let λ be non singular. Then, the composition $\lambda \circ \mu$ is singular if and only if μ is singular. (ii) Let μ be non singular. Then, the composition $\lambda \circ \mu$ is singular if and only if λ is singular.*

Proof: (i) Let μ be singular. Thus, there exist $x_0 \neq 0$ such that $\mu(x_0) = 0$. Then, $(\lambda \circ \mu)(x_0) = 0$, and hence, the composition $\lambda \circ \mu$ is singular. Let the composition $\lambda \circ \mu$ be singular, i. e. there exist $x_0 \neq 0$ such that $(\lambda \circ \mu)(x_0) = \lambda(\mu(x_0)) = 0$. Since λ is non singular it follows that $\mu(x_0) = 0$, thus, μ is singular. (ii) This part is similar. □

A linear mapping $\lambda : \mathbb{H} \to \mathbb{H}$ also defines a linear mapping $\tilde{\lambda} : \mathbb{R}^4 \to \mathbb{R}^4$ which can be defined by a matrix.

THEOREM 2.4. *Let $\lambda : \mathbb{H} \to \mathbb{H}$ be a quaternionic linear mapping. Then, there is a matrix $\mathbf{M} \in \mathbb{R}^{4\times 4}$ such that $\lambda(x) = \mathbf{Mx}$ where the quaternions $x, \lambda(x)$ have to be identified with the column vectors $\mathbf{x}, \mathbf{Mx} \in \mathbb{R}^4$, respectively, and where*

$$\mathbf{M} := (\, \lambda(\mathcal{N}_1), \lambda(\mathcal{N}_2), \lambda(\mathcal{N}_3), \lambda(\mathcal{N}_4) \,)\,. \tag{2.4}$$

The entries $\lambda(\mathcal{N}_j),\ j = 1, 2, 3, 4$ must be read as column vectors in $\mathbb{R}^4$.

Proof: Let $x := (x_1, x_2, x_3, x_4) \in \mathbb{H}$ and put $x = \sum_{j=1}^4 x_j \mathcal{N}_j$. Then, using the linearity, $\lambda(x) = \sum_{j=1}^4 x_j \lambda(\mathcal{N}_j)$. And this expression equals $\mathbf{Mx}$. □

This theorem is of importance because it allows us in a concrete case to reduce the given linear equation to a linear matrix equation in four variables for which solution techniques are known. This will be the topic of the next section.

LEMMA 2.5. *The linear mapping λ is singular if and only if $\det \mathbf{M} = 0$, where $\mathbf{M}$ is defined in (2.4)*

Proof: Follows from Definition 2.1 and Theorem 2.4. □

Let $a_1, a_2, a_3, a_4 \in \mathbb{H}$ be any collection of four quaternions and define

$$\lambda(x) := \sum_{j=1}^{4} x_j a_j, \text{ where } x := (x_1, x_2, x_3, x_4) \in \mathbb{H}.$$

Then, λ is a quaternionic linear mapping and the above matrix is $\mathbf{M} = (\, a_1, a_2, a_3, a_4 \,)$ with columns $a_j, j = 1, 2, 3, 4$ and thus, $\mathbf{M}$ could be any matrix in $\mathbb{R}^{4\times 4}$. However, this linear mapping has no direct relation to quaternionic algebra and we do not want to consider this type of mapping in our investigation.

3. Quaternionic linear mappings of type $\lambda(\mathbf{x}) := \sum_{\mathbf{j=1}}^{\mathbf{m}} \mathbf{b_j x c_j}$. We are interested in studying the linear mappings of type

$$\lambda_m(x) := \sum_{j=1}^{m} b_j x c_j, \tag{3.1}$$

defined for any fixed positive integer m and for $2m$ quaternionic constants$b_j, c_j \in \mathbb{H},\ j = 1, 2, \ldots, m$. The function λ_m will loose its property of linearity if in (2.2) the real multipliers γ, δ are replaced with non real multipliers. Each term $b_j x c_j$ in the definition (3.1) will be called *middle term* if neither b_j nor c_j is real. It is reasonable to assume that $b_j c_j \neq 0$ for all $j = 1, 2, \ldots, m$. The particular difficulty resulting from the existence of middle terms arises from the fact that they are not linear with respect to the defining constants b_j, c_j. It is not difficult to find the corresponding matrix representation of λ_m by applying Theorem 2.4. For this purpose we first repeat the multiplication rule for quaternions, since it is used very frequently. Let $a, b \in \mathbb{H}$ and $a := (a_1, a_2, a_3, a_4),\ b := (b_1, b_2, b_3, b_4)$. Then

$$\begin{aligned} ab = (&a_1b_1 - a_2b_2 - a_3b_3 - a_4b_4, a_1b_2 + a_2b_1 + a_3b_4 - a_4b_3, \\ &a_1b_3 - a_2b_4 + a_3b_1 + a_4b_2, a_1b_4 + a_2b_3 - a_3b_2 + a_4b_1). \end{aligned} \tag{3.2}$$

Second, we introduce two mappings $\imath_1 : \mathbb{H} \to \mathbb{R}^{4\times4},\ \imath_2 : \mathbb{H} \to \mathbb{R}^{4\times4}$, and put $\mathbf{A} := \imath_1(a),\ \tilde{\mathbf{A}} := \imath_2(a)$, where

$$\mathbf{A} := (\, a\mathbb{N}_1, a\mathbb{N}_2, a\mathbb{N}_3, a\mathbb{N}_4 \,) = \begin{pmatrix} a_1 & -a_2 & -a_3 & -a_4 \\ a_2 & a_1 & -a_4 & a_3 \\ a_3 & a_4 & a_1 & -a_2 \\ a_4 & -a_3 & a_2 & a_1 \end{pmatrix} \in \mathbb{R}^{4\times4}, \tag{3.3}$$

$$\tilde{\mathbf{A}} := (\, \mathbb{N}_1 a, \mathbb{N}_2 a, \mathbb{N}_3 a, \mathbb{N}_4 a \,) = \begin{pmatrix} a_1 & -a_2 & -a_3 & -a_4 \\ a_2 & a_1 & a_4 & -a_3 \\ a_3 & -a_4 & a_1 & a_2 \\ a_4 & a_3 & -a_2 & a_1 \end{pmatrix} \in \mathbb{R}^{4\times4}. \tag{3.4}$$

In the above definition the quantities $a\mathbb{N}_j, \mathbb{N}_j a,\ j = 1, 2, 3, 4$ have to be read as column vectors. We will denote all matrices of the form given in (3.3) by $\mathbb{H}_{\mathbb{R}}$ and those of (3.4) by $\mathbb{H}_{\mathbb{P}}$. The first mapping $\imath_1$ is a well known isomorphism between $\mathbb{H}$ and $\mathbb{H}_{\mathbb{R}}$, the second mapping $\imath_2$ could be called *pseudo isomorphism* between $\mathbb{H}$ and $\mathbb{H}_{\mathbb{P}}$ because it reverses the order of the multiplication. We also call the elements of $\mathbb{H}_{\mathbb{P}}$ *pseudo quaternions*. More details are in the following theorem.

THEOREM 3.1. *Let $\imath_1, \imath_2$ be defined as in (3.3),(3.4), respectively. Let $a, b \in \mathbb{H}$ and let $\mathbf{b} \in \mathbb{R}^4$ be the column vector corresponding to b. Then,*

$$\imath_1(a)\imath_2(b) = \imath_2(b)\imath_1(a) \text{ for all } a, b \in \mathbb{H}, \tag{3.5}$$

$$\imath_2(ab) = \imath_2(b)\imath_2(a) \in \mathbb{H}_{\mathbb{P}} \text{ for all } a, b \in \mathbb{H}, \tag{3.6}$$

$$\imath_2(a)\,\mathbf{b} = ba \in \mathbb{H} \text{ for all } a, b \in \mathbb{H}, \tag{3.7}$$

$$\imath_2(a^{-1}) = (\imath_2(a))^{-1} = (\imath_2(a))^{\mathrm{T}}/|a|^2 \in \mathbb{H}_{\mathbb{P}} \text{ for all } a \in \mathbb{H}\backslash\{0\}, \tag{3.8}$$

$$\imath_2(a) = \mathbf{0} \Leftrightarrow \imath_2(a) \text{ is singular } \Leftrightarrow a = 0 \text{ where } \mathbf{0} = \text{zero matrix}, \tag{3.9}$$

$$\imath_1(a) = \imath_2(a) \Leftrightarrow a \in \mathbb{R}, \tag{3.10}$$

where in (3.7) one has to read the right hand side ba as column vector.

Proof: By evaluation and comparison. □

The pseudo quaternions behave almost like quaternions. We only have to change the multiplication rule according to (3.6). The elements of $\mathbb{H}_{\mathbb{P}}$ commute with all elements in $\mathbb{H}_{\mathbb{R}}$, however, the product $\imath_1(a)\imath_2(b)$ is in general neither in $\mathbb{H}_{\mathbb{P}}$ nor in $\mathbb{H}_{\mathbb{R}}$. Nevertheless, the first column of $\imath_1(a)\imath_2(b)$ contains the product ab. We could define a new algebra $\tilde{\mathbb{H}} := \mathbb{R}^4$ by keeping the addition rules of $\mathbb{H}$ and by introducing a new multiplication rule in $\tilde{\mathbb{H}}$, denoted by $a \star b$, namely

$$a \star b = ba,$$

where ba is the ordinary product in $\mathbb{H}$. In this new algebra, $\imath_2$ is an isomorphism between $\tilde{\mathbb{H}}$ and $\mathbb{H}_{\mathbb{P}}$, since $\imath_2(a \star b) = \imath_2(ba) = \imath_2(a)\imath_2(b)$, using (3.6) from Theorem 3.1. It should be observed that $\tilde{\mathbf{A}}$ is not a real polynomial in $\mathbf{A}$, since all such polynomials will remain in $\mathbb{H}_{\mathbb{R}}$. There is another well known isomorphism between $\mathbb{H}$ and certain complex (2×2)-matrices which will not be used, however.

From the above representation it is clear how to recover a quaternion from the corresponding matrix. Thus, it is also possible to introduce inverse mappings $\imath_1^{-1} : \mathbb{H}_{\mathbb{R}} \to \mathbb{H},\ \imath_2^{-1} : \mathbb{H}_{\mathbb{P}} \to \mathbb{H}$. Since $\imath_1$ defines an isomorphism between quaternions and certain real (4×4) matrices it is possible to associate notions known from matrix theory with quaternions. Let $a := (a_1, a_2, a_3, a_4)$, then:

$$\det(a) := \det(\imath_1(a)) = |a|^4, \tag{3.11}$$

$$\mathrm{tr}(a) := \mathrm{tr}(\imath_1(a)) = 4a_1, \tag{3.12}$$

$$\mathrm{eig}(a) := \mathrm{eig}(\imath_1(a)) = [\sigma_+, \sigma_+, \sigma_-, \sigma_-], \text{ where} \tag{3.13}$$

$$\sigma_+ = a_1 + \sqrt{a_2^2 + a_3^2 + a_4^2}\,\mathbf{i}, \quad \sigma_- = \overline{\sigma_+}. \tag{3.14}$$

$$|a| = ||\imath_1(a)||_2 = \sqrt{\sum_{j=1}^{4} a_j^2}, \tag{3.15}$$

$$\mathrm{cond}(a) := \mathrm{cond}(\imath_1(a)) = 1, \quad a \neq 0, \tag{3.16}$$

where $\det, \mathrm{tr}, \mathrm{eig}, \mathrm{cond}$ refer to *determinant, trace, collection of eigenvalues* (in square brackets [ ]), *condition*, respectively. It should be noted however, that a general theory of determinants for quaternionic valued square matrices cannot exist. See [2, FAN, 2003]. Let $\lambda : \mathbb{H} \to \mathbb{H}$ be any linear mapping in the sense of (2.2) and let $y := \lambda(x)$. Since $\mathbb{H}$ is isomorphic to $\mathbb{H}_\mathbb{R}$ the mapping λ induces a linear mapping $\Lambda : \mathbb{H}_\mathbb{R} \to \mathbb{H}_\mathbb{R}$ defined by the following diagram where we have put $X := \imath_1(x), Y := \imath_1(y)$:

$$\begin{array}{ccc} x & \stackrel{\lambda}{\longrightarrow} & y \\ \downarrow \imath_1 & & \downarrow \imath_1 \\ X & \stackrel{\Lambda}{\longrightarrow} & Y \end{array}. \tag{3.17}$$

Now, let us return to the mapping λ_m defined in (3.1). We first only determine the matrix $\mathbf{M}$ which defines the linear mapping for $m = 1$. Thus, we investigate only the mapping $\lambda_1(x) := bxc$ and determine $\mathbf{M}$ such that $\lambda_1(x) = \mathbf{M}\mathbf{x}$ where the column vector $\mathbf{x}$ and the matrix product $\mathbf{M}\mathbf{x}$ must be identified with the corresponding quaternion. The result is put into the following lemma which also shows the connection to $\imath_2$.

LEMMA 3.2. *Let $x, b, c \in \mathbb{H}$ and let $\imath_1, \imath_2$ be defined as in (3.3),(3.4), respectively. Define the linear mapping $\lambda_1 : \mathbb{H} \to \mathbb{H}$ by*

$$\lambda_1(x) := bxc$$

and identify the quaternions $x, y := \lambda_1(x)$ with the corresponding column vectors $\mathbf{x}, \mathbf{y}$, respectively. Also put $\imath_1(b) := \mathbf{B}, \imath_2(c) := \tilde{\mathbf{C}}$. Then,

$$\mathbf{y} := \lambda_1(x) = \mathbf{M}\mathbf{x} \text{ where } \mathbf{M} := \imath_1(b)\imath_2(c) = \mathbf{B}\tilde{\mathbf{C}}. \tag{3.18}$$

Proof: Using (3.7) of Theorem 3.1 we have $z := \imath_2(c)\,\mathbf{x} = xc$. Since $bz = bxc$, the proof is complete. □

It should be noted that the above Lemma has an analogue in matrix mappings $\mathbf{AXB} \to \mathbf{C}$ where Kronecker's product is employed. See, [4, Lemma 4.3.1].

