International Journal of Mathematics and
Computer Science, **10**(2015), no. 2, 139–156

On the Higher Order Sylvester quaternion equation $aq^n + q^n b = c$

Andreas Boukas[1], Anargyros Fellouris[2]

[1]Centro Vito Volterra
Università di Roma Tor Vergata
via Columbia 2, 00133 Roma, Italy.

[2]Department of Mathematics
Faculty of Applied Sciences
National Technical University of Athens
Zografou Campus, 157 80, Athens, Greece.

email: andreasboukas@yahoo.com, afellou@math.ntua.gr

(Received April 14, 2015, Accepted June 21, 2015)

Abstract

We study the existence of quaternion solutions q of the Higher Order Sylvester equation $aq^n + q^n b = c$ where $n = 2, 3, ...$ and a, b, c are quaternions. In the cases when solutions exist an explicit formula for them is provided using both the Hamilton and matrix representation of quaternions. The recursive solution of polynomial quaternion equations and in particular of the Higher Order Sylvester equation using the Newton-Raphson scheme for non-linear systems of equations is also discussed.

1 Introduction

A set F equipped with two operations " $+$ " and " $\cdot$ " is called a *skew-field* if the pair $(F, +)$ is an abelian group and the pair $(F^*, \cdot)$ is a non-abelian group. If the pair $(F^*, \cdot)$ is an abelian group then F is a *field*. The operations $+$ and

Key words and phrases: Sylvester equation, quaternion, quaternion equation, n-th root of a quaternion, Newton-Raphson method, non-linear systems of equations.
AMS (MOS) Subject Classifications: 11R52, 12E15, 12Y05, 15A66.
ISSN 1814-0432, 2015, http://ijmcs.future-in-tech.net

$\cdot$ must be compatible, i.e., connected through distributivity, and the additive and the multiplicative identities $\mathbf{0}$ and $\mathbf{1}$ must be distinct. The notation F^* means $F \setminus \{\mathbf{0}\}$. For example, $\mathbb{R}$ with the usual addition and multiplication of real numbers is a field. Similarly $\mathbb{R}^2$ is a field with the usual vector addition in $\mathbb{R}^2$ and multiplication

$$(a, b) \cdot (c, d) = (ac - bd, ad + bc)$$

corresponding to the identification of the pairs (a, b) and (c, d) with the complex numbers $a + ib$ and $c + id$ respectively.

In an effort to extend the properties of a field to $\mathbb{R}^n$, where $n > 2$, in 1843 William Rowan Hamilton while studying the case $n = 4$ identified the vector $(a, b, c, d) \in \mathbb{R}^4$ with the *quaternion* $a\mathbf{1} + b\mathbf{i} + c\mathbf{j} + d\mathbf{k}$ and defined the (in general noncommutative product) of the vectors $(a, b, c, d), (A, B, C, D) \in \mathbb{R}^4$, as the vector in $\mathbb{R}^4$ that is obtained by writing (a, b, c, d) and (A, B, C, D) as the quaternions $a\mathbf{1} + b\mathbf{i} + c\mathbf{j} + d\mathbf{k}$ and $A\mathbf{1} + B\mathbf{i} + C\mathbf{j} + D\mathbf{k}$ respectively, and multiplying them using the distributive property and the rules

$$\mathbf{i^2} = \mathbf{j^2} = \mathbf{k^2} = \mathbf{ijk} = \mathbf{-1}$$

Equipped with the usual vector addition and quaternion multiplication $\mathbb{R}^4$ becomes a skew-field denoted by $\mathbb{H}$. The conjugate, the modulus, and the multiplicative inverse of the quaternion $q = a\mathbf{1} + b\mathbf{i} + c\mathbf{j} + d\mathbf{k}$ are defined, respectively, by

$$\bar{q} = a\mathbf{1} - b\mathbf{i} - c\mathbf{j} - d\mathbf{k}\ ,\ \ |q| = \sqrt{a^2 + b^2 + c^2 + d^2}\ ,\ \ q^{-1} = \frac{\bar{q}}{|q|^2}\ ,\ |q| \neq 0$$

If $|q| = 1$ then q is a *unit quaternion.* If $a = 0$ then $q = b\mathbf{i} + c\mathbf{j} + d\mathbf{k}$ is an *imaginary quaternion.*

Quaternions are used in the description of rotations in $\mathbb{R}^3$, i.e., in the description of the elements of the Lie group $SO(3)$. If $u = b\mathbf{i} + c\mathbf{j} + d\mathbf{k} = (b, c, d) \in \mathbb{R}^3$ is a unit imaginary quaternion and $\theta \in \mathbb{R}$ is an angle, then the *conjugation* $q \mapsto t^{-1}qt$, where $t = \cos\theta + u \sin\theta = \cos\theta\mathbf{1} + b\sin\theta\mathbf{i} + c\sin\theta\mathbf{j} + d\sin\theta\mathbf{k}$, describes the rotation of the vector $q = q_1\mathbf{i} + q_2\mathbf{j} + q_3\mathbf{k} = (q_1, q_2, q_3) \in \mathbb{R}^3$ by an angle 2θ around the axis u passing through the origin.

Each quaternion $q = a\mathbf{1} + b\mathbf{i} + c\mathbf{j} + d\mathbf{k}$ can be represented as the (2×2) complex matrix

$$q = \begin{pmatrix} a + id & -b - ic \\ b - ic & a - id \end{pmatrix} = \begin{pmatrix} z & -w \\ \bar{w} & \bar{z} \end{pmatrix} \tag{1.1}$$

or as

$$q = \begin{pmatrix} a+ib & c+id \\ -c+id & a-ib \end{pmatrix} = \begin{pmatrix} z & w \\ -\bar{w} & \bar{z} \end{pmatrix} \tag{1.2}$$

where $i^2 = -1$.

In recent years there has been an increasing interest in the extension and the formulation in the context of quaternions of problems and results usually obtained within real and complex numbers, see for example [11, 8] and the references within. Of particular interest has been the study of quaternion equations extending classical algebraic equations such as the *Algebraic Riccati Equation* and especially the *Sylvester Equation* $AX + XB = C$ of control theory [11, 12, 13, 2] where A, B, C, X are real $(n \times n)$ matrices. In view of the representation of quaternions as complex (2×2) matrices, the quaternion version $aq + qb = c$ of the classical Sylvester equation is an extension to the case of (2×2) complex matrices of the form (1.1) and (1.2). Notice that in the above matrix representation: $|q|^2 = \det q$, $\bar{q}$ is the Hermitian adjoint of q and q^{-1} is the matrix inverse of q. For general quadratic quaternion equations see, for example, [6] and [11]).

