Quaternion Equations and Quaternion Polynomial Matrices

by

Liji Huang

A Thesis submitted to the Faculty of Graduate Studies of

The University of Manitoba

in partial fulfilment of the requirements of the degree of

MASTER OF SCIENCE

Department of Mathematics

University of Manitoba

Winnipeg

Contents

Abstract

In this thesis, we mainly consider finding formula solutions to quaternion equations and calculating the generalized inverses of quaternion polynomial matrices.

In the first chapter, we give a condensed review of the history of quaternions.

The topic of the second chapter is solving quadratic and other quaternion equations. Many mathematicians have worked on solving various quaternion equations. However, due to the complex nature of quaternions, the current "best" non-numerical result is a special case of quadratic quaternion equation presented in (Au-Yeung, 2003), which pales in comparison to what is known about equations in complex numbers. In this chapter, we present a new way to solve not only almost all of the quaternion equations that have been solved so far, including Au-Yeung's result, but also several more complicated ones. This method can also be easily implemented into computer algebra systems such as Maple. Examples will be included.

The topic of the third chapter is calculating the generalized inverses of quaternion polynomial matrices. It has been shown that a real (Karampetakis, 1997) or complex (Stanimirović et al., 2007) polynomial matrix has the generalized inverse that can be calculated using the Leverrier-Faddeev algorithm (Faddeeva, 1959). It is natural to ask whether a quaternion polynomial matrix has a generalized inverse, and if yes, how it can be calculated. In this chapter, we first give conditions that a quaternion polynomial matrix must satisfy in order to have the generalized inverse. We then

present an interpolation method that can be used to calculate the generalized inverse of a given quaternion polynomial matrix.

In the appendix, we include our quaternion polynomial matrix Maple module as well as some examples.

Acknowledgments

First and foremost, I would like to express my most sincere gratitude to my thesis advisor, Dr. Zhang Yang. I greatly appreciate his encouragement and dedicated support. Countless hours of discussion with Dr. Zhang helped me overcome many obstacles during the writing up of this thesis. Without his wisdom and motivation this thesis would not have been possible.

I greatly appreciate the faculty and staff in the Department of Mathematics for their teaching and support during my two years of being a graduate student in mathematics.

I greatly thank my thesis committee members, Dr. Guenter Krause, Dr. Xikui Wang and Dr. Joe Williams, for reading my thesis and giving me feedback.

On behalf of myself and my wife, I would like to thank Dr. Zhang, the Faculty of Graduate Studies, the Faculty of Science, and the Department of Mathematics, for financially supporting me through my graduate studies.

I would like to thank the University of Manitoba library system for providing a rich supply of online and offline publications. I would also like to thank mapleprimes.com and latex-community.org for all of the excellent technical support I received from them, and their continued voluntary service to current and future graduate students and academics all over the world.

Finally, I would like to express my appreciation to my family and friends, for all of their support and love.

Dedicated to my wife Zifang.

1. Introduction

1.1. History of Quaternions

Quaternions were first discovered by Sir William Rowan Hamilton of Ireland on October 16^{th}, 1843, in a "flash of genius", on the Brougham Bridge. It is uncommon that both the date and location of a major mathematical discovery are known, but we precisely know these because at the end of Hamilton's life, in a letter to his second son, Archibald Henry, he wrote:

Every morning in the early part of [October 1843], on my coming down to breakfast, your (then) little brother, William Edwin, and yourself, used to ask me, "Well, papa, can you multiply triplets?" Whereto I was always obliged to reply, with a sad shake of the head: "No, I can only add and subtract them".

But on the 16th day of the same month—which happened to be Monday, and a Council day of the Royal Irish Academy—I was walking in to attend and preside, and your mother was walking with me along the Royal Canal, to which she had perhaps driven; and although she talked with me now and then, yet an undercurrent of thought was going on in my mind which gave at last a result, whereof it is not too much to say that I felt at once the importance. An electric circuit seemed to close; and a spark flashed forth the herald (as I foresaw immediately) of many long years to come of definitely directed thought and work by myself, if spared, and, at all *events, on the part of others if I should even be allowed to live long enough distinctly to communicate the discovery. Nor could I resist the impulse—unphilosophical as it may have been—to cut with a knife on a stone of Brougham Bridge, as we passed it, the fundamental formula with the symbols* **i**, **j**, **k**:

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

which contains the Solution of the Problem, but, of course, the inscription has long since mouldered away (Derbyshire, 2006).

It was a breakthrough discovery in the sense that after the step from one-dimensional real numbers to two-dimensional complex numbers, the next step was not the natural one to three-dimensional super-complex numbers but to four-dimensional ones, i.e., quaternions. Hamilton embarked on an eight-year journey to develop an algebra for bracketed triplets but ended up discovering quaternions in a moment of euphoria. The Dutch mathematician Van der Waerden called Hamilton's inspiration "*the leap into the fourth dimension*" (van der Waerden, 1985). The Scottish mathematical physicist Peter Guthrie Tait, the leading quaternions exponent of his time, claimed "*...no figure, nor even model, can be more expressive or intelligible than a quaternion equation*" (Tait, 1890). However, Tait's lifetime friend, William Thomson, 1^{st} Baron Kelvin, disliked quaternions and would have nothing to do with them (Smith and Wise, 1989). Tait's other great friend, the Scottish mathematical physicist James Clerk Maxwell, remained unconvinced of quaternion's value in actual calculation, as he put it "*...quaternions...is a mathematical method, but it is a method of thinking, and not, at least for the present generation, a method of saving thought*" (Maxwell, 1873).

Furthermore, the American scientist Josiah Willard Gibbs did not find quaternions suited to his needs and rejected the quaternionic analysis in favor of his own invention: vector analysis (Hastings, 2010; Gibbs et al., 1902). The English electrical engineer Oliver Heaviside, who independently formulated vector analysis similar to Gibbs's system, wrote "*...it is impossible to think in quaternions, you can only pretend to do it*" (Nahin, 2002). The British mathematician Arthur Cayley gave a poetic illustration on the matter "*...I compare a quaternion formula to a pocket-map, a capital thing to put in one's pocket, but which for use must be unfolded; the formula, to be understood, must be translated into coordinates*" (Cayley, 1894). Even the English mathematician Alexander McAulay, a quaternion advocate, stated that "*...not much advance in physics has been made by the aid of quaternions*" (McAulay, 1892). Indeed, in the century that has passed since McAulay wrote, it can be argued that quaternions faded away to a minor role in mathematics and physics while the Gibbs-Heaviside vector algebra became the everyday work tool of all modern engineers and scientists. In 1892, Heaviside, hinting subtly that quaternions are what he called "*a positive evil of no inconsiderable magnitude*", summarized the matter in this way, and probably the best way "*...the invention of quaternions must be regarded as a most remarkable feat of human ingenuity...but to find out quaternions required a genius*" (Nahin).

Today, quaternions still play a role in quantum physics (Adler, 1995; Finkelstein et al., 1962) and other branches of physics (Churchill, 1990). Quaternions are now used throughout the aerospace industry for attitude control of aircraft and spacecraft (Kuipers, 2002). Some people claim that quaternions are an important aspect in modern computer graphics for calculations involving three-dimensional rotations because quaternions use less memory, compose faster, and are naturally suited for efficient interpolation (Hanson, 2005). Although, like in the academic circle of physics over a century ago, there are controversies regarding the advantages of quaternion orientation representations in today's computer graphics and modeling community

(Hanson, 2005). Quaternions are also attracting more and more attention in mathematics. There has been a new surge of interest in quaternions in the mathematics world—a quick search in the American Mathematical Society's database shows that more than half of the 3232 publications containing the word "quaternion" or "quaternionic" in the title have been published after the year of 1996.

1.2. Background for Quaternion Equations

This section attempts to give a general and brief review on what has been done on quaternion equations, §2.1 offers a more detailed introduction to quaternions themselves.

Let x be a quaternion indeterminate and let a_i be quaternions. In 1941, Niven looked into the equation:

$$x^m + a_1 x^{m-1} + a_2 x^{m-2} + \cdots + a_m = 0, \tag{1.1}$$

where $a_m \neq 0$. He showed that 1.1 has at least one quaternion root.

In the following year, Niven completely solved the equation

$$x^m - a_m = 0.$$

In 1944, Eilenberg and Niven showed that even the most general equation with a unique highest term,

$$a_1 x a_2 x \cdots x a_m + \phi(x) = 0,$$

where $\phi(x)$ is a sum of finite number of monomials of the form $b_1 x b_2 x \cdots x b_k$, $k < m$, has at least one quaternion root. This is also known as the "fundamental theorem of algebra" for quaternions.

In 1944, Johnson gave necessary and sufficient conditions for the equation

$$x a_1 + a_2 x + a_3 = 0$$

to have a quaternion solution. His result was extended by Tian to include

$$\overline{x} a_1 + a_2 x + a_3 = 0$$

and

$$\overline{x}a_1x + a_2 = 0,$$

where $\overline{x}$ denotes the quaternion conjugate of x.

Very little was done to obtain formula solutions of quaternion equations after Johnson's result, until, in 2002, Huang and So solved the special quadratic equation:

$$x^2 + a_1x + a_2 = 0. \tag{1.2}$$

Then in the following year, Au-Yeung solved another special quadratic equation, which included both Johnson's and Huang and So's results as special cases:

$$x^2 + a_1x + xa_2 + a_3 = 0. \tag{1.3}$$

Au-Yeung's paper was the latest non-numerical attempt to find the exact roots of a non-linear quaternion equation.

In 2010, Tian studied and solved two pairs of linear quaternion equations

$$\begin{cases} xa_1y = a_2, \\ ya_2x = a_1, \end{cases} \tag{1.4}$$

and

$$\begin{cases} \overline{x}a_1y = a_2, \\ \overline{y}a_2x = a_1, \end{cases} \tag{1.5}$$

where y is also a quaternion indeterminate. In §2.3, we present a new method to finding the solutions to 1.3, 1.4 and 1.5. In addition, we give solutions to several previously unsolved equations and least norm problems, including, for example, the "two-sided" quadratic homogeneous equation:

$$x^2 + a_1xa_2 = 0,$$

and the least norm problem:

$$\min_{x \in \mathbb{H}} |a_1 x - x a_2 - a_3| .$$

In §A.2, we present some maple codes to solve these quaternion equations and problems.

1.3. Background for Quaternion Polynomial Matrices

In 1955, Penrose introduced the term *generalized inverse* that exists for any matrix with complex elements. In 1965, Decell gave a method for computing the generalized inverse of a constant complex matrix based on Leverrier-Faddeev's algorithm (Faddeeva, 1959), which recursively computes coefficients of the characteristic polynomial of the matrix. This method was later extended by Karampetakis to real polynomial matrices. In 2006, Stanimirović and Petković applied an interpolation method to the result of Karampetakis. Then in the following year, Stanimirović et al. extended Stanimirović and Petković's result to include complex polynomial matrices with interpolations at real number data points.

In Chapter 3, we first give conditions that a quaternion polynomial matrix must satisfy in order to have a generalized inverse. Then we present an interpolation method that can be used to calculate the generalized inverse of a given quaternion polynomial matrix.

In §A.2, we include our Maple codes for calculating the generalized inverse of a given quaternion matrix or quaternion polynomial matrix. We also include our Maple codes that can be used to do Newton and Lagrange quaternionic interpolation, which are closely related to the interpolation of the generalized inverse of a quaternion polynomial matrix.

2. Quaternion Equations

2.1. Quaternions and Equivalence Classes

In this section, we will outline some definitions and properties of quaternions. First, recall the definition of quaternions:

Definition 2.1.1. Let $\mathbb{R}$ denote the field of the real numbers. Let $\mathbb{H}$ be a four-dimensional vector space over $\mathbb{R}$ with basis 1, $\mathbf{i}$, $\mathbf{j}$ and $\mathbf{k}$. A quaternion is a vector

$$x = x_0 \cdot 1 + x_1\mathbf{i} + x_2\mathbf{j} + x_3\mathbf{k} \in \mathbb{H}$$

with real coefficients x_0, x_1, x_2 and x_3. Throughout this thesis, for convenience, if $h \in \mathbb{H}$, then we will assume that $h = h_0 + h_1\mathbf{i} + h_2\mathbf{j} + h_3\mathbf{k}$ where $h_i \in \mathbb{R}$, $0 \leq i \leq 3$. We denote the real part, h_0, by $\operatorname{Re} h$ and the imaginary part, $h_1\mathbf{i} + h_2\mathbf{j} + h_3\mathbf{k}$, by $\operatorname{Im} h$.

Addition and multiplication in $\mathbb{H}$ can be defined as follows:

Definition 2.1.2. Let x, $y \in \mathbb{H}$. Then quaternion addition is defined as:

$$x + y = (x_0 + y_0) + (x_1 + y_1)\,\mathbf{i} + (x_2 + y_2)\,\mathbf{j} + (x_3 + y_3)\,\mathbf{k}.$$

Quaternion multiplication is defined as:

$$\begin{aligned}xy =& x_0y_0 - x_1y_1 - x_2y_2 - x_3y_3 \\ &+ (x_0y_1 + x_1y_0 + x_2y_3 - x_3y_2)\,\mathbf{i} \\ &+ (x_0y_2 - x_1y_3 + x_2y_0 + x_3y_1)\,\mathbf{j} \\ &+ (x_0y_3 + x_1y_2 - x_2y_1 + x_3y_0)\,\mathbf{k}.\end{aligned}$$

Remark 2.1.3. By direct calculation, we obtain:

$$\mathbf{ii} = \mathbf{jj} = \mathbf{kk} = -1,$$
$$\mathbf{ij} = -\mathbf{ji} = \mathbf{k}, \qquad \mathbf{jk} = -\mathbf{kj} = \mathbf{i}, \qquad \mathbf{ki} = -\mathbf{ik} = \mathbf{j}.$$

It is also worth noting that with the above equations, the multiplication defined in Definition 2.1.2 follows.

Lemma 2.1.4. *Quaternion multiplication is not commutative in general.* $\mathbb{H}$ *is a division ring with center* $\mathbb{R}$*.*

Proof. By direct calculation. □

Next, conjugate and norm for quaternions are defined as follows:

Definition 2.1.5. Let $h \in \mathbb{H}$. Then $\overline{h}$ denotes the conjugate of h, $\overline{h} = h_0 - h_1\mathbf{i} - h_2\mathbf{j} - h_3\mathbf{k}$, and $|h|$ denotes the norm of h, $|h| = \sqrt{h\overline{h}} = \sqrt{h_0^2 + h_1^2 + h_2^2 + h_3^2}$.

Lemma 2.1.6. *Let* x, $y \in \mathbb{H}$*. Then* $|x|$ *is non-negative and* $|\cdot|$ *is a norm on* $\mathbb{H}$*, i.e.,*

$$|x| = 0 \Leftrightarrow x = 0,$$
$$|x + y| \leq |x| + |y|\,,$$
$$|xy| = |yx| = |x|\,|y|\,.$$

Proof. We will only show that $|x+y| \leq |x| + |y|$. The rest of the lemma is true by direct calculation. Since

$$\begin{aligned}
|x+y|^2 &= \left|(x_0+y_0) + \mathbf{i}\,(x_1+y_1) + \mathbf{j}\,(x_2+y_2) + \mathbf{k}\,(x_3+y_3)\right|^2 \\
&= \left(\sqrt{(x_0+y_0)^2 + (x_1+y_1)^2 + (x_2+y_2)^2 + (x_3+y_3)^2}\right)^2 \\
&= (x_0+y_0)^2 + (x_1+y_1)^2 + (x_2+y_2)^2 + (x_3+y_3)^2 \\
&= x_0^2 + x_1^2 + x_2^2 + x_3^2 + y_0^2 + y_1^2 + y_2^2 + y_3^2 + 2x_0y_0 + 2x_1y_1 + 2x_2y_2 + 2x_3y_3 \\
&= |x|^2 + |y|^2 + 2x_0y_0 + 2x_1y_1 + 2x_2y_2 + 2x_3y_3,
\end{aligned}$$

it suffices to show that

$$|2x_0y_0 + 2x_1y_1 + 2x_2y_2 + 2x_3y_3| \leq 2\,|x|\,|y|$$

$$\Leftrightarrow$$

$$|x_0y_0 + x_1y_1 + x_2y_2 + x_3y_3| \leq |x|\,|y|$$

$$\Leftrightarrow$$

$$(x_0y_0 + x_1y_1 + x_2y_2 + x_3y_3)^2 \leq |x|^2\,|y|^2$$

$$\Leftrightarrow$$

$$(x_0y_0 + x_1y_1 + x_2y_2 + x_3y_3)^2 \leq \left(x_0^2 + x_1^2 + x_2^2 + x_3^2\right)\left(y_0^2 + y_1^2 + y_2^2 + y_3^2\right),$$

which is true by the Cauchy-Schwartz inequality. □

Next, inverse and conjugacy class for quaternions are defined as follows:

Lemma 2.1.7. *For $0 \neq h \in \mathbb{H}$, h^{-1}, the inverse of h, is uniquely determined by $h^{-1} = \frac{\overline{h}}{|h|^2}$. Furthermore, we have $|h^{-1}| = \frac{1}{|h|}$.*

Proof. By direct calculation. □

Definition 2.1.8. Two quaternions x and y are said to be similar, or in the same conjugacy class, if these exists a quaternion $v \neq 0$ such that $v^{-1}xv = y$, and we write $x \backsim y$.

Lemma 2.1.9. *$x \backsim y$ implies $|x| = |y|$.*

Proof. By Lemma 2.1.7,

$$x \backsim y \Rightarrow v^{-1}xv = y \Rightarrow |y| = \left|v^{-1}xv\right| = \frac{1}{|v|}\,|x|\,|v| = |x|\,.$$

□

Next, we state and prove an important result regarding similar quaternions.

Theorem 2.1.10. *$x \backsim y$ if and only if Re x = Re y and $|$Im $x| = |$Im $y|$.*

Proof. We only prove the forward direction of the theorem. The backward direction is straightforward.