THEOREM 3.3. *Let λ_m be defined as in (3.1) and let $\imath_1, \imath_2$ be defined as in (3.3),(3.4), respectively. Then, there is a matrix $\mathbf{M} \in \mathbb{R}^{4\times 4}$ such that*

$$\lambda_m(x) = \mathbf{M}\mathbf{x}, \text{ where } \mathbf{M} := \sum_{j=1}^{m} \mathbf{M}_j, \text{ and } \mathbf{M}_j := \imath_1(b_j)\imath_2(c_j), \tag{3.19}$$

where the quaternions $x, \lambda_m(x)$ have to be identified with the corresponding column vectors $\mathbf{x}, \mathbf{M}\mathbf{x}$, respectively.

Proof: Follows from Lemma 3.2, formula (3.18). □

This theorem is the concrete form of Theorem 2.4. It allows us to solve all equations of the form $\lambda_m(x) = e$ by applying matrix techniques. However, it does not include information on the question whether we are dealing with singular or non singular cases. It should be mentioned, that the above matrix $\mathbf{M}$ is in general not normal.

For later use we introduce a new notion, namely that of equivalence between two quaternions. In this connection, algebraists (see [12, V. D. WAERDEN, 1960, p. 35]) usually use the term *conjugate*, which, however, for quaternions is not a good choice.

DEFINITION 3.4. Two quaternions $a, b \in \mathbb{H}$ will be called *equivalent*, if there is an $h \in \mathbb{H}\backslash\{0\}$ such that $b = h^{-1}ah$ (or $hb - ah = 0$). Equivalent quaternions a, b will be denoted by $a \sim b$. The set

$$[a] := \{s : s := h^{-1}ah, \; h \in \mathbb{H}\}$$

will be called *equivalence class of* a. Let $a = (a_1, a_2, a_3, a_4)$. The complex number

$$\sigma_+ := a_1 + \sqrt{a_2^2 + a_3^2 + a_4^2}\,\mathbf{i} \in [a].$$

will be called *complex representative of* $[a]$.

LEMMA 3.5. *The above defined notion of equivalence defines an equivalence relation. Two quaternions a, b are equivalent if and only if*

$$\Re a = \Re b, \quad |a| = |b|. \tag{3.20}$$

Proof: [6, JANOVSKÁ & OPFER, 2003]. □

It turns out, that the set of all linear mappings λ_m separates into two classes. One class consists of all mappings λ_m with $m \leq 2$ for which a complete answer to all reasonable questions can be found. All other λ_m, namely those with
$m \geq 3$ belong to a class for which we can gather only incomplete information. We introduce some formal simplification. Let $\mu(x) := b_m^{-1} x c_1^{-1}$. This linear mapping is non singular and thus, (put $\tilde{b}_j = b_m^{-1} b_j$, $\tilde{c}_j = c_j c_1^{-1}$, $j = 2, 3, \ldots, m-1$)

$$(\mu \circ \lambda_m)(x) = b_m^{-1}(\sum_{j=1}^{m} b_j x c_j) c_1^{-1} = \begin{cases} b_1 x + \sum_{j=2}^{m-1} \tilde{b}_j x \tilde{c}_j + x c_m & \text{for } m \geq 2, \\ x & \text{for } m = 1, \end{cases}$$

will be singular if and only if λ_m is singular. See Lemma 2.3. Without loss of generality we can, therefore, investigate

$$\lambda_m(x) := ax + \sum_{j=1}^{m-2} b_j x c_j + xd, \quad m \geq 2, \tag{3.21}$$

which reduces to

$$\lambda_2(x) := ax + xd \quad \text{for } m = 2, \text{ and for } m = 3 \text{ we write simply} \tag{3.22}$$

$$\lambda_3(x) := ax + bxc + xd. \tag{3.23}$$

These two equations will be treated in the sequel. The investigation of the mapping λ_3 already shows all difficulties of λ_m for $m \geq 3$.

4. Sylvester's equation. The equation defined by (3.22), namely $\lambda_2(x) := ax + xd = e$ is ordinarily referred to as *Sylvester's equation* and λ_2 will be called correspondingly *Sylvester's function* or *mapping*. We will show that Sylvester's equation has an elementary solution in terms of quaternions.

THEOREM 4.1. *Let $a := (a_1, a_2, a_3, a_4)$, $d := (d_1, d_2, d_3, d_4)$. The function defined by $\lambda_2(x) := ax + xd$ is non singular if and only if $\sum_{j=2}^{4}(a_j^2 - d_j^2) \neq 0$ or $a_1 + d_1 \neq 0$. If λ_2 is non singular, the solution of $\lambda_2(x) = e$ is*

$$x = f_l^{-1}(e + a^{-1} e \overline{d}), \quad f_l := 2\Re d + a + |d|^2 a^{-1} \quad \text{if } a \neq 0, \text{ or} \tag{4.1}$$

$$x = (e + \overline{a} e d^{-1}) f_r^{-1}, \quad f_r := 2\Re a + d + |a|^2 d^{-1} \quad \text{if } d \neq 0. \tag{4.2}$$

Proof: The mapping λ_2 is described by a matrix $\mathbf{M}$ whose form is given in Theorem 3.3:

$$\mathbf{M} := \imath_1(a) + \imath_2(d) := \begin{pmatrix} \sigma_1 & -\sigma_2 & -\sigma_3 & -\sigma_4 \\ \sigma_2 & \sigma_1 & -\delta_4 & \delta_3 \\ \sigma_3 & \delta_4 & \sigma_1 & -\delta_2 \\ \sigma_4 & -\delta_3 & \delta_2 & \sigma_1 \end{pmatrix}, \tag{4.3}$$

where $\sigma := (\sigma_1, \sigma_2, \sigma_3, \sigma_4) := a + d$, $\delta := (\delta_1, \delta_2, \delta_3, \delta_4) := a - d$ and where $\imath_1, \imath_2$ are defined in (3.3), (3.4). The determinant of this matrix is (see [8])

$$\det(\mathbf{M}) = \sigma_1^2(|\sigma|^2 + \delta_2^2 + \delta_3^2 + \delta_4^2) + (\sigma_2\delta_2 + \sigma_3\delta_3 + \sigma_4\delta_4)^2. \tag{4.4}$$

It does not vanish if and only if the above mentioned conditions are met. In order to find the solution, let $\mu(x) := a^{-1} x \overline{d} + x$. The composite mapping $\mu \circ \lambda_2$ is $(\mu \circ \lambda_2)(x) = \mu(\lambda_2(x)) = a^{-1}\lambda_2(x)\overline{d} + \lambda_2(x) =$

$(2\Re d + a + |d|^2 a^{-1})x$. This yields (4.1). The same technique with $\mu(x) := \overline{a}xd^{-1} + x$ yields (4.2). And the mapping μ is non singular if and only if the mapping λ_2 is non singular. Only a little later, we will see that the two employed μs for composition are apart from a factor forms of the inverse mapping of λ_2. See also [8, JANOVSKÁ & OPFER, Lemma 2]. □

COROLLARY 4.2. *Let a, d be arbitrary quaternions. Then, Sylvester's equation $\lambda_2(x) := ax + xd$ will be singular if and only if*

$$|a| = |d| \text{ and } \Re a + \Re d = 0, \tag{4.5}$$

or in other words if and only if a and $-d$ are equivalent. The null space of λ_2 is

$$\mathcal{N}_2 = \begin{cases} \{0\} & \text{if (4.5) is not valid,} \\ \mathbb{H} & \text{if (4.5) is valid and } a, d \in \mathbb{R}, \\ \text{2 dim, real subspace of } \mathbb{H} & \text{if (4.5) is valid and } a \notin \mathbb{R} \text{ or } d \notin \mathbb{R}. \end{cases} \tag{4.6}$$

Proof: Let $a := (a_1, a_2, a_3, a_4), d := (d_1, d_2, d_3, d_4)$. From Theorem 4.1, it follows that λ_2 is singular if and only of $a_1 + d_1 = 0$ and $\sum_{j=2}^4 (a_j^2 - d_j^2) = 0$. The first condition implies $a_1^2 = d_1^2$, therefore, $\sum_{j=1}^4 (a_j^2 - d_j^2) = 0$, which is equivalent to $|a|^2 = |d|^2$. An alternative proof is: Let $ax + xd = 0$ for some $x \neq 0$. Then, according to Definition 3.4 the two quantities $a, -d$ are equivalent and the result follows from Lemma 3.5. In order to find the null space, we have to investigate the matrix $\mathbf{M}$ defined in (4.3), only with $\sigma_1 = 0$. Becausse of $\mathbf{M} + \mathbf{M}^{\mathrm{T}} = \mathbf{0}$, the matrix can have only even rank and the ranks zero or four are not possible. See also [6]. □

It seams reasonable also to compute the null space (kernel) of $\lambda_2 := ax + xd$ explicitly. For the definition see (2.3).

TABLE 4.3. Various cases for kernel of Sylvester's equation.

Case	σ_2	δ_2	σ_3	δ_3	σ_4	δ_4
(ia)	0	0	$\neq 0$	$\neq 0$	$\neq 0$	$\neq 0$
(ib)	0	0	0	$\neq 0$	0	$\neq 0$
(ic)	0	0	0	$\neq 0$	$\neq 0$	0
(id)	0	0	$\neq 0$	0	0	$\neq 0$
(ie)	0	0	$\neq 0$	0	$\neq 0$	0
(iia)	0	$\neq 0$	$\neq 0$	$\neq 0$	$\neq 0$	$\neq 0$
(iib)	0	$\neq 0$	0	$\neq 0$	0	$\neq 0$
(iic)	0	$\neq 0$	0	$\neq 0$	$\neq 0$	0
(iid)	0	$\neq 0$	$\neq 0$	0	0	$\neq 0$
(iie)	0	$\neq 0$	$\neq 0$	0	$\neq 0$	0
(iiia)	$\neq 0$	0	$\neq 0$	$\neq 0$	$\neq 0$	$\neq 0$
(iiib)	$\neq 0$	0	0	$\neq 0$	0	$\neq 0$
(iiic)	$\neq 0$	0	0	$\neq 0$	$\neq 0$	0
(iiid)	$\neq 0$	0	$\neq 0$	0	0	$\neq 0$
(iiie)	$\neq 0$	0	$\neq 0$	0	$\neq 0$	0

THEOREM 4.4. *Let $a, d \in \mathbb{H}$ be given such that $\lambda_2(x) := ax + xd$ is singular and that at least one of the two quaternions is not real. Put*

$$\sigma := a + d, \quad \delta := a - d \quad \text{with components} \quad \sigma_j,\ \delta_j,\ j = 1, 2, 3, 4.$$

(i) Let the following matrix $\mathbf{m}$ be well defined and non singular:

$$\mathbf{m} := \frac{1}{\sigma_2} \begin{pmatrix} -\delta_4 & \delta_3 \\ \sigma_3 & \sigma_4 \end{pmatrix}. \tag{4.7}$$

Then,

$$\mathbf{m}^{-1} := \frac{-1}{\delta_2} \begin{pmatrix} -\sigma_4 & \delta_3 \\ \sigma_3 & \delta_4 \end{pmatrix}, \tag{4.8}$$

and the null space (kernel) of λ_2 is

$$\mathcal{N}_2 := \mathrm{kernel}(\lambda_2) := \left\{ \mathbf{x} = \begin{pmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{pmatrix} : \mathbf{x}_1 + \mathbf{m}\mathbf{x}_2 = \mathbf{0} \text{ or } \mathbf{m}^{-1}\mathbf{x}_1 + \mathbf{x}_2 = \mathbf{0} \right\}. \tag{4.9}$$

TABLE 4.5. The kernel of Sylvester's equation for various cases.

Case	x_1	x_2	x_3	x_4
(ia)	$-\frac{\delta_4}{\sigma_3}x_2$	arbitrary	$-\frac{\sigma_4}{\sigma_3}x_4$	arbitrary
(ib)	arbitrary	0	$\frac{\delta_3}{\delta_4}x_4$	arbitrary
(ic)	$\frac{\delta_3}{\sigma_4}x_2$	arbitrary	arbitrary	0
(id)	$-\frac{\delta_4}{\sigma_3}x_2$	arbitrary	0	arbitrary
(ie)	0	arbitrary	$-\frac{\sigma_4}{\sigma_3}x_4$	arbitrary
(iia)	$\frac{\delta_3x_2-\delta_2x_3}{\sigma_4}$	arbitrary	$-\frac{\sigma_4}{\sigma_3}x_4$	arbitrary
(iib)	arbitrary	$\frac{\delta_2}{\delta_3}x_3$	$\frac{\delta_3}{\delta_4}x_4$	arbitrary
(iic)	$\frac{\delta_3x_2-\delta_2x_3}{\sigma_4}$	arbitrary	arbitrary	0
(iid)	$\frac{-\delta_4x_2+\delta_2x_4}{\sigma_3}$	arbitrary	0	arbitrary
(iie)	$\frac{\delta_2}{\sigma_3}x_4$	arbitrary	$-\frac{\sigma_4}{\sigma_3}x_4$	arbitrary
(iiia)	$-\frac{\delta_4}{\sigma_3}x_2$	arbitrary	$\frac{\sigma_2x_1+\delta_3x_4}{\delta_4}$	arbitrary
(iiib)	$\frac{\delta_4x_3-\delta_3x_4}{\sigma_2}$	0	arbitrary	arbitrary
(iiic)	$\frac{\delta_3}{\sigma_4}x_2$	$-\frac{\sigma_4}{\sigma_2}x_4$	arbitrary	arbitrary
(iiid)	$-\frac{\delta_4}{\sigma_3}x_2$	$-\frac{\sigma_3}{\sigma_2}x_3$	arbitrary	arbitrary
(iiie)	0	arbitrary	$-\frac{\sigma_2x_2+\sigma_4x_4}{\sigma_3}$	arbitrary

(ii) If the matrix $\mathbf{m}$ *is either singular* ($\sigma_3\delta_3+\sigma_4\delta_4=0$) *or not well defined* ($\sigma_2=0$) *we summarize the result in Table 4.5 which refers to various cases which are listed in Table 4.3.*

Proof: In matrix terms we have to solve $\mathbf{Mx}=\mathbf{0}$, where $\mathbf{M}$ is given in (4.3) with the additional property that its determinant is vanishing which by (4.4) is equivalent to $\sigma_1=0$ and $\sigma_2\delta_2+\sigma_3\delta_3+\sigma_4\delta_4=0$. We partition $\mathbf{M}$ as follows:

$$\mathbf{M} := \left(\begin{array}{cc|cc} 0 & -\sigma_2 & -\sigma_3 & -\sigma_4 \\ \sigma_2 & 0 & -\delta_4 & \delta_3 \\ \hline \sigma_3 & \delta_4 & 0 & -\delta_2 \\ \sigma_4 & -\delta_3 & \delta_2 & 0 \end{array}\right) =: \begin{pmatrix} \mathbf{M}_{\mathrm{ul}} & \mathbf{M}_{\mathrm{ur}} \\ \mathbf{M}_{\mathrm{ll}} & \mathbf{M}_{\mathrm{lr}} \end{pmatrix}. \tag{4.10}$$

By putting $\mathbf{x} := \begin{pmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{pmatrix}$ with $\mathbf{x}_1,\mathbf{x}_2\in\mathbb{R}^2$, we can write the system $\mathbf{Mx}=\mathbf{0}$ into the form

$$\begin{aligned} \mathbf{M}_{\mathrm{ul}}\mathbf{x}_1+\mathbf{M}_{\mathrm{ur}}\mathbf{x}_2 &= \mathbf{0}, \\ \mathbf{M}_{\mathrm{ll}}\mathbf{x}_1+\mathbf{M}_{\mathrm{lr}}\mathbf{x}_2 &= \mathbf{0}. \end{aligned} \tag{4.11}$$

We distinguish between the following two cases.