In this note we start in Section 2 with a proof of the classic quaternion power formula using the matrix representation of a quaternion. We also provide a *vector form* of that formula. In Sections 3 and 4 we study the solutions of the *Higher Order Sylvester Quaternion Equation*

$$aq^n + q^n b = c \tag{1.3}$$

by first reducing it to the first order quaternion Sylvester equation $ap+pb = c$ through the transformation $p = q^n$ and then proceeding in two ways:

- In section 3 we obtain the solution to $ap+pb = c$ using the formula given in [3] for the (unique) solution of the Sylvester equation $ap + pb = c$ in Hamilton's quaternion representation. The validity of this formula is discussed in [11], Section 2.3 and in [3] Theorem 2.3. We then use the results of the classic paper [9] for the n-th root of a quaternion to get the solutions of $aq^n + q^n b = c$.

- In section 4 we study $ap + pb = c$ using the matrix representation of quaternions and Gaussian elimination and then solve $aq^n + q^n b = c$ using the theory for finding roots of matrices.

Finally, several authors have addressed the numerical solution of polynomial quaternion equations, see for example [5], [15], [11] and the references within. In section 5 we briefly discuss the use of the Newton-Raphson method for the numerical solution of general polynomial quaternion equations and in particular of $aq^n + q^n b = c$. The Newton-Raphson method was used and studied in detail in [4] in order to find the n-th root of a quaternion.

Throughout this paper, numerical and complicated symbolic calculations were done with the use of Mathematica version 9.

2 Non-negative integer powers of a quaternion

The following Proposition is known [11]. Here we provide a proof based on the matrix representation of quaternions.

Proposition 2.1. *Let* $n \in 0, 1, 2, ...$ *and* $q = a\mathbf{1} + b\mathbf{i} + c\mathbf{j} + d\mathbf{k} = a\mathbf{1} + \vec{q} \in \mathbb{H}$. *Then*

$$q^n = |q|^n \left(\cos(n\theta) + \sin(n\theta)\,\hat{\vec{q}}\right) \tag{2.1}$$

where

$$\theta = \arg(a + i|\vec{q}|) \ , \ \hat{\vec{q}} = \frac{1}{|\vec{q}|}\vec{q} \tag{2.2}$$

Proof. Let q have the matrix form (1.2). Since the set of matrices of the form (1.2) is closed under matrix multiplication and addition, q^n will also have the matrix and equivalent Hamilton representation

$$q^n = \begin{pmatrix} Z & W \\ -\bar{W} & \bar{Z} \end{pmatrix} \ , \ q^n = A\mathbf{1} + B\mathbf{i} + C\mathbf{j} + D\mathbf{k}$$

where

$$A = \frac{Z + \bar{Z}}{2} \ , \ B = \frac{Z - \bar{Z}}{2i} \ , \ C = \frac{W + \bar{W}}{2} \ , \ D = \frac{W - \bar{W}}{2i}$$

Computing the matrix products we find that

$$Z = \frac{(a - i|\vec{q}|)^n + (a + i|\vec{q}|)^n}{2} - \frac{b}{2|\vec{q}|}((a - i|\vec{q}|)^n - (a + i|\vec{q}|)^n)$$

$$W = \frac{c + id}{2i|\vec{q}|}((a + i|\vec{q}|)^n - (a - i|\vec{q}|)^n)$$

If $z = |z|e^{i\theta}$ is a complex number then

$$z^n + \bar{z}^n = 2|z|^n \cos(n\theta) \ , \ z^n - \bar{z}^n = 2i|z|^n \sin(n\theta)$$

Thus

$$Z = |a + i|\vec{q}||^n \cos(n\theta) + i\frac{b}{|\vec{q}|}|a + i|\vec{q}||^n \sin(n\theta) = |q|^n \cos(n\theta) + i\frac{b|q|^n}{|\vec{q}|}\sin(n\theta)$$

and

$$W = \frac{c|q|^n}{|\vec{q}|}\sin(n\theta) + i\frac{d|q|^n}{|\vec{q}|}\sin(n\theta)$$

where

$$\theta = \arg(a + i|\vec{q}|)$$

Thus

$$q^n = |q|^n \cos(n\theta) + \frac{b|q|^n}{|\vec{q}|}\sin(n\theta)\mathbf{i} + \frac{c|q|^n}{|\vec{q}|}\sin(n\theta)\mathbf{j} + \frac{d|q|^n}{|\vec{q}|}\sin(n\theta)\mathbf{k}$$

i.e.,

$$q^n = |q|^n \left(\cos(n\theta) + \sin(n\theta)\,\frac{1}{|\vec{q}|}\vec{q}\right) = |q|^n \left(\cos(n\theta) + \sin(n\theta)\,\hat{\vec{q}}\right)$$

□

Remark 2.2. *The result in the above Proposition is in agreement with the Lemma in Niven's paper [9] which states that "any positive integral power of a quaternion* $q = a\mathbf{1} + b\mathbf{i} + c\mathbf{j} + d\mathbf{k}$ *has coefficients of* $\mathbf{i}, \mathbf{j}$ *and* $\mathbf{k}$ *which are proportional to* b, c *and* d*".*

Corollary 2.3. *In the notation of Proposition 2.1*

$$q^n = \sum_{k=0}^{n} \binom{n}{k} a^k |\vec{q}|^{n-k} \left(\cos\left(\frac{(n-k)\pi}{2}\right) + \sin\left(\frac{(n-k)\pi}{2}\right)\hat{\vec{q}}\right) \tag{2.3}$$

Proof. Using Vieta's formulas [14]

$$\sin n\theta = \sum_{k=0}^{n} \binom{n}{k} \cos^k\theta \sin^{n-k}\theta \sin\left(\frac{(n-k)\pi}{2}\right)$$

$$\cos n\theta = \sum_{k=0}^{n} \binom{n}{k} \cos^k\theta \sin^{n-k}\theta \cos\left(\frac{(n-k)\pi}{2}\right)$$

and the fact that

$$\theta = \arg(a + i|\vec{q}|) \implies \cos\theta = \frac{a}{|q|} \ , \ \sin\theta = \frac{|\vec{q}|}{|q|}$$

we find

$$q^n = |q|^n(\sum_{k=0}^{n} \binom{n}{k} (\frac{a}{|q|})^k (\frac{|\vec{q}|}{|q|})^{n-k} \cos(\frac{(n-k)\pi}{2}) + \sum_{k=0}^{n} \binom{n}{k} (\frac{a}{|q|})^k (\frac{|\vec{q}|}{|q|})^{n-k} \sin(\frac{(n-k)\pi}{2}) \frac{1}{|\vec{q}|}\vec{q})$$

$$= \sum_{k=0}^{n} \binom{n}{k} a^k |\vec{q}|^{n-k} \cos\left(\frac{(n-k)\pi}{2}\right) + \sum_{k=0}^{n} \binom{n}{k} a^k |\vec{q}|^{n-k} \sin\left(\frac{(n-k)\pi}{2}\right) \frac{1}{|\vec{q}|}\vec{q}$$

$$= \sum_{k=0}^{n} \binom{n}{k} a^k |\vec{q}|^{n-k} \left(\cos\left(\frac{(n-k)\pi}{2}\right) + \sin\left(\frac{(n-k)\pi}{2}\right) \frac{1}{|\vec{q}|}\vec{q}\right)$$

$$= \sum_{k=0}^{n} \binom{n}{k} a^k |\vec{q}|^{n-k} \left(\cos\left(\frac{(n-k)\pi}{2}\right) + \sin\left(\frac{(n-k)\pi}{2}\right) \hat{\vec{q}}\right)$$