If $x = 0$ or $y = 0$, then the claim is true trivially. Assuming $x \neq 0$ and $y \neq 0$, we will first show that $x_0 = y_0$. Let $xv = vy$ where $v \neq 0$. That is,

$$\begin{aligned} &(x_0 + x_1\mathbf{i} + x_2\mathbf{j} + x_3\mathbf{k})\,(v_0 + v_1\mathbf{i} + v_2\mathbf{j} + v_3\mathbf{k}) \\ =&\,(v_0 + v_1\mathbf{i} + v_2\mathbf{j} + v_3\mathbf{k})\,(y_0 + y_1\mathbf{i} + y_2\mathbf{j} + y_3\mathbf{k})\,. \end{aligned} \tag{2.1}$$

After expanding 2.1 and comparing coefficients, we get:

$$\begin{cases} x_0v_0 - x_1v_1 - x_2v_2 - x_3v_3 = y_0v_0 - y_1v_1 - y_2v_2 - y_3v_3 = A, \\ x_0v_1 + x_1v_0 + x_2v_3 - x_3v_2 = y_1v_0 + y_0v_1 + y_3v_2 - y_2v_3 = B, \\ x_2v_0 + x_0v_2 + x_3v_1 - x_1v_3 = y_2v_0 + y_0v_2 + y_1v_3 - y_3v_1 = C, \\ x_0v_3 + x_3v_0 + x_1v_2 - x_2v_1 = y_0v_3 + y_3v_0 + y_2v_1 - y_1v_2 = D, \end{cases} \tag{2.2}$$

which is equivalent to:

$$\begin{cases} v_0\left(x_0-y_0\right)=v_1\left(x_1-y_1\right)+v_2\left(x_2-y_2\right)+v_3\left(x_3-y_3\right), \\ v_0\left(x_1-y_1\right)=-v_1\left(x_0-y_0\right)+v_2\left(x_3+y_3\right)-v_3\left(x_2+y_2\right), \\ v_0\left(x_2-y_2\right)=-v_1\left(x_3+y_3\right)-v_2\left(x_0-y_0\right)+v_3\left(x_1+y_1\right), \\ v_0\left(x_3-y_3\right)=v_1\left(x_2+y_2\right)-v_2\left(x_1+y_1\right)-v_3\left(x_0-y_0\right). \end{cases} \tag{2.3}$$

Multiplying the four equations of 2.3 by (x_0+y_0), (x_1+y_1), (x_2+y_2) and (x_3+y_3) respectively, then adding them, we get

$$x_0\left(v_1y_1+v_2y_2+v_3y_3\right)=y_0\left(v_1x_1+v_2x_2+v_3x_3\right)$$

$$\Leftrightarrow$$

$$x_0\left(y_0v_0-A\right)=y_0\left(x_0v_0-A\right)$$

$$\Leftrightarrow$$

$$Ax_0=Ay_0.$$

If $x_0=y_0$, we are done. Assume $x_0\neq y_0$. Then $A=0$.

Next, multiplying the four equations of 2.3 by (x_1+y_1), $-(x_0+y_0)$, $-(x_3-y_3)$ and (x_2-y_2) respectively, then adding them, we get

$$x_0\left(y_1v_0-y_2v_3+y_3v_2\right)=y_0\left(x_2v_3-x_3v_2+x_1v_0\right)$$

$$\Leftrightarrow$$

$$x_0\left(B-y_0v_1\right)=y_0\left(B-x_0v_1\right)$$

$$\Leftrightarrow$$

$$Bx_0=By_0.$$

Since $x_0\neq y_0$, $B=0$.

Next, multiplying the four equations of 2.3 by (x_2+y_2), (x_3-y_3), $-(x_0+y_0)$ and

$-(x_1 - y_1)$ respectively, then adding them, we get

$$x_0 (y_2 v_0 + y_1 v_3 - y_3 v_1) = y_0 (x_2 v_0 + x_3 v_1 - x_1 v_3)$$

$$\Leftrightarrow$$

$$x_0 (C - y_0 v_2) = y_0 (C - x_0 v_2)$$

$$\Leftrightarrow$$

$$C x_0 = C y_0.$$

Since $x_0 \neq y_0$, $C = 0$.

Lastly, multiplying the four equations of 2.3 by $(x_3 + y_3)$, $-(x_2 - y_2)$, $(x_1 - y_1)$ and $-(x_0 + y_0)$ respectively, then adding them, we get

$$x_0 (y_3 v_0 + y_2 v_1 - y_1 v_2) = y_0 (x_3 v_0 + x_1 v_2 - x_2 v_1)$$

$$\Leftrightarrow$$

$$x_0 (D - y_0 v_3) = y_0 (D - x_0 v_3)$$

$$\Leftrightarrow$$

$$D x_0 = D y_0.$$

Since $x_0 \neq y_0$, $D = 0$. That is, $A = B = C = D = 0 \Rightarrow v = 0$ or $x = y = 0$. Contradiction. Therefore, $x_0 = y_0$. Then by Lemma 2.1.9, $x_1^2 + x_2^2 + x_3^2 = y_1^2 + y_2^2 + y_3^2$, that is, $|\mathrm{Im}\, x| = |\mathrm{Im}\, y|$. $\square$

2.2. General Quaternion Quadratic Equations

Let a_i, b_i, $c \in \mathbb{H}$ where $i \in \mathbb{N}$. We write $a_i = a_{i0} + a_{i1}\mathbf{i} + a_{i2}\mathbf{j} + a_{i3}\mathbf{k}$ and $b_i = b_{i0} + b_{i1}\mathbf{i} + b_{i2}\mathbf{j} + b_{i3}\mathbf{k}$ for easier notation. In this section, we focus on the general quadratic equation:

$$x^2 + \sum_{i=1}^{k} a_i x b_i + c = 0, \tag{2.4}$$

where $k \in \mathbb{N}$.

2.4 is equivalent to the real non-linear system:

$$\begin{cases} (2x_0 + m_{11})\,x_1 + m_{12}x_2 + m_{13}x_3 = -c_1 + x_0 n_1, & (1) \\ m_{21}x_1 + (2x_0 + m_{22})\,x_2 + m_{23}x_3 = -c_2 + x_0 n_2, & (2) \\ m_{31}x_1 + m_{32}x_2 + (2x_0 + m_{33})\,x_3 = -c_3 + x_0 n_3, & (3) \\ x_0^2 - x_1^2 - x_2^2 - x_3^2 + d_1 x_0 + d_2 x_1 + d_3 x_2 + d_4 x_3 + c_0 = 0, & (4) \end{cases} \tag{2.5}$$

where m_{ij}, n_i and d_i are given by:

$$M = (m_{ij})_{3\times 3}$$
$$= \sum_{i=1}^{k} \begin{pmatrix} a_{i0}b_{i0} - a_{i1}b_{i1} + a_{i2}b_{i2} + a_{i3}b_{i3} & a_{i0}b_{i3} - a_{i1}b_{i2} - a_{i2}b_{i1} - a_{i3}b_{i0} & -a_{i0}b_{i2} - a_{i1}b_{i3} + a_{i2}b_{i0} - a_{i3}b_{i1} \\ -a_{i0}b_{i3} - a_{i1}b_{i2} - a_{i2}b_{i1} + a_{i3}b_{i0} & a_{i0}b_{i0} + a_{i1}b_{i1} - a_{i2}b_{i2} + a_{i3}b_{i3} & a_{i0}b_{i1} - a_{i1}b_{i0} - a_{i2}b_{i3} - a_{i3}b_{i2} \\ a_{i0}b_{i2} - a_{i1}b_{i3} - a_{i2}b_{i0} - a_{i3}b_{i1} & -a_{i0}b_{i1} + a_{i1}b_{i0} - a_{i2}b_{i3} - a_{i3}b_{i2} & a_{i0}b_{i0} + a_{i1}b_{i1} + a_{i2}b_{i2} - a_{i3}b_{i3} \end{pmatrix},$$

$$N = (n_i)_{3\times 1} = \begin{pmatrix} \sum_{i=1}^{k} (-a_{i0}b_{i1} - a_{i1}b_{i0} - a_{i2}b_{i3} + a_{i3}b_{i2}) \\ \sum_{i=1}^{k} (-a_{i0}b_{i2} + a_{i1}b_{i3} - a_{i2}b_{i0} - a_{i3}b_{i1}) \\ \sum_{i=1}^{k} (-a_{i0}b_{i3} - a_{i1}b_{i2} + a_{i2}b_{i1} - a_{i3}b_{i0}) \end{pmatrix},$$

and

$$D = (d_i)_{4\times 1} = \begin{pmatrix} \sum_{i=1}^{k} (a_{i0}b_{i0} - a_{i1}b_{i1} - a_{i2}b_{i2} - a_{i3}b_{i3}) \\ \sum_{i=1}^{k} (-a_{i0}b_{i1} - a_{i1}b_{i0} + a_{i2}b_{i3} - a_{i3}b_{i2}) \\ \sum_{i=1}^{k} (-a_{i0}b_{i2} - a_{i1}b_{i3} - a_{i2}b_{i0} + a_{i3}b_{i1}) \\ \sum_{i=1}^{k} (-a_{i0}b_{i3} + a_{i1}b_{i2} - a_{i2}b_{i1} - a_{i3}b_{i0}) \end{pmatrix}.$$

Treating x_0 as a constant, let $\check{M}$ be the coefficient matrix of (1), (2) and (3) of 2.5:

$$\begin{pmatrix} 2x_0 + m_{11} & m_{12} & m_{13} \\ m_{21} & 2x_0 + m_{22} & m_{23} \\ m_{31} & m_{32} & 2x_0 + m_{33} \end{pmatrix}.$$

Direct calculation yields:

$$\begin{aligned} \sigma(x_0) =& \det \check{M} \\ =& 8x_0^3 + 4(m_{11} + m_{22} + m_{33})x_0^2 \\ &+ 2(m_{11}m_{22} + m_{11}m_{33} - m_{12}m_{21} - m_{13}m_{31} - m_{23}m_{32} + m_{22}m_{33})x_0 \\ &+ m_{11}m_{22}m_{33} - m_{11}m_{23}m_{32} - m_{12}m_{21}m_{33} \\ &+ m_{13}m_{21}m_{32} + m_{12}m_{23}m_{31} - m_{13}m_{22}m_{31}. \end{aligned}$$

We continue the discussion by cases:

Case 1. We assume that $\det \check{M} \neq 0$.

Applying Cramer's rule to (1), (2) and (3) of 2.5, we obtain x_1, x_2 and x_3 as rational polynomials of x_0 as follows:

$$x_1 = \frac{p_1(x_0)}{\sigma(x_0)}, \tag{2.6}$$

$$x_2 = \frac{p_2(x_0)}{\sigma(x_0)} \tag{2.7}$$

and

$$x_3 = \frac{p_3(x_0)}{\sigma(x_0)}, \tag{2.8}$$

where

$$\begin{aligned} p_1(x_0) =& 4n_1x_0^3 + 2\left(n_1\left(m_{22}+m_{33}\right) - n_2m_{12} - n_3m_{13} - 2c_1\right)x_0^2 \\ &+ 2\left(m_{12}c_2 + m_{13}c_3 - c_1\left(m_{22}+m_{33}\right)\right)x_0 \\ &+ n_1\left(m_{22}m_{33} - m_{23}m_{32}\right)x_0 + n_2\left(m_{13}m_{32} - m_{12}m_{33}\right)x_0 \\ &+ n_3\left(m_{12}m_{23} - m_{13}m_{22}\right)x_0 + c_1\left(m_{23}m_{32} - m_{33}m_{22}\right) \\ &+ c_2\left(m_{12}m_{33} - m_{13}m_{32}\right) + c_3\left(m_{13}m_{22} - m_{12}m_{23}\right), \end{aligned}$$

$$\begin{aligned} p_2(x_0) =& 4n_2x_0^3 + 2\left(n_2\left(m_{11}+m_{33}\right) - n_1m_{21} - n_3m_{23} - 2c_2\right)x_0^2 \\ &+ 2\left(m_{21}c_1 + m_{23}c_3 - c_2\left(m_{11}+m_{33}\right)\right)x_0 \\ &+ n_1\left(m_{23}m_{31} - m_{21}m_{33}\right)x_0 + n_2\left(m_{11}m_{33} - m_{13}m_{31}\right)x_0 \\ &+ n_3\left(m_{13}m_{21} - m_{11}m_{23}\right)x_0 + c_1\left(m_{21}m_{33} - m_{23}m_{31}\right) \\ &+ c_2\left(m_{13}m_{31} - m_{11}m_{33}\right) + c_3\left(m_{11}m_{23} - m_{13}m_{21}\right) \end{aligned}$$

and

$$\begin{aligned} p_3(x_0) =& 4n_3x_0^3 + 2\left(n_3\left(m_{11}+m_{22}\right) - n_1m_{31} - n_2m_{32} - 2c_3\right)x_0^2 \\ &+ 2\left(m_{31}c_1 + m_{32}c_2 - c_3\left(m_{11}+m_{22}\right)\right)x_0 \\ &+ n_1\left(m_{21}m_{32} - m_{22}m_{31}\right)x_0 + n_2\left(m_{12}m_{31} - m_{11}m_{32}\right)x_0 \\ &+ n_3\left(m_{11}m_{22} - m_{12}m_{21}\right)x_0 + c_1\left(m_{22}m_{31} - m_{21}m_{32}\right) \\ &+ c_2\left(m_{11}m_{32} - m_{12}m_{31}\right) + c_3\left(m_{12}m_{21} - m_{11}m_{22}\right). \end{aligned}$$

Substituting 2.6, 2.7 and 2.8 back into (4) of 2.5, we obtain:

$$\begin{aligned} 0 =& x_0^2 + d_1 x_0 + c_0 - \frac{p_1(x_0)^2 + p_2(x_0)^2 + p_3(x_0)^2}{\sigma(x_0)^2} \\ &+ \frac{d_2 p_1(x_0) + d_3 p_2(x_0) + d_4 p_3(x_0)}{\sigma(x_0)}. \end{aligned}$$

Multiplication of the above equation by $\sigma(x_0)^2$ yields:

$$\begin{aligned} 0 =& \sigma(x_0)^2 \left(x_0^2 + d_1 x_0 + c_0\right) - p_1(x_0)^2 - p_2(x_0)^2 - p_3(x_0)^2 \\ &+ \sigma(x_0)\left(d_2 p_1(x_0) + d_3 p_2(x_0) + d_4 p_3(x_0)\right). \end{aligned} \tag{2.9}$$

2.9 is an 8^{th} order equation of x_0. After solving 2.9 for a real solution x_0, we substitute its value back into 2.6, 2.7 and 2.8 to get x_1, x_2 and x_3 respectively to obtain a solution to 2.4. The prime difficulty here is that, by Galois theory, there is no formula solution for general eighth degree equations (Morandi, 1996). Furthermore, there does not seem to be a general formula to factor 2.9. However, as we will show in the next section, under certain conditions, 2.9 can be solved.

Case 2. We assume that $\det \check{M} = 0$.

Then x_0 is known: it is equal to a (fixed) real root (which always exists) of the cubic polynomial $\sigma(x_0)$. Let $\check{N}$ be the 3×1 vector:

$$\begin{pmatrix} -c_1 + x_0 n_1 \\ -c_2 + x_0 n_2 \\ -c_3 + x_0 n_3 \end{pmatrix}.$$

It is clear that 2.5 has a solution only if the last row of the reduced row echelon form of the augmented matrix $\left(\check{M} \mid \check{N}\right)$ is a zero row. Under these conditions, (1), (2) and (3) of 2.5, now viewed as a linear system with unknowns x_1, x_2 and x_3, has infinitely many solutions of the form:

$$\begin{cases} x_1 = s_1 y + t_1 z + w_1, \\ x_2 = s_2 y + t_2 z + w_2, \\ x_3 = s_3 y + t_3 z + w_3, \end{cases} \tag{2.10}$$

where s_i, t_i and $w_i \in \mathbb{R}$, $1 \leq i \leq 3$, and y, $z \in \mathbb{R}$ are parameters. We plug 2.10 into (4) of 2.5 and obtain:

$$\begin{aligned} 0 = & -(s_1 y + t_1 z + w_1)^2 - (s_2 y + t_2 z + w_2)^2 - (s_3 y + t_3 z + w_3)^2 \\ & + d_2 (s_1 y + t_1 z + w_1) + d_3 (s_2 y + t_2 z + w_2) + d_4 (s_3 y + t_3 z + w_3) \\ & + x_0^2 + d_1 x_0 + c_0. \end{aligned} \tag{2.11}$$

The parameters y and z must satisfy 2.11, therefore 2.11 can be seen as a quadratic equation of, WLOG, y:

$$y^2 + f(z)\, y + g(z) = 0, \tag{2.12}$$

where $f(z)$ is a linear polynomial of z and $g(z)$ is a quadratic polynomial of z. We can take any $z \in \mathbb{R}$ such that the discriminant of 2.12, $f(z)^2 - 4g(z)$, a quadratic polynomial of z, is non-negative. We then solve for $y = \frac{-f(z) \pm \sqrt{f(z)^2 - 4g(z)}}{2}$. Finally, we substitute the value of y and z back into to 2.10 to get x_1, x_2 and x_3 to obtain a solution to 2.4.

2.3. Special Quaternion Equations and problems

In this section, we give solutions to some specific quaternion equations and least norm problems.

2.3.1. A quadratic equation

In this subsection, we consider the quadratic equation that was studied and solved in (Au-Yeung, 2003):

$$t^2 + \alpha t + t\beta + \gamma = 0, \tag{2.13}$$

where α, β and $\gamma \in \mathbb{H}$ and t is a quaternion indeterminate. Here I present a somewhat simpler solution.

Let $x = t + \frac{\alpha_0+\beta_0}{2}$. Then 2.13 becomes:

$$\begin{aligned} 0 &= \left(x - \frac{\alpha_0+\beta_0}{2}\right)^2 + \alpha\left(x - \frac{\alpha_0+\beta_0}{2}\right) + \left(x - \frac{\alpha_0+\beta_0}{2}\right)\beta + \gamma \\ &= x^2 + ax + xb + c, \end{aligned} \tag{2.14}$$

where $a = \mathrm{Im}\ \alpha$, $b = \mathrm{Im}\ \beta$ and $c = \left(\frac{\alpha_0+\beta_0}{2}\right)^2 - (\alpha+\beta)\frac{\alpha_0+\beta_0}{2} + \gamma$. The advantage of 2.14 over 2.13 is that the left and right coefficients of the indeterminate are both pure imaginary.

2.14 is equivalent to the real non-linear system:

$$\begin{cases} 2x_0x_1 + (-a_3+b_3)\,x_2 + (a_2-b_2)\,x_3 = (-a_1-b_1)\,x_0 - c_1, & (1) \\ (a_3-b_3)\,x_1 + 2x_0x_2 + (-a_1+b_1)\,x_3 = (-a_2-b_2)\,x_0 - c_2, & (2) \\ (-a_2+b_2)\,x_1 + (a_1-b_1)\,x_2 + 2x_0x_3 = (-a_3-b_3)\,x_0 - c_3, & (3) \\ x_1^2 + x_2^2 + x_3^2 + (a_1+b_1)\,x_1 + (a_2+b_2)\,x_2 + (a_3+b_3)\,x_3 = x_0^2 + c_0. & (4) \end{cases} \tag{2.15}$$

We first show how to solve 2.15 if $a = b$. When $a = b$, 2.15 becomes:

$$\begin{cases} 2x_0\left(x_1 + a_1\right) = -c_1, & (1) \\ 2x_0\left(x_2 + a_2\right) = -c_2, & (2) \\ 2x_0\left(x_3 + a_3\right) = -c_3, & (3) \\ x_1^2 + x_2^2 + x_3^2 + 2a_1x_1 + 2a_2x_2 + 2a_3x_3 = x_0^2 + c_0. & (4) \end{cases} \tag{2.16}$$

Case 1. If $c = c_0 \in \mathbb{R}$, then 2.16 can have solutions only if $x_0 = 0$ or $\begin{cases} x_1 = & -a_1, \\ x_2 = & -a_2, \\ x_3 = & -a_3, \end{cases}$ is part of the solution.