1. Assume that all four submatrices are non singular. Then

$$\mathbf{x}_1+\mathbf{M}_{\mathrm{ul}}^{-1}\mathbf{M}_{\mathrm{ur}}\mathbf{x}_2=\mathbf{0},$$
$$\mathbf{x}_1+\mathbf{M}_{\mathrm{ll}}^{-1}\mathbf{M}_{\mathrm{lr}}\mathbf{x}_2=\mathbf{0}.$$

Since we already know that $\operatorname{rank}\mathbf{M} = 2$, we have

$$\mathbf{m} := \mathbf{M}_{\rm ul}^{-1}\mathbf{M}_{\rm ur} = \mathbf{M}_{\rm ll}^{-1}\mathbf{M}_{\rm lr} = \frac{1}{\sigma_2}\begin{pmatrix} -\delta_4 & \delta_3 \\ \sigma_3 & \sigma_4 \end{pmatrix}.$$

Using the second equation, it is easy to find the inverse of $\mathbf{m}$:

$$\mathbf{m}^{-1} = -\frac{1}{\delta_2}\begin{pmatrix} -\sigma_4 & \delta_3 \\ \sigma_3 & \delta_4 \end{pmatrix}.$$

For the kernel we, thus, have the given formula (4.9).

2. Assume that at least one of the submatrices is singular. Then, necessarily $\sigma_2\delta_2 = \sigma_3\delta_3 + \sigma_4\delta_4 = 0$. If $\sigma_2 = \delta_2 = 0$, then $\mathbf{M}_{\rm ur} \neq \mathbf{0}, \mathbf{M}_{\rm ll} \neq \mathbf{0}$, since the case $\mathbf{M}_{\rm ur} = \mathbf{0} \Leftrightarrow \mathbf{M}_{\rm ll} = \mathbf{0} \Leftrightarrow a, d \in \mathbb{R}$ was excluded. There are 15 cases which we have listed above in Table 4.3. Let us put $\mathbf{x}_1 = (x_1, x_2)^{\rm T}, \mathbf{x}_2 = (x_3, x_4)^{\rm T}$. Then in the cases (ia) to (ie) of Table 4.3 it follows from (4.11) that

$$\mathbf{M}_{\rm ur}\mathbf{x}_2 = \begin{pmatrix} -\sigma_3 x_3 - \sigma_4 x_4 \\ -\delta_4 x_3 + \delta_3 x_4 \end{pmatrix} = \mathbf{0},\ \mathbf{M}_{\rm ll}\mathbf{x}_1 = \begin{pmatrix} \sigma_3 x_1 + \delta_4 x_2 \\ \sigma_4 x_1 - \delta_3 x_2 \end{pmatrix} = \mathbf{0}. \tag{4.12}$$

The cases (iia) to (iie) are governed by the following set of equations

$$\begin{aligned} \mathbf{M}_{\rm ur}\mathbf{x}_2 = \begin{pmatrix} -\sigma_3 x_3 - \sigma_4 x_4 \\ -\delta_4 x_3 + \delta_3 x_4 \end{pmatrix} &= \mathbf{0}, \\ \mathbf{m}^{-1}\mathbf{x}_1 + \mathbf{x}_2 = \frac{-1}{\delta_2}\begin{pmatrix} -\sigma_4 x_1 + \delta_3 x_2 \\ \sigma_3 x_1 + \delta_4 x_2 \end{pmatrix} + \begin{pmatrix} x_3 \\ x_4 \end{pmatrix} &= \mathbf{0}, \end{aligned} \tag{4.13}$$

where it should be observed, that $\mathbf{m}^{-1}$ coincides with the formerly introduced matrix, though it is singular. The cases (iiia) to (iiie) follow

$$\begin{aligned} \mathbf{M}_{\rm ll}\mathbf{x}_1 = \begin{pmatrix} \sigma_3 x_1 + \delta_4 x_2 \\ \sigma_4 x_1 - \delta_3 x_2 \end{pmatrix} &= \mathbf{0}, \\ \mathbf{x}_1 + \mathbf{m}\mathbf{x}_2 = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \frac{1}{\sigma_2}\begin{pmatrix} -\delta_4 x_3 + \delta_3 x_4 \\ \sigma_3 x_3 + \sigma_4 x_4 \end{pmatrix} &= \mathbf{0}. \end{aligned} \tag{4.14}$$

Here, it should also be noted, that $\mathbf{m}$ is singular.

The 15 cases are treated in the sequel one by one. Explicitly occurring coefficients are supposed to be non zero.

$$\begin{pmatrix} -\sigma_3 x_3 - \sigma_4 x_4 \\ -\delta_4 x_3 + \delta_3 x_4 \end{pmatrix} = \mathbf{0}, \quad \begin{pmatrix} \sigma_3 x_1 + \delta_4 x_2 \\ \sigma_4 x_1 - \delta_3 x_2 \end{pmatrix} = \mathbf{0}. \tag{4.15}$$

$$\begin{pmatrix} -0 * x_3 - 0 * x_4 \\ -\delta_4 x_3 + \delta_3 x_4 \end{pmatrix} = \mathbf{0}, \quad \begin{pmatrix} 0 * x_1 + \delta_4 x_2 \\ 0 * x_1 - \delta_3 x_2 \end{pmatrix} = \mathbf{0}. \tag{4.16}$$

$$\begin{pmatrix} -0 * x_3 - \sigma_4 x_4 \\ -0 * x_3 + \delta_3 x_4 \end{pmatrix} = \mathbf{0}, \quad \begin{pmatrix} 0 * x_1 + 0 * x_2 \\ \sigma_4 x_1 - \delta_3 x_2 \end{pmatrix} = \mathbf{0}. \tag{4.17}$$

$$\begin{pmatrix} -\sigma_3 x_3 - 0 * x_4 \\ -\delta_4 x_3 + 0 * x_4 \end{pmatrix} = \mathbf{0}, \quad \begin{pmatrix} \sigma_3 x_1 + \delta_4 x_2 \\ 0 * x_1 - 0 * x_2 \end{pmatrix} = \mathbf{0}. \tag{4.18}$$

$$\begin{pmatrix} -\sigma_3 x_3 - \sigma_4 x_4 \\ -0 * x_3 + 0 * x_4 \end{pmatrix} = \mathbf{0}, \quad \begin{pmatrix} \sigma_3 x_1 + 0 * x_2 \\ \sigma_4 x_1 - 0 * x_2 \end{pmatrix} = \mathbf{0}. \tag{4.19}$$

$$\begin{pmatrix} -\sigma_3 x_3 - \sigma_4 x_4 \\ -\delta_4 x_3 + \delta_3 x_4 \end{pmatrix} = \mathbf{0}, \quad \frac{-1}{\delta_2}\begin{pmatrix} -\sigma_4 x_1 + \delta_3 x_2 \\ \sigma_3 x_1 + \delta_4 x_2 \end{pmatrix} + \begin{pmatrix} x_3 \\ x_4 \end{pmatrix} = \mathbf{0}. \tag{4.20}$$

$$\begin{pmatrix} -0 * x_3 - 0 * x_4 \\ -\delta_4 x_3 + \delta_3 x_4 \end{pmatrix} = \mathbf{0}, \quad \frac{-1}{\delta_2}\begin{pmatrix} -0 * x_1 + \delta_3 x_2 \\ 0 * x_1 + \delta_4 x_2 \end{pmatrix} + \begin{pmatrix} x_3 \\ x_4 \end{pmatrix} = \mathbf{0}. \tag{4.21}$$

$$\begin{pmatrix} -0 * x_3 - \sigma_4 x_4 \\ -0 * x_3 + \delta_3 x_4 \end{pmatrix} = \mathbf{0}, \quad \frac{-1}{\delta_2}\begin{pmatrix} -\sigma_4 x_1 + \delta_3 x_2 \\ 0 * x_1 + 0 * x_2 \end{pmatrix} + \begin{pmatrix} x_3 \\ x_4 \end{pmatrix} = \mathbf{0}. \tag{4.22}$$

$$(4.23)\qquad \begin{pmatrix} -\sigma_3x_3-0*x_4\\ -\delta_4x_3+0*x_4\end{pmatrix}=\mathbf{0},\quad \frac{-1}{\delta_2}\begin{pmatrix} -0*x_1+0*x_2\\ \sigma_3x_1+\delta_4x_2\end{pmatrix}+\begin{pmatrix} x_3\\ x_4\end{pmatrix}=\mathbf{0}.$$

$$(4.24)\qquad \begin{pmatrix} -\sigma_3x_3-\sigma_4x_4\\ -0*x_3+0*x_4\end{pmatrix}=\mathbf{0},\quad \frac{-1}{\delta_2}\begin{pmatrix} -\sigma_4x_1+0*x_2\\ \sigma_3x_1+0*x_2\end{pmatrix}+\begin{pmatrix} x_3\\ x_4\end{pmatrix}=\mathbf{0}.$$

$$(4.25)\qquad \begin{pmatrix} \sigma_3x_1+\delta_4x_2\\ \sigma_4x_1-\delta_3x_2\end{pmatrix}=\mathbf{0},\quad \begin{pmatrix} x_1\\ x_2\end{pmatrix}+\frac{1}{\sigma_2}\begin{pmatrix} -\delta_4x_3+\delta_3x_4\\ \sigma_3x_3+\sigma_4x_4\end{pmatrix}=\mathbf{0}.$$

$$(4.26)\qquad \begin{pmatrix} 0*x_1+\delta_4x_2\\ 0*x_1-\delta_3x_2\end{pmatrix}=\mathbf{0},\quad \begin{pmatrix} x_1\\ x_2\end{pmatrix}+\frac{1}{\sigma_2}\begin{pmatrix} -\delta_4x_3+\delta_3x_4\\ 0*x_3+0*x_4\end{pmatrix}=\mathbf{0}.$$

$$(4.27)\qquad \begin{pmatrix} 0*x_1+0*x_2\\ \sigma_4x_1-\delta_3x_2\end{pmatrix}=\mathbf{0},\quad \begin{pmatrix} x_1\\ x_2\end{pmatrix}+\frac{1}{\sigma_2}\begin{pmatrix} -0*x_3+\delta_3x_4\\ 0*x_3+\sigma_4x_4\end{pmatrix}=\mathbf{0}.$$

$$(4.28)\qquad \begin{pmatrix} \sigma_3x_1+\delta_4x_2\\ 0*x_1-0*x_2\end{pmatrix}=\mathbf{0},\quad \begin{pmatrix} x_1\\ x_2\end{pmatrix}+\frac{1}{\sigma_2}\begin{pmatrix} -\delta_4x_3+0*x_4\\ \sigma_3x_3+0*x_4\end{pmatrix}=\mathbf{0}.$$

$$(4.29)\qquad \begin{pmatrix} \sigma_3x_1+0*x_2\\ \sigma_4x_1-0*x_2\end{pmatrix}=\mathbf{0},\quad \begin{pmatrix} x_1\\ x_2\end{pmatrix}+\frac{1}{\sigma_2}\begin{pmatrix} -0*x_3+0*x_4\\ \sigma_3x_3+\sigma_4x_4\end{pmatrix}=\mathbf{0}.$$

The solutions of these systems are summarized in Table 4.5. □

The two formulas (4.1), (4.2) for x produce the same value if $ad\neq 0$. However, for numerical purposes one would prefer the first formula if $|d|\le|a|$ and otherwise the second formula. These formulas are inversion formulas for the operator λ_2 and we could also write

$$\lambda_2^{-1}(x)=f_l^{-1}(x+a^{-1}x\overline{d})=(x+\overline{a}xd^{-1})f_r^{-1},\quad ad\neq 0.$$

We observe, that λ_2^{-1} is formally different from λ_2. However,

$$af_l\lambda_2^{-1}(x)=ax+x\overline{d},\quad \lambda_2^{-1}(x)f_rd=\overline{a}x+xd.$$

If we have a look at the solution formulas (4.1), (4.2), then it is clear that the corresponding λ_2 is singular if and only if $f_l=0$ or $f_r=0$. This can be shown directly.

LEMMA 4.6. *Let* $a:=(a_1,a_2,a_3,a_4), d:=(d_1,d_2,d_3,d_4)$ *and* $ad\neq 0$. *Sylvester's function* λ_2, *defined by* $\lambda_2(x):=ax+xd$ *is singular if and only if* $f_l=0$ $(f_r=0)$, *where* $f_l:=2\Re d+a+|d|^2a^{-1}$ $(f_r:=2\Re a+d+|a|^2d^{-1})$.

Proof: (a) Let (4.5) be valid. Then, $f_l=-2\Re a+a+|a|^2a^{-1}=-2\Re a+2\Re a=0$. (b) Assume that $f_l=0$. Put $q^2:=|d|^2/|a^2|$. Then, $f_l=2\Re d+a+q^2\overline{a}=(2d_1+(1+q^2)a_1,(1-q^2)a_2,(1-q^2)a_3,(1-q^2)a_4)=(0,0,0,0)$. If $q^2=1$, then, (4.5) follows. Let $q^2\neq 1$. Then, $a_2=a_3=a_4=0$ and $a_1\neq 0$. The remaining equation is $f_l=2d_1+(1+q^2)a_1=\frac{(a_1+d_1)^2+d_2^2+d_3^2+d_4^2}{a_1}=0$. Thus, d is also real and $a_1+d_1=0$, which implies that $q^2=1$, a contradiction. The case $q^2\neq 1$ is impossible. The proof for f_r is almost the same. □

The above solution formulas (4.1), (4.2) put us in the position to solve another type of equation.