□

3 Solution of $aq^n + q^n b = c$ using Hamilton's formulation

Letting $p = q^n$ in $aq^n + q^n b = c$ it becomes $ap + pb = c$ which is the classical quaternion Sylvester equation. Assuming that $ab \neq \mathbf{0}$, it was shown in [3] that a quaternion solution p exists if and only if $|a| \neq |b|$ and/or $\text{Re } a \neq -\text{Re } b$. The solution is unique and is given by either one of the following formulas

$$p = \left(2\text{Re } b + a + |b|^2 a^{-1}\right)^{-1} \left(c + a^{-1} c \bar{b}\right) \tag{3.1}$$

$$p = \left(c + \bar{a} c \, b^{-1}\right) \left(2\text{Re } a + b + |a|^2 b^{-1}\right)^{-1} \tag{3.2}$$

Using, for example, the first one of these formulas and assuming that

$$a = a_1\mathbf{1} + a_2\mathbf{i} + a_3\mathbf{j} + a_4\mathbf{k} \ , \ b = b_1\mathbf{1} + b_2\mathbf{i} + b_3\mathbf{j} + b_4\mathbf{k} \ , \ c = c_1\mathbf{1} + c_2\mathbf{i} + c_3\mathbf{j} + c_4\mathbf{k}$$

we find that

$$p = \left(2\text{Re } b + a + |b|^2 a^{-1}\right)^{-1} \left(c + a^{-1} c \bar{b}\right) = p_1\mathbf{1} + p_2\mathbf{i} + p_3\mathbf{j} + p_4\mathbf{k} \tag{3.3}$$

where

$$p_1 = \frac{p_{11}}{p_{12}}\ ,\ p_2 = \frac{p_{21}}{p_{22}}\ ,\ p_3 = \frac{p_{31}}{p_{32}}\ ,\ p_4 = \frac{p_{41}}{p_{42}} \tag{3.4}$$

with

$$p_{11} = a_1^3c_1+a_1a_2^2c_1+a_1a_3^2c_1+a_1a_4^2c_1+3a_1^2b_1c_1+a_2^2b_1c_1+a_3^2b_1c_1+a_4^2b_1c_1+3a_1b_1^2c_1+b_1^3c_1$$
$$-2a_1a_2b_2c_1-2a_2b_1b_2c_1+a_1b_2^2c_1+b_1b_2^2c_1-2a_1a_3b_3c_1-2a_3b_1b_3c_1+a_1b_3^2c_1+b_1b_3^2c_1-2a_1a_4b_4c_1-2a_4b_1b_4c_1$$
$$+a_1b_4^2c_1+b_1b_4^2c_1+a_1^2a_2c_2+a_2^3c_2+a_2a_3^2c_2+a_2a_4^2c_2+2a_1a_2b_1c_2+a_2b_1^2c_2+a_1^2b_2c_2-a_2^2b_2c_2-a_3^2b_2c_2$$
$$-a_4^2b_2c_2+2a_1b_1b_2c_2+b_1^2b_2c_2-a_2b_2^2c_2+b_2^3c_2-2a_1a_4b_3c_2-2a_4b_1b_3c_2-a_2b_3^2c_2+b_2b_3^2c_2+2a_1a_3b_4c_2$$
$$+2a_3b_1b_4c_2-a_2b_4^2c_2+b_2b_4^2c_2+a_1^2a_3c_3+a_2^2a_3c_3+a_3^3c_3+a_3a_4^2c_3+2a_1a_3b_1c_3+a_3b_1^2c_3+2a_1a_4b_2c_3$$
$$+2a_4b_1b_2c_3-a_3b_2^2c_3+a_1^2b_3c_3-a_2^2b_3c_3-a_3^2b_3c_3-a_4^2b_3c_3+2a_1b_1b_3c_3+b_1^2b_3c_3+b_2^2b_3c_3-a_3b_3^2c_3$$
$$+b_3^3c_3-2a_1a_2b_4c_3-2a_2b_1b_4c_3-a_3b_4^2c_3+b_3b_4^2c_3+a_1^2a_4c_4+a_2^2a_4c_4+a_3^2a_4c_4+a_4^3c_4+2a_1a_4b_1c_4$$
$$+a_4b_1^2c_4-2a_1a_3b_2c_4-2a_3b_1b_2c_4-a_4b_2^2c_4+2a_1a_2b_3c_4+2a_2b_1b_3c_4-a_4b_3^2c_4+a_1^2b_4c_4-a_2^2b_4c_4$$
$$-a_3^2b_4c_4 - a_4^2b_4c_4 + 2a_1b_1b_4c_4 + b_1^2b_4c_4 + b_2^2b_4c_4 + b_3^2b_4c_4 - a_4b_4^2c_4 + b_4^3c_4$$

$$p_{21} = -a_1^2a_2c_1-a_2^3c_1-a_2a_3^2c_1-a_2a_4^2c_1-2a_1a_2b_1c_1-a_2b_1^2c_1-a_1^2b_2c_1+a_2^2b_2c_1+a_3^2b_2c_1+a_4^2b_2c_1$$
$$-2a_1b_1b_2c_1-b_1^2b_2c_1+a_2b_2^2c_1-b_2^3c_1-2a_1a_4b_3c_1-2a_4b_1b_3c_1+a_2b_3^2c_1-b_2b_3^2c_1+2a_1a_3b_4c_1+2a_3b_1b_4c_1$$
$$+a_2b_4^2c_1-b_2b_4^2c_1+a_1^3c_2+a_1a_2^2c_2+a_1a_3^2c_2+a_1a_4^2c_2+3a_1^2b_1c_2+a_2^2b_1c_2+a_3^2b_1c_2+a_4^2b_1c_2$$
$$+3a_1b_1^2c_2+b_1^3c_2-2a_1a_2b_2c_2-2a_2b_1b_2c_2+a_1b_2^2c_2+b_1b_2^2c_2+2a_1a_3b_3c_2+2a_3b_1b_3c_2+a_1b_3^2c_2+b_1b_3^2c_2$$
$$+2a_1a_4b_4c_2+2a_4b_1b_4c_2+a_1b_4^2c_2+b_1b_4^2c_2+a_1^2a_4c_3+a_2^2a_4c_3+a_3^2a_4c_3+a_4^3c_3+2a_1a_4b_1c_3+a_4b_1^2c_3$$
$$-2a_1a_3b_2c_3-2a_3b_1b_2c_3-a_4b_2^2c_3-2a_1a_2b_3c_3-2a_2b_1b_3c_3-a_4b_3^2c_3-a_1^2b_4c_3+a_2^2b_4c_3+a_3^2b_4c_3+a_4^2b_4c_3$$
$$-2a_1b_1b_4c_3-b_1^2b_4c_3-b_2^2b_4c_3-b_3^2b_4c_3-a_4b_4^2c_3-b_4^3c_3-a_1^2a_3c_4-a_2^2a_3c_4-a_3^3c_4-a_3a_4^2c_4$$
$$-2a_1a_3b_1c_4-a_3b_1^2c_4-2a_1a_4b_2c_4-2a_4b_1b_2c_4+a_3b_2^2c_4+a_1^2b_3c_4-a_2^2b_3c_4-a_3^2b_3c_4-a_4^2b_3c_4+2a_1b_1b_3c_4$$
$$+b_1^2b_3c_4 + b_2^2b_3c_4 + a_3b_3^2c_4 + b_3^3c_4 - 2a_1a_2b_4c_4 - 2a_2b_1b_4c_4 + a_3b_4^2c_4 + b_3b_4^2c_4$$