1. When $x_0 = 0$ is part of the solution, (4) of 2.16 becomes:

$$x_1^2 + x_2^2 + x_3^2 + 2a_1x_1 + 2a_2x_2 + 2a_3x_3 = c$$

$$\Leftrightarrow$$

$$\sum_{i=1}^{3}\left(x_i + a_i\right)^2 = c + \sum_{i=1}^{3} a_i^2 = c - a^2. \tag{2.17}$$

   2.17 can only hold if $c - a^2 \geq 0$. In which case, x_1, x_2 and x_3 can be any real numbers that satisfies 2.17.

2. When $\begin{cases} x_1 = & -a_1, \\ x_2 = & -a_2, \\ x_3 = & -a_3, \end{cases}$ is part of the solution, (4) of 2.16 becomes:

$$-a_1^2 - a_2^2 - a_3^2 = x_0^2 + c$$

$$\Leftrightarrow$$

$$a^2 - c = x_0^2.$$

The above equation can only hold if and $a^2 - c \geq 0$. In which case, $x_0 = \pm\sqrt{a^2 - c}$.

Case 2. If $c \neq \mathbb{R}$, then $x_0 = 0$ can never be part of the solution. We obtain, using the method demonstrated in §2.2, a quartic equation for x_0:

$$x_0^4 + \left(c_0 + a_1^2 + a_2^2 + a_3^2\right) x_0^2 - \frac{1}{4}\left(c_1^2 + c_2^2 + c_3^2\right) = 0.$$

Replacing x_0^2 by y, we obtain a quadratic equation for y:

$$f(y) = y^2 + \left(c_0 + a_1^2 + a_2^2 + a_3^2\right) y - \frac{1}{4}\left(c_1^2 + c_2^2 + c_3^2\right) = 0.$$

The solutions of f are:

$$y = \frac{\pm\sqrt{\left(c_0 + a_1^2 + a_2^2 + a_3^2\right)^2 + \left(c_1^2 + c_2^2 + c_3^2\right)} - \left(c_0 + a_1^2 + a_2^2 + a_3^2\right)}{2}.$$

Since $c \neq \mathbb{R}$, $c_1^2 + c_2^2 + c_3^2 > 0$, and therefore

$$\frac{\sqrt{\left(c_0 + a_1^2 + a_2^2 + a_3^2\right)^2 + \left(c_1^2 + c_2^2 + c_3^2\right)} - \left(c_0 + a_1^2 + a_2^2 + a_3^2\right)}{2} > 0$$

and

$$\frac{-\sqrt{\left(c_0 + a_1^2 + a_2^2 + a_3^2\right)^2 + \left(c_1^2 + c_2^2 + c_3^2\right)} - \left(c_0 + a_1^2 + a_2^2 + a_3^2\right)}{2} < 0.$$

So

$$\begin{aligned} x_0 &= \pm\sqrt{\frac{\sqrt{\left(c_0 + a_1^2 + a_2^2 + a_3^2\right)^2 + \left(c_1^2 + c_2^2 + c_3^2\right)} - \left(c_0 + a_1^2 + a_2^2 + a_3^2\right)}{2}} \\ &= \sqrt{\frac{\sqrt{\left(c_0 - a^2\right)^2 - \left(\operatorname{Im} c\right)^2} - \left(c_0 - a^2\right)}{2}}, \end{aligned}$$

and

$$\begin{cases} x_1 = & \frac{-c_1}{2x_0} - a_1, \\ x_2 = & \frac{-c_2}{2x_0} - a_2, \\ x_3 = & \frac{-c_3}{2x_0} - a_3. \end{cases}$$

From now on, we can and we will assume that $a \neq b$. The determinant of the coefficient matrix of (1), (2) and (3) of 2.15, $\check{M} = \begin{pmatrix} 2x_0 & -a_3+b_3 & a_2-b_2 \\ a_3-b_3 & 2x_0 & -a_1+b_1 \\ -a_2+b_2 & a_1-b_1 & 2x_0 \end{pmatrix}$, is:

$$\begin{aligned} \sigma(x_0) =& \det \check{M} \\ =& 8x_0^3 + 2\left((a_1-b_1)^2 + (a_2-b_2)^2 + (a_3-b_3)^2\right) x_0 \\ =& 2x_0 \left(4x_0^2 + (a_1-b_1)^2 + (a_2-b_2)^2 + (a_3-b_3)^2\right). \end{aligned}$$

2.15 can have infinitely many solutions only if $\sigma(x_0) = 0$, this would mean that $x_0 = 0$.

First, when $x_0 = 0$, (4) of 2.16 becomes

$$x_1^2 + x_2^2 + x_3^2 + 2a_1x_1 + 2a_2x_2 + 2a_3x_3 = c$$

$$\Leftrightarrow$$

$$\sum_{i=1}^{3} (x_i + a_i)^2 = c + \sum_{i=1}^{3} a_i^2 = c - a^2.$$

The above equation can only hold if $c - a^2 \geq 0$.

Next, when $x_0 = 0$, the augmented matrix of (1), (2) and (3) of 2.15 is

$$A = \begin{pmatrix} 0 & -a_3+b_3 & a_2-b_2 & -c_1 \\ a_3-b_3 & 0 & -a_1+b_1 & -c_2 \\ -a_2+b_2 & a_1-b_1 & 0 & -c_3 \end{pmatrix}.$$

We will show that A is the augmented matrix of a consistent linear system if and only if

$$\sum_{i=1}^{3} c_i \left(a_i - b_i\right) = 0.$$

We will show the forward direction first.

Case 1. If $a_i - b_i \neq 0$ for $1 \leq i \leq 3$. We apply Gaussian elimination to A and obtain ($\doteq$ means row equivalent) :

$$A \doteq \begin{pmatrix} a_3 - b_3 & 0 & -a_1 + b_1 & -c_2 \\ 0 & -a_3 + b_3 & a_2 - b_2 & -c_1 \\ 0 & 0 & 0 & -\frac{\sum_{i=1}^{3} c_i (a_i - b_i)}{a_3 - b_3} \end{pmatrix}.$$

Since A is the augmented matrix of a consistent linear system, $\sum_{i=1}^{3} c_i \left(a_i - b_i\right) = 0$.

Case 2. If $a_1 - b_1 = 0$, $a_i - b_i \neq 0$ for $i = 2, 3$, then

$$A = \begin{pmatrix} 0 & -a_3 + b_3 & a_2 - b_2 & -c_1 \\ a_3 - b_3 & 0 & 0 & -c_2 \\ -a_2 + b_2 & 0 & 0 & -c_3 \end{pmatrix}.$$

Since A is the augmented matrix of a consistent linear system, the last two rows give:

$$\begin{cases} \left(a_3 - b_3\right) x_1 = & -c_2 \\ \left(-a_2 + b_2\right) x_1 = & -c_3 \end{cases}$$

$$\Rightarrow$$

$$\left(a_2 - b_2\right) c_2 + \left(a_3 - b_3\right) c_3 = 0$$

$$\Rightarrow$$

$$\sum_{i=1}^{3} c_i \left(a_i - b_i\right) = 0.$$

Similarly, we can show the same result when $a_2 - b_2 = 0$, $a_i - b_i \neq 0$ for $i = 1, 3$ and $a_3 - b_3 = 0$, $a_i - b_i \neq 0$ for $i = 1, 2$.

Case 3. If $a_1 - b_1 \neq 0$, $a_i - b_i = 0$ for $i = 2, 3$, then

$$A = \begin{pmatrix} 0 & 0 & 0 & -c_1 \\ 0 & 0 & -a_1 + b_1 & -c_2 \\ 0 & a_1 - b_1 & 0 & -c_3 \end{pmatrix}.$$

Since A is the augmented matrix of a consistent linear system, the first row gives $c_1 = 0$, and therefore $\sum_{i=1}^{3} c_i (a_i - b_i) = 0$. Similarly, we can show the same result when $a_2 - b_2 \neq 0$, $a_i - b_i = 0$ for $i = 1, 3$ and $a_3 - b_3 \neq 0$, $a_i - b_i = 0$ for $i = 1, 2$.

For the converse, we assume that $\sum_{i=1}^{3} c_i (a_i - b_i) = 0$.

Case 1. If $a_3 - b_3 \neq 0$, then we apply Gaussian elimination to A to obtain:

$$\begin{aligned} A &\doteq \begin{pmatrix} a_3 - b_3 & 0 & -a_1 + b_1 & -c_2 \\ 0 & -a_3 + b_3 & a_2 - b_2 & -c_1 \\ 0 & 0 & 0 & -\frac{\sum_{i=1}^{3} c_i (a_i - b_i)}{a_3 - b_3} \end{pmatrix} \\ &= \begin{pmatrix} a_3 - b_3 & 0 & -a_1 + b_1 & -c_2 \\ 0 & -a_3 + b_3 & a_2 - b_2 & -c_1 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \end{aligned}$$

A is always the augmented matrix of a consistent linear system as it has the solution set

$$\begin{cases} x_1 = & \frac{-c_2 + (a_1 - b_1)s}{a_3 - b_3}, \\ x_2 = & \frac{c_1 + (a_2 - b_2)s}{a_3 - b_3}, \\ x_3 = & s, \end{cases}$$

where s is a parameter. In order for 2.15 to have a solution, we must

have $c - a^2 \geq 0$ and s must be a solution of

$$\begin{aligned}
0 =& \left(\frac{-c_2 + (a_1 - b_1)\, s}{a_3 - b_3}\right)^2 + \left(\frac{c_1 + (a_2 - b_2)\, s}{a_3 - b_3}\right)^2 + s^2 - c_0 \\
&+ (a_1 + b_1)\left(\frac{-c_2 + (a_1 - b_1)\, s}{a_3 - b_3}\right) + (a_2 + b_2)\left(\frac{c_1 + (a_2 - b_2)\, s}{a_3 - b_3}\right) + (a_3 + b_3)\, s \\
&\Leftrightarrow \\
0 =& \left(\sum_{i=1}^{3} (a_i - b_i)^2\right) s^2 + \mathcal{O}(s), \qquad (2.18)
\end{aligned}$$

where $\mathcal{O}(s)$ is a complicated linear polynomial of s. The right hand side of 2.18 is not identically equal to zero as the coefficient of s^2 is nonzero, and thus it can have at most 2 solutions and therefore 2.15 can have at most 2 solutions.

Case 2. If $a_3 - b_3 = 0$ and $a_2 - b_2 = 0$, then $a_1 - b_1 \neq 0$ because $a \neq b$. $\sum_{i=1}^{3} c_i (a_i - b_i) = 0$ implies that $c_1 = 0$ and we have:

$$A = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -a_1 + b_1 & -c_2 \\ 0 & a_1 - b_1 & 0 & -c_3 \end{pmatrix}.$$

This system is always consistent as it has the solution set

$$\begin{cases} x_1 = & s, \\ x_2 = & \frac{c_2}{a_1 - b_1}, \\ x_3 = & \frac{-c_3}{a_1 - b_1}, \end{cases}$$

where s is a parameter. In order for 2.15 to have a solution, we must

have $c - a^2 \geq 0$ and s must be a solution of

$$\begin{aligned}
0 =& s^2 + \left(\frac{c_2}{a_1 - b_1}\right)^2 + \left(\frac{-c_3}{a_1 - b_1}\right)^2 - c_0 \\
&+ (a_1 + b_1)\, s + (a_2 + b_2)\left(\frac{c_2}{a_1 - b_1}\right) + (a_3 + b_3)\left(\frac{-c_3}{a_1 - b_1}\right) \\
&\Leftrightarrow \\
0 =& (a_1 - b_1)^2\, s^2 + \mathcal{O}(s)\,, \qquad (2.19)
\end{aligned}$$

where $\mathcal{O}(s)$ is a complicated linear polynomial of s. The right hand side of 2.19 is not identically equal to zero as the coefficient of s^2 is nonzero, and therefore it can have at most 2 solutions and therefore 2.15 can have at most 2 solutions.

Case 3. If $a_3 - b_3 = 0$ and $a_2 - b_2 \neq 0$, then

$$A = \begin{pmatrix} 0 & 0 & a_2 - b_2 & -c_1 \\ 0 & 0 & -a_1 + b_1 & -c_2 \\ -a_2 + b_2 & a_1 - b_1 & 0 & -c_3 \end{pmatrix}.$$

$\sum_{i=1}^{3} c_i\,(a_i - b_i) = 0$ implies that $c_2\,(a_2 - b_2) = -c_1\,(a_1 - b_1)$ and therefore the first two rows are consistent. Furthermore, this system is always consistent as it has the solution set

$$\begin{cases} x_1 = & \frac{c_3 + (a_1 - b_1)s}{a_2 - b_2}, \\ x_2 = & s, \\ x_3 = & \frac{-c_1}{a_2 - b_2}, \end{cases}$$

where s is a parameter. In order for 2.15 to have a solution, we must

have $c - a^2 \geq 0$ and s must be a solution of

$$\begin{aligned}
0 = & \left(\frac{c_3 + (a_1 - b_1)\, s}{a_2 - b_2}\right)^2 + s^2 + \left(\frac{-c_1}{a_2 - b_2}\right)^2 - c_0 \\
& + (a_1 + b_1) \left(\frac{c_3 + (a_1 - b_1)\, s}{a_2 - b_2}\right) + (a_2 + b_2)\, s + (a_3 + b_3) \left(\frac{-c_1}{a_2 - b_2}\right) \\
\Leftrightarrow & \\
0 = & \left((a_1 - b_1)^2 + (a_2 - b_2)^2\right) s^2 + \mathcal{O}(s)\,, \qquad (2.20)
\end{aligned}$$

where $\mathcal{O}(s)$ is a complicated linear polynomial of s. The right hand side of 2.20 is not identically equal to zero as the coefficient of s^2 is nonzero, and therefore it can have at most 2 solutions and therefore 2.15 can have at most 2 solutions.

This concludes our discussion for possible scenarios where 2.15 can have infinitely many solutions.

When $x_0 \neq 0$, we have $\sigma(x_0) \neq 0$. Applying Cramer's rule to (1), (2) and (3) of 2.15, we obtain x_1, x_2 and x_3 as rational polynomials of x_0 as follows:

$$x_1 = \frac{p_1(x_0)}{\sigma(x_0)}, \qquad (2.21)$$

$$x_2 = \frac{p_2(x_0)}{\sigma(x_0)} \qquad (2.22)$$

and

$$x_3 = \frac{p_3(x_0)}{\sigma(x_0)}, \qquad (2.23)$$

where

$$\begin{aligned}
p_1(x_0) = &-4(a_1+b_1)x_0^3 + 4(a_2b_3 - a_3b_2 - c_1)x_0^2 \\
&+ (b_1 - a_1)\left(a_1^2 + a_2^2 + a_3^2 - b_1^2 - b_2^2 - b_3^2\right)x_0 \\
&+ 2\left(c_2(b_3 - a_3) + c_3(a_2 - b_2)\right)x_0 \\
&+ (b_1 - a_1)\left(c_1(a_1 - b_1) + c_2(a_2 - b_2) + c_3(a_3 - b_3)\right),
\end{aligned}$$

$$\begin{aligned}
p_2(x_0) = &-4(a_2+b_2)x_0^3 + 4(a_3b_1 - a_1b_3 - c_2)x_0^2 \\
&+ (b_2 - a_2)\left(a_1^2 + a_2^2 + a_3^2 - b_1^2 - b_2^2 - b_3^2\right)x_0 \\
&+ 2\left(c_1(a_3 - b_3) + c_3(b_1 - a_1)\right)x_0 \\
&+ (b_2 - a_2)\left(c_1(a_1 - b_1) + c_2(a_2 - b_2) + c_3(a_3 - b_3)\right)
\end{aligned}$$

and

$$\begin{aligned}
p_3(x_0) = &-4(a_3+b_3)x_0^3 + 4(a_1b_2 - a_2b_1 - c_3)x_0^2 \\
&+ (b_3 - a_3)\left(a_1^2 + a_2^2 + a_3^2 - b_1^2 - b_2^2 - b_3^2\right)x_0 \\
&+ 2\left(c_1(b_2 - a_2) + c_2(a_1 - b_1)\right)x_0 \\
&+ (b_3 - a_3)\left(c_1(a_1 - b_1) + c_2(a_2 - b_2) + c_3(a_3 - b_3)\right).
\end{aligned}$$

Substituting 2.21, 2.23 and 2.23 back into (4) of 2.15, we obtain:

$$\begin{aligned}
0 = &-x_0^2 - c_0 + \frac{p_1(x_0)^2 + p_2(x_0)^2 + p_3(x_0)^2}{\sigma(x_0)^2} \\
&+ \frac{(a_1+b_1)p_1(x_0) + (a_2+b_2)p_2(x_0) + (a_3+b_3)p_3(x_0)}{\sigma(x_0)}.
\end{aligned}$$

Multiplication of the above equation by $\sigma(x_0)^2$ yields a 6^{th} degree equation of x_0:

$$\begin{aligned}
0 =& 16x_0^6 + 8\left(a_1^2+a_2^2+a_3^2+b_1^2+b_2^2+b_3^2+2c_0\right)x_0^4\\
&+4c_0\left((a_1-b_1)^2+(a_2-b_2)^2+(a_3-b_3)^2\right)x_0^2\\
&+\left(a_1^2+a_2^2+a_3^2-b_1^2-b_2^2-b_3^2\right)^2 x_0^2\\
&+(8a_2b_3-8a_3b_2-4c_1)\,c_1x_0^2\\
&+(8a_3b_1-8a_1b_3-4c_2)\,c_2x_0^2\\
&+(8a_1b_2-8a_2b_1-4c_3)\,c_3x_0^2\\
&-\left(\sum_{i=1}^{3}c_i(a_i-b_i)\right)^2.
\end{aligned}$$

Replacing x_0^2 by y, we obtain a cubic equation of y:

$$\begin{aligned}
0 =& f(y)\\
=&16y^3+8\left(a_1^2+a_2^2+a_3^2+b_1^2+b_2^2+b_3^2+2c_0\right)y^2\\
&+4c_0\left((a_1-b_1)^2+(a_2-b_2)^2+(a_3-b_3)^2\right)y\\
&+\left(a_1^2+a_2^2+a_3^2-b_1^2-b_2^2-b_3^2\right)^2 y\\
&+\left((8a_2b_3-8a_3b_2)\,c_1-4c_1^2\right)y\\
&+\left((8a_3b_1-8a_1b_3)\,c_2-4c_2^2\right)y\\
&+\left((8a_1b_2-8a_2b_1)\,c_3-4c_3^2\right)y\\
&-\left(\sum_{i=1}^{3}c_i(a_i-b_i)\right)^2\\
=&16y^3+D_2y^2+D_1y+D_0. \qquad (2.24)
\end{aligned}$$

We next show that f has exactly one strictly positive solution.