COROLLARY 4.7. *Let* A,B,C,D,E *be given quaternions with* $ABCD\neq 0$. *Define the function* $\Lambda_2:\mathbb{H}\to\mathbb{H}$ *by*

$$(4.30)\qquad \Lambda_2(x):=AxB+CxD,\ \textit{and solve}\ \Lambda_2(x)=E,\quad x\in\mathbb{H}.$$

Then, the function Λ_2 *is non singular if and only if*

$$(4.31)\qquad \Re(A^{-1}C)+\Re(BD^{-1})\neq 0\ \textit{or}\ \sum_{j=2}^{4}((A^{-1}C)_j^2-(BD^{-1})_j^2)\neq 0$$

where the subscript j defines the component number. If (4.31) is true, the solution of $\Lambda_2(x) = E$ is

(4.32) $$x = F_l^{-1}(A^{-1}ED^{-1} + C^{-1}E\overline{B}/|D|^2), \text{ where}$$
(4.33) $$F_l := 2\Re(BD^{-1}) + A^{-1}C + |BD^{-1}|^2C^{-1}A \text{ or}$$
(4.34) $$x = (A^{-1}ED^{-1} + \overline{C}EB^{-1}/|A|^2)F_r^{-1}, \text{ where}$$
(4.35) $$F_r := 2\Re(A^{-1}C) + BD^{-1} + |A^{-1}C|^2DB^{-1}.$$

Proof: From $\Lambda_2(x) = E$ it follows by multiplication from the left by A^{-1} and from the right by D^{-1} that

$$xBD^{-1} + A^{-1}Cx = A^{-1}ED^{-1}.$$

Put

$$a := A^{-1}C, \quad d := BD^{-1}, \quad e := A^{-1}ED^{-1},$$

then we have exactly equation (3.22) and Theorem 4.1 yields the desired result. □

The expression for x in formulas (4.32), (4.34) is an expression for the inverse mapping of Λ_2. Therefore, we summarize the result in the following theorem.

THEOREM 4.8. *Let Λ_2 be defined as in (4.30). Then, the inverse of Λ_2, if it exists, has the same form as Λ_2.*
Proof: Compare the form (4.30) of Λ_2 with the solution formulas (4.32), (4.34). □

The function λ_3, defined in (3.23), depends on four quaternions a, b, c, d, i. e. on sixteen unrelated real numbers. If we reduce this number to twelve in a specific way we are able to characterize the non singular cases of λ_3.

LEMMA 4.9. *Let λ_3 be defined by $\lambda_3(x) := ax + bxc + xd$. Put $a = (a_1, a_2, a_3, a_4)$ and analogously for b, c, d. Let $abcd \neq 0$. If (i) $b = a$ and $c \neq -1$ or if (ii) $c = d$ and $b \neq -1$, then λ_3 is singular if and only if*

(4.36) $$r|a| = |d| \text{ and } \Re a + \Re\{d(1+\overline{c})/r^2\} = 0 \text{ where } r := |1+c|, \text{ for (i)},$$
(4.37) $$\rho|d| = |a| \text{ and } \Re d + \Re\{(1+\overline{b})a/|\rho^2\} = 0 \text{ where } \rho := |1+b|, \text{ for (ii)}.$$

If λ_3 is non singular, then the solution of $\lambda_3(x) = e$ is

(4.38) $$x = \begin{cases} f_1^{-1}(e(1+\overline{c}) + a^{-1}e\overline{d}) & \text{for (i)}, \\ f_2^{-1}((1+b)^{-1}e + a^{-1}e\overline{d}) & \text{for (ii)}, \end{cases}$$

where

(4.39) $$\begin{cases} f_1 := 2\Re\{d(1+\overline{c})\} + |1+c|^2a + |d|^2a^{-1} & \text{for (i)}, \\ f_2 := 2\Re d + (1+b)^{-1}a + |d|^2a^{-1}(1+b) & \text{for (ii)}. \end{cases}$$

Proof: (i) We have $\lambda_3(x) := ax + axc + xd = ax(1+c) + xd$. Multiplying from the right by $(1+c)^{-1}$ yields $\lambda_3(x)(1+c)^{-1} = ax + xd(1+c)^{-1}$. This equation has the form of Sylvester's equation and it is according to Corollary 4.2 singular if and only if $|a| = |d(1+c)^{-1}|$ and $\Re a + \Re\{d(1+c)^{-1}\} = 0$. Put $r := |1+c| > 0$, then $(1+c)^{-1} = (1+\overline{c})/r^2$. The given conditions can be written as $\Re a + \Re\{d(1+\overline{c})/r^2\} = a_1 + [(1+c_1)d_1 + c_2d_2 + c_3d_3 + c_4d_4]/r^2 = 0$ and $r|a| = |d|$. (ii) We have $\lambda_3(x) = ax + bxd + xd = ax + (b+1)xd$. Similarly to (i) we obtain $\Re\{(1+b)^{-1}a\} + \Re d = 0$ and $\rho|d| = |a|$, where $\rho := |1+b|$. For the solution we used formula (4.1) from Theorem 4.1. □

How can we find quaternions which lead to singular functions which obey the equations (4.36), (4.37). The answer for (4.36) is:

$$\text{Choose } c, d \in \mathbb{H}\backslash\mathbb{R} \text{ at random, determine } r^2 := |1+c|^2,$$
$$\text{put } a_1 := -\frac{1}{r^2}((1+c_1)d_1 + c_2d_2 + c_3d_3 + c_4d_4),$$
$$\text{choose } a_2, a_3, a_4 \text{ such that } r^2|a|^2 = |d|^2.$$

EXAMPLE 4.10. *The above construction may result in:*

$$\begin{aligned} c &:= (-1,0,-6,3), \quad d := (-4,9,5,-2), \quad r^2 := |1+c|^2 = 45,\\ a_1 &:= 4/5,\\ a &:= (4,3\sqrt{2},3\sqrt{2},3\sqrt{2})/5. \end{aligned}$$

Now check (4.36):

$$\begin{aligned} |d|^2 &= 126, \quad r^2|a|^2 = 45\cdot 14/5 = 126 = |d|^2,\\ r^2 a_1 &+ (1+c_1)d_1 + c_2d_2 + c_3d_3 + c_4d_4 = 45\cdot 4/5 - 36 = 0. \end{aligned}$$

The following relation between some of the numbers given in (4.1), (4.2) will be used later.

LEMMA 4.11. *Let $a,d\in\mathbb{H}\backslash\{0\}$ and define f_l, f_r as in (4.1), (4.2). Then*

$$|a||f_l| = |d||f_r|. \tag{4.40}$$

Proof: We show that $|a|^2|f_l|^2 = |af_l|^2 = (af_l)\overline{(af_l)} = (df_r)\overline{(df_r)}$. We have $(af_l)\overline{(af_l)} = (2(\Re d)a + a^2 + |d|^2)(2(\Re d)\overline{a} + \overline{a}^2 + |d|^2) = 4(\Re d)^2|a|^2 + |a|^4 + |d|^4 + 4\Re a\Re d(|a|^2+|d|^2) + 2|d|^2\Re a^2$. If we switch a and d, then the two sides of (4.40) also switch. Therefore, it is sufficient to show that $|a|^2(4(\Re d)^2 - 2\Re d^2) + |d|^2(2\Re a^2 - 4(\Re a)^2) = 0$. If we employ the multiplication rule (3.2), we obtain $4(\Re d)^2 - 2\Re d^2 = 2|d|^2$ and, correspondingly, $2\Re a^2 - 4(\Re a)^2 = -2|a|^2$. □

Thus, a complete solution of Sylvester's equation has been obtained. An application will be treated in the next section. It should be noted that the solution formulas offered by [10, Theorem 1, formula (10)] and [5, Corollary 6, formula (37)] have both the form of sums and differ from those given here in Theorem 4.1.

5. The centralizers of a and of $\imath_1(\mathbf{a})$. We begin with the definition of the algebraic notion *centralizer*. It makes sense only in non commutative algebras.

DEFINITION 5.1. Let $a\in\mathbb{H}$ be a quaternion. The *centralizer of* a will be denoted by $C(a)$ and is defined by the set

$$C(a) := \{h\in\mathbb{H} : ah - ha = 0\}. \tag{5.1}$$

Let $\mathbf{A}\in\mathbb{R}^{n\times n}$ be a square matrix with real elements. The *centralizer of* $\mathbf{A}$ will be denoted by $C(\mathbf{A})$ and is defined by the set

$$C(\mathbf{A}) := \{\mathbf{H}\in\mathbb{R}^{n\times n} : \mathbf{AH} - \mathbf{HA} = \mathbf{0}\}. \tag{5.2}$$

Both sets are not only vector spaces over $\mathbb{R}$ with finite dimension, but they also form an algebra. In the following we are going to determine $C(a)$ and $C(\imath_1(a))$, where the real (4×4) matrix $\imath_1(a)$ is defined in (3.3). For the general case of matrices we refer to [3, HORN & JOHNSON, Section 3.2] and [4, HORN & JOHNSON, Section 4.4].

If $a\in\mathbb{R}$ (which includes the case $a=0$) we have $C(a)=\mathbb{H}$ and $C(\imath_1(a)) = \mathbb{R}^{4\times 4}$. Thus, this case is of no interest. Therefore, we always assume that $a\in\mathbb{H}\backslash\mathbb{R}$, when we want to determine $C(a)$ and $C(\imath_1(a))$. We define two sets of polynomials $P(a)\subset\mathbb{H}$, and $P(\mathbf{A})\subset\mathbb{R}^{n\times n}$ by

$$P(a) := \{v\in\mathbb{H} : v := \sum_{j=\mu}^{\nu}\alpha_j a^j,\ \alpha_j\in\mathbb{R}\}, \tag{5.3}$$

$$P(\mathbf{A}) := \{\mathbf{V}\in\mathbb{R}^{n\times n} : \mathbf{V} := \sum_{j=\mu}^{\nu}\alpha_j \mathbf{A}^j,\ \alpha_j\in\mathbb{R}\}, \tag{5.4}$$

where μ,ν are variable integers with $\mu\le\nu$. If $\mathbf{A}$ is not invertible we have the additional constraint $0\le\mu\le\nu$ when defining $P(\mathbf{A})$. We have the following evident inclusions:

$$P(a)\subset C(a), \quad P(\mathbf{A})\subset C(\mathbf{A}). \tag{5.5}$$

LEMMA 5.2. *Let $a \in \mathbb{H}\backslash\{0\}$ and $\mathbf{A} \in \mathbb{R}^{n\times n}$ be given. Let $b \in C(a)$ and let $\mathbf{B} \in C(\mathbf{A})$ with invertible $b, \mathbf{B}$. Then,*

$$P(b) \subset C(a), \quad P(\mathbf{B}) \subset C(\mathbf{A}).$$

Proof: We show that $b^j \in C(a)$ for all integers j. The real span of all b^j is $P(b)$ and since $C(a)$ is a linear space, we have $P(b) \subset C(a)$. The proof will be by induction. Let $ba = ab$ which is equivalent to $a = b^{-1}ab$. Assume that $b^j \in C(a)$, which means, that $a = b^{-j}ab^j$. Multiplying by b^{-1} from the left and by b from the right implies $a = b^{-1}ab = b^{-j-1}ab^{j+1}$, thus, $b^{j+1} \in C(a)$. The proof for matrices is the same. □

THEOREM 5.3. *Let $a \in \mathbb{H}\backslash\{\mathbb{R}\}$ be given. Then,*

$$C(a) = P(a) \text{ and } \dim C(a) = 2,$$

where $C(a)$ is the centralizer of a and $P(a)$ is the polynomial defined in (5.3).

Proof: We have to solve Sylvester's equation

$$\lambda_2(x) := ax - xa = 0. \tag{5.6}$$

Apparently, λ_2 is by Corollary 4.2 singular and by definition, the centralizer $C(a)$ is the set of solutions, which coincides with the null space (or kernel) of λ_2. According to Theorem 3.3, equation (5.6) is equivalent to the real matrix equation

$$(\imath_1(a) - \imath_2(a))\mathbf{x} := \mathbf{M}\mathbf{x} := 2\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -a_4 & a_3 \\ 0 & a_4 & 0 & -a_2 \\ 0 & -a_3 & a_2 & 0 \end{pmatrix}\mathbf{x} = \mathbf{0},$$

where the column vector $\mathbf{x}$ has to be identified with the quaternion x. The matrix $\mathbf{M}$ has the property that $\mathbf{M} + \mathbf{M}^{\mathrm{T}} = \mathbf{0}$, which implies that the rank must be even. Since by assumption at least one of the three components a_2, a_3, a_4 does not vanish, the rank of $\mathbf{M}$ must be two. This proves the last statement of the lemma. Because of the first general inclusion (5.5) we only need to show that the space $P(a)$ has (at least) dimension two. This is true, since the two elements $1 = a^0$ and a belong to $P(a)$ and since, by assumption, a is not real. □

COROLLARY 5.4. *Let $a \in \mathbb{H}\backslash\mathbb{R}$, then all matrices*

$$\mathbf{S} := [a^{j_1}, a^{j_2}, a^{j_3}, a^{j_4}] \in \mathbb{R}^{4\times 4}$$

with $j_1, j_2, j_3, j_4 \in \mathbb{Z}$ have rank at most two.

Proof: Follows from the previous lemma. □

The surprising consequence is contained in the following corollary.