$$p_{31} = -a_1^2a_2c_1-a_2^3c_1-a_2a_3^2c_1-a_2a_4^2c_1-2a_1a_2b_1c_1-a_2b_1^2c_1-a_1^2b_2c_1+a_2^2b_2c_1+a_3^2b_2c_1+a_4^2b_2c_1$$
$$-2a_1b_1b_2c_1-b_1^2b_2c_1+a_2b_2^2c_1-b_2^3c_1-2a_1a_4b_3c_1-2a_4b_1b_3c_1+a_2b_3^2c_1-b_2b_3^2c_1+2a_1a_3b_4c_1+2a_3b_1b_4c_1$$
$$+a_2b_4^2c_1-b_2b_4^2c_1+a_1^3c_2+a_1a_2^2c_2+a_1a_3^2c_2+a_1a_4^2c_2+3a_1^2b_1c_2+a_2^2b_1c_2+a_3^2b_1c_2+a_4^2b_1c_2$$
$$+3a_1b_1^2c_2+b_1^3c_2-2a_1a_2b_2c_2-2a_2b_1b_2c_2+a_1b_2^2c_2+b_1b_2^2c_2+2a_1a_3b_3c_2+2a_3b_1b_3c_2+a_1b_3^2c_2+b_1b_3^2c_2$$
$$+2a_1a_4b_4c_2+2a_4b_1b_4c_2+a_1b_4^2c_2+b_1b_4^2c_2+a_1^2a_4c_3+a_2^2a_4c_3+a_3^2a_4c_3+a_4^3c_3+2a_1a_4b_1c_3+a_4b_1^2c_3$$
$$-2a_1a_3b_2c_3-2a_3b_1b_2c_3-a_4b_2^2c_3-2a_1a_2b_3c_3-2a_2b_1b_3c_3-a_4b_3^2c_3-a_1^2b_4c_3+a_2^2b_4c_3+a_3^2b_4c_3+a_4^2b_4c_3$$

$-2a_1b_1b_4c_3-b_1^2b_4c_3-b_2^2b_4c_3-b_3^2b_4c_3-a_4b_4^2c_3-b_4^3c_3-a_1^2a_3c_4-a_2^2a_3c_4-a_3^3c_4-a_3a_4^2c_4$
$-2a_1a_3b_1c_4-a_3b_1^2c_4-2a_1a_4b_2c_4-2a_4b_1b_2c_4+a_3b_2^2c_4+a_1^2b_3c_4-a_2^2b_3c_4-a_3^2b_3c_4-a_4^2b_3c_4$
$+2a_1b_1b_3c_4$
$+b_1^2b_3c_4+b_2^2b_3c_4+a_3b_3^2c_4+b_3^3c_4-2a_1a_2b_4c_4-2a_2b_1b_4c_4+a_3b_4^2c_4+b_3b_4^2c_4$

$p_{41}=-a_1^2a_4c_1-a_2^2a_4c_1-a_3^2a_4c_1-a_4^3c_1-2a_1a_4b_1c_1-a_4b_1^2c_1-2a_1a_3b_2c_1-2a_3b_1b_2c_1$
$+a_4b_2^2c_1+2a_1a_2b_3c_1+2a_2b_1b_3c_1+a_4b_3^2c_1-a_1^2b_4c_1+a_2^2b_4c_1+a_3^2b_4c_1+a_4^2b_4c_1-2a_1b_1b_4c_1-b_1^2b_4c_1$
$-b_2^2b_4c_1-b_3^2b_4c_1+a_4b_4^2c_1-b_4^3c_1+a_1^2a_3c_2+a_2^2a_3c_2+a_3^3c_2+a_3a_4^2c_2+2a_1a_3b_1c_2+a_3b_1^2c_2$
$-2a_1a_4b_2c_2-2a_4b_1b_2c_2-a_3b_2^2c_2-a_1^2b_3c_2+a_2^2b_3c_2+a_3^2b_3c_2+a_4^2b_3c_2-2a_1b_1b_3c_2-b_1^2b_3c_2-b_2^2b_3c_2$
$-a_3b_3^2c_2-b_3^3c_2-2a_1a_2b_4c_2-2a_2b_1b_4c_2-a_3b_4^2c_2-b_3b_4^2c_2-a_1^2a_2c_3-a_2^3c_3-a_2a_3^2c_3-a_2a_4^2c_3$
$-2a_1a_2b_1c_3-a_2b_1^2c_3+a_1^2b_2c_3-a_2^2b_2c_3-a_3^2b_2c_3-a_4^2b_2c_3+2a_1b_1b_2c_3+b_1^2b_2c_3+a_2b_2^2c_3$
$+b_2^3c_3-2a_1a_4b_3c_3-2a_4b_1b_3c_3+a_2b_3^2c_3+b_2b_3^2c_3-2a_1a_3b_4c_3-2a_3b_1b_4c_3+a_2b_4^2c_3+b_2b_4^2c_3+a_1^3c_4$
$+a_1a_2^2c_4+a_1a_3^2c_4+a_1a_4^2c_4+3a_1^2b_1c_4+a_2^2b_1c_4+a_3^2b_1c_4+a_4^2b_1c_4+3a_1b_1^2c_4+b_1^3c_4+2a_1a_2b_2c_4$
$+2a_2b_1b_2c_4+a_1b_2^2c_4+b_1b_2^2c_4+2a_1a_3b_3c_4+2a_3b_1b_3c_4+a_1b_3^2c_4+b_1b_3^2c_4-2a_1a_4b_4c_4$
$-2a_4b_1b_4c_4+a_1b_4^2c_4+b_1b_4^2c_4$