Since $x_0 \neq 0$, $D_0 = -\left(\sum_{i=1}^{3}c_i(a_i-b_i)\right)^2 < 0$. By Descartes' rule of signs, f has at least one strictly positive solution. Furthermore, f can have multiple positive solutions only if $D_2 < 0$ and $D_1 > 0$. We show that this is impossible.

If $D_2 < 0$, then $c_0 < -\frac{a_1^2+a_2^2+a_3^2+b_1^2+b_2^2+b_3^2}{2}$. Therefore,

$$\begin{aligned}
&4c_0\left((a_1-b_1)^2+(a_2-b_2)^2+(a_3-b_3)^2\right)+\left(a_1^2+a_2^2+a_3^2-b_1^2-b_2^2-b_3^2\right)^2\\
&\quad+(2a_2b_3-2a_3b_2)^2+(2a_3b_1-2a_1b_3)^2+(2a_1b_2-2a_2b_1)^2\\
<&-2\left(a_1^2+a_2^2+a_3^2+b_1^2+b_2^2+b_3^2\right)\left((a_1-b_1)^2+(a_2-b_2)^2+(a_3-b_3)^2\right)\\
&\quad+\left(a_1^2+a_2^2+a_3^2-b_1^2-b_2^2-b_3^2\right)^2\\
&\quad+(2a_2b_3-2a_3b_2)^2+(2a_3b_1-2a_1b_3)^2+(2a_1b_2-2a_2b_1)^2\\
=&-\left((a_1-b_1)^2+(a_2-b_2)^2+(a_3-b_3)^2\right)^2<0.
\end{aligned}$$

This gives:

$$\begin{aligned}
&4c_0\left((a_1-b_1)^2+(a_2-b_2)^2+(a_3-b_3)^2\right)+\left(a_1^2+a_2^2+a_3^2-b_1^2-b_2^2-b_3^2\right)^2\\
<&-(2a_2b_3-2a_3b_2)^2-(2a_3b_1-2a_1b_3)^2-(2a_1b_2-2a_2b_1)^2.
\end{aligned}$$

Then D_1 can be simplified as follows:

$$\begin{aligned}
D_1=&4c_0\left((a_1-b_1)^2+(a_2-b_2)^2+(a_3-b_3)^2\right)+\left(a_1^2+a_2^2+a_3^2-b_1^2-b_2^2-b_3^2\right)^2\\
&\quad+(8a_2b_3-8a_3b_2)c_1-4c_1^2+(8a_3b_1-8a_1b_3)c_2-4c_2^2+(8a_1b_2-8a_2b_1)c_3-4c_3^2\\
<&-(2a_2b_3-2a_3b_2)^2-(2a_3b_1-2a_1b_3)^2-(2a_1b_2-2a_2b_1)^2\\
&\quad+(8a_2b_3-8a_3b_2)c_1-4c_1^2+(8a_3b_1-8a_1b_3)c_2-4c_2^2+(8a_1b_2-8a_2b_1)c_3-4c_3^2\\
=&-(2a_2b_3-2a_3b_2-2c_1)^2-(2a_3b_1-2a_1b_3-2c_2)^2-(2a_1b_2-2a_2b_1-2c_3)^2\\
\leq&0.
\end{aligned}$$

Therefore, f has exactly one strictly positive solution z. We can plug $x_0=\pm\sqrt{z}$ into 2.21, 2.23 and 2.23 to get x_1, x_2 and x_3 respectively.

In all of the above discussion, solutions to 2.13 can be obtained by calculating

$t = x - \frac{\alpha_0+\beta_0}{2} = \left(x_0 - \frac{\alpha_0+\beta_0}{2}\right) + x_1\mathbf{i} + x_2\mathbf{j} + x_3\mathbf{k}$, that is, $\begin{cases} t_0 = & x_0 - \frac{\alpha_0+\beta_0}{2}, \\ t_1 = & x_1, \\ t_2 = & x_2, \\ t_3 = & x_3, \end{cases}$ for each x.

We are now able to give a theorem that sums up our result so far in this subsection.

Theorem 2.3.1. *The solutions of the quadratic equation*

$$t^2 + \alpha t + t\beta + \gamma = 0, \tag{2.25}$$

where α, β and $\gamma \in \mathbb{H}$ and t is a quaternion indeterminate, can be obtained by formulas according to the following cases, let $a = Im\ \alpha$, $b = Im\ \beta$ and $c = \left(\frac{\alpha_0+\beta_0}{2}\right)^2 - (\alpha+\beta)\frac{\alpha_0+\beta_0}{2} + \gamma$. In the following, we call

$$\begin{cases} t_0 = & \sqrt{s} - \frac{\alpha_0+\beta_0}{2}, \\ t_1 = & \frac{p_1(\sqrt{s})}{\sigma(\sqrt{s})}, \\ t_2 = & \frac{p_2(\sqrt{s})}{\sigma(\sqrt{s})}, \\ t_3 = & \frac{p_3(\sqrt{s})}{\sigma(\sqrt{s})}, \end{cases} \quad \textit{and} \quad \begin{cases} t_0 = & -\sqrt{s} - \frac{\alpha_0+\beta_0}{2}, \\ t_1 = & \frac{p_1(-\sqrt{s})}{\sigma(-\sqrt{s})}, \\ t_2 = & \frac{p_2(-\sqrt{s})}{\sigma(-\sqrt{s})}, \\ t_3 = & \frac{p_3(-\sqrt{s})}{\sigma(-\sqrt{s})}, \end{cases}$$

where s is the unique positive root of 2.24, the standard solutions.

1. *$a = b$.*

   a) *$c \in \mathbb{R}$.*

      i. *$c - a^2 > 0$. $t_0 = -\frac{\alpha_0+\beta_0}{2}$ and (t_1, t_2, t_3) can be any real number triple that satisfies the equation $\sum_{i=1}^{3}(t_i + a_i)^2 = c - a^2$. This is the only case where 2.25 has infinitely many solutions.*

ii. $a^2 - c \geq 0.$

$$\begin{cases} t_0 = & \pm\sqrt{a^2-c} - \frac{\alpha_0+\beta_0}{2}, \\ t_1 = & -a_1, \\ t_2 = & -a_2, \\ t_3 = & -a_3. \end{cases}$$

b) $c \notin \mathbb{R}.$

$$\begin{cases} t_0 = & s - \frac{\alpha_0+\beta_0}{2}, \\ t_1 = & -a_1 - \frac{c_1}{2s}, \\ t_2 = & -a_2 - \frac{c_2}{2s}, \\ t_3 = & -a_2 - \frac{c_2}{2s}, \end{cases}$$

or

$$\begin{cases} t_0 = & -s - \frac{\alpha_0+\beta_0}{2}, \\ t_1 = & -a_1 + \frac{c_1}{2s}, \\ t_2 = & -a_2 + \frac{c_2}{2s}, \\ t_3 = & -a_2 + \frac{c_2}{2s}, \end{cases}$$

where $= \sqrt{\frac{\sqrt{(c_0-a^2)^2-(Im\ c)^2}-(c_0-a^2)}{2}}.$

2. $a \neq b$*. Standard solutions always exist. Furthermore,*

a) $\sum_{i=1}^{3} c_i\,(a_i - b_i) = 0$ *and* $c - a^2 \geq 0.$

i. $a_3 \neq b_3$. *Possible solutions must be of the form*

$$\begin{cases} t_0 = & -\frac{\alpha_0+\beta_0}{2}, \\ t_1 = & \frac{-c_2+(a_1-b_1)s}{a_3-b_3}, \\ t_2 = & \frac{c_1+(a_2-b_2)s}{a_3-b_3}, \\ t_3 = & s, \end{cases}$$

where s is a real number that satisfies the quadratic equation 2.18

ii. $a_3 = b_3$ *and* $a_2 = b_2$. *Possible solutions must be of the form*

$$\begin{cases} t_0 = & -\frac{\alpha_0+\beta_0}{2}, \\ t_1 = & s, \\ t_2 = & \frac{c_2}{a_1-b_1}, \\ t_3 = & \frac{-c_3}{a_1-b_1}, \end{cases}$$

where s is a real number that satisfies the quadratic equation 2.19

iii. $a_3 = b_3$ *and* $a_2 \neq b_2$. *Possible solutions must be of the form*

$$\begin{cases} t_0 = & -\frac{\alpha_0+\beta_0}{2}, \\ t_1 = & \frac{c_3+(a_1-b_1)s}{a_2-b_2}, \\ t_2 = & s, \\ t_3 = & \frac{-c_1}{a_2-b_2}, \end{cases}$$

where s is a real number that satisfies the quadratic equation 2.20

b) Otherwise. No additional solutions exist.

Example 2.3.2. Find all the solutions of

$$t^2 + (1 - 2\mathbf{i} - 2\mathbf{j} - 2\mathbf{k})\, t + t\,(1 + \mathbf{i} - 3\mathbf{j} + \mathbf{k}) + (-1 - 3\mathbf{i} + \mathbf{j} - 2\mathbf{k}) = 0. \tag{2.26}$$

The following steps have been implemented into a single procedure in Maple. For details, see the “AY” procedure of §A.2.

Let $x = t + 1$. Then 2.26 becomes

$$x^2 + (-2\mathbf{i} - 2\mathbf{j} - 2\mathbf{k})\,x + x\,(\mathbf{i} - 3\mathbf{j} + \mathbf{k}) + (-2 - 2\mathbf{i} + 6\mathbf{j} - \mathbf{k}) = 0. \tag{2.27}$$

We have

$$\sum_{i=1}^{3} c_i\,(a_i - b_i) = (-2)\cdot(-2-1) + 6\cdot(-2+3) + (-1)\cdot(-2-1) = 15 \neq 0.$$

By Theorem 2.3.1, there are only standard solutions and we obtain a cubic polynomial of $y = x_0^2$:

$$f\,(y) = 16y^3 + 152y^2 - 251y - 225.$$

$f\,(y)$ has one positive root

$$y = \frac{13}{6}\sqrt{13}\sin\left(\frac{1}{3}\arctan\left(\frac{18}{85634}\sqrt{10091901}\right) + \frac{1}{6}\pi\right) - \frac{19}{6} = s.$$

The two solutions to 2.26 are

$$\left(\sqrt{s} - 1\right) + \frac{p_1\left(\sqrt{s}\right)}{\sigma\left(\sqrt{s}\right)}\mathbf{i} + \frac{p_2\left(\sqrt{s}\right)}{\sigma\left(\sqrt{s}\right)}\mathbf{j} + \frac{p_3\left(\sqrt{s}\right)}{\sigma\left(\sqrt{s}\right)}\mathbf{k} \approx 0.41 + 0.8\mathbf{i} + 1.14\mathbf{k}$$

and

$$\left(-\sqrt{s} - 1\right) + \frac{p_1\left(-\sqrt{s}\right)}{\sigma\left(-\sqrt{s}\right)}\mathbf{i} + \frac{p_2\left(-\sqrt{s}\right)}{\sigma\left(-\sqrt{s}\right)}\mathbf{j} + \frac{p_3\left(-\sqrt{s}\right)}{\sigma\left(-\sqrt{s}\right)}\mathbf{k} \approx -2.41 + 0.87\mathbf{i} + 1.65\mathbf{j} - 1.93\mathbf{k}$$

These answers are verified to be correct by direct claculation.

2.3.2. Another quadratic equation

In this subsection, we consider another quadratic equation:

$$x^2 + axb + c = 0, \tag{2.28}$$

where a and $b \in \mathbb{H}$ with $a_0 = b_0 = 0$, $c \in \mathbb{R}$ and x is a quaternion indeterminate. We assume that $ab \neq 0$. Otherwise, 2.28 is reduced to $x^2 + c = 0$, which was studied and solved in (Niven, 1942). We also assume that $c \neq 0$. Otherwise, 2.28 is reduced to $x^2 + axb = 0$, which we will discuss in §2.3.3.

2.28 is equivalent to the real non-linear system:

$$\begin{cases} (2x_0 + m_{11})\, x_1 + m_{12}x_2 + m_{13}x_3 = (-a_2b_3 + a_3b_2)\, x_0, & (1) \\ m_{21}x_1 + (2x_0 + m_{22})\, x_2 + m_{23}x_3 = (a_1b_3 - a_3b_1)\, x_0, & (2) \\ m_{31}x_1 + m_{32}x_2 + (2x_0 + m_{33})\, x_3 = (-a_1b_2 + a_2b_1)\, x_0, & (3) \\ x_0^2 - x_1^2 - x_2^2 - x_3^2 + d_1x_0 + d_2x_1 + d_3x_2 + d_4x_3 + c = 0, & (4) \end{cases} \tag{2.29}$$

where m_{ij} and d_i are given by:

$$\begin{aligned} M &= (m_{ij})_{3\times 3} \\ &= \begin{pmatrix} -a_1b_1 + a_2b_2 + a_3b_3 & -a_1b_2 - a_2b_1 & -a_1b_3 - a_3b_1 \\ -a_1b_2 - a_2b_1 & a_1b_1 - a_2b_2 + a_3b_3 & -a_2b_3 - a_3b_2 \\ -a_1b_3 - a_3b_1 & -a_2b_3 - a_3b_2 & a_1b_1 + a_2b_2 - a_3b_3 \end{pmatrix}, \end{aligned}$$

and

$$D = (d_i)_{4\times 1} = \begin{pmatrix} -a_1b_1 - a_2b_2 - a_3b_3 \\ a_2b_3 - a_3b_2 \\ -a_1b_3 + a_3b_1 \\ a_1b_2 - a_2b_1 \end{pmatrix}.$$

We already discussed in §2.2 how to solve 2.28 if the determinant of the coefficient matrix, $\check{M} = \begin{pmatrix} 2x_0 + m_{11} & m_{12} & m_{13} \\ m_{21} & 2x_0 + m_{22} & m_{23} \\ m_{31} & m_{32} & 2x_0 + m_{33} \end{pmatrix}$, is zero. We can and will assume that $\det \check{M} \neq 0$.

$$\begin{aligned}\sigma(x_0) =& \det \check{M} \\ =& (b_1a_1 + a_2b_2 + a_3b_3 + 2x_0)\left(4x_0^2 - \left(a_1^2 + a_2^2 + a_3^2\right)\left(b_1^2 + b_2^2 + b_3^2\right)\right) \neq 0,\end{aligned}$$

therefore, $4x_0^2 - (a_1^2 + a_2^2 + a_3^2)(b_1^2 + b_2^2 + b_3^2) \neq 0$. Applying Cramer's rule to (1), (2) and (3) of 2.29, we obtain x_1 as a rational polynomial of x_0,

$$\begin{aligned}x_1 =& \frac{[x_0(a_3b_2 - a_2b_3)](4x_0^2 - (a_1^2 + a_2^2 + a_3^2)(b_1^2 + b_2^2 + b_3^2))}{(b_1a_1 + a_2b_2 + a_3b_3 + 2x_0)(4x_0^2 - (a_1^2 + a_2^2 + a_3^2)(b_1^2 + b_2^2 + b_3^2))} \\ =& \frac{x_0(a_3b_2 - a_2b_3)}{b_1a_1 + a_2b_2 + a_3b_3 + 2x_0} = p_1(x_0). \end{aligned} \tag{2.30}$$

Similarly,

$$x_2 = \frac{x_0(a_1b_3 - a_3b_1)}{b_1a_1 + a_2b_2 + a_3b_3 + 2x_0} = p_2(x_0) \tag{2.31}$$

and

$$x_3 = \frac{x_0(a_2b_1 - a_1b_2)}{b_1a_1 + a_2b_2 + a_3b_3 + 2x_0} = p_3(x_0). \tag{2.32}$$

Substituting 2.30, 2.31 and 2.32 back into (4) of 2.29, then multiplying through by $(b_1a_1 + a_2b_2 + a_3b_3 + 2x_0)^2$, we obtain a quartic equation of x_0,

$$f(x_0) = x_0^4 + \left(c - \frac{3}{4}\alpha\right)x_0^2 + \frac{1}{4}\beta(4c - \alpha)x_0 + \frac{1}{4}c\beta^2 = 0, \tag{2.33}$$

where

$$\alpha = \left(a_1^2 + a_2^2 + a_3^2\right)\left(b_1^2 + b_2^2 + b_3^2\right)$$

and

$$\beta = a_1b_1 + a_2b_2 + a_3b_3.$$

The discriminant of f is $\triangle_f = \frac{1}{256}\beta^2\left(\alpha-\beta^2\right)\left(\left(\alpha+4c\right)\left(3\alpha-4c\right)^3 - 1024c^3\beta^2\right)$. Since $\alpha-\beta^2 \geq 0$ by the Cauchy–Schwartz inequality, $\triangle_f < 0$ if an only if $\left(\alpha+4c\right)\left(3\alpha-4c\right)^3 - 1024c^3\beta^2 < 0$.

Once we solve f for x_0, we substitute its value back into 2.30, 2.31 and 2.32 to get x_1, x_2 and x_3 respectively to obtain a solution to 2.28.

Example 2.3.3. Find all the solutions of

$$x^2 + \left(-\mathbf{i} - 3\mathbf{j} + 2\mathbf{k}\right)x\left(-2\mathbf{i} + \mathbf{j} - \mathbf{k}\right) + 21 = 0. \tag{2.34}$$

The following steps have been implemented into a single procedure in Maple. For details, see the "AXBc" procedure of §A.2.