COROLLARY 5.5. *Given $a \in \mathbb{H}\backslash\mathbb{R}$ and $j \in \mathbb{Z}$. Then we can always find two real numbers α, β such that*

$$a^j = \alpha + \beta a. \tag{5.7}$$

Proof: Let $a := (a_1, a_2, a_3, a_4)$. We will give a proof by induction. Assume first that $j \geq 0$. Equation (5.7) is apparently true for $j = 0$ and $j = 1$. Multiplying equation (5.7) by a yields $a^{j+1} = \alpha a + \beta a^2$. Computing a^2 with the rule given in (3.2) yields $a^{j+1} = \beta(a_1^2 - 2a_1 - a_2^2 + a_3^2 + a_4^2) + (\alpha + 2\beta a_1)a = \tilde{\alpha} + \tilde{\beta}a$. This shows that (5.7) is valid for all $j \geq 0$ and all $a \in \mathbb{H}$. Put $a = b^{-1}$. Then $b^{-j} = \alpha + \beta b^{-1} = \alpha + \beta\overline{b}/|b|^2$ where $\overline{b} := (b_1, -b_2, -b_3, -b_4)$ is the quaternionic conjugate of $b := (b_1, b_2, b_3, b_4)$. Now it is easy to see that in general we have $\alpha + \beta\overline{a} = \alpha + 2\beta a_1 - \beta a = \tilde{\alpha} + \tilde{\beta}a$. Therefore, equation (5.7) is also valid for negative j. □

To find an explicit formula for α, β let $a^j =: (A_1, A_2, A_3, A_4)$ and

$$a_k := \arg\max_{l\geq 2} |a_l|.$$

Since a is not real we have $a_k \neq 0$ and

$$\beta := A_k/a_k, \quad \alpha := A_1 - \beta a_1.$$

EXAMPLE 5.6. Take $a := (2,3,5,7)$, then $a^{10} = (2\,773\,885\,841, 1\,363\,602\,108, 2\,272\,670\,180, 3\,181\,738\,252)$, $\arg\max_{l\ge 2}|a_l| = a_4 = 7$, $\beta = A_4/a_4 = 3\,181\,738\,252/7 = 454\,534\,036$, $\alpha = A_1 - \beta a_1 = 2\,773\,885\,841 - 2\beta = 1\,864\,817\,769$. Finally, $a^{10} = \alpha + \beta a = 1\,864\,817\,769(1,0,0,0) + 454\,534\,036(2,3,5,7)$. In particular, all powers of $e := (1,1,1,1)$ have the form $e^j = (A_1, A_2, A_2, A_2)$, $j \in \mathbb{Z}$.

It is clear that Corollary 5.5 is also valid for the isomorphic image $\mathbf{A} := \imath_1(a)$ of a. And this is reflected also in the first part of the next theorem which is based on matrix theory.

We turn now to the determination of $C(\imath_1(a))$. Though $\imath_1(a)$ is isomorphic to a by the mapping $\imath_1$, the centralizers $C(a)$ and $C(\imath_1(a))$ - as will be shown - are not isomorphic by this mapping. Here it is useful to apply the general matrix theory. See [4, HORN & JOHNSON, Section 4.4]. Let $a := (a_1, a_2, a_3, a_4)$. From (3.13), (3.14) it follows that the *characteristic polynomial* $\chi_{\mathbf{A}}$ of $\mathbf{A} := \imath_1(a)$ is

$$\chi_{\mathbf{A}}(\lambda) := (\lambda - \sigma_+)^2(\lambda - \sigma_-)^2 = (\lambda^2 - 2a_1\lambda + |a|^2)^2.$$

However, the *minimal polynomial* $q_{\mathbf{A}}$ of $\mathbf{A}$ is

$$q_{\mathbf{A}}(\lambda) := (\lambda - \sigma_+)(\lambda - \sigma_-) = (\lambda^2 - 2a_1\lambda + |a|^2).$$

This follows from $q_{\mathbf{A}}(a) := (a^2 - 2a_1 a + a\overline{a}) = a(a - 2a_1 + \overline{a}) = 0$. The geometric multiplicity of σ_+, σ_- is two. Therefore, the matrix $\mathbf{A}$ is *derogatory*. See [3, Section 3.2]. This follows also from the fact that we have already found one element of the centralizer $C(\mathbf{A})$, namely $\tilde{\mathbf{A}}$, defined in (3.4), which is not a polynomial in $\mathbf{A}$. See [4, Corollary 4.4.18].

THEOREM 5.7. *Let $a \in \mathbb{H}\backslash\mathbb{R}$ and $\mathbf{A} := \imath_1(a)$. Then,*

$$\dim P(\mathbf{A}) = \deg(q_{\mathbf{A}}) = 2, \quad \dim C(\mathbf{A}) > 4. \tag{5.8}$$

Proof: [4, Theorem 4.4.17]. □

Actually, we have already shown in Theorem 3.1 that $\mathbb{H}_{\mathbb{P}} \subset C(\mathbf{A})$. Thus, the centralizer $C(\mathbf{A})$ contains at least the real span of $\mathbb{H}_{\mathbb{P}} \cup P(\mathbf{A})$. Since all elements of $\mathbb{H}_{\mathbb{P}}$ commute with all elements of $\mathbb{H}_{\mathbb{R}}$, and since $\mathbb{H}_{\mathbb{R}} \cap \mathbb{H}_{\mathbb{P}} = \imath_1(\mathbb{R})$, the centralizer $C(\mathbf{A})$ is different from (larger than) the set of polynomials $P(\mathbf{A})$. In order to find whether other elements than those contained in $\mathbb{H}_{\mathbb{P}}$ belonging to $C(\mathbf{A})$, we solve $\mathbf{AX} - \mathbf{XA} = \mathbf{0}$ for $\mathbf{X} \in \mathbb{R}^{4\times 4}$.

LEMMA 5.8. *Let $C(\mathbf{A})$ be the centralizer of $\mathbf{A} := \imath_1(a)$ for $a \in \mathbb{H}\backslash\mathbb{R}$. Then,*

$$\mathbf{B} \in C(\mathbf{A}) \Leftrightarrow \mathbf{B} \in C(\mathbf{A}^{\mathrm{T}}),$$
$$\mathbf{B} \in C(\mathbf{A}) \Leftrightarrow \mathbf{B}^{\mathrm{T}} \in C(\mathbf{A}).$$

In other words, the centralizers of $\mathbf{A}$ and $\mathbf{A}^{\mathrm{T}}$ are the same and if $\mathbf{B}$ belongs to $C(\mathbf{A})$ then also $\mathbf{B}^{\mathrm{T}}$ belongs to $C(\mathbf{A})$ and v. v.

Proof: $C(\mathbf{A}) := \{\mathbf{X} : \mathbf{AX} - \mathbf{XA} = \mathbf{0}\}$. The matrix $\mathbf{A}$ has the property that $\mathbf{A} + \mathbf{A}^{\mathrm{T}} = 2\Re(a)\mathbf{I}$, where $\mathbf{I}$ is the (4×4) identity matrix. Inserting $\mathbf{A} = 2\Re(a)\mathbf{I} - \mathbf{A}^{\mathrm{T}}$ into the defining equation yields
$(2\Re(a)\mathbf{I} - \mathbf{A}^{\mathrm{T}})\mathbf{X} = \mathbf{X}(2\Re(a)\mathbf{I} - \mathbf{A}^{\mathrm{T}}) \Leftrightarrow \mathbf{A}^{\mathrm{T}}\mathbf{X} = \mathbf{X}\mathbf{A}^{\mathrm{T}} \Leftrightarrow \mathbf{A}\mathbf{X}^{\mathrm{T}} = \mathbf{X}^{\mathrm{T}}\mathbf{A}$. □

In explicit terms, the system $\mathbf{AX} - \mathbf{XA}$ reduces to the linear system of 16 equations:

$$\sum_{l=1}^{4}(a_{jl}x_{lk} - a_{lk}x_{jl}) = 0, \quad j,k = 1,2,3,4, \tag{5.9}$$

where one has to identify the elements $(a_{jk}), j,k = 1,2,3,4$ by the elements a_1, a_2, a_3, a_4 according to (3.3), p. 3. At this stage it would have been possible to apply Kronecker's product to find the equivalent system in 16 unknowns. See [4, HORN & JOHNSON, CHAPTER 4, LEMMA 4.3.1]. In the (16×16)-matrix $\mathbf{G}$ corresponding to (5.9) only the quantities $\pm a_j, j = 2,3,4$ and the number zero occur. In order to reduce the typographical size of $\mathbf{G}$ we put j instead of a_j and $-j$ instead of $-a_j$. We number the unknowns rather than by (x_{jk}), row-wise[1] by $x_1, x_2, \ldots, x_{16}$ and denote the vector of the unknowns by $\mathbf{x}$. Equation (5.9) takes then the final form $\mathbf{Gx} = \mathbf{0}$, where

[1] According to Lemma 5.8 it does not matter whether we enumerate the elements of $\mathbf{X}$ row-wise or column-wise.

(5.10)
$$\mathbf{G} :=$$

$$\begin{pmatrix}
0 & -2 & -3 & -4 & -2 & 0 & 0 & 0 & -3 & 0 & 0 & 0 & -4 & 0 & 0 & 0\\
2 & 0 & 0 & 0 & 0 & -2 & -3 & -4 & -4 & 0 & 0 & 0 & 3 & 0 & 0 & 0\\
3 & 0 & 0 & 0 & 4 & 0 & 0 & 0 & 0 & -2 & -3 & -4 & -2 & 0 & 0 & 0\\
4 & 0 & 0 & 0 & -3 & 0 & 0 & 0 & 2 & 0 & 0 & 0 & 0 & -2 & -3 & -4\\
2 & 0 & -4 & 3 & 0 & -2 & 0 & 0 & 0 & -3 & 0 & 0 & 0 & -4 & 0 & 0\\
0 & 2 & 0 & 0 & 2 & 0 & -4 & 3 & 0 & -4 & 0 & 0 & 0 & 3 & 0 & 0\\
0 & 3 & 0 & 0 & 0 & 4 & 0 & 0 & 2 & 0 & -4 & 3 & 0 & -2 & 0 & 0\\
0 & 4 & 0 & 0 & 0 & -3 & 0 & 0 & 0 & 2 & 0 & 0 & 2 & 0 & -4 & 3\\
3 & 4 & 0 & -2 & 0 & 0 & -2 & 0 & 0 & 0 & -3 & 0 & 0 & 0 & -4 & 0\\
0 & 0 & 2 & 0 & 3 & 4 & 0 & -2 & 0 & 0 & -4 & 0 & 0 & 0 & 3 & 0\\
0 & 0 & 3 & 0 & 0 & 0 & 4 & 0 & 3 & 4 & 0 & -2 & 0 & 0 & -2 & 0\\
0 & 0 & 4 & 0 & 0 & 0 & -3 & 0 & 0 & 0 & 2 & 0 & 3 & 4 & 0 & -2\\
4 & -3 & 2 & 0 & 0 & 0 & 0 & -2 & 0 & 0 & 0 & -3 & 0 & 0 & 0 & -4\\
0 & 0 & 0 & 2 & 4 & -3 & 2 & 0 & 0 & 0 & 0 & -4 & 0 & 0 & 0 & 3\\
0 & 0 & 0 & 3 & 0 & 0 & 0 & 4 & 4 & -3 & 2 & 0 & 0 & 0 & 0 & -2\\
0 & 0 & 0 & 4 & 0 & 0 & 0 & -3 & 0 & 0 & 0 & 2 & 4 & -3 & 2 & 0
\end{pmatrix}.$$

Matrix $\mathbf{G}$ has the property that $\mathbf{G}+\mathbf{G}^{\mathrm{T}} = \mathbf{0}$, which implies that the rank of $\mathbf{G}$ is even. See [8, Theorem 1]. Some preliminary, numerical tests suggest that the matrix $\mathbf{G}$ has rank eight. We can obtain more precise information.

LEMMA 5.9. *Let J, K be vectors with positive, strictly increasing integer entries with $\max J, K \le 16$. By $\mathbf{G}_{J,K}$ we denote the submatrix $\mathbf{G}_{j\in J, k\in K}$ of $\mathbf{G}$. By $\mu:\nu$, $\mu\le\nu$ we denote the integer vector $\mu, \mu+1, \ldots, \nu$. We find that*

$$D_1 := \det \mathbf{G}_{1:8,1:8} = \det \mathbf{G}_{9:16,9:16} = \det \mathbf{G}_{1:8,9:16} = \det \mathbf{G}_{9:16,1:8} = (a_3^2 + a_4^2)^4, \tag{5.11}$$

$$D_2 := \det \mathbf{G}_{5:12,5:12} = \det \mathbf{G}_{5:12,1:4\cup 13:16} = \det \mathbf{G}_{1:4\cup 13:16,5:12} = \det \mathbf{G}_{1:4\cup 13:16,1:4\cup 13:16} = (a_2^2 + a_3^2)^4, \tag{5.12}$$

and $\mathbf{G}_{5:12,5:12}$ is the only matrix in the class $\mathbf{G}_{j:j+7,k:k+7}$, $j,k = 1,2,\ldots,9$ which attains the given determinant. In addition, there is no submatrix in this class with determinant $(a_2^2 + a_3^2 + a_4^2)^4$.

Proof: Since the (16×16) matrix $\mathbf{G}$ contains (at least) 160 zeros and 96 entries with $\pm a_j$, $j = 2,3,4$ it is possible to compute the determinants of all given submatrices of $\mathbf{G}$. Some assistance from `MAPLE` is acknowledged. □

In order to show that the whole matrix $\mathbf{G}$ is of rank eight, there must be some dependencies between the matrices mentioned in Lemma 5.9.

LEMMA 5.10. *Let $a \in \mathbb{H}\backslash\mathbb{R}$. We consider the following two partitions of $\mathbf{G}$, where $\mathbf{G}$ is defined in (5.10):*

$$\mathbf{G} =: \begin{pmatrix} \mathbf{G}_{\mathrm{ul}} & \mathbf{G}_{\mathrm{ur}} \\ \mathbf{G}_{\mathrm{ll}} & \mathbf{G}_{\mathrm{lr}} \end{pmatrix}, \tag{5.13}$$

$$\mathbf{G} =: \left(\begin{array}{c|c|c} \mathbf{G}_4 & \mathbf{G}_3 & \mathbf{G}_4 \\ \hline \mathbf{G}_2 & \mathbf{G}_1 & \mathbf{G}_2 \\ \hline \mathbf{G}_4 & \mathbf{G}_3 & \mathbf{G}_4 \end{array}\right), \tag{5.14}$$

where $\mathbf{G}_{\mathrm{ul}} = \mathbf{G}_{1:8,1:8}$ (upper left), $\mathbf{G}_{\mathrm{ur}} = \mathbf{G}_{1:8,9:16}$ (upper right), $\mathbf{G}_{\mathrm{ll}} = \mathbf{G}_{9:16,1:8}$ (lower left), $\mathbf{G}_{\mathrm{lr}} = \mathbf{G}_{9:16,9:16}$ (lower right), where we have used the notation for submatrices introduced in Lemma 5.9. With the same notation we have $\mathbf{G}_1 := \mathbf{G}_{5:12,5:12}$, $\mathbf{G}_2 := \mathbf{G}_{5:12,1:4\cup 13:16}$, $\mathbf{G}_3 := \mathbf{G}_{1:4\cup 13:16,5:12}$, $\mathbf{G}_4 := \mathbf{G}_{1:4\cup 13:16,1:4\cup 13:16}$.