$p_{12}=p_{22}=p_{32}=p_{42}=a_1^4+2a_1^2a_2^2+a_2^4+2a_1^2a_3^2+2a_2^2a_3^2$
$+a_3^4+2a_1^2a_4^2+2a_2^2a_4^2+2a_3^2a_4^2+a_4^4+4a_1^3b_1+4a_1a_2^2b_1+4a_1a_3^2b_1+4a_1a_4^2b_1+6a_1^2b_1^2$
$+2a_2^2b_1^2+2a_3^2b_1^2+2a_4^2b_1^2+4a_1b_1^3+b_1^4+2a_1^2b_2^2-2a_2^2b_2^2-2a_3^2b_2^2-2a_4^2b_2^2+4a_1b_1b_2^2$
$+2b_1^2b_2^2+b_2^4+2a_1^2b_3^2-2a_2^2b_3^2-2a_3^2b_3^2-2a_4^2b_3^2+4a_1b_1b_3^2+2b_1^2b_3^2+2b_2^2b_3^2+b_3^4$
$+2a_1^2b_4^2-2a_2^2b_4^2-2a_3^2b_4^2-2a_4^2b_4^2+4a_1b_1b_4^2+2b_1^2b_4^2+2b_2^2b_4^2+2b_3^2b_4^2+b_4^4$

For $q=p^{1/n}\in\mathbb{H}$ there are two cases as in [9]:

1. $\operatorname{Im} p=p_2\mathbf{i}+p_3\mathbf{j}+p_4\mathbf{k}\neq 0\mathbf{i}+0\mathbf{j}+0\mathbf{k}$. In this case there are exactly n distinct quaternion solutions q of (1.3).

2. $\operatorname{Im} p=p_2\mathbf{i}+p_3\mathbf{j}+p_4\mathbf{k}=0\mathbf{i}+0\mathbf{j}+0\mathbf{k}$, i.e., $p=p_1\in\mathbb{R}$. Here there are two subcases:

   (a) $n=2$ and $p=p_1>0$. There are only two quaternion solutions q of (1.3), namely $q=\pm\sqrt{p_1}$ (the usual square roots of a real number).

(b) $n \neq 2$ and/or $p = p_1 < 0$. There are infinitely many quaternion solutions q of (1.3).

In each case, the quaternion solutions q of (1.3) are obtained as follows: In Case (1): Let $t = \mathrm{Tr}\, p := 2p_1$ and $N = |p|$. Since p satisfies the *minimal equation* $\xi^2 - t\xi + N^2 = 0$ (note: here there is a misprint in [9] where the square in N^2 is missing) and p is not real, following [9], we let

$$\lambda := \frac{t + i\sqrt{-\Delta}}{2} = p_1 + i\sqrt{|p|^2 - p_1^2} = p_1 + i\sqrt{p_2^2 + p_3^2 + p_4^2} \tag{3.5}$$

be the root with positive imaginary coordinate. Then, for fixed $n \in \{2, 3, ...\}$, writing as in [9]

$$\lambda = r^n(\cos n\theta + i \sin n\theta)\ ,\ 0 < n\theta < \pi \tag{3.6}$$

where

$$r^n = |\lambda| = |p| \implies r = |p|^{\frac{1}{n}} = (p_1^2 + p_2^2 + p_3^2 + p_4^2)^{\frac{1}{2n}} > 0 \tag{3.7}$$

and

$$n\theta = \arg\lambda \implies \theta = \frac{1}{n}\arg\lambda = \begin{cases} \frac{1}{n}\arctan\left(\frac{\sqrt{p_2^2+p_3^2+p_4^2}}{p_1}\right) & \text{if } \mathrm{Re}\,\lambda = \mathrm{p}_1 > 0 \\ \frac{1}{n}\left(\arctan\left(\frac{\sqrt{p_2^2+p_3^2+p_4^2}}{p_1}\right) + \pi\right) & \text{if } \mathrm{Re}\,\lambda = \mathrm{p}_1 < 0 \end{cases} \tag{3.8}$$

and using Theorem 1 of [9] we find that the n quaternion solutions q of (1.3) are given by

$$q = p^{\frac{1}{n}} = \frac{1}{r^{n-1}}\frac{\sin\left(\theta + \frac{2\pi K}{n}\right)}{\sin n\theta}p + r\left(\cos\left(\theta + \frac{2\pi K}{n}\right) - \cot n\theta \sin\left(\theta + \frac{2\pi K}{n}\right)\right)\ ,\ K = 0, 1, ..., n-1 \tag{3.9}$$

In Case (2), using Theorem 2 of [9] we find that the quaternion solutions q of (1.3) are given by

$$q = p^{\frac{1}{n}} = r\cos\frac{s\pi + 2\pi K}{n}\mathbf{1} + y\mathbf{i} + z\mathbf{j} + w\mathbf{k}\ ,\ K = 0, 1, ..., n-1 \tag{3.10}$$

where r is in (3.7) and $y, z, w \in \mathbb{R}$ are, for each choice of K, any real numbers satisfying

$$y^2 + z^2 + w^2 = r^2 \sin^2\frac{s\pi + 2\pi K}{n}$$

where

$$s = \begin{cases} 0 & \text{if } p = p_1 > 0 \\ 1 & \text{if } p = p_1 < 0 \end{cases}$$

As pointed out in [9], if $s = 0$ and $n = 2$ then

$$y^2 + z^2 + w^2 = r^2 \sin^2 \frac{s\pi + 2\pi K}{n} \implies y^2 + z^2 + w^2 = 0 \implies y = z = w = 0$$

and the quaternion solutions q of (1.3)

$$q = p^{\frac{1}{2}} = r \cos \pi K \mathbf{1} = (-1)^K r \ , \ K = 0, 1 \tag{3.11}$$

4 Solution of $aq^n + q^n b = c$ using matrix representation

Letting $p = q^n$ where, as in (1.1),

$$p = \begin{pmatrix} z & -w \\ \bar{w} & \bar{z} \end{pmatrix}, \ a = \begin{pmatrix} A & -B \\ \bar{B} & \bar{A} \end{pmatrix}, \ b = \begin{pmatrix} C & -D \\ \bar{D} & \bar{C} \end{pmatrix}, \ c = \begin{pmatrix} E & -Z \\ \bar{Z} & \bar{E} \end{pmatrix} \tag{4.1}$$

$aq^n + q^n b = c$ is reduced to the linear complex system of equations

$$\begin{pmatrix} A + C & 0 & -\bar{D} & -B \\ -D & -B & -A - \bar{C} & 0 \\ \bar{B} & \bar{D} & 0 & \bar{A} + C \\ 0 & \bar{A} + \bar{C} & -\bar{B} & -D \end{pmatrix} \begin{pmatrix} z \\ \bar{z} \\ w \\ \bar{w} \end{pmatrix} = \begin{pmatrix} E \\ -Z \\ \bar{Z} \\ \bar{E} \end{pmatrix} \tag{4.2}$$