2.34 is equivalent to

$$\begin{cases} \left(2x_0-7\right)x_1 - 5x_2 + 3x_3 = -x_0, & (1) \\ -5x_1 + \left(2x_0+3\right)x_2 - 5x_3 = 5x_0, & (2) \\ 3x_1 - 5x_2 + \left(2x_0+1\right)x_3 = 7x_0, & (3) \\ x_1^2 + x_2^2 + x_3^2 - x_1 + 5x_2 + 7x_3 = x_0^2 + 21. & (4) \end{cases} \tag{2.35}$$

When the determinant of the coefficient matrix, $\begin{pmatrix} 2x_0-7 & -5 & 3 \\ -5 & 2x_0+3 & -5 \\ 3 & -5 & 2x_0+1 \end{pmatrix}$, is 0, that is, when $\sigma\left(x_0\right) = 8x_0^3 - 12x_0^2 - 168x_0 + 252 = 0$, 2.35 has no solutions because the last row of the reduced row echelon form of $\begin{pmatrix} 2x_0-7 & -5 & 3 & -x_0 \\ -5 & 2x_0+3 & -5 & 5x_0 \\ 3 & -5 & 2x_0+1 & 7x_0 \end{pmatrix}$ is not a zero row. Otherwise, by 2.33, we obtain a quartic polynomial of x_0,

$$f\left(x_0\right) = x_0^4 - 42x_0^2 + \frac{189}{4}$$

$f\left(x_0\right)$ has four positive roots $x_0 = \pm\frac{1}{2}\sqrt{84 \pm 30\sqrt{7}}$. Therefore, by 2.30, 2.31 and

2.32, the four solutions to 2.34 are

$$\frac{\sqrt{84+30\sqrt{7}}}{2}+p_1\left(\frac{\sqrt{84+30\sqrt{7}}}{2}\right)\mathbf{i}+p_2\left(\frac{\sqrt{84+30\sqrt{7}}}{2}\right)\mathbf{j}+p_3\left(\frac{\sqrt{84+30\sqrt{7}}}{2}\right)\mathbf{k}$$

$$\approx 6.39-0.65\mathbf{i}+3.27\mathbf{j}+4.57\mathbf{k},$$

$$\frac{\sqrt{84-30\sqrt{7}}}{2}+p_1\left(\frac{\sqrt{84-30\sqrt{7}}}{2}\right)\mathbf{i}+p_2\left(\frac{\sqrt{84-30\sqrt{7}}}{2}\right)\mathbf{j}+p_3\left(\frac{\sqrt{84-30\sqrt{7}}}{2}\right)\mathbf{k}$$

$$\approx 1.08+1.27\mathbf{i}-6.34\mathbf{j}-8.87\mathbf{k},$$

$$\frac{-\sqrt{84+30\sqrt{7}}}{2}+p_1\left(\frac{-\sqrt{84+30\sqrt{7}}}{2}\right)\mathbf{i}+p_2\left(\frac{-\sqrt{84+30\sqrt{7}}}{2}\right)\mathbf{j}+p_3\left(\frac{-\sqrt{84+30\sqrt{7}}}{2}\right)\mathbf{k}$$

$$\approx -6.39-0.40\mathbf{i}+2.02\mathbf{j}+2.83\mathbf{k},$$

and

$$\frac{-\sqrt{84-30\sqrt{7}}}{2}+p_1\left(\frac{-\sqrt{84-30\sqrt{7}}}{2}\right)\mathbf{i}+p_2\left(\frac{-\sqrt{84-30\sqrt{7}}}{2}\right)\mathbf{j}+p_3\left(\frac{-\sqrt{84-30\sqrt{7}}}{2}\right)\mathbf{k}$$

$$\approx -1.08-0.21\mathbf{i}+1.04\mathbf{j}+1.46\mathbf{k}$$

These answers are verified to be correct by direct claculation.

2.3.3. A homogeneous quadratic equation

In this subsection, we consider the currently unsolved homogeneous quadratic equation:

$$x^2 + axb = 0, \tag{2.36}$$

where a and $b \in \mathbb{H}$ and x is the quaternion unknown. We assume that $a \notin \mathbb{R}$ and $b \notin \mathbb{R}$. Otherwise, 2.36 is reduced to $x\,(x + ab) = 0$ or $(x + ab)\,x = 0$.

2.36 is equivalent to the real non-linear system:

$$\begin{cases} (2x_0 + m_{11})\,x_1 + m_{12}x_2 + m_{13}x_3 = x_0 n_1, & (1) \\ m_{21}x_1 + (2x_0 + m_{22})\,x_2 + m_{23}x_3 = x_0 n_2, & (2) \\ m_{31}x_1 + m_{32}x_2 + (2x_0 + m_{33})\,x_3 = x_0 n_3, & (3) \\ x_0^2 - x_1^2 - x_2^2 - x_3^2 + d_1 x_0 + d_2 x_1 + d_3 x_2 + d_4 x_3 + c_0 = 0, & (4) \end{cases} \tag{2.37}$$

where m_{ij}, n_i and d_i are given by:

$$M = (m_{ij})_{3\times 3}$$
$$= \begin{pmatrix} a_0b_0 - a_1b_1 + a_2b_2 + a_3b_3 & a_0b_3 - a_1b_2 - a_2b_1 - a_3b_0 & -a_0b_2 - a_1b_3 + a_2b_0 - a_3b_1 \\ -a_0b_3 - a_1b_2 - a_2b_1 + a_3b_0 & a_0b_0 + a_1b_1 - a_2b_2 + a_3b_3 & a_0b_1 - a_1b_0 - a_2b_3 - a_3b_2 \\ a_0b_2 - a_1b_3 - a_2b_0 - a_3b_1 & -a_0b_1 + a_1b_0 - a_2b_3 - a_3b_2 & a_0b_0 + a_1b_1 + a_2b_2 - a_3b_3 \end{pmatrix},$$

$$N = (n_i)_{3\times 1} = \begin{pmatrix} -a_0b_1 - a_1b_0 - a_2b_3 + a_3b_2 \\ -a_0b_2 + a_1b_3 - a_2b_0 - a_3b_1 \\ -a_0b_3 - a_1b_2 + a_2b_1 - a_3b_0 \end{pmatrix},$$

and

$$D = (d_i)_{4\times 1} = \begin{pmatrix} a_0b_0 - a_1b_1 - a_2b_2 - a_3b_3 \\ -a_0b_1 - a_1b_0 + a_2b_3 - a_3b_2 \\ -a_0b_2 - a_1b_3 - a_2b_0 + a_3b_1 \\ -a_0b_3 + a_1b_2 - a_2b_1 - a_3b_0 \end{pmatrix}.$$

We already showed in §2.2 how to solve such an equation when the determinant of the coefficient matrix, $\check{M} = \left[\begin{array}{c|c|c} 2x_0 + m_{11} & m_{12} & m_{13} \\ m_{21} & 2x_0 + m_{22} & m_{23} \\ m_{31} & m_{32} & 2x_0 + m_{33} \end{array}\right]$, is zero. We can and will assume that $\det \check{M} \neq 0$. Direct calculation yields:

$$\begin{aligned}\sigma(x_0) =& \det \check{M} \\ =& 8x_0^3 + 4(m_{11} + m_{22} + m_{33})x_0^2 \\ &+ 2\left(m_{11}m_{22} + m_{11}m_{33} - m_{12}m_{21} - m_{13}m_{31} - m_{23}m_{32} + m_{22}m_{33}\right)x_0 \\ &+ m_{11}m_{22}m_{33} - m_{11}m_{23}m_{32} - m_{12}m_{21}m_{33} \\ &+ m_{13}m_{21}m_{32} + m_{12}m_{23}m_{31} - m_{13}m_{22}m_{31}.\end{aligned}$$

Applying Cramer's rule to (1), (2) and (3) of 2.37, we obtain x_1, x_2 and x_3 as rational polynomials of x_0 as follows:

$$x_1 = \frac{p_1(x_0)}{\sigma(x_0)}, \tag{2.38}$$

$$x_2 = \frac{p_2(x_0)}{\sigma(x_0)} \tag{2.39}$$

and

$$x_3 = \frac{p_3(x_0)}{\sigma(x_0)}, \tag{2.40}$$

where

$$\begin{aligned}p_1(x_0) =& 4n_1x_0^3 + 2\left(n_1\left(m_{22}+m_{33}\right) - n_2m_{12} - n_3m_{13}\right)x_0^2\\ &+ n_1\left(m_{22}m_{33} - m_{23}m_{32}\right)x_0\\ &+ n_2\left(m_{13}m_{32} - m_{12}m_{33}\right)x_0\\ &+ n_3\left(m_{12}m_{23} - m_{13}m_{22}\right)x_0,\end{aligned}$$

$$\begin{aligned}p_2(x_0) =& 4n_2x_0^3 + 2\left(n_2\left(m_{11}+m_{33}\right) - n_1m_{21} - n_3m_{23}\right)x_0^2\\ &+ n_1\left(m_{23}m_{31} - m_{21}m_{33}\right)x_0\\ &+ n_2\left(m_{11}m_{33} - m_{13}m_{31}\right)x_0\\ &+ n_3\left(m_{13}m_{21} - m_{11}m_{23}\right)x_0\end{aligned}$$

and

$$\begin{aligned}p_3(x_0) =& 4n_3x_0^3 + 2\left(n_3\left(m_{11}+m_{22}\right) - n_1m_{31} - n_2m_{32}\right)x_0^2\\ &+ n_1\left(m_{21}m_{32} - m_{22}m_{31}\right)x_0\\ &+ n_2\left(m_{12}m_{31} - m_{11}m_{32}\right)x_0\\ &+ n_3\left(m_{11}m_{22} - m_{12}m_{21}\right)x_0.\end{aligned}$$

Substituting 2.38, 2.39 and 2.40 back into (4) of 2.37, we obtain:

$$\begin{aligned}0 =& x_0^2 + d_1x_0\\ &- \frac{p_1(x_0)^2 + p_2(x_0)^2 + p_3(x_0)^2}{\sigma(x_0)^2} + \frac{d_2p_1(x_0) + d_3p_2(x_0) + d_4p_3(x_0)}{\sigma(x_0)},\end{aligned}$$

Multiplication of the above equation by $\sigma(x_0)^2$ yields an 8^{th} degree polynomial of

x_0,

$$\begin{aligned} 0 =& f\left(x_0\right) \\ =& \sigma\left(x_0\right)^2\left(x_0^2+d_1 x_0\right)-d_1^2 x_0^2-p_2\left(x_0\right)^2-p_3\left(x_0\right)^2 \\ &+\sigma\left(x_0\right)\left(d_2 p_1\left(x_0\right)+d_3 p_2\left(x_0\right)+d_4 p_3\left(x_0\right)\right). \end{aligned} \tag{2.41}$$

2.41 can be factored into the product of x_0, a cubic and a quartic polynomial of x_0,

$$f\left(x_0\right)=x_0 h\left(x_0\right) g\left(x_0\right),$$

where the cubic polynomial is

$$\begin{aligned} h\left(x_0\right) =& x_0^3+2 a_0 b_0 x_0^2 \\ &+\left(\frac{4}{3} a_0^2 b_0^2-\frac{1}{12}\left(a_0^2-3 a_1^2-3 a_2^2-3 a_3^2\right)\left(b_0^2-3 b_1^2-3 b_2^2-3 b_3^2\right)\right) x_0 \\ &+\frac{1}{4} B\left(a_0 b_0-a_1 b_1-a_2 b_2-a_3 b_3\right), \end{aligned} \tag{2.42}$$

and the quartic polynomial is

$$g\left(x_0\right)=x_0^4+2 a_0 b_0 x_0^3+\frac{1}{2} A x_0^2+\frac{1}{2} a_0 b_0 B x_0+\frac{1}{16} B^2, \tag{2.43}$$

with

$$A=4 a_0^2 b_0^2-\left(a_1^2+a_2^2+a_3^2-a_0^2\right)\left(b_1^2+b_2^2+b_3^2-b_0^2\right)$$

and

$$B=\left(a_0^2+a_1^2+a_2^2+a_3^2\right)\left(b_0^2+b_1^2+b_2^2+b_3^2\right).$$

Furthermore, $g\left(x_0\right)$ can be factored in $\mathbb{R}$ as

$$g\left(x_0\right)=g_1\left(x_0\right) g_2\left(x_0\right),$$

where

$$g_1(x_0) = x_0^2 + sx_0 + \frac{1}{4}B$$

and

$$g_2(x_0) = x_0^2 + tx_0 + \frac{1}{4}B,$$

with

$$s = a_0b_0 + \sqrt{(b_1^2 + b_2^2 + b_3^2)(a_1^2 + a_2^2 + a_3^2)}$$

and

$$t = a_0b_0 - \sqrt{(b_1^2 + b_2^2 + b_3^2)(a_1^2 + a_2^2 + a_3^2)}.$$

The discriminants of g_1 and g_2 are

$$\begin{aligned}\triangle_{g_1} =& - b_3^2a_0^2 - a_1^2b_0^2 - a_2^2b_0^2 - a_3^2b_0^2 - b_2^2a_0^2 - a_0^2b_1^2 \\ &+ 2b_0a_0\sqrt{(b_1^2 + b_2^2 + b_3^2)(a_1^2 + a_2^2 + a_3^2)} \\ =& - \left(a_0\sqrt{(b_1^2 + b_2^2 + b_3^2)} - b_0\sqrt{(a_1^2 + a_2^2 + a_3^2)}\right)^2 \leq 0\end{aligned}$$

and

$$\begin{aligned}\triangle_{g_2} =& - b_3^2a_0^2 - a_1^2b_0^2 - a_2^2b_0^2 - a_3^2b_0^2 - b_2^2a_0^2 - a_0^2b_1^2 \\ &- 2b_0a_0\sqrt{(b_1^2 + b_2^2 + b_3^2)(a_1^2 + a_2^2 + a_3^2)} \\ =& - \left(a_0\sqrt{(b_1^2 + b_2^2 + b_3^2)} + b_0\sqrt{(a_1^2 + a_2^2 + a_3^2)}\right)^2 \leq 0.\end{aligned}$$

We continue the discussion by cases:

Case 1. When $a_0 = b_0 = 0$.

Then $h(x_0)$, $g(x_0)$ and $\sigma(x_0)$ can be simplified to:

$$\begin{aligned} h(x_0) =& x_0^3 - \frac{1}{12}\left(3a_1^2 + 3a_2^2 + 3a_3^2\right)\left(3b_1^2 + 3b_2^2 + 3b_3^2\right)x_0 \\ & - \frac{1}{4}\left(a_1^2 + a_2^2 + a_3^2\right)\left(b_1^2 + b_2^2 + b_3^2\right)\left(a_1b_1 + a_2b_2 + a_3b_3\right), \end{aligned} \tag{2.44}$$

$$g(x_0) = \left(x_0^2 - \frac{1}{4}\left(a_1^2 + a_2^2 + a_3^2\right)\left(b_1^2 + b_2^2 + b_3^2\right)\right)^2 \tag{2.45}$$

and

$$\sigma(x_0) = 4\left(2x_0 + a_1b_1 + a_2b_2 + a_3b_3\right)\left(x_0^2 - \frac{1}{4}\left(a_1^2 + a_2^2 + a_3^2\right)\left(b_1^2 + b_2^2 + b_3^2\right)\right). \tag{2.46}$$

Note that we already assumed that $\sigma(x_0) \neq 0$, and therefore 2.45 does not contribute any roots. We next show that 2.44 always has three real solutions. The discriminant of h is

$$\begin{aligned} \triangle_h =& \frac{27}{16}\left(a_1^2 + a_2^2 + a_3^2\right)^2\left(b_1^2 + b_2^2 + b_3^2\right)^2 \\ & \times \left(\left(a_1b_2 - a_2b_1\right)^2 + \left(a_1b_3 - a_3b_1\right)^2 + \left(a_2b_3 - a_3b_2\right)^2\right) \\ \geq& 0. \end{aligned}$$

So the solution to $f(x_0)$ are 0 and the 3 real roots of 2.44.

Case 2. Otherwise. $\triangle_{g_1} < 0$ unless $\frac{a_0}{b_0} = \frac{\sqrt{\left(a_1^2+a_2^2+a_3^2\right)}}{\sqrt{\left(b_1^2+b_2^2+b_3^2\right)}}$, that is, $a_0 = cb_0$ and $a_1^2 + a_2^2 + a_3^2 = c^2\left(b_1^2 + b_2^2 + b_3^2\right)$ for some $0 < c$. Under these conditions, $\triangle_{g_2} < 0$ and therefore g_2 does not contribute any solutions. However, g_1 can be simplified to

$$x_0^2 + c\left(b_0^2 + b_1^2 + b_2^2 + b_3^2\right)x_0 + \frac{1}{4}c^2\left(b_0^2 + b_1^2 + b_2^2 + b_3^2\right)^2. \tag{2.47}$$

Similarly, $\triangle_{g_2} < 0$ unless $\frac{a_0}{b_0} = -\frac{\sqrt{\left(a_1^2+a_2^2+a_3^2\right)}}{\sqrt{\left(b_1^2+b_2^2+b_3^2\right)}} = c$ for some $c < 0$. Under these conditions, $\triangle_{g_1} < 0$ and therefore g_1 does not contribute any solutions. However g_2 can also be simplified to 2.47. So the roots of $f(x_0)$ are 0, the real root(s) of 2.42 and $-\frac{a_0\left(b_0^2+b_1^2+b_2^2+b_3^2\right)}{2b_0}$ which has multiplicity 2 and is contributed by 2.47.

Once we obtain x_0, we substitute its value back into 2.38, 2.39 and 2.40 to get x_1, x_2 and x_3 respectively to obtain a solution to 2.36.

Example 2.3.4. Find all the solutions of

$$x^2 + (-19 + 15\mathbf{i} - \mathbf{j} + 11\mathbf{k})\, x\, (9 + 10\mathbf{i} - 4\mathbf{j} + 10\mathbf{k}) = 0. \tag{2.48}$$

The following steps have been implemented into a single procedure in Maple. For details, see the "AXB" procedure of §A.2.

2.48 is equivalent to

$$\begin{cases} (2x_0 - 207)\, x_1 - 219x_2 - 345x_3 = 21x_0, & (1) \\ 359x_1 + (2x_0 + 85)\, x_2 - 271x_3 = -27x_0, & (2) \\ -175x_1 + 379x_2 + (2x_0 - 127)\, x_3 = 141x_0, & (3) \\ x_1^2 + x_2^2 + x_3^2 - 89x_1 + 107x_2 - 41x_3 = x_0^2 - 435x_0. & (4) \end{cases} \tag{2.49}$$

When the determinant of the coefficient matrix,

$$\begin{pmatrix} 2x_0 - 207 & -219 & -345 \\ 359 & 2x_0 + 85 & -271 \\ -175 & 379 & 2x_0 - 127 \end{pmatrix}$$

is 0, that is, when $\sigma(x_0) = 8x_0^3 - 996x_0^2 + 237708x_0 - 91470060 = 0$, 2.49 has no solutions because the last row of the reduced row echelon form of

$$\begin{pmatrix} 2x_0 - 207 & -219 & -345 & 21x_0 \\ 359 & 2x_0 + 85 & -271 & -27x_0 \\ -175 & 379 & 2x_0 - 127 & 141x_0 \end{pmatrix}$$

is not a zero row. Otherwise, by 2.42, we obtain a cubic polynomial of x_0,

$$f\left(x_0\right) = x^3 - 342x^2 + 6858x - 22867515$$

$f\left(x_0\right)$ has one real root:

$$\begin{aligned} x_0 =& \frac{3}{2}\left(3710932 + 4\sqrt{853947883289}\right)^{\frac{1}{3}} + \frac{7140}{\left(3710932 + 4\sqrt{853947883289}\right)^{\frac{1}{3}}} + 114 \\ =& \delta. \end{aligned}$$

Therefore, by 2.38, 2.39 and 2.40, the solutions to 2.48 are 0 and

$$\begin{aligned} &\delta + \frac{p_1\left(\delta\right)}{\sigma\left(\delta\right)}\mathbf{i} + \frac{p_2\left(\delta\right)}{\sigma\left(\delta\right)}\mathbf{j} + \frac{p_3\left(\delta\right)}{\sigma\left(\delta\right)}\mathbf{k} \\ \approx& 443.03 + 62.3\mathbf{i} - 7.35\mathbf{j} + 100.33\mathbf{k}. \end{aligned}$$

This answer is verified to be correct by direct claculation.