(i) Let $a_3^2+a_4^2>0$. Then, with

$$\mathbf{P} := \begin{pmatrix} 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 \end{pmatrix} \tag{5.15}$$

we have

$$\left(\mathbf{G}_{\mathrm{ll}} \quad \mathbf{G}_{\mathrm{lr}} \right) = \left(\mathbf{G}_{\mathrm{ul}} \quad \mathbf{G}_{\mathrm{ur}} \right) \mathbf{P}. \tag{5.16}$$

(ii) Let $a_2^2+a_3^2>0$. Then, with

$$\mathbf{Q} := \begin{pmatrix} 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 \end{pmatrix}. \tag{5.17}$$

we have

$$\left(\mathbf{G}_3 \quad \mathbf{G}_4 \right) = \left(\mathbf{G}_1 \quad \mathbf{G}_2 \right) \mathbf{Q}. \tag{5.18}$$

Both matrices $\mathbf{P}$ and $\mathbf{Q}$ are essentially permutation matrices, with the additional property that $\mathbf{P}+\mathbf{P}^{\mathrm{T}}=\mathbf{0}$, which implies that $\mathbf{P}^{-1}=-\mathbf{P}$ and thus, $\mathbf{P}^2=-\mathbf{I}$ and the same for $\mathbf{Q}$.

Proof: Since $a\notin\mathbb{R}$ (at least) one of the cases (i) or (ii) applies. The mentioned properties of $\mathbf{P},\mathbf{Q}$ are obvious. The application of $\mathbf{P}$ in the form $\mathbf{BP}=\mathbf{C}$ (with arbitrary $\mathbf{B},\mathbf{C}\in\mathbb{R}^{8\times 8}$) has the following effect: Put $\mathbf{P}=:(\mathbf{P}_{jk}),\ j,k=1,2,\ldots,8$. If

$$\mathbf{P}_{jk} = \begin{cases} 1 & \text{then, column } j \text{ of } \mathbf{B} \text{ becomes column } k \text{ of } \mathbf{C}, \\ -1 & \text{then, column } j \text{ of } \mathbf{B} \text{ becomes after multiplication by } -1 \\ & \text{column } k \text{ of } \mathbf{C}. \end{cases}$$

Using this rule, one can verify (5.16) and (5.18). □

LEMMA 5.11. *Let $a\in\mathbb{H}\backslash\mathbb{R}$ and $\mathbf{G}\in\mathbb{R}^{16\times 16}$ be defined as in (5.10). Then, $\operatorname{rank}\mathbf{G}=8$ and the solutions of the homogeneous equation $\mathbf{Gx}=\mathbf{0}$ are:*

(i) If $a_3^2+a_4^2>0$ the solutions of

$$\mathbf{G}_{\mathrm{ul}}\mathbf{x}_1+\mathbf{G}_{\mathrm{ur}}\mathbf{x}_2=\mathbf{0}, \tag{5.19}$$

and we can find all solutions by selecting $\mathbf{x}_2:=(x_9,x_{10},\ldots,x_{16})^{\mathrm{T}}$ arbitrarily and solve (5.19) for $\mathbf{x}_1:=(x_1,x_2,\ldots,x_8)^{\mathrm{T}}$ or vice versa.

(ii) If $a_2^2+a_3^2>0$ the solutions of

$$\mathbf{G}_1\mathbf{y}_1+\mathbf{G}_2\mathbf{y}_2=\mathbf{0}, \tag{5.20}$$

and we can find all solutions by selecting $\mathbf{y}_2:=(x_1,\ldots,x_4,x_{13},\ldots,x_{16})^{\mathrm{T}}$ arbitrarily and solve (5.20) for $\mathbf{y}_1:=(x_5,x_6,\ldots,x_{12})^{\mathrm{T}}$ or vice versa. The final solution is then $\mathbf{x}=(x_1,x_2,\ldots,x_{16})^{\mathrm{T}}$. The four submatrices employed are defined in (5.13), (5.14). □

Proof: (i) Let $a_3^2+a_4^2>0$. Then by Lemma 5.9 all four submatrices $\mathbf{G}_{\mathrm{ul}},\mathbf{G}_{\mathrm{ur}},\mathbf{G}_{\mathrm{ll}},\mathbf{G}_{\mathrm{lr}}$ are non singular and thus $\operatorname{rank}\mathbf{G}\geq 8$. By (5.16) the lower half of $\mathbf{G}$ depends linearly on the upper half, therefore $\operatorname{rank}\mathbf{G}=8$ and the solution can be found by using only the upper eight equations. (ii) Let $a_2^2+a_3^2>0$. Then by Lemma 5.9 all four submatrices $\mathbf{G}_1,\mathbf{G}_2,\mathbf{G}_3,\mathbf{G}_4$ are non singular and thus $\operatorname{rank}\mathbf{G}\geq 8$. By (5.18) the middle part (rows 5 to 12) of $\mathbf{G}$ depends linearly on the remaining part, thus $\operatorname{rank}\mathbf{G}=8$ and solving $\mathbf{Gx}=\mathbf{0}$, reduces to the given equation. □

It should be noted that the (8×16) matrix $(\mathbf{G}_{\rm ul},\mathbf{G}_{\rm ur})$ has exactly one additional partition into two (8×8) matrices such that the determinants of these matrices are also $(a_3^2+a_4^2)^4$. For this case we have to select the columns $1,2,5,6,9,10,$
$13,14$ for one matrix and the columns $3,4,7,8,11,12,15,16$ for the other matrix.

THEOREM 5.12. *Let $a\in\mathbb{H}\backslash\mathbb{R}$ be given and $\mathbf{A}:=\imath_1(a)$, where $\imath_1(a)$ is defined in (3.3), p. 3. The centralizer of $\mathbf{A}$ is*

(5.21)
$$C(\mathbf{A}):=\left\{\mathbf{X}\in\mathbb{R}^{4\times 4}:\mathbf{X}:=\begin{pmatrix}\mathbf{x}_1(1:4)\\ \mathbf{x}_1(5:8)\\ \mathbf{x}_2(1:4)\\ \mathbf{x}_2(5:8)\end{pmatrix}\right\}$$

where $\mathbf{x}=(x_1,x_2,\ldots,x_{16})^{\rm T}$ is defined in Lemma 5.11, and

(5.22)
$$\dim C(\mathbf{A})=8.$$

Proof:. The centralizer is the null space of $\mathbf{AX}-\mathbf{XA}$ and this has dimension eight according to Lemma 5.11. □

If we choose $\mathbf{x}_1$ according to the first two columns (rows) of $\imath_2(b)$, p. 3, with an arbitrary $b\in\mathbb{H}$, then, the formula for $\mathbf{x}_2$ furnishes the last two columns (rows) of $\imath_2(b)$. This is another proof of (3.5) in Theorem 3.1, p. 3. If we compare Theorems 5.12 and 5.3 we see, that $C(\mathbf{A}):=C(\imath_1(a))$ is considerably larger than $\imath_1(C(a))$. In the next theorem we show how to construct a basis for the centralizer $C(\mathbf{A})$.

THEOREM 5.13. *Let $a=(a_1,a_2,a_3,a_4)\in\mathbb{H}\backslash\mathbb{R}$ be given. Define*

$$A:=a_2a_4;\ B:=a_2a_3;\ C:=a_3^2;\ D:=a_3a_4;\ E:=a_4^2;\ F:=a_2^2.$$

Let $d_1:=a_3^2+a_4^2>0$. Then, the following eight matrices $I_1,I_2,\ldots,I_8$ define a basis for $C(\mathbf{A})$:

$$I_1:=\begin{pmatrix}d_1&0&A&-B\\0&0&B&A\\0&0&C&D\\0&0&D&E\end{pmatrix},\quad I_2:=\begin{pmatrix}0&0&-B&-A\\d_1&0&A&-B\\0&0&D&E\\0&0&-C&-D\end{pmatrix},$$

$$I_3:=\begin{pmatrix}0&0&-C&-D\\0&0&-D&-E\\d_1&0&A&-B\\0&0&B&A\end{pmatrix},\quad I_4:=\begin{pmatrix}0&0&-D&-E\\0&0&C&D\\0&0&-B&-A\\d_1&0&A&-B\end{pmatrix},$$

$$I_5:=\begin{pmatrix}0&d_1&-B&-A\\0&0&A&-B\\0&0&D&-C\\0&0&E&-D\end{pmatrix},\quad I_6:=\begin{pmatrix}0&0&-A&B\\0&d_1&-B&-A\\0&0&E&-D\\0&0&-D&C\end{pmatrix},$$

$$I_7:=\begin{pmatrix}0&0&-D&C\\0&0&-E&D\\0&d_1&-B&-A\\0&0&A&-B\end{pmatrix},\quad I_8:=\begin{pmatrix}0&0&-E&D\\0&0&D&-C\\0&0&-A&B\\0&d_1&-B&-A\end{pmatrix}.$$

Let $d_2:=a_2^2+a_3^2>0$. Then the following eight matrices $J_1,J_2,\ldots,J_8$ define a basis for $C(\mathbf{A})$:

$$J_1:=\begin{pmatrix}-D&d_2&0&A\\-F&0&0&-B\\-B&0&0&-C\\-A&0&0&-D\end{pmatrix},\quad J_2:=\begin{pmatrix}F&0&0&B\\-D&d_2&0&A\\-A&0&0&-D\\B&0&0&C\end{pmatrix},$$

$$J_3:=\begin{pmatrix}B&0&0&C\\A&0&0&D\\-D&d_2&0&A\\-F&0&0&-B\end{pmatrix},\quad J_4:=\begin{pmatrix}A&0&0&D\\-B&0&0&-C\\F&0&0&B\\-D&d_2&0&A\end{pmatrix},$$

$$J_5:=\begin{pmatrix}A&0&d_2&D\\-B&0&0&F\\-C&0&0&B\\-D&0&0&A\end{pmatrix},\quad J_6:=\begin{pmatrix}B&0&0&-F\\A&0&d_2&D\\-D&0&0&A\\C&0&0&-B\end{pmatrix},$$

$$J_7 := \begin{pmatrix} C & 0 & 0 & -B \\ D & 0 & 0 & -A \\ A & 0 & d_2 & D \\ -B & 0 & 0 & F \end{pmatrix}, \quad J_8 := \begin{pmatrix} D & 0 & 0 & -A \\ -C & 0 & 0 & B \\ B & 0 & 0 & -F \\ A & 0 & d_2 & D \end{pmatrix}.$$

Proof: If we put I_1 to I_8 row-wise into one 8×16 matrix and the same with J_1 to J_8, we see that these two matrices have rank eight because they both contain a multiple of the (8×8) identity matrix. Thus, the basis elements are linearly independent. □

The entries A to F in the above 16 basis elements will be integer, if $a \in \mathbb{H}$ has integer components. This is the reason why the entries d_1, d_2 (instead of 1) were chosen. Note that all basis elements are neither quaternions nor pseudo quaternions. Note also, that $\imath_2(a)$ does not commute with the above defined basis elements which implies that the centralizers $C(\imath_1(a))$ and $C(\imath_2(a))$ are different.

EXAMPLE 5.14. Let $\mathbf{A} := \imath_1(a)$ be given, where $a \notin \mathbb{R}$. Since $\mathbf{A} \in C(\mathbf{A})$ there must be a representation of $\mathbf{A}$ with respect to the given basis elements. Let as before $d_1 := a_3^2 + a_4^2, d_2 := a_2^2 + a_3^2$. The representation is (in the first case $d_1 > 0$, in the second case $d_2 > 0$) given by

$$\mathbf{A} = \begin{cases} \big(a_1 I_1 + a_2 I_2 + a_3 I_3 + a_4 I_4 - a_2 I_5 + a_1 I_6 + a_4 I_7 - a_3 I_8\big)/d_1, \\ \big(-a_2 J_1 + a_1 J_2 + a_4 J_3 - a_3 J_4 - a_3 J_5 - a_4 J_6 + a_1 J_7 + a_2 J_8\big)/d_2. \end{cases}$$

Let $a \in \mathbb{R}$. Then $\mathbf{A} := \imath_1(a)$ is a multiple of the identity matrix and thus, in this case $C(\mathbf{A}) = \mathbb{R}^{4\times 4}$ and the above representation is not valid.

EXAMPLE 5.15. The centralizer $C(\mathbf{A})$ has the property that $\mathbf{B} \in C(\mathbf{A})$ implies $\mathbf{B}^{\mathrm{T}} \in C(\mathbf{A})$. See Lemma 5.8. Let a representation $\mathbf{B} = \sum_{j=1}^{8} \alpha_j I_j$ for $d_1 > 0$ or $\mathbf{B} := \sum_{j=1}^{8} \beta_j J_j$ for $d_2 > 0$ be given. One might be interested in finding the corresponding representation for $\mathbf{B}^{\mathrm{T}}$. What is needed is a representation $I_k^{\mathrm{T}} = \sum_{j=1}^{8} \gamma_j^{(k)} I_j$ for $d_1 > 0$ and $J_k^{\mathrm{T}} = \sum_{j=1}^{8} \delta_j^{(k)} J_j$ for $d_2 > 0$. Both equations can be reduced to the linear (8×8) systems

$$\mathbf{BI}\,\boldsymbol{\gamma}^{(k)} = \mathbf{rI}^{(k)}, \quad \mathbf{BJ}\,\boldsymbol{\delta}^{(k)} = \mathbf{rJ}^{(k)},$$

where the two matrices $\mathbf{BI}$, $\mathbf{BJ}$ and the 16 right hand sides are defined as follows:

$$\begin{aligned} \mathbf{BI}(2(m-1)+n,k) &:= I_k(n,m);\ m=1,2,3,4;\ n=1,2, \\ \mathbf{BJ}(2(m-1)+n-1,k) &:= J_k(n,m);\ m=1,2,3,4;\ n=2,3,\ k=1,2,\ldots,8; \\ \mathbf{rI}^{(k)}(2(m-1)+n) &:= I_k(m,n);\ m=1,2,3,4;\ n=1,2, \\ \mathbf{rJ}^{(k)}(2(m-1)+n-1) &:= J_k(m,n);\ m=1,2,3,4;\ n=2,3,\ k=1,2,\ldots,8. \end{aligned}$$

If we put all eight solutions column-wise in one matrix $\boldsymbol{\gamma}$ for $d_1 > 0$ and $\boldsymbol{\delta}$ for $d_2 > 0$ we find that $\boldsymbol{\gamma} d_1 = \mathbf{\Pi}\,\mathbf{BI}$, $\boldsymbol{\delta} d_2 = \mathbf{\Pi}\,\mathbf{BJ}$ where $\mathbf{\Pi}$ is the same permutation matrix in all cases which has the following effect: Row 1 and 8 of $\mathbf{BI}$ and of $\mathbf{BJ}$ are unchanged, otherwise make the following permutations: Row $2 \to 5$, row $3 \to 2$, row $4 \to 6$, row $5 \to 3$, row $6 \to 7$, row $7 \to 4$. Then, the wanted solutions are in columns 1 to 8. Let us return to our original problem of finding a representation for $\mathbf{B}^{\mathrm{T}}$, when a representation for $\mathbf{B}$ is given. Let $\mathbf{B}^{\mathrm{T}} = \sum_{j=1}^{8} \alpha_j^{\mathrm{new}} I_j$ for $d_1 > 0$, $\mathbf{B}^{\mathrm{T}} := \sum_{j=1}^{8} \beta_j^{\mathrm{new}} J_j$ for $d_2 > 0$. Then, $\boldsymbol{\alpha}^{\mathrm{new}} = \mathbf{\Pi}\,\mathbf{BI}\,\boldsymbol{\alpha}/d_1$, $\boldsymbol{\beta}^{\mathrm{new}} = \mathbf{\Pi}\,\mathbf{BJ}\,\boldsymbol{\beta}/d_2$.