The system has a unique solution if and only if the determinant of the coefficient matrix M is nonzero, i.e., if and only if

$$\begin{aligned} \det M = & -A|A|^2C - |A|^2C^2 - |A|^4 - A|A|^2C - A|B|^2C - |B|^2C^2 - 2|A|^2|B|^2 \\ & -|B|^2C\bar{A} - |B|^4 - A^2|C|^2 - AC|C|^2 - A|A|^2\bar{C} - 2|A|^2|C|^2 \\ & -C|C|^2\bar{A} - |A|^2\bar{A}\bar{C} - |C|^2\bar{A}^2 - A|B|^2\bar{C} - |B|^2\bar{A}\bar{C} \\ & -A|C|^2\bar{C} - |C|^-|A|^2\bar{C}^2 - |C|^2\bar{A}\bar{C} - |B|^2\bar{C}^2 - A^2|D|^2 \\ & -AC|D|^2 - C|D|^2\bar{A} - |D|^2\bar{A}^2 + 2|B|^2|D|^2 \\ & -A|D|^2\bar{C} - 2|C|^2|D|^2 - |D|^2\bar{A}\bar{C} - |D|^4 \neq 0 \end{aligned} \tag{4.3}$$

The unique solution is

$$z = \frac{z_1}{z_2}\ ,\ w = \frac{w_1}{w_2} \tag{4.4}$$

where

$$z_1 = |A|^2CE+|A|^2E\bar{A}+|B|^2CE+|B|^2E\bar{A}+A|C|^2E+|A|^2E\bar{C}+|C|^2E\bar{A}+E\bar{A}^2\bar{C}+|C|^2E\bar{C}+E\bar{A}\bar{C}^2$$
$$+A|D|^2E+CZ\bar{A}\bar{D}+Z\bar{A}^2\bar{D}-|B|^2Z\bar{D}+|D|^2E\bar{C}+|C|^2Z\bar{D}+Z\bar{A}\bar{C}\bar{D}+|D|^2Z\bar{D}-AB\bar{D}\bar{E}-BC\bar{D}\bar{E}$$
$$-B\bar{A}\bar{D}\bar{E}-B\bar{C}\bar{D}\bar{E}+|A|^2B\bar{Z}+B|B|^2\bar{Z}+AB\bar{C}\bar{Z}+B\bar{A}\bar{C}\bar{Z}+B\bar{C}^2\bar{Z}-B|D|^2\bar{Z}$$

$$w_1 = -CDE\bar{A}+|A|^2CZ+C^2Z\bar{A}-DEE\bar{A}^2+|A|^2Z\bar{A}$$
$$+CZ\bar{A}^2+|B|^2DEE+|B|^2Z\bar{A}-|C|^2DE+A|C|^2Z+C|C|^2Z-DE\bar{A}\bar{C}+|A|^2Z\bar{C}+|C|^2Z\bar{A}$$
$$+|B|^2Z\bar{C}-D|D|^2E+A|D|^2Z+C|D|^2Z-ABC\bar{E}-BC^2\bar{E}-|A|^2B\bar{E}-BC\bar{A}\bar{E}-B|B|^2\bar{E}+B|D|^2\bar{E}$$
$$-ABD\bar{Z}-BCD\bar{Z}-BD\bar{A}\bar{Z}-BD\bar{C}\bar{Z})$$

and

$$z_2 = w_2 = \det M$$

We now consider, from the matrix point of view, the question of finding $q = p^{1/n}$. The eigenvalues of the matrix

$$p = \begin{pmatrix} z & -w \\ \bar{w} & \bar{z} \end{pmatrix}$$

are

$$\lambda = \operatorname{Re} z \pm ih \tag{4.5}$$

where

$$h = \sqrt{\det p - (\operatorname{Re} z)^2} = \sqrt{(\operatorname{Im} z)^2 + |w|^2} \tag{4.6}$$

If $h \neq 0$, meaning $\operatorname{Im} z \neq 0$ and/or $w \neq 0$, then the eigenvalues are distinct and p can be diagonalized in the form

$$p = PDP^{-1} \tag{4.7}$$

where, assuming $w \neq 0$,

$$P = \begin{pmatrix} -w & -w \\ i(h - \operatorname{Im} z) & -i(h + \operatorname{Im} z) \end{pmatrix}\ ,\ D = \begin{pmatrix} \operatorname{Re} z + ih & 0 \\ 0 & \operatorname{Re} z - ih \end{pmatrix}, P^{-1} = \begin{pmatrix} -\frac{h+\operatorname{Im} z}{2hw} & -\frac{i}{2h} \\ \frac{\operatorname{Im} z - h}{2hw} & \frac{i}{2h} \end{pmatrix} \tag{4.8}$$

Thus there are exactly n quaternion solutions q of (1.3) given by

$$q = p^{1/n} = PD^{1/n}P^{-1} \tag{4.9}$$

where

$$D^{1/n} = \begin{pmatrix} (\operatorname{Re} z + ih)^{1/n} & 0 \\ 0 & (\operatorname{Re} z - ih)^{1/n} \end{pmatrix} \tag{4.10}$$

and the $n-th$ complex roots of the main diagonal elements are computed for the same value of the parameter $k = 0, 1, ..., n-1$ appearing in the formula for complex roots, so that diagonal elements remain conjugate. If $h \neq 0$ and $w = 0$ then

$$p = \begin{pmatrix} z & 0 \\ 0 & \bar{z} \end{pmatrix}$$

has the distinct eigenvalues z and $\bar{z}$ and there are exactly n quaternion solutions q of (1.3) given by

$$q = p^{\frac{1}{n}} = \begin{pmatrix} z^{\frac{1}{n}} & 0 \\ 0 & \bar{z}^{\frac{1}{n}} \end{pmatrix} \tag{4.11}$$

where the parameter $k = 0, 1, ..., n-1$ appearing in the formula for complex roots is chosen as before. If $h = 0$ then $w = 0$ and $z \in \mathbb{R}$. Thus

$$p = z \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \Longrightarrow q = z^{\frac{1}{n}} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}^{\frac{1}{n}}$$

Then, as in Section 3, Case (2): if $n = 2$ and $p = p_1 > 0$, i.e., if $z = p_1$ then

$$p = p_1 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

and there are only two quaternion solutions q of (1.3), namely $q = \pm\sqrt{p_1}\ \mathbf{I}^{1/2}$ corresponding to the two unique quaternion square roots of the identity matrix

$$\mathbf{I}^{1/2} = \begin{pmatrix} \pm 1 & 0 \\ 0 & \pm 1 \end{pmatrix}$$

If $n = 2$ and $p = p_1 = z < 0$ then there are infinitely many quaternion solutions q of (1.3) given by

$$q = |p_1|^{1/2}(-\mathbf{I})^{1/2} \tag{4.12}$$

where $|p_1|^{1/2}$ is the positive real square root of $|p_1|$ and $(-\mathbf{I})^{1/2}$ is anyone of the infinitely many quaternion square roots of the matrix $-\mathbf{I}$ of the form

$$\begin{pmatrix} ib & c+id \\ -c+id & -ib \end{pmatrix} \ , \ b,c,d \in \mathbb{R} \ , \ b^2+c^2+d^2=1 \tag{4.13}$$