2.3.4. A pair of quaternion equations

Two nonzero quaternions a and b are said to be semisimilar if the following system has a solution:

$$\begin{cases} yax = b, \\ xby = a, \end{cases} \tag{2.50}$$

where x and y are nonzero quaternion indeterminates. 2.50 was studied and solved in (Tian, 2010). In this subsection, we present a somewhat simpler solution.

2.50 can be written as

$$\begin{cases} y = bx^{-1}a^{-1}, \\ y = b^{-1}x^{-1}a, \end{cases}$$

this gives:

$$bx^{-1}a^{-1} = b^{-1}x^{-1}a \tag{2.51}$$

$$\Rightarrow$$

$$|b|\left|a^{-1}\right| = \left|b^{-1}\right||a| \Rightarrow |a| = |b|.$$

Therefore, a and b must satisfy the following equation for 2.55 to have solutions,

$$a_0^2 + a_1^2 + a_2^2 + a_3^2 = b_0^2 + b_1^2 + b_2^2 + b_3^2. \tag{2.52}$$

2.51 is equivalent to $xb^2 = a^2x$, which is equivalent to the real linear homogeneous system:

$$\begin{cases} m_{11}x_1 + m_{12}x_2 + m_{13}x_3 + m_{14}x_0 = 0, \\ m_{21}x_1 + m_{22}x_2 + m_{23}x_3 + m_{24}x_0 = 0, \\ m_{31}x_1 + m_{32}x_2 + m_{33}x_3 + m_{34}x_0 = 0, \\ m_{41}x_1 + m_{42}x_2 + m_{43}x_3 + m_{44}x_0 = 0. \end{cases} \tag{2.53}$$

where m_{ij} is given by the 4×4 skew-symmetric matrix:

$$M = (m_{ij})_{4\times 4}$$
$$\begin{pmatrix} d & -2\left(a_0a_3 + b_0b_3\right) & 2\left(a_0a_2 + b_0b_2\right) & 2\left(a_0a_1 - b_0b_1\right) \\ 2\left(a_0a_3 + b_0b_3\right) & d & -2\left(a_0a_1 + b_0b_1\right) & 2\left(a_0a_2 - b_0b_2\right) \\ -2\left(a_0a_2 + b_0b_2\right) & 2\left(a_0a_1 + b_0b_1\right) & d & 2\left(a_0a_3 - b_0b_3\right) \\ -2\left(a_0a_1 - b_0b_1\right) & -2\left(a_0a_2 - b_0b_2\right) & -2\left(a_0a_3 - b_0b_3\right) & d \end{pmatrix},$$

where the diagonal entries are $d = a_0^2 - a_1^2 - a_2^2 - a_3^2 - b_0^2 + b_1^2 + b_2^2 + b_3^2$. By 2.52,

$$d = a_0^2 - b_0^2 + a_0^2 - b_0^2 = 2a_0^2 - 2b_0^2.$$

Therefore, M becomes:

$$\begin{pmatrix} a_0^2 - b_0^2 & -\left(a_0a_3 + b_0b_3\right) & a_0a_2 + b_0b_2 & a_0a_1 - b_0b_1 \\ a_0a_3 + b_0b_3 & a_0^2 - b_0^2 & -\left(a_0a_1 + b_0b_1\right) & a_0a_2 - b_0b_2 \\ -\left(a_0a_2 + b_0b_2\right) & a_0a_1 + b_0b_1 & a_0^2 - b_0^2 & a_0a_3 - b_0b_3 \\ -\left(a_0a_1 - b_0b_1\right) & -\left(a_0a_2 - b_0b_2\right) & -\left(a_0a_3 - b_0b_3\right) & a_0^2 - b_0^2 \end{pmatrix}.$$

We continue the discussion by cases.

Case 1. If $a_0 = b_0$ or $a_0 = -b_0$, then $\det M = 0$. Therefore, the homogeneous system 2.53 has a solution set with an infinite number of solutions $(x_0,\, x_1,\, x_2,\, x_3)$. Each such x gives a solution to 2.50,

$$\begin{cases} x = x, \\ y = bx^{-1}a^{-1}. \end{cases}$$

Case 2. Otherwise, we show that 2.50 has no solutions. Let $A = a_0^2 + a_1^2 + a_2^2 + a_3^2$ and $B = b_0^2 + b_1^2 + b_2^2 + b_3^2$. Direct calculation yields:

$$\begin{aligned} \det M =& a_0^2 b_0^2 \left(A - B\right) \left(A - B - 4a_0^2 + 4b_0^2\right) \\ &+ \left(a_0 + b_0\right)\left(a_0 - b_0\right)\left(a_0A - b_0B\right)\left(a_0A + b_0B\right). \end{aligned}$$

$A = B$ by 2.52 and therefore,

$$\begin{aligned}\det M =& 0 + (a_0 + b_0)(a_0 - b_0) A (a_0 - b_0) A (a_0 + b_0) \\ =& (a_0 + b_0)^2 (a_0 - b_0)^2 A^2.\end{aligned}$$

Since $a \neq 0$, $A \neq 0$ and therefore $\det M \neq 0$. This implies that the homogeneous system 2.53 has only trivial solutions $(x_0, x_1, x_2, x_3) = (0, 0, 0, 0)$. A contradiction. Therefore, 2.50 has no solutions.

Example 2.3.5. Find all (x, y) such that the following system holds

$$\begin{cases} y(2 + 5\mathbf{i} + 12\mathbf{j} - 5\mathbf{k})\, x = 2 + 4\mathbf{i} - 3\mathbf{j} + 13\mathbf{k}, \\ x(2 + 4\mathbf{i} - 3\mathbf{j} + 13\mathbf{k})\, y = 2 + 5\mathbf{i} + 12\mathbf{j} - 5\mathbf{k}. \end{cases} \tag{2.54}$$

The following steps have been implemented into a single procedure in Maple. For details, see the “SemiSim” procedure of §A.2.

Since $\operatorname{Re}(2 + 4\mathbf{i} - 3\mathbf{j} + 13\mathbf{k}) = 2 = \operatorname{Re}(2 + 5\mathbf{i} + 12\mathbf{j} - 5\mathbf{k})$ and

$$\begin{aligned}|2 + 4\mathbf{i} - 3\mathbf{j} + 13\mathbf{k}| =& 2^2 + 4^2 + (-3)^2 + 13^2 = 198 \\ =& 2^2 + 5^2 + 12^2 + (-5)^2 = |2 + 5\mathbf{i} + 12\mathbf{j} - 5\mathbf{k}|,\end{aligned}$$

2.54 has infinitely many solutions. First we solve the linear system

$$\begin{cases} -16x_2 + 18x_3 - 2x_0 = 0, \\ 16x_1 - 18x_3 - 30x_0 = 0, \\ 18x_1 - 18x_2 - 36x_0 = 0, \\ 2x_1 + 30x_2 - 36x_3 = 0. \end{cases}$$

Direct calculation yields:

$$\begin{cases} x_0 = -8t_2 + 9t_3, \\ x_1 = -15t_2 + 18t_3, \\ x_2 = t_2, \\ x_3 = t_3, \end{cases}$$

where t_2 and t_3 are arbitrary real numbers such that $x = x_0 + x_1\mathbf{i} + x_2\mathbf{j} + x_3\mathbf{k} \neq 0$.

Then we have a solution set with an infinite number of solutions of the form:

$$\begin{cases} x = x, \\ y = (2 + 5\mathbf{i} + 12\mathbf{j} - 5\mathbf{k})\, x^{-1} \,(2 + 4\mathbf{i} - 3\mathbf{j} + 13\mathbf{k})^{-1}. \end{cases}$$

For example, when $t_2 = 2$ and $t_3 = 3$, we have:

$$\begin{cases} x = -11 + 24\mathbf{i} + 2\mathbf{j} + 3\mathbf{k}, \\ y = -\frac{11}{710} - \frac{12}{355}\mathbf{i} - \frac{1}{355}\mathbf{j} - \frac{3}{710}\mathbf{k}. \end{cases}$$

Direct calculation shows that $(x,\, y)$ satisfies 2.54.

2.3.5. Another pair of quaternion equations

Two nonzero quaternions a and b are said to be consemisimilar if the following system has a solution:

$$\begin{cases} \overline{y}ax = b, \\ \overline{x}by = a, \end{cases} \tag{2.55}$$

where x and y are nonzero quaternion indeterminates. 2.55 was studied and solved in (Tian, 2010). In this subsection, we present a somewhat simpler solution.

Since $\overline{st} = \overline{t}\,\overline{s}$ and $\overline{s}^{-1} = \overline{s^{-1}}$ for all s, $t \in \mathbb{H}$, 2.55 can be written as

$$\begin{cases} y = \overline{a}^{-1}\overline{x}^{-1}\overline{b}, \\ y = b^{-1}\overline{x}^{-1}a, \end{cases}$$

this gives:

$$\overline{a}^{-1}\overline{x}^{-1}\overline{b} = b^{-1}\overline{x}^{-1}a \tag{2.56}$$

$$\Rightarrow$$

$$\left|\overline{a}^{-1}\right|\left|\overline{b}\right| = \left|b^{-1}\right||a| \Rightarrow |a| = |b|.$$

Therefore a and b must satisfy the following equation in order for 2.55 to have solutions,

$$a_0^2 + a_1^2 + a_2^2 + a_3^2 = b_0^2 + b_1^2 + b_2^2 + b_3^2. \tag{2.57}$$

2.56 is equivalent to $\overline{x}^{-1}\overline{b}a^{-1} = \overline{a}b^{-1}\overline{x}^{-1}$. On the other hand, we have:

$$a^{-1} = \frac{a_0 - a_1\mathbf{i} - a_2\mathbf{j} - a_3\mathbf{k}}{a_0^2 + a_1^2 + a_2^2 + a_3^2} = \frac{\overline{a}}{a_0^2 + a_1^2 + a_2^2 + a_3^2}$$

and

$$b^{-1} = \frac{b_0 - b_1\mathbf{i} - b_2\mathbf{j} - b_3\mathbf{k}}{b_0^2 + b_1^2 + b_2^2 + b_3^2} = \frac{\overline{b}}{b_0^2 + b_1^2 + b_2^2 + b_3^2}.$$

By 2.57, 2.56 is equivalent to $\overline{b}\overline{a}\overline{x} = \overline{x}\overline{a}\overline{b} \Leftrightarrow xab = bax$, which is equivalent to the following system:

$$\begin{cases} m_{11}x_1 + m_{12}x_2 + m_{13}x_3 + m_{14}x_0 = 0, \\ m_{21}x_1 + m_{22}x_2 + m_{23}x_3 + m_{24}x_0 = 0, \\ m_{31}x_1 + m_{32}x_2 + m_{33}x_3 + m_{34}x_0 = 0, \\ m_{41}x_1 + m_{42}x_2 + m_{43}x_3 + m_{44}x_0 = 0, \end{cases} \tag{2.58}$$

where m_{ij} is given by the 4×4 skew-symmetric matrix:

$$\begin{aligned} M &= (m_{ij})_{4\times4} \\ &= \begin{pmatrix} 0 & a_0b_3 + a_3b_0 & -(a_0b_2 + a_2b_0) & a_2b_3 - a_3b_2 \\ -(a_0b_3 + a_3b_0) & 0 & a_0b_1 + a_1b_0 & -(a_1b_3 - a_3b_1) \\ a_0b_2 + a_2b_0 & -(a_0b_1 + a_1b_0) & 0 & a_1b_2 - a_2b_1 \\ -(a_2b_3 - a_3b_2) & a_1b_3 - a_3b_1 & -(a_1b_2 - a_2b_1) & 0 \end{pmatrix}. \end{aligned}$$

Direct calculation yields $\det M = 0$. Therefore, the homogeneous system 2.58 has a solution set with an infinite number of solutions (x_0, x_1, x_2, x_3). Each such x gives a solution to 2.55:

$$\begin{cases} x = & x, \\ y = & b^{-1}\overline{x}^{-1}a. \end{cases}$$

Example 2.3.6. Find all (x, y) such that the following system holds:

$$\begin{cases} \overline{y}\,(1 - 25\mathbf{i} + 8\mathbf{j} + 15\mathbf{k})\,x = 17 + \mathbf{i} - 24\mathbf{j} - 7\mathbf{k}, \\ \overline{x}\,(17 + \mathbf{i} - 24\mathbf{j} - 7\mathbf{k})\,y = 1 - 25\mathbf{i} + 8\mathbf{j} + 15\mathbf{k}. \end{cases} \tag{2.59}$$

The following steps have been implemented into a single procedure in Maple. For details, see the “CSSim” procedure of §A.2.

Since

$$\begin{aligned}|17+\mathbf{i}-24\mathbf{j}-7\mathbf{k}| =& 17^2+1^2+(-24)^2+(-7)^2=915\\ =& 1^2+(-25)^2+8^2+15^2=|1-25\mathbf{i}+8\mathbf{j}+15\mathbf{k}|\,,\end{aligned}$$

2.59 has infinitely many solutions. First we solve the linear system:

$$\begin{cases}248x_2-112x_3-304x_0=0,\\ -248x_1-424x_3+160x_0=0,\\ 112x_1+424x_2-592x_0=0,\\ 304x_1-160x_2+592x_3=0,\end{cases}$$

to obtain:

$$\begin{cases}x_0=-\frac{31}{20}t_1-\frac{53}{20}t_3,\\ x_1=t_1,\\ x_2=\frac{19}{10}t_1+\frac{37}{10}t_3,\\ x_3=t_3,\end{cases}$$

where t_1 and t_3 are arbitrary real numbers such that $x=x_0+x_1\mathbf{i}+x_2\mathbf{j}+x_3\mathbf{k}\neq 0$. Then we have a solution set with an infinite number of solutions of the form:

$$\begin{cases}x= & x,\\ y= & (17+\mathbf{i}-24\mathbf{j}-7\mathbf{k})^{-1}\,\overline{x}^{-1}\,(1-25\mathbf{i}+8\mathbf{j}+15\mathbf{k})\,.\end{cases}$$

For example, when $t_1=1$ and $t_3=3$, we have:

$$\begin{cases}x=-\frac{19}{2}+\mathbf{i}+13\mathbf{j}+3\mathbf{k},\\ y=\frac{2}{1077}+\frac{56}{1077}\mathbf{i}-\frac{4}{359}\mathbf{j}-\frac{32}{1077}\mathbf{k}.\end{cases}$$

Direct calculation shows that $(x,\,y)$ satisfies 2.59.

2.3.6. A least norm problem

The following least norm problem was proposed in (Tian, 2010),

$$\min_{x\in\mathbb{H}} |ax - xb - c|,$$

where a, b and $c \in \mathbb{H}$ and x is a quaternion indeterminate that minimizes the value of $|ax - xb - c|$. In this subsection, we present a solution.

We consider:

$$ax - xb = c. \tag{2.60}$$

2.60 is equivalent to the real linear homogeneous system:

$$\begin{cases} m_{11}x_1 + m_{12}x_2 + m_{13}x_3 + m_{14}x_0 = c_1, \\ m_{21}x_1 + m_{22}x_2 + m_{23}x_3 + m_{24}x_0 = c_2, \\ m_{31}x_1 + m_{32}x_2 + m_{33}x_3 + m_{34}x_0 = c_3, \\ m_{41}x_1 + m_{42}x_2 + m_{43}x_3 + m_{44}x_0 = c_0, \end{cases} \tag{2.61}$$

where m_{ij} is given by:

$$M = (m_{ij})_{4\times 4} = \begin{pmatrix} a_0 - b_0 & -a_3 - b_3 & a_2 + b_2 & a_1 - b_1 \\ a_3 + b_3 & a_0 - b_0 & -a_1 - b_1 & a_2 - b_2 \\ -a_2 - b_2 & a_1 + b_1 & a_0 - b_0 & a_3 - b_3 \\ -a_1 + b_1 & -a_2 + b_2 & -a_3 + b_3 & a_0 - b_0 \end{pmatrix}.$$

Direct calculation yields:

$$\begin{aligned} \det M = &\left(a_1^2 + a_2^2 + a_3^2 - b_1^2 - b_2^2 - b_3^2\right)^2 \\ &+ (a_0 - b_0)^2 \left((a_0 - b_0)^2 + 2a_1^2 + 2a_2^2 + 2a_3^2 + 2b_1^2 + 2b_2^2 + 2b_3^2\right). \end{aligned}$$

We continue the discussion by cases:

Case 1. $a \sim b$. Then $\det M = 0$. Therefore, we can find the least norm solution of 2.61 by calculating $\begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_0 \end{pmatrix} = M^{\dagger}c'$ where $M^{\dagger}$ is the Moore-Penrose pseudoinverse of M and $c' = \begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ c_0 \end{pmatrix}$ (Bjõrck, 1996). This means that $x = y_0 + y_1\mathbf{i} + y_2\mathbf{j} + y_3\mathbf{k}$ satisfies $|ax - xb - c| \leq |az - zb - c|$ for all $z \in \mathbb{H}$.

Case 2. Otherwise. Then $\det M \neq 0$. Therefore, we can solve the linear system 2.61 to obtain a unique solution $\begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_0 \end{pmatrix}$ such that 2.61 holds. This means that $x = y_0 + y_1\mathbf{i} + y_2\mathbf{j} + y_3\mathbf{k}$ satisfies $|ax - xb - c| = 0$.

Example 2.3.7. Solve the following least norm problem:

$$\min_{x\in\mathbb{H}} \left|(5 - 10\mathbf{i} - 5\mathbf{j} + 2\mathbf{k})\, x - x\,(3 - 4\mathbf{i} - 4\mathbf{j} - 8\mathbf{k}) - (-9 - 2\mathbf{i} + 10\mathbf{j} - 2\mathbf{k})\right|. \quad (2.62)$$

The following steps have been implemented into a single procedure in Maple. For details, see the “LN1” procedure of §A.2.

Since $5 - 10\mathbf{i} - 5\mathbf{j} + 2\mathbf{k} \nsim 3 - 4\mathbf{i} - 4\mathbf{j} - 8\mathbf{k}$, there exists an exact solution. We solve the linear system:

$$\begin{cases} 2x_1 + 6x_2 - 9x_3 - 6x_0 = -2, \\ -6x_1 + 2x_2 + 14x_3 - x_0 = 10, \\ 9x_1 - 14x_2 + 2x_3 + 10x_0 = -2, \\ 6x_1 + x_2 - 10x_3 + 2x_0 = -9, \end{cases}$$

to obtain:

$$\begin{cases} x_0 = -\frac{3364}{2905}, \\ x_1 = \frac{128}{415}, \\ x_2 = -\frac{1073}{2905}, \\ x_3 = \frac{2372}{2905}. \end{cases}$$

Direct calculation shows that $x = -\frac{3364}{2905} + \frac{128}{415}\mathbf{i} - \frac{1073}{2905}\mathbf{j} + \frac{2372}{2905}\mathbf{k}$ satisfies 2.62.