6. The general equation. We now turn to the general equation (3.23), p. 5:

$$\lambda_3(x) := ax + bxc + xd = e, \quad a,b,c,d,e \in \mathbb{H}.$$

It is not quite obvious how to find an explicit solution to this equation in terms of quaternions and to find conditions under which a unique solution exists. The cases $ad = 0$ or $bc = 0$ reduce the above equation immediately to Sylvester's equation. Thus, we may assume that $ad \neq 0$ and $bc \neq 0$.

It would be sufficient to reduce λ_3 to one form of Sylvester's mapping. either in the original form (3.22) $\lambda_2(x) = e$ or an equivalent form $\Lambda_2(x) = E$, defined in (4.30). But this seems to be impossible. Sylvester's equation was solved by applying the composite mapping $\mu \circ \lambda_2$ where μ was essentially the inverse mapping of λ_2. See proof of Theorem 4.1. However, we can not hope to guess the inverse of λ_3.

LEMMA 6.1. *Let $b, c, b', c' \in \mathbb{H}\backslash\mathbb{R}$ be given. Assume that it is possible to find $b'', c'' \in \mathbb{H}$ such that*

(6.1) $$\Lambda_2(x) := bxc + b'xc' = b''xc'' \text{ for all } x \in \mathbb{H}.$$

Then, either b and b' are real multiples of b'', or c and c' are real multiples of c''.

Proof: Multiply from the right by c''^{-1}, then, the right hand side $b''x$ is linear with respect to multiplication from the right, but $\Lambda_2(x)c''^{-1} = bxcc''^{-1} + b'xc'c''^{-1}$ would be linear in this sense only if both $cc''^{-1}, c'c''^{-1}$ would be real. By multiplication with b''^{-1} from the left we would obtain that $b''^{-1}b$ and $b''^{-1}b'$ must both be real. □

COROLLARY 6.2. *Let $b, c, b', c' \in \mathbb{H}\backslash\mathbb{R}$ and assume that both b, b' are not real multiples of the same quaternion and that also c, c' are not real multiples of the same but possibly another quaternion. Then, an equation of the form (6.1) is impossible.*

Proof: Apply the previous lemma. □

Because of the previous results, we will first develop a method which guarantees a unique solution under a certain condition and which also includes an algorithm to approximate this solution. The tool is Banach's fixed point theorem.

6.1. Application of Banach's fixed point theorem. We repeat Banach's fixed point thoerem for completeness in a form which fits to our situation.

THEOREM 6.3. *(Banach, 1932) Let $(X, \| \quad \|)$ be a real Banach space and let $f : X \to X$ be a contractive mapping, i. e. there is a constant $\kappa < 1$ such that*

$$||f(x) - f(y)|| \le \kappa ||x - y|| \text{ for all } x, y \in X. \tag{6.2}$$

Then f has a unique fixed point ξ defined by $\xi := f(\xi)$, and this fixed point can be approximated by the sequence

$$x_{j+1} = f(x_j); \quad j = 0, 1, \ldots, \quad x_0 \text{ arbitrary} \tag{6.3}$$

and there is the following error estimate which also shows the convergence speed:

$$||x_j - \xi|| \le \frac{\kappa^j}{1-\kappa} ||x_1 - x_0||, \quad j \ge 1. \tag{6.4}$$

Proof: [1, BANACH]. □

As we see, this theorem guarantees existence, uniqueness of a fixed point, and it contains a method to approximate that fixed point and also gives an error estimate for the approximate solution. We apply this theorem to the mapping

$$f : \mathbb{H} \to \mathbb{H}, \quad f(x) := b^{-1}(e - ax - xd)c^{-1}, \quad bc \ne 0, \tag{6.5}$$

where $f(x) = x$ is equivalent to the original equation $l(x) := ax + bxc + xd = e$. The mapping defined by f has the property that

$$|f(x) - f(y)| \le \kappa |x - y|, \quad \kappa := (|a| + |d|)/(|b||c|). \tag{6.6}$$

Thus, Banach's fixed point theorem can be applied if $\kappa < 1$. There is the following simple application.

COROLLARY 6.4. *Let $|a| \le k, |d| \le k, |b| \ge k, |c| \ge k$ for some positive constant k and otherwise arbitrary quaternions a, b, c, d. Then, $\lambda_3(x) := ax + bxc + xd$ is non singular if $k > 2$. This conclusion is in general not true if $k = 2$.*

Proof: In this case we have $\kappa := (|a| + |d|)/(|b||c|) \le 2k/k^2 = 2/k < 1$ and Banach's fixed point theorem is applicable. In order to show that $k = 2$ does not necessarily yield non singular mappings we choose all components of a, b, c, d to be ± 1, which implies $k = |a| = |b| = |c| = |d| = 2$. There are 2^{16} such cases and a computer search reveals that 2560 of these cases are singular. Two of these cases are

$$\begin{aligned} a &:= (\ \ 1,\ \ 1,\ \ 1,\ \ 1),\ b := (\ \ 1,\ \ 1,\ \ 1, -1),\\ c &:= (-1,\ \ 1,\ \ 1,\ \ 1),\ d := (\ \ 1,\ \ 1, -1, -1),\\ a &:= (\ \ 1,\ \ 1, -1, -1),\ b := (\ \ 1,\ \ 1, -1,\ \ 1),\\ c &:= (-1, -1, -1,\ \ 1),\ d := (\ \ 1,\ \ 1, -1, -1). \end{aligned}$$

□

EXAMPLE 6.5. Define the following quantities:

$$a := (1, -1, 1, 1),\ b := (0, 1, 2, 3),$$
$$c := (-1, 1, -1, -1),\ d := (1, 1, 1, 1),$$
$$e := (16, -8, 8, 4), x := (1, 2, 2, 1). \text{ Then,}$$

$$\lambda_3(x) := ax + bxc + xd = e. \tag{6.7}$$

For the data of Example 6.5 we have $\kappa \approx 0.5345$. If we start with $x_0 = 0$, the iterate x_{33} differs from the solution $\xi := x$ by at most $2 \cdot 10^{-14}$ in each component and the error estimate (6.4) yields

$$|x_{33} - \xi| \le 2.2645 \cdot 10^{-9}|x_1 - x_0| = 6.0522 \cdot 10^{-9}.$$

Now, the condition $\kappa < 1$ may not be valid. Then a common trick is to iterate the inverse mapping of (6.3), i. e. we switch j and $j+1$. In our case we obtain $x_j := b^{-1}(e - ax_{j+1} - x_{j+1}d)c^{-1}$, which is equivalent to $ax_{j+1} + x_{j+1}d = e - bx_jc, \quad j = 0, 1, \ldots, \quad x_0$ arbitrary. Thus, in each step Sylvester's equation has to be solved. In the present case two new fixed point equations evolve by applying the solution formulas (4.1), (4.2), namely

$$F_l(x) := f_l^{-1}\big(e - bxc + a^{-1}(e - bxc)\overline{d}\big), \tag{6.8}$$
$$F_r(x) := \big(e - bxc + \overline{a}(e - bxc)d^{-1}\big)f_r^{-1}, \tag{6.9}$$

where f_l, f_r were already defined in (4.1), (4.2). Since the two formulas (4.1), (4.2) define the same x, we have $F_l(x) = F_r(x)$ for all x. We find

$$|F_l(x) - F_l(y)| \le \kappa_l|x - y|, \quad |F_r(x) - F_r(y)| \le \kappa_r|x - y|,$$

where

$$\kappa_l := \frac{|b||c|(1 + |a^{-1}||d|)}{|f_l|} = \frac{|b||c|(1 + |d^{-1}||a|)}{|f_r|} =: \kappa_r. \tag{6.10}$$

The middle identity follows from (4.40) of Lemma 4.11. Let us put

$$\tilde{\kappa} := \kappa_l = \kappa_r.$$

Thus, Banach's fixed point theorem can be applied to the (identical) functions F_l, F_r defined in (6.8), (6.9) if $\tilde{\kappa} < 1$. Let us use Example 6.5 again. Then we find $\tilde{\kappa} \approx 3.7417$ and we cannot apply Banach's fixed point theorem. This is no surprise, since the former κ is $\kappa \approx 0.5345$ and Banach's theorem works for (6.5).

Corollary 6.4 can be given the following qualitative form. Fix $|a| + |d|$. Then, sufficiently large $|b||c|$ guarantee that λ_3 is non singular. The opposite is true as well.

COROLLARY 6.6. *Fix $a, d \in \mathbb{H}\backslash\mathbb{R}$. Then, a sufficiently small product $|b||c|$ guarantees the non singularity of λ_3.*

Proof: The condition $\kappa_l < 1$ (see (6.10)) can be written as $|b||c|(|a| + |d|) < |a|\,|f_l|$. Now, a look at the definition (4.1) shows that $|f_l|$ depends only on a, d. □

We summarize our results.

THEOREM 6.7. *Let $a, b, c, d \in \mathbb{H}\backslash\mathbb{R}$ and $\lambda_3(x) := ax + bxc + xd$. Define κ by (6.6) and $\tilde{\kappa} := \kappa_l = \kappa_r$ by (6.10). If $\kappa < 1$ or if $\tilde{\kappa} < 1$ then λ_3 is non singular. If $\kappa < 1$ the unique solution of $\lambda_3(x) = e$ can be approximated by forming $x_{j+1} = f(x_j),\ j = 0, 1, \ldots,\ x_0$ arbitrary, where f is defined in (6.5) and an error estimate can be found in (6.4). If $\tilde{\kappa} < 1$, then the corresponding functions to be used for iteration are either F_l or F_r defined in (6.8), (6.9) and for the error estimate (6.4) is again applied. For the product we have $\kappa\tilde{\kappa} \ge 1$ and it depends only on a and d.*

Proof: Apart from $\kappa\tilde{\kappa} \ge 1$ everything has been shown already. Now,

$$\kappa\tilde{\kappa} = \frac{|a| + |d|}{|b||c|}\,\frac{|b||c|(1 + |a^{-1}||d|)}{|f_l|} = \frac{(|a| + |d|)^2}{|a||f_l|}.$$

Since $|a||f_l| = |2a\Re d + a^2 + |d|^2| \le 2|a||\Re d| + |a|^2 + |d|^2 \le (|a| + |d|)^2$ (because of $|\Re d| \le |d|$) the proof is complete. □

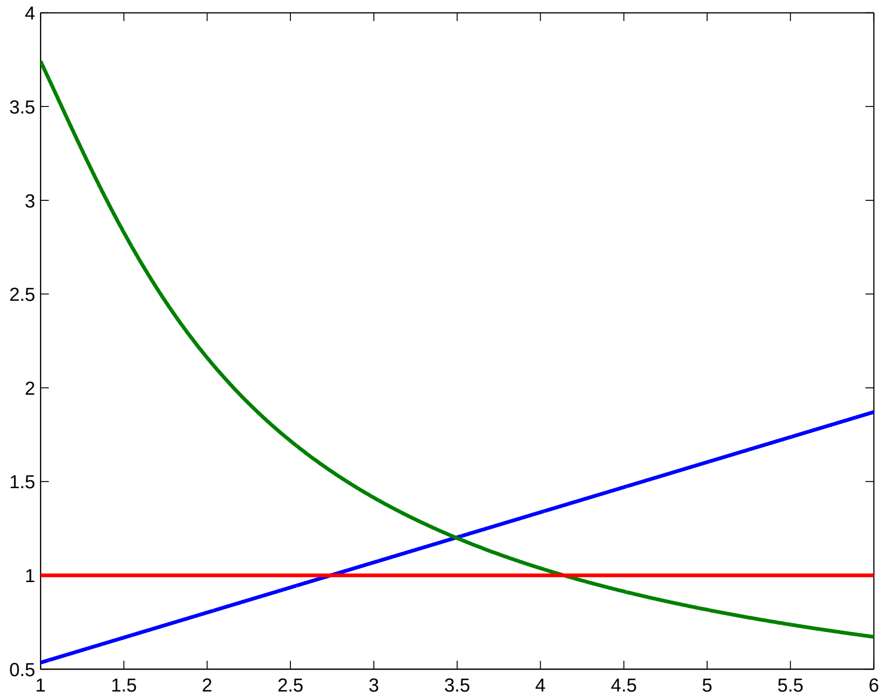

FIGURE 6.8. Values of κ and $\tilde{\kappa}$ for Example 6.5 with $d = \lambda(1,1,1,1), \lambda \in [1,6]$

We have developed two fixed point equations with two different contraction constants, $\kappa, \tilde{\kappa}$ for the same fixed point. It would be very favorable if at least one of the two contraction constants would be less than one. But this is not the case. We take the given example (p. 19) and vary d in the form λd with positive λ. The result is shown in Figure 6.8. The ascending, straight line represents κ as a function over λ, the other, curved line represents $\tilde{\kappa}$ over λ. For $\lambda < \lambda_0 \approx 2.74$ we have $\kappa < 1$, for $\lambda > \lambda_1 \approx 4.14$ we have $\tilde{\kappa} < 1$, but for $\lambda \in [\lambda_0, \lambda_1]$ we have $\kappa \geq 1, \tilde{\kappa} \geq 1$. Thus, there is no hope that all equations of the considered type (3.23) can be solved by Banach's fixed point theorem. In addition, $\kappa\tilde{\kappa} \geq 1$ implies, that it is impossible that simultaneously $\kappa < 1$ and $\tilde{\kappa} < 1$.