If $n \neq 2$ there are infinitely many quaternion solutions q of (1.3) of the form

$$q = p^{1/n} = \begin{pmatrix} r\cos\frac{s\pi+2\pi K}{n} + i\psi & \zeta + i\omega \\ -\zeta + i\omega & r\cos\frac{s\pi+2\pi K}{n} - i\psi \end{pmatrix} \ , \ K \in \{0,1,...,n-1\} \tag{4.14}$$

where

$$r = |p_1|^{\frac{1}{n}} > 0 \tag{4.15}$$

and $\psi, \zeta, \omega \in \mathbb{R}$ are, for each choice of K, any real numbers satisfying

$$\psi^2 + \zeta^2 + \omega^2 = r^2 \sin^2 \frac{s\pi + 2\pi K}{n} \tag{4.16}$$

where

$$s = \begin{cases} 0 & \text{if } p = p_1 > 0 \\ 1 & \text{if } p = p_1 < 0 \end{cases}$$

Example 1.

For $n = 3$ and

$$a = 1+3i-4j+k = \begin{pmatrix} 1+i & -3+4i \\ 3+4i & 1-i \end{pmatrix} \ , \ b = -2i+2j+2k = \begin{pmatrix} 2i & 2-2i \\ -2-2i & -2i \end{pmatrix}$$

$$c = -1+6i+0j+k = \begin{pmatrix} -1+i & -6 \\ 6 & -1-i \end{pmatrix}$$

in the notation of Section 3, Case (1), using the decimal approximation provided by Mathematica 9 for convenience, we find

$$\lambda = 0.835165+2.29315i \, , \, r = 1.34636 \, , \, \theta = 0.407176 \, , \, p = 0.835165+1.28571i-1.65934j-0.923077k$$

and the solutions q_1, q_2, q_3 to the 3-rd order Sylvester equation $aq^3 + q^3 b = c$ are

$$q_1 = 1.23628 + 0.298941i - 0.385813j - 0.214625k$$

$$q_2 = -1.07989 + 0.450817i - 0.581824j - 0.323664k$$

$$q_3 = -0.156393 - 0.749759i + 0.967637j + 0.538288k$$

In the notation of Section 4, solving the system (4.2) which has the form

$$\begin{pmatrix} 1+3i & 0 & 2+2i & -3+4i \\ 2-2i & -3+4i & -1+i & 0 \\ 3+4i & -2-2i & 0 & 1+i \\ 0 & 1-3i & -3-4i & 2-2i \end{pmatrix} \begin{pmatrix} z \\ \bar{z} \\ w \\ \bar{w} \end{pmatrix} = \begin{pmatrix} -1+i \\ -6 \\ 6 \\ -1-i \end{pmatrix}$$

we find

$\det M = -273$, $z = 0.835165 - 0.923077i$, $w = 1.28571 - 1.65934i$, $h = 2.29315$

$$P = \begin{pmatrix} -1.28571 + 1.65934i & -1.28571 + 1.65934i \\ 3.21623i & -1.37007i \end{pmatrix}$$

$$P^{-1} = \begin{pmatrix} -0.0871635 - 0.112493i & -0.218041i \\ -0.204615 - 0.264076i & 0.218041i \end{pmatrix}$$

$$D = \begin{pmatrix} 0.835165 + 2.29315i & 0 \\ 0 & 0.835165 - 2.29315i \end{pmatrix}$$

Using the three different cubic roots of D and equation (4.9) we obtain the roots q_1, q_2, q_3 as before, but in matrix representation. For example, we find

$$q_1 = \begin{pmatrix} 1.23628 - 0.214625i & -0.298941 + 0.385813i \\ 0.298941 + 0.385813i & 1.23628 + 0.214625i \end{pmatrix}$$

5 Remark on the Newton-Raphson method

Letting

$$q = \begin{pmatrix} z & -w \\ \bar{w} & \bar{z} \end{pmatrix} = \begin{pmatrix} x_1 + ix_2 & -x_3 - ix_4 \\ x_3 - ix_4 & x_1 - ix_2 \end{pmatrix} \tag{5.1}$$

using the fact that the set of matrices of the form (5.1) is closed under matrix addition and multiplication, after carrying out all matrix operations and equating the real and imaginary parts of the entries of the resulting (2×2) matrix to zero, every bilateral quaternion polynomial equation

$$\sum_{k=0}^{n} \left(a_k q^k b_k + c_k q^k d_k \right) = 0 \tag{5.2}$$

in matrix representation, can be reduced to a system of only four nonlinear equations

$$f_i(x_1, x_2, x_3, x_4) = 0 \ , \ \ i = 1, 2, 3, 4 \tag{5.3}$$

or in vector notation

$$F(X) = 0 \ , \ X = (x_1, x_2, x_3, x_4)^T \ , \ F = (f_1, f_2, f_3, f_4)^T \tag{5.4}$$

Traditionally, such systems are solved recursively using the Newton-Raphson method for nonlinear systems. The method produces a sequence

$$X^{(k+1)} = X^{(k)} + H^{(k)} \ , \ k = 0, 1, 2, ... \tag{5.5}$$

where the correction term H is computed by solving using Gaussian elimination the *Jacobian system*

$$F'(X^{(k)})H^{(k)} = -F(X^{(k)}) \tag{5.6}$$

assuming that the (4×4) Jacobian matrix

$$F'(X) = \left(\frac{\partial f_i}{\partial x_j} \right)_{1 \le i,j \le 4} \tag{5.7}$$

is nonsingular at least in a neighborhood of the solution (see, for example, [10] and [1]). Convergence of the method depends on making a good initial guess $X(0)$. In the case of the higher order Sylvester equation $aq^n + q^n b = c$ where a, b, c are as in (4.1), letting

$$M(x_1, x_2, x_3, x_4) = aq^n + q^n b - c$$

viewed as a (2×2) matrix $M = (M_{ij})_{1 \le i,j \le 2}$, we have

$$f_1(x_1, x_2, x_3, x_4) = \text{Re } M_{11}(x_1, x_2, x_3, x_4) \ , \ f_2(x_1, x_2, x_3, x_4) = \text{Im } M_{11}(x_1, x_2, x_3, x_4)$$

$$f_3(x_1, x_2, x_3, x_4) = \text{Re } M_{12}(x_1, x_2, x_3, x_4) \ , \ f_4(x_1, x_2, x_3, x_4) = \text{Im } M_{12}(x_1, x_2, x_3, x_4)$$

The solution of nonlinear systems of equations with the use of Mathematica version 9 is done through the command FindRoot which uses a damped or under-relaxed Newton-Raphson method to alleviate poor initial approximations $X(0)$ to the solution X of the system. The default values for FindRoot are 100 iterations with a precision goal of two decimal places. In the example below we have changed the precision goal to six and found, in each case, the first iteration at which the method converged. In each case one should also check the accuracy of the method by computing $f_i(x_1, x_2, x_3, x_4)$, $i = 1, 2, 3, 4$, for the values of $X = (x_1, x_2, x_3, x_4)^T$ to which the method has converged, making sure that the answer is zero.