2.3.7. Another least norm problem

The following least norm problem was proposed in (Tian, 2010):

$$\min_{x\in\mathbb{H}} |ax - \bar{x}b - c|, \tag{2.63}$$

where a, b and $c \in \mathbb{H}$ and x is a quaternion indeterminate that minimizes the value of $|ax - \bar{x}b - c|$. In this subsection, we present a solution.

We consider:

$$ax - \bar{x}b - c = 0. \tag{2.64}$$

2.64 is equivalent to the real linear homogeneous system:

$$\begin{cases} m_{11}x_1 + m_{12}x_2 + m_{13}x_3 + m_{14}x_0 = c_1, \\ m_{21}x_1 + m_{22}x_2 + m_{23}x_3 + m_{24}x_0 = c_2, \\ m_{31}x_1 + m_{32}x_2 + m_{33}x_3 + m_{34}x_0 = c_3, \\ m_{41}x_1 + m_{42}x_2 + m_{43}x_3 + m_{44}x_0 = c_0, \end{cases} \tag{2.65}$$

where m_{ij} is given by:

$$M = (m_{ij})_{4\times 4} = \begin{pmatrix} a_0 + b_0 & -a_3 + b_3 & a_2 - b_2 & a_1 - b_1 \\ a_3 - b_3 & a_0 + b_0 & -a_1 + b_1 & a_2 - b_2 \\ -a_2 + b_2 & a_1 - b_1 & a_0 + b_0 & a_3 - b_3 \\ -a_1 - b_1 & -a_2 - b_2 & -a_3 - b_3 & a_0 - b_0 \end{pmatrix}.$$

Direct calculation yields:

$$\begin{aligned} \det M = &\left((a_0 + b_0)^2 + (a_1 - b_1)^2 + (a_2 - b_2)^2 + (a_3 - b_3)^2\right) \\ &\times \left(a_0^2 + a_1^2 + a_2^2 + a_3^2 - b_0^2 - b_1^2 - b_2^2 - b_3^2\right). \end{aligned}$$

We continue the discussion by cases:

Case 1. $a \sim b$. Then $\det M = 0$. Therefore, we can find the least norm solution of 2.65 by calculating $\begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_0 \end{pmatrix} = M^\dagger c'$ where $M^\dagger$ is the Moore-Penrose pseudoinverse of M and $c' = \begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ c_0 \end{pmatrix}$ (Bjõrck, 1996). This means that $x = y_0 + y_1\mathbf{i} + y_2\mathbf{j} + y_3\mathbf{k}$ satisfies $|ax - \bar{x}b - c| \le |az - \overline{z}b - c|$ for all $z \in \mathbb{H}$.

Case 2. Otherwise. Then $\det M \neq 0$. Therefore, we can solve the linear system 2.65 to obtain a unique solution $\begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_0 \end{pmatrix}$ such that 2.65 holds. This means that $x = y_0 + y_1\mathbf{i} + y_2\mathbf{j} + y_3\mathbf{k}$ satisfies $|ax - \bar{x}b - c| = 0$.

Example 2.3.8. Solve the following least norm problem:

$$\min_{x\in\mathbb{H}} \left|(6 - 8\mathbf{i} + \mathbf{j} + 5\mathbf{k})\, x - \overline{x}\,(6 + \mathbf{i} + 5\mathbf{j} - 8\mathbf{k}) - (-3 + \mathbf{i} + \mathbf{j} - 5\mathbf{k})\right|. \tag{2.66}$$

The following steps have been implemented into a single procedure in Maple. For details, see the “LN2” procedure of §A.2.

Since $6 - 8\mathbf{i} + \mathbf{j} + 5\mathbf{k} \sim 6 + \mathbf{i} + 5\mathbf{j} - 8\mathbf{k}$, we can only find a least square solution. We calculate $M^\dagger c'$ where $M^\dagger$ is the Moore-Penrose pseudoinverse of

$$M = \begin{pmatrix} 12 & -13 & -4 & -9 \\ 13 & 12 & 9 & -4 \\ 4 & -9 & 12 & 13 \\ 7 & -6 & 3 & 0 \end{pmatrix} \text{ and } c' = \begin{pmatrix} 1 \\ 1 \\ -5 \\ -3 \end{pmatrix} \text{ to obtain:}$$

$$\begin{cases} x_0 = -\frac{39}{205}, \\ x_1 = -\frac{119}{7380}, \\ x_2 = \frac{1133}{8610}, \\ x_3 = -\frac{2519}{17220}. \end{cases}$$

Direct calculation shows that $x = -\frac{39}{205} - \frac{119}{7380}\mathbf{i} + \frac{1133}{8610}\mathbf{j} - \frac{2519}{17220}\mathbf{k}$ satisfies 2.66.

3. Quaternion Polynomial Matrices

3.1. Generalized Inverse

Definition 3.1.1. $\mathbb{H}[x]$ denotes the set of quaternion polynomials

$$\left\{ \sum_{i=0}^{n} a_i x^i \mid a_i \in \mathbb{H},\, a_n \neq 0,\, n \in \mathbb{N} \right\},$$

where x commutes element-wise with $\mathbb{H}$. $\mathbb{H}[x]^{m\times n}$ denotes the space of $m \times n$ matrices with entries from $\mathbb{H}[x]$.

Remark 3.1.2. If no ambiguity arises, the indeterminate x will be omitted and we write $p \in \mathbb{H}[x]$ or $A \in \mathbb{H}[x]^{m\times n}$ and so on. I_m denotes the $m \times m$ identity matrix. $0_{m\times n}$ denotes the $m \times n$ zero matrix. When the dimension is unspecified, a matrix is of appropriate dimension.

Definition 3.1.3. Let $A \in \mathbb{H}[x]^{m\times n}$ and $\sum_{i=0}^{n} p_i x^i = p_0 + p_1 x + \cdots + p_n x^n = p \in \mathbb{H}[x]$ where $p_i \in \mathbb{H}$. We give a list of some notations as follows:

- $\overline{p} = \sum_{i=0}^{n} \overline{p_i} x^i$, $\overline{A}$ denotes the conjugate of A, $\left(\overline{A}\right)_{ij} = \overline{(A_{ij})}$.
- If $A = P + Q\mathbf{j}$ where $P, Q \in \mathbb{C}[x]^{m\times n}$, then $\chi_A = \begin{pmatrix} P & Q \\ -\overline{Q} & \overline{P} \end{pmatrix} \in \mathbb{C}[x]^{2m\times 2n}$ denotes the complex adjoint of A.
- $A^T \in \mathbb{H}[x]^{n\times m}$ denotes the transpose of A, $\left(A^T\right)_{ij} = A_{ji}$.

- $A^* \in \mathbb{H}[x]^{n\times m}$ denotes the conjugate transpose of A, $(A^*)_{ij} = \overline{(A_{ji})}$. In particular, for all $p \in \mathbb{H}[x]$, $p^* = \overline{p}$.
- $A^\dagger \in \mathbb{H}[x]^{n\times m}$, a generalized inverse of A (Penrose, 1955), if it exists, denotes the solution of the following system of equations,

$$AXA = A \tag{3.1}$$

$$XAX = X \tag{3.2}$$

$$(AX)^* = AX \tag{3.3}$$

$$(XA)^* = XA \tag{3.4}$$

We first show some elementary properties that will be used often throughout this chapter.

Lemma 3.1.4. *Let $A \in \mathbb{H}[x]^{m_1\times n}$ and $B \in \mathbb{H}[x]^{n\times m_2}$. Then $(AB)^* = B^*A^*$.*

Proof. By direct calculation. □

Corollary 3.1.5. *For $A \in \mathbb{H}[x]^{m\times n}$, let $B = AA^*$. Then $B = B^*$.*

Proof. $B^* = (AA^*)^* = (A^*)^* A^* = AA^* = B$. □

Lemma 3.1.6. *Let $A \in \mathbb{H}[x]^{m\times n}$ have a generalized inverse $A^\dagger$. Then $(A^*)^\dagger = \left(A^\dagger\right)^*$.*

Proof. We show that $\left(A^\dagger\right)^*$ satisfies the defining relations of a generalized inverse stated in Definition 3.1.3,

$$\begin{aligned}
A^*\left(A^\dagger\right)^* A^* &= \left(AA^\dagger A\right)^* = A^*,\\
\left(A^\dagger\right)^* A^* \left(A^\dagger\right)^* &= \left(A^\dagger A A^\dagger\right)^* = \left(A^\dagger\right)^*,\\
\left[A^*\left(A^\dagger\right)^*\right]^* &= \left[\left(A^\dagger A\right)^*\right]^* = A^\dagger A = \left(A^\dagger A\right)^* = A^*\left(A^\dagger\right)^*,\\
\left[\left(A^\dagger\right)^* A^*\right]^* &= \left[\left(AA^\dagger\right)^*\right]^* = AA^\dagger = \left(AA^\dagger\right)^* = \left(A^\dagger\right)^* A^*.
\end{aligned}$$

Therefore, $(A^*)^\dagger = \left(A^\dagger\right)^*$. $\square$

Lemma 3.1.7. *(Adaption of (Penrose, 1955)) Let $A \in \mathbb{H}[x]^{m\times n}$ have a generalized inverse $A^\dagger$. Then relations satisfied by $A^\dagger$ include: $A^\dagger\left(A^\dagger\right)^* A^* = A^\dagger = A^*\left(A^\dagger\right)^* A^\dagger$ and $A^\dagger A A^* = A^* = A^* A A^\dagger$.*

Proof. First,

$$\begin{aligned} A^\dagger\left(A^\dagger\right)^* A^* =& A^\dagger\left(AA^\dagger\right)^* = A^\dagger A A^\dagger = A^\dagger \\ =& A^\dagger A A^\dagger = \left(A^\dagger A\right)^* A^\dagger = A^*\left(A^\dagger\right)^* A^\dagger. \end{aligned}$$

Next,

$$\begin{aligned} A^\dagger A A^* &= \left(A^\dagger A\right)^* A^* = \left(AA^\dagger A\right)^* = A^* \\ &= \left(AA^\dagger A\right)^* = A^*\left(A^\dagger\right)^* A^* = A^*\left(AA^\dagger\right)^* = A^* A A^\dagger. \end{aligned}$$

$\square$

We next show that $A^\dagger$ is uniquely defined for quaternion polynomial matrices.

Lemma 3.1.8. *(Adaption of (Penrose, 1955)) Let $A \in \mathbb{H}[x]^{m\times n}$ have a generalized inverse $A^\dagger$. Then $A^\dagger$ is unique.*

Proof. We first show that 3.2 and 3.3 are equivalent to

$$XX^*A^* = X. \tag{3.5}$$

Substituting 3.3 into 3.2, we obtain $X = XAX = X(AX)^* = XX^*A^*$.

Conversely, 3.5 implies that:

$$AXX^*A^* = AX \Rightarrow (AXX^*A^*)^* = (AX)^* \Rightarrow AXX^*A^* = (AX)^*.$$

So $AX = (AX)^*$ and therefore $X = XX^*A^* = X(AX)^* = XAX$. Similarly 3.1 and 3.4 can be replaced by the equation:

$$XAA^* = A^*. \tag{3.6}$$

Therefore, 3.5 and 3.6 are equivalent to 3.1, 3.2, 3.3 and 3.4.

Likewise the following two equations are also equivalent to 3.1, 3.2, 3.3 and 3.4,

$$X^*A^*A = A \tag{3.7}$$

$$A^*X^*X = X. \tag{3.8}$$

Suppose $A^\dagger$ satisfies 3.5 and 3.6, B satisfies 3.7 and 3.8. Then

$$\begin{aligned} A^\dagger =& A^\dagger \left(A^\dagger\right)^* A^* = A^\dagger \left(A^\dagger\right)^* (B^*A^*A)^* \\ =& A^\dagger \left(A^\dagger\right)^* A^*AB = A^\dagger AB = A^\dagger AA^*B^*B = A^*B^*B = B. \end{aligned}$$

Therefore, if it exists, $A^\dagger$ is uniquely defined. □

Due to the non-commutativity of quaternions, there are two types of eigenvalues: right eigenvalues and left eigenvalues. Right eigenvalues have been studied extensively (Brenner, 1951; Lee, 1949; Baker, 1999). We will work with right eigenvalues towards our main result. For convenience, we will just use the term "eigenvalue" from now on.

Definition 3.1.9. $A \in \mathbb{H}^{m\times m}$ is unitary if $AA^* = A^*A = I$.

Lemma 3.1.10. *(Zhang, 1997) $A \in \mathbb{H}^{m\times m}$ is hermitian, that is, $A = A^*$, if and only if there exists a unitary matrix $U \in \mathbb{H}^{m\times m}$ such that $U^*AU = \begin{pmatrix} d_1 & & \\ & \ddots & \\ & & d_m \end{pmatrix}$, where d_i are the eigenvalues of A.*

Proof. Corollary 6.2 of (Zhang, 1997). □

Adapting ideas from (Puystjens and de Smet, 1980), we will give conditions that quaternion polynomial matrices must satisfy to have generalized inverses. We need to prove some lemmas first.

Lemma 3.1.11. *(Adaption of (Puystjens and de Smet, 1980)) Let* $A \in \mathbb{H}[x]^{m\times n}$ *have the generalized inverse* $A^{\dagger}$*. If* $U \in \mathbb{H}^{m\times m}$ *is a unitary matrix, then* $(UA)^{\dagger} = A^{\dagger}U^{*}$*.*

Proof. We show that $A^{\dagger}U^{*}$ satisfies the defining relations of the generalized inverse stated in Definition 3.1.3,

$$
\begin{aligned}
UA\left(A^{\dagger}U^{*}\right)UA &= UAA^{\dagger}A = UA,\\
A^{\dagger}U^{*}\left(UA\right)A^{\dagger}U^{*} &= A^{\dagger}AA^{\dagger}U^{*} = A^{\dagger}U^{*},\\
\left(UAA^{\dagger}U^{*}\right)^{*} &= U\left(AA^{\dagger}\right)^{*}U^{*} = UAA^{\dagger}U^{*},\\
\left(A^{\dagger}U^{*}UA\right)^{*} &= \left(A^{\dagger}A\right)^{*} = A^{\dagger}A = A^{\dagger}U^{*}UA.
\end{aligned}
$$

Therefore, $(UA)^{\dagger} = A^{\dagger}U^{*}$. □

Lemma 3.1.12. $A \in \mathbb{H}^{m\times n}$ *induces a homomorphism* $\mathbb{H}[x]^{n\times 1}$ *to* $\mathbb{H}[x]^{m\times 1}$*, that is, for all* P*,* $Q \in \mathbb{H}[x]^{n\times 1}$*,* $A(P+Q) = AP + AQ \in \mathbb{H}[x]^{m\times 1}$*.*

Proof. By direct calculation. □

Lemma 3.1.13. *(Adaption of (Puystjens and de Smet, 1980)) Let* $A \in \mathbb{H}[x]^{m\times n}$ *have the generalized inverse* $A^{\dagger}$*. Consider* A *as a homomorphism from* $\mathbb{H}[x]^{n\times 1}$ *to* $\mathbb{H}[x]^{m\times 1}$*. Then* $ImageA = ImageAA^{*} = ImageAA^{\dagger}$ *and* $ImageA^{*} = ImageA^{*}A = ImageA^{\dagger}A$*.*

Proof. Let $P \in \mathbb{H}[x]^{n\times 1}$. Then

$$
AP = AA^{\dagger}AP = A\left(A^{\dagger}A\right)^{*}P = AA^{*}\left[\left(A^{\dagger}\right)^{*}P\right],
$$

where $AP \in \mathbb{H}[x]^{m\times 1}$ and $\left(A^{\dagger}\right)^{*} P \in \mathbb{H}[x]^{m\times 1}$. Therefore, Image$A \subseteq$ ImageAA^{*} and Image$A \subseteq$ Image$AA^{\dagger}$.

For any $Q \in \mathbb{H}[x]^{m\times 1}$, it is clear that $AA^{*}Q = A(A^{*}Q)$ where $A^{*}Q \in \mathbb{H}[x]^{n\times 1}$ and that $AA^{\dagger}Q = A\left(A^{\dagger}Q\right)$ where $A^{\dagger}Q \in \mathbb{H}[x]^{n\times 1}$. Thus, Image$AA^{*} \subseteq$ Image$A$ and Image$AA^{\dagger} \subseteq$ Image$A$.

Therefore, Image$A$ = ImageAA^{*} and Image$A$ = Image$AA^{\dagger}$, i.e., Image$A$ = ImageAA^{*} = Image$AA^{\dagger}$.

Similarly we can show that ImageA^{*} = Image$A^{*}A$ = Image$A^{\dagger}A$. $\square$

Lemma 3.1.14. *(Adaption of (Puystjens and de Smet, 1980)) If $E \in \mathbb{H}[x]^{m\times m}$ is a symmetric projection, that is, $E = E^2 = E^{*}$, then $E \in \mathbb{H}^{m\times m}$.*

Proof. If $f_1, \cdots, f_m$ are the entries in the first row of E, with $f_1 = \overline{f_1}$, then

$$f_1\overline{f_1} + \sum_{i=2}^{m} f_i\overline{f_i} = f_1 \Rightarrow f_1^2 + \sum_{i=2}^{m} f_i\overline{f_i} = f_1.$$

Since $f_1 = \overline{f_1}$, the leading coefficient of f_1^2 is a positive real number. Note that the leading coefficient of $\sum_{i=2}^{m} f_i\overline{f_i}$ is also a positive real number. Thus,

$$\begin{aligned}\deg\left(f_1^2\right) \geq \deg f_1 = \deg\left(f_1^2 + \sum_{i=2}^{m} f_i\overline{f_i}\right) &= \max\left\{\deg\left(f_1^2\right),\ \deg\left(\sum_{i=2}^{m} f_i\overline{f_i}\right)\right\} \\ &\geq \deg\left(f_1^2\right).\end{aligned}$$

This shows that $f_1 \in \mathbb{H}$. Furthermore, $0 = \deg f_1 = \deg\left(\sum_{i=2}^{m} f_i\overline{f_i}\right)$ and therefore $f_i \in \mathbb{H}$ for all $1 \leq i \leq m$.

The same can be done for the other rows of E. Therefore, $E \in \mathbb{H}^{m\times m}$. $\square$

Now we are ready to give conditions that quaternion polynomial matrices must satisfy in order to have generalized inverses.