7. Some further problems and observations. We mention some problems in connection with the general form of the linear mapping λ_m. If we look at Theorem 4.8 then it appears likely that this theorem is only a sample of a more general theorem.

CONJECTURE 7.1. *Let λ_m be defined as in (3.1), p. 2 and let λ_m be non singular. Then there are $2m$ quaternions $b_j', c_j', j = 1, 2, \ldots, m$ such that the inverse of λ_m has the representation*

$$\lambda_m^{-1}(x) := \sum_{j=1}^{m} b_j' x c_j',$$

and the inverse of the matrix $\mathbf{M}$, defined in (3.19), can be written as

$$\mathbf{M}^{-1} := \sum_{j=1}^{m} \imath_1(b_j')\imath_2(c_j'),$$

where the mappings $\imath_1, \imath_2$ are defined in the beginning, p. 3, by (3.3), (3.4).

It is clear that the two statements in the above conjecture are equivalent.

Since λ_m is defined by $2m$ quaternions one might ask whether interpolation problems $\lambda_m(x_k) = y_k, k = 1, 2, \ldots, 2m$ with arbitrary quaternions y_k and pairwise distinct quaternions x_k have a solution.

Another, connected question is the following: Given λ_m, can one recover the defining $2m$ constants by function evaluations with arbitrary, but finitely many samples. Let us treat two simple examples.

The first example is Sylvester's equation $\lambda_2(x) = ax + xd$ where none of the two constants a, d is real. The interpolation problem for two points has the form

$$\begin{cases} xa + ay = A, \\ xb + by = B, \end{cases} \tag{7.1}$$

where $a, b \in \mathbb{H}\backslash\{0\}$ and $a \neq b$, and A, B arbitrary. We have changed the notation, named the unknowns a, b now x, y and the two points which replace the former x by a, b. If we apply the former matrix technique, this is equivalent to the (8×8) matrix equation

$$\begin{pmatrix} \tilde{\mathbf{A}} & \mathbf{A} \\ \tilde{\mathbf{B}} & \mathbf{B} \end{pmatrix} \begin{pmatrix} \mathbf{x} \\ \mathbf{y} \end{pmatrix} = \begin{pmatrix} A \\ B \end{pmatrix}, \tag{7.2}$$

where $\mathbf{A} := \imath_1(a), \tilde{\mathbf{A}} := \imath_2(a), \mathbf{B} := \imath_1(b), \tilde{\mathbf{B}} := \imath_2(b)$. However, this matrix does not have full rank. The first and the fifth column are the same. Actually, the rank is even 6, thus, the interpolation problem has solutions only for very specifically chosen A, B.

As a second example let us take $\lambda_1(x) := bxc$. The interpolation problem for two points with the same change of notation as above reads

$$\begin{cases} xay = A, \\ xby = B. \end{cases} \tag{7.3}$$

This is a non linear equation which in matrix terms reads

$$\mathbf{X}\tilde{\mathbf{Y}}\mathbf{a} = A, \quad \mathbf{X}\tilde{\mathbf{Y}}\mathbf{b} = B, \quad \mathbf{X} := \imath_1(x), \tilde{\mathbf{Y}} := \imath_2(y).$$

This problem will in general have no solution.

Now let us treat the recovery problem. We compute $p_j := \lambda_2(\mathcal{N}_j) = a\mathcal{N}_j + \mathcal{N}_j d,\ j = 1, 2, 3, 4$. If we have a look at formulas (3.3), (3.4), we deduce from the first three samples that we can recover the last three components of a and of d. But all four samples are not enough to recover the real part of a and d. We only find its sum.

The recovery problem also suffers from the non linearity in the coefficients. If we compute $q_j := \lambda_1(\mathcal{N}_j) = b\mathcal{N}_j c$ we obtain sums of products which are, however, not sufficient to find the coefficients b, c.

8. Appendix 1: The centralizer of $\tilde{\mathbf{A}} := \imath_2(\mathbf{a})$. The same techniques which were used to determine the centralizer $C(\mathbf{A})$ of $\mathbf{A} := \imath_1(a)$ (Theorem 5.13) can be used for determining the centralizer $C(\tilde{\mathbf{A}})$ of $\tilde{\mathbf{A}} := \imath_2(a)$. See (3.4) on p. 3 for $\imath_2$ and Definition 5.1 on p. 11 for centralizer.

THEOREM 8.1. *Let $a = (a_1, a_2, a_3, a_4) \in \mathbb{H}\backslash\mathbb{R}$ be given. Define*

$$d_1 := a_3^2 + a_4^2,\ d_2 := a_2^2 + a_3^2,$$
$$A := a_2 a_4,\ B := a_2 a_3,\ C := a_3^2,\ D := a_3 a_4,\ E := a_4^2,\ F := a_2^2.$$

The centralizer $C(\imath_2(a))$ has dimension eight and a basis is given below. If a has integer entries, than all given basis elements have also integer entries.

Case 1: $d_1 > 0$:

$$\tilde{I}_1 := \begin{pmatrix} d_1 & 0 & -A & B \\ 0 & 0 & B & A \\ 0 & 0 & C & D \\ 0 & 0 & D & E \end{pmatrix}, \quad \tilde{I}_2 := \begin{pmatrix} 0 & 0 & -B & -A \\ d_1 & 0 & -A & B \\ 0 & 0 & -D & -E \\ 0 & 0 & C & D \end{pmatrix},$$

$$\tilde{I}_3 := \begin{pmatrix} 0 & 0 & -C & -D \\ 0 & 0 & D & E \\ d_1 & 0 & -A & B \\ 0 & 0 & -B & -A \end{pmatrix}, \quad \tilde{I}_4 := \begin{pmatrix} 0 & 0 & -D & -E \\ 0 & 0 & -C & -D \\ 0 & 0 & B & A \\ d_1 & 0 & -A & B \end{pmatrix},$$

$$\tilde{I}_5 := \begin{pmatrix} 0 & d_1 & -B & -A \\ 0 & 0 & -A & B \\ 0 & 0 & -D & C \\ 0 & 0 & -E & D \end{pmatrix}, \quad \tilde{I}_6 := \begin{pmatrix} 0 & 0 & A & -B \\ 0 & d_1 & -B & -A \\ 0 & 0 & E & -D \\ 0 & 0 & -D & C \end{pmatrix},$$

$$\tilde{I}_7 := \begin{pmatrix} 0 & 0 & D & -C \\ 0 & 0 & -E & D \\ 0 & d_1 & -B & -A \\ 0 & 0 & A & -B \end{pmatrix}, \quad \tilde{I}_8 := \begin{pmatrix} 0 & 0 & E & -D \\ 0 & 0 & D & -C \\ 0 & 0 & -A & B \\ 0 & d_1 & -B & -A \end{pmatrix}.$$

Case 2: $d_2 > 0$*:*

$$\tilde{J}_1 := \begin{pmatrix} D & d_2 & 0 & A \\ -F & 0 & 0 & B \\ -B & 0 & 0 & C \\ -A & 0 & 0 & D \end{pmatrix}, \quad \tilde{J}_2 := \begin{pmatrix} F & 0 & 0 & -B \\ D & d_2 & 0 & A \\ A & 0 & 0 & -D \\ -B & 0 & 0 & C \end{pmatrix},$$

$$\tilde{J}_3 := \begin{pmatrix} B & 0 & 0 & -C \\ -A & 0 & 0 & D \\ D & d_2 & 0 & A \\ F & 0 & 0 & -B \end{pmatrix}, \quad \tilde{J}_4 := \begin{pmatrix} A & 0 & 0 & -D \\ B & 0 & 0 & -C \\ -F & 0 & 0 & B \\ D & d_2 & 0 & A \end{pmatrix},$$

$$\tilde{J}_5 := \begin{pmatrix} -A & 0 & d_2 & D \\ -B & 0 & 0 & -F \\ -C & 0 & 0 & -B \\ -D & 0 & 0 & -A \end{pmatrix}, \quad \tilde{J}_6 := \begin{pmatrix} B & 0 & 0 & F \\ -A & 0 & d_2 & D \\ D & 0 & 0 & A \\ -C & 0 & 0 & -B \end{pmatrix},$$

$$\tilde{J}_7 := \begin{pmatrix} C & 0 & 0 & B \\ -D & 0 & 0 & -A \\ -A & 0 & d_2 & D \\ B & 0 & 0 & F \end{pmatrix}, \quad \tilde{J}_8 := \begin{pmatrix} D & 0 & 0 & A \\ C & 0 & 0 & B \\ -B & 0 & 0 & -F \\ -A & 0 & d_2 & D \end{pmatrix}.$$

9. Appendix 2: A list of singular cases for $\lambda_3(\mathbf{x}) := \mathbf{ax} + \mathbf{bxc} + \mathbf{xd}$. According to Theorem 3.3, formula (3.19) $\lambda_3(x)$ has the matrix equivalent

$$\lambda_3(x) = (\imath_1(a) + \imath_1(b)\imath_2(c) + \imath_2(d))\mathbf{x} =: \mathbf{M}\mathbf{x}.$$

We list some cases where the matrix $\mathbf{M}$ is singular. We found these examples by searching random quaternions a, b, c, d with integer entries uniformly distributed in $[-10, 10]$. There was about one example in 10^6 cases. We used the random number generator `rand` of `matlab` version `7.2.0.283`.

1. $a = (-9, 4, 10, -2), b = (-2, -2, -2, 1), c = (0, -3, -2, 1), d = (0, 2, 3, 4)$;
2. $a = (4, 7, 10, 9), b = (-2, 2, -3, 5), c = (1, 0, 0, -2), d = (2, 3, 8, 5)$;
3. $a = (-7, 3, -10, -3), b = (1, 2, 1, -7), c = (1, 0, 1, -2), d = (-6, 2, -4, -7)$;
4. $a = (4, 6, 1, 2), b = -(1, 1, 1, 1), c = (-1, -1, 3, 2), d = (-5, -7, -4, 3)$;
5. $a = (-3, -2, -3, 10), b = (-3, 1, -4, -3), c = (2, 0, 1, -2), d = (0, 8, 3, -8)$;
6. $a = -(6, 8, 8, 3), b = (-2, -1, 1, -1), c = -(3, 2, 4, 3), d = (-7, -4, 1, -3)$;
7. $a = (-8, 2, 5, -9), b = (5, -2, -2, -4), c = (1, -1, 2, 2), d = (-7, 5, -7, 7)$;
8. $a = (8, -6, 7, -10), b = (-3, -4, 1, 3), c = (0, 2, 3, -2), d = (7, 5, 8, 0)$;
9. $a = (2, 4, -2, 3), b = (-4, -3, 0, 3), c = (1, 0, 3, 0), d = (5, 7, 9, -4)$;
10. $a = (10, 9, 6, 5), b = (-2, 2, 2, -2), c = (0, 0, 1, 5), d = (2, -3, 6, -1)$;
11. $a = (-6, 7, -5, -4), b = (5, 1, -6, 5), c = (0, 0, 1, -1), d = (-1, 4, -5, 8)$;
12. $a = (8, 2, 1, -9), b = (3, -4, -7, 2), c = (0, 1, 0, 2), d = (2, -1, -8, -5)$;
13. $a = (7, 0, 5, 0), b = (-5, -2, -1, 3), c = (0, 1, -1, 0), d = (-10, 4, -5, -1)$;
14. $a = (4, 4, 5, -6), b = (0, 0, -1, 1), c = (7, 4, -2, 6), d = (4, 0, 6, 3)$;
15. $a = (0, 6, 4, -3), b = (0, 3, -2, 0), c = (1, -1, -1, 4), d = (-10, 1, -8, 1)$;
16. $a = (-5, -4, -5, 5), b = (-1, 10, -1, 0), c = (1, -1, -1, 1), d = -(5, 5, 8, 3)$.

In the non singular cases, the determinant is usually large. But within the many investigated cases we found one case where the determinant is equal to one. This is

$$a = (6, -3, -4, -7), b = (-4, 4, -1, 5), c = (1, -1, -2, -1), d = (7, -4, -9, 7).$$

Acknowledgment. The authors acknowledge with pleasure the support of the Grant Agency of the Czech Republic (grant No. 201/06/0356). The work is a part of the research project MSM 6046137306 financed by MSMT, Ministry of Education, Youth and Sports, Czech Republic. The second author would like to thank Prof. KP Hadeler (Phoenix, AZ and Tübingen) for a fruitful discussion on the subject of Section 3.

REFERENCES

[1] S. BANACH, Théorie des opérateurs linéaires, Warschau, 1932.
[2] J. FAN, *Determinants and multiplicative functionals on quaternionic matrices*, Linear Algebra Appl. **369** (2003), 193–201.
[3] R. A. HORN & C. R. JOHNSON, Matrix Analysis, Cambridge University Press, Cambridge, 1992, 561p.
[4] R. A. HORN & C. R. JOHNSON, Topics in Matrix Analysis, Cambridge University Press, Cambridge, 1991, 607p.

[5] L. HUANG, *The matrix equation $AXB - GXD = E$ over the quaternion field*, Linear Algebra Appl. **234** (1996), 197–208.
[6] D. JANOVSKÁ & G. OPFER, *Givens' transformation applied to quaternion valued vectors*, BIT **43** (2003), Suppl., 991–1002.
[7] D. JANOVSKÁ & G. OPFER, *Fast Givens Transformation for Quaternionic Valued Matrices Applied to Hessenberg Reductions*, Electron. Trans. Numer. Anal. **20** (2005), 1–26.
http://etna.math.kent.edu/vol.20.2005/pp1-26.dir/pp1-26.html.
[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] 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] O. ORE, *Linear equations in non-commutative fields*, Ann. of Math., 2nd series, **32** (1931), 463–477.
[12] B. L. VAN DER WAERDEN, Algebra, 5. Aufl., Springer, Berlin, Göttingen, Heidelberg, 1960, 292 p.