Example 2.

As in Example 1, for $n = 3$ and

$$a = 1+3i-4j+k = \begin{pmatrix} 1+i & -3+4i \\ 3+4i & 1-i \end{pmatrix}, \; b = -2i+2j+2k = \begin{pmatrix} 2i & 2-2i \\ -2-2i & -2i \end{pmatrix}$$

$$c = -1+6i+0j+k = \begin{pmatrix} -1+i & -6 \\ 6 & -1-i \end{pmatrix}$$

we find

$$f_1(x_1,x_2,x_3,x_4) = 1+x_1^3-9x_1^2x_2-3x_1x_2^2+3x_2^3-3x_1^2x_3+x_2^2x_3-3x_1x_3^2+3x_2x_3^2+x_3^3+6x_1^2x_4$$
$$-2x_2^2x_4-2x_3^2x_4-3x_1x_4^2+3x_2x_4^2+x_3x_4^2-2x_4^3$$
$$f_2(x_1,x_2,x_3,x_4) = -1+3x_1^3+3x_1^2x_2-9x_1x_2^2-x_2^3+18x_1^2x_3-6x_2^2x_3-9x_1x_3^2-x_2x_3^2-6x_3^3+15x_1^2x_4$$
$$-5x_2^2x_4-5x_3^2x_4-9x_1x_4^2-x_2x_4^2-6x_3x_4^2-5x_4^3$$
$$f_3(x_1,x_2,x_3,x_4) = 6-x_1^3+18x_1^2x_2+3x_1x_2^2-6x_2^3-3x_1^2x_3+x_2^2x_3+3x_1x_3^2-6x_2x_3^2+x_3^3$$
$$-3x_1^2x_4+x_2^2x_4+x_3^2x_4+3x_1x_4^2-6x_2x_4^2+x_3x_4^2+x_4^3$$
$$f_4(x_1,x_2,x_3,x_4) = 2x_1^3+15x_1^2x_2-6x_1x_2^2-5x_2^3+3x_1^2x_3-x_2^2x_3-6x_1x_3^2-5x_2x_3^2-x_3^3$$
$$-3x_1^2x_4+x_2^2x_4+x_3^2x_4-6x_1x_4^2-5x_2x_4^2-x_3x_4^2+x_4^3$$
$$\det F'(X) = 2457(-3x_1^4-2x_1^2(x_2^2+x_3^2+x_4^2)+(x_2^2+x_3^2+x_4^2)^2)^2$$

For an initial guess $X(0) = (1,0,0,0)^T$, after 6 iterations, using the FindRoot command the method converged to a six decimal point accuracy to

$$X = (1.23628, -0.214625, 0.298941, -0.385813)^T$$

For an initial guess $X(0) = (-1,0,0,0)^T$, after 8 iterations, the method converged to

$$X = (-1.07989, -0.323664, 0.450817, -0.581824)^T$$

For an initial guess $X(0) = (0,0,1,0)^T$, after 13 iterations, the method converged to

$$X = (-0.156393, 0.538288, -0.749759, 0.967637)^T$$

in agreement with the results of Example 1. For the initial choice $X(0) = (0,0,0,0)^T$ the method run into a singular Jacobian. A slight variation $X(0) = (0.01,0,0,0)^T$ provided convergence to

$$X = (1.23628, -0.214625, 0.298941, -0.385813)^T$$

after 11 iterations. A totally random choice $X(0) = (-10,10,15,20)^T$ converged after 13 iterations to

$$X = (-0.156393, 0.538288, -0.749759, 0.967637)^T$$

In this example, the overall performance of the method (and of Mathematica 9) was very good.

Acknowledgement. We thank the referee for remarks that helped make this paper better.

References

[1] W. Cheney, D. Kincaid, *Numerical Analysis, Mathematics of Scientific Computing,* Brooks/Cole Publishing Company, Pacific Grove, California, 1991.

[2] J. Helmstetter, The quaternionic equation ax+xb=c, *Advances in Applied Clifford Algebras,* **22**, no. 4, (2012), 1055-1059.

[3] D. Janovska, G. Opfer, Linear equations in quaternionic variables, *Mitt. Math. Ges. Hamburg,* **27**, (2008), 223-234.

[4] D. Janovska, G. Opfer, Computing quaternionic roots by Newtons method, *Trans. Numer. Anal.,* **26**, (2007), 82102.

[5] D. Janovska, G. Opfer, A note on the computation of all zeros of simple quaternionic polynomials, *SIAM J. Numer. Anal.,* **48**, no. 1, (2010), 244-256.

[6] N. L. Matiadou, P. Papadopoulos, Quaternionic Quadratic Formulas, *Journal of Multidisciplinary Engineering Science and Technology,* **2**, no. 3, (2015), 334-338.

[7] D. Mierzejewski, Linear Manifolds in Sets of Solutions of Quaternionic Polynomial Equations of Several types, *Adv. Appl. Clifford Algebras,* **21**, (2011), 417-428.

[8] J. P. Morais, S. Georgiev, W. Sprößig, *Real Quaternionic Calculus Handbook,* Birkhäuser, Springer Basel, 2014.

[9] I. Niven, The roots of a quaternion, *American Mathematical Monthly,* **49**, no. 6, (2004), 379-389.

[10] J. M. Ortega, W. C. Rheinboldt, *Iterative solution of nonlinear equations in several variables,* Classics in applied mathematics 30, Society for Industrial and Applied Mathematics, 1987.

[11] L. Rodman, *Topics in Quaternion Linear Algebra,* Princeton Series in Applied Mathematics, Princeton University Press, Princeton and Oxford, 2014.

[12] http://en.wikipedia.org/wiki/Sylvester_equation

[13] J. Sylvester, Sur lequations en matrices $px = xq$, *C. R. Acad. Sc. Paris,* **99**, (2), (1884), 6771, 115116.

[14] http://en.wikipedia.org/wiki/List_of_trigonometric_identities

[15] S. De Leo, G. Ducati, V. Leonardi, Zeros of unilateral quaternionic polynomials, *Electronic Journal of Linear Algebra,* ELA Volume 15, Article 25 (2006).