Theorem 3.1.15. *(Adaption of (Puystjens and de Smet, 1980)) Let* $A \in \mathbb{H}[x]^{m\times n}$*. Then* A *has the generalized inverse* $A^\dagger$ *if and only if* $A = U\begin{pmatrix} A_1 & A_2 \\ 0 & 0 \end{pmatrix}$ *with* $U \in \mathbb{H}^{m\times m}$ *unitary and* $A_1A_1^* + A_2A_2^*$ *a unit in* $\mathbb{H}[x]^{r\times r}$ *with* $r \leq \min\{m, n\}$*. Moreover,*

$$A^\dagger = \begin{pmatrix} A_1^*\left(A_1A_1^* + A_2A_2^*\right)^{-1} & 0 \\ A_2^*\left(A_1A_1^* + A_2A_2^*\right)^{-1} & 0 \end{pmatrix} U^*.$$

Proof. If A has the generalized inverse $A^\dagger$, then

$$AA^\dagger = AA^\dagger AA^\dagger = \left(AA^\dagger\right)^2 = \left(AA^\dagger\right)^*.$$

By Lemma 3.1.14, $AA^\dagger \in \mathbb{H}^{m\times m}$. $AA^\dagger$ is hermitian and hence, by Lemma 3.1.10, there exists a unitary matrix $U \in \mathbb{H}^{m\times m}$ such that $U^*AA^\dagger U = D$ where D is diagonal. Since

$$D^2 = U^*AA^\dagger UU^*AA^\dagger U = U^*AA^\dagger AA^\dagger U = U^*AA^\dagger U = D,$$

the diagonal entries of D are either 1 or 0. Therefore, we can rearrange the rows of U so that $D = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$ with $r \leq \min\{m, n\}$.

Set $A' = U^*A$. By Lemma 3.1.11, A' has its own generalized inverse $A'^\dagger$ and $A'A'^\dagger = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$. Set $A' = \begin{pmatrix} A_1 & A_2 \\ A_3 & A_4 \end{pmatrix}$, for arbitrary quaternion polynomial matrices $A_1 \in \mathbb{H}[x]^{r\times r}$, $A_2 \in \mathbb{H}[x]^{r\times(n-r)}$, $A_3 \in \mathbb{H}[x]^{(m-r)\times r}$ and $A_4 \in \mathbb{H}[x]^{(m-r)\times(n-r)}$. Since $A' = A'A'^\dagger A' = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} A_1 & A_2 \\ A_3 & A_4 \end{pmatrix} = \begin{pmatrix} A_1 & A_2 \\ 0 & 0 \end{pmatrix}$, we must have $A' = \begin{pmatrix} A_1 & A_2 \\ 0 & 0 \end{pmatrix}$ and therefore $A'A'^* = \begin{pmatrix} A_1A_1^* + A_2A_2^* & 0 \\ 0 & 0 \end{pmatrix}$. Similarly, $A'^\dagger = \begin{pmatrix} B_1 & 0 \\ B_2 & 0 \end{pmatrix}$ for some B_1 and B_2.

By Lemma 3.1.13,

$$\text{Image}A'A'^{*} = \text{Image}A' = \text{Image}A'A'^{\dagger} = \text{Image}\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}.$$

This implies the surjectivity of $A_1A_1^* + A_2A_2^*$ on $\mathbb{H}[x]^{r\times 1}$. Therefore, $A_1A_1^* + A_2A_2^*$ is a unit in $\mathbb{H}[x]^{r\times r}$ and

$$A = UA' = U\begin{pmatrix} A_1 & A_2 \\ 0 & 0 \end{pmatrix}.$$

Next, we have that:

$$\begin{aligned} A'^{\dagger} =& A'^{\dagger}\left(A'^{\dagger}\right)^{*}A'^{*} = A'^{\dagger}\left(A'^{*}\right)^{\dagger}A'^{*} = A'^{*}\left(A'A'^{*}\right)^{\dagger} \\ =& \begin{pmatrix} A_1^* & 0 \\ A_2^* & 0 \end{pmatrix}\begin{pmatrix} \left(A_1A_1^* + A_2A_2^*\right)^{-1} & 0 \\ 0 & 0 \end{pmatrix} \\ =& \begin{pmatrix} A_1^*\left(A_1A_1^* + A_2A_2^*\right)^{-1} & 0 \\ A_2^*\left(A_1A_1^* + A_2A_2^*\right)^{-1} & 0 \end{pmatrix}, \end{aligned}$$

which gives:

$$A^{\dagger} = \begin{pmatrix} A_1^*\left(A_1A_1^* + A_2A_2^*\right)^{-1} & 0 \\ A_2^*\left(A_1A_1^* + A_2A_2^*\right)^{-1} & 0 \end{pmatrix}U^*.$$

We will show the converse by direct computation.

Let $A = U\begin{pmatrix} A_1 & A_2 \\ 0 & 0 \end{pmatrix}$ be in $\mathbb{H}[x]^{m\times n}$ with $U \in \mathbb{H}^{m\times m}$ unitary and $A_1A_1^* + A_2A_2^*$ a unit in $\mathbb{H}[x]^{r\times r}$ with $r \leq \min\{m, n\}$. Let

$$B = \begin{pmatrix} A_1^*\left(A_1A_1^* + A_2A_2^*\right)^{-1} & 0 \\ A_2^*\left(A_1A_1^* + A_2A_2^*\right)^{-1} & 0 \end{pmatrix}U^*.$$

We show that $A^{\dagger}U^*$ satisfies the defining relations of the generalized inverse stated in Definition 3.1.3,

- the first condition:

$$
\begin{aligned}
&ABA\\
=&U\begin{pmatrix} A_1 & A_2 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} A_1^*\left(A_1A_1^*+A_2A_2^*\right)^{-1} & 0 \\ A_2^*\left(A_1A_1^*+A_2A_2^*\right)^{-1} & 0 \end{pmatrix}U^*U\begin{pmatrix} A_1 & A_2 \\ 0 & 0 \end{pmatrix}\\
=&U\begin{pmatrix} A_1 & A_2 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} A_1^*\left(A_1A_1^*+A_2A_2^*\right)^{-1} & 0 \\ A_2^*\left(A_1A_1^*+A_2A_2^*\right)^{-1} & 0 \end{pmatrix}\begin{pmatrix} A_1 & A_2 \\ 0 & 0 \end{pmatrix}\\
=&U\begin{pmatrix} A_1 & A_2 \\ 0 & 0 \end{pmatrix}=A,
\end{aligned}
$$

- the second condition:

$$
\begin{aligned}
&BAB\\
=&\begin{pmatrix} A_1^*\left(A_1A_1^*+A_2A_2^*\right)^{-1} & 0 \\ A_2^*\left(A_1A_1^*+A_2A_2^*\right)^{-1} & 0 \end{pmatrix}U^*U\begin{pmatrix} A_1 & A_2 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} A_1^*\left(A_1A_1^*+A_2A_2^*\right)^{-1} & 0 \\ A_2^*\left(A_1A_1^*+A_2A_2^*\right)^{-1} & 0 \end{pmatrix}U^*\\
=&\begin{pmatrix} A_1^*\left(A_1A_1^*+A_2A_2^*\right)^{-1} & 0 \\ A_2^*\left(A_1A_1^*+A_2A_2^*\right)^{-1} & 0 \end{pmatrix}\begin{pmatrix} A_1 & A_2 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} A_1^*\left(A_1A_1^*+A_2A_2^*\right)^{-1} & 0 \\ A_2^*\left(A_1A_1^*+A_2A_2^*\right)^{-1} & 0 \end{pmatrix}U^*\\
=&\begin{pmatrix} A_1^*\left(A_1A_1^*+A_2A_2^*\right)^{-1} & 0 \\ A_2^*\left(A_1A_1^*+A_2A_2^*\right)^{-1} & 0 \end{pmatrix}U^*=B,
\end{aligned}
$$

- the third condition:

$$
\begin{aligned}
(AB)^*=&\left(U\begin{pmatrix} A_1 & A_2 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} A_1^*\left(A_1A_1^*+A_2A_2^*\right)^{-1} & 0 \\ A_2^*\left(A_1A_1^*+A_2A_2^*\right)^{-1} & 0 \end{pmatrix}U^*\right)^*\\
=&\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}^*=\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}=AB,
\end{aligned}
$$

- the last condition:

$$(BA)^* = \left(\begin{pmatrix} A_1^* \left(A_1A_1^* + A_2A_2^*\right)^{-1} & 0 \\ A_2^* \left(A_1A_1^* + A_2A_2^*\right)^{-1} & 0 \end{pmatrix} U^*U \begin{pmatrix} A_1 & A_2 \\ 0 & 0 \end{pmatrix} \right)^*$$

$$= \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}^* = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} = BA.$$

Therefore, $B = \begin{pmatrix} A_1^* \left(A_1A_1^* + A_2A_2^*\right)^{-1} & 0 \\ A_2^* \left(A_1A_1^* + A_2A_2^*\right)^{-1} & 0 \end{pmatrix} U^* = A^\dagger$. The uniqueness of $A^\dagger$ is the result of Lemma 3.1.8. □

Lemma 3.1.16. *Let $A \in \mathbb{H}^{m\times m}$ be hermitian. Then the eigenvalues of A are real.*

Proof. Let $\lambda \in \mathbb{H}$ be an eigenvalue of A with corresponding eigenvector $0 \neq X = \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix}$ such that $AX = X\lambda$. Then $X^*AX = X^*X\lambda$. So $X^*AX = \lambda^*X^*X$ and therefore

$$X^*X\lambda = \lambda^*X^*X = \left(X^*X\lambda\right)^*$$

$$\Rightarrow$$

$$\left(\overline{x_1}, \cdots \overline{x_m}\right) \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} \lambda = \left(\sum \overline{x_i}x_i\right)\lambda = \left(\left(\sum \overline{x_i}x_i\right)\lambda\right)^*$$

$$= \lambda^* \left(\sum \overline{x_i}x_i\right)^* = \lambda^* \left(\sum \overline{x_i}x_i\right).$$

Note that $0 \neq \sum \overline{x_i}x_i \in \mathbb{R}[x]$. Division of the above equation by $\sum \overline{x_i}x_i$ gives $\lambda = \lambda^* \Rightarrow \lambda \in \mathbb{R}$. □

Theorem 3.1.17. *(Zhang, 1997)(Cayley-Hamilton theorem for Quaternion Matrices) Let $A \in \mathbb{H}^{m\times m}$. $f_A(\lambda) = \det\left(\lambda I_{2m} - \chi_A\right)$ is called the characteristic polyno-*

mial of A, where λ is a complex indeterminate. Then $f_A(A) = 0$. Furthermore, $f_A(\lambda_0) = 0$ if and only if λ_0 is an eigenvalue of A.

Proof. Theorem 8.1 of (Zhang, 1997). $\square$

Definition 3.1.18. Let $A \in \mathbb{H}[x]^{m\times n}$ have the generalized inverse $A^{\dagger}$. For any $x \in \mathbb{R}$, set $B = AA^{*}$. Then $f_B(\lambda) = \det(\lambda I_{2m} - \chi_B)$ is called the characteristic polynomial of B. By Lemma 3.1.16, λ can be assumed to be an indeterminate that enjoys the following: $\lambda = \overline{\lambda}$ and λ commutes element-wise with $\mathbb{H}[x]$.

Lemma 3.1.19. *Let $A \in \mathbb{H}[x]^{m\times n}$ have the generalized inverse $A^{\dagger}$. For any $x \in \mathbb{R}$, set $B = AA^{*}$. Then $f_B(\lambda) = g(\lambda)^2$ where $g(\lambda) \in (\mathbb{R}[x])[\lambda]$.*

Proof. We first show that $f_B(\lambda) \in (\mathbb{R}[x])[\lambda]$. Since

$$\det\left((\lambda I_{2m} - \chi_B)^T\right) = \det(\lambda I_{2m} - \chi_B) = \det\left((\lambda I_{2m} - \chi_B)^{*}\right)$$

$$\Rightarrow$$

$$\det(\lambda I_{2m} - \chi_B) = \det\overline{(\lambda I_{2m} - \chi_B)} = \overline{\det(\lambda I_{2m} - \chi_B)},$$

we have that:

$$\det(\lambda I_{2m} - \chi_B) = f_B(\lambda) \in (\mathbb{R}[x])[\lambda]. \tag{3.9}$$

Next, we show that $f_B(\lambda) = g(\lambda)^2$ where $g(\lambda) \in (\mathbb{C}[x])[\lambda]$. Let $B = P + Q\mathbf{i}$. For any fixed $1 \le i,\ j \le m$, we have $B_{ij} = a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k}$ where a, b, c and $d \in \mathbb{R}[x]$. Since B is hermitian, $B_{ji} = a - b\mathbf{i} - c\mathbf{j} - d\mathbf{k}$ and therefore $P_{ij} = a + b\mathbf{i}$, $P_{ji} = a - b\mathbf{i}$ and $Q_{ij} = c + d\mathbf{i}$, $Q_{ji} = -c - d\mathbf{i}$. So $P^T = \overline{P}$ and $Q = -Q^T$. Therefore,

$$\chi_B = \begin{pmatrix} P & Q \\ -\overline{Q} & \overline{P} \end{pmatrix} = \begin{pmatrix} P & Q \\ -\overline{Q} & P^T \end{pmatrix}$$

$$\Rightarrow$$

$$\lambda I_{2m} - \chi_B = \begin{pmatrix} \lambda I_m - P & Q \\ -\overline{Q} & \lambda I_m - P^T \end{pmatrix}.$$

Next, we have:

$$\begin{pmatrix} I_m & -I_m \\ 0 & I_m \end{pmatrix} \begin{pmatrix} I_m & 0 \\ I_m & I_m \end{pmatrix} \begin{pmatrix} I_m & -I_m \\ 0 & I_m \end{pmatrix} \begin{pmatrix} \lambda I_m - P & Q \\ -\overline{Q} & \lambda I_m - P^T \end{pmatrix}$$
$$= \begin{pmatrix} \overline{Q} & P^T - \lambda I_m \\ \lambda I_m - P & Q \end{pmatrix}$$

and

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

Therefore,

$$f_B(\lambda) = \det \begin{pmatrix} \lambda I_m - P & Q \\ -\overline{Q} & \lambda I_m - P^T \end{pmatrix} = \det \begin{pmatrix} \overline{Q} & P^T - \lambda I_m \\ \lambda I_m - P & Q \end{pmatrix}.$$

Note that:

$$-\begin{pmatrix} \overline{Q} & P^T - \lambda I_m \\ \lambda I_m - P & Q \end{pmatrix}^T = \begin{pmatrix} \overline{Q} & P^T - \lambda I_m \\ \lambda I_m - P & Q \end{pmatrix},$$

which implies that $\begin{pmatrix} \overline{Q} & P^T - \lambda I_m \\ \lambda I_m - P & Q \end{pmatrix}$ is skew-symmetric. By (Muir, 2003), the determinant of $\begin{pmatrix} \overline{Q} & P^T - \lambda I_m \\ \lambda I_m - P & Q \end{pmatrix}$, also called its Pfaffian, can be written as the square of a polynomial in its entries. Therefore, $f_B(\lambda) = g(\lambda)^2$ where $g(\lambda) \in (\mathbb{C}[x])[\lambda]$.

Finally we show that $g(\lambda) \in (\mathbb{R}[x])[\lambda]$. Suppose otherwise. Then $g(\lambda) = a(\lambda) + b(\lambda)\mathbf{i}$ where a and $b \in (\mathbb{R}[x])[\lambda]$ with $b(\lambda) \neq 0$. By 3.9, $g(\lambda)^2 \in (\mathbb{R}[x])[\lambda]$. So $a(\lambda) = 0$ and therefore $f_B(\lambda) = g(\lambda)^2 = (b(\lambda)\mathbf{i})^2 = -b(\lambda)^2$

where $b(\lambda) \in (\mathbb{R}[x])[\lambda]$. For a fixed $x \in \mathbb{R}$, let $\lambda' \in \mathbb{R}$ be large enough such that $\lambda' I_{2m} - \chi_B \in \mathbb{C}^{2m\times 2m}$ is diagonally dominant with non-negative diagonal entries and that $(b(x))(\lambda') \neq 0$. Since $\lambda' I_{2m} - \chi_B$ is also hermitian, $\lambda' I_{2m} - \chi_B$ is positive definite (Horn and Johnson, 1990). But $\det(\lambda' I_{2m} - \chi_B) = -((b(x))(\lambda'))^2 < 0$. Contradiction. So, $b = 0$ and therefore $f_B(\lambda) = g(\lambda)^2$ where $g(\lambda) \in (\mathbb{R}[x])[\lambda]$. $\square$

Corollary 3.1.20. *Let $A \in \mathbb{H}[x]^{m\times n}$ have the generalized inverse $A^\dagger$. For any $x \in \mathbb{R}$, set $B = AA^*$ and $f_B(\lambda) = g(\lambda)^2$. Then $g(B) = 0$. We will call $g(\lambda)$ the generalized characteristic polynomial of A.*

Proof. $g(\lambda) \in (\mathbb{R}[\alpha])[\lambda]$ by Lemma 3.1.19, so $\chi_{g(B)} = g(\chi_B)$. Next, $f_B(\chi_B) = 0$ by the Cayley-Hamilton theorem for complex polynomial matrices (Horn and Johnson, 1990), so $g(\chi_B) = 0$. Therefore $0 = g(\chi_B) = \chi_{g(B)}$, that is, $g(B) = 0$. $\square$

3.2. Leverrier-Faddeev Method

Lemma 3.2.1. *Let $A \in \mathbb{H}[x]^{m \times n}$ have the generalized inverse $A^{\dagger}$. Set $B = AA^{*}$. Then $B^{\dagger} = (A^{*})^{\dagger} A^{\dagger}$ and $B^{\dagger}B = BB^{\dagger}$.*

Proof. First we show that $(A^{*})^{\dagger} A^{\dagger}$ satisfies the defining relations of the generalized inverse stated in Definition 3.1.3,

- the first condition:

$$\begin{aligned} B\left[(A^{*})^{\dagger} A^{\dagger}\right] B =& AA^{*} (A^{*})^{\dagger} A^{\dagger} AA^{*} \\ =& AA^{*} (A^{*})^{\dagger} A^{*} = AA^{*} = B, \end{aligned}$$

- the second condition:

$$\begin{aligned} \left[(A^{*})^{\dagger} A^{\dagger}\right] B \left[(A^{*})^{\dagger} A^{\dagger}\right] &= (A^{*})^{\dagger} A^{\dagger} AA^{*} (A^{*})^{\dagger} A^{\dagger} \\ &= (A^{*})^{\dagger} A^{*} (A^{*})^{\dagger} A^{\dagger} = (A^{*})^{\dagger} A^{\dagger}, \end{aligned}$$

- the third condition:

$$\begin{aligned} \left(B\left[(A^{*})^{\dagger} A^{\dagger}\right]\right)^{*} =& \left(AA^{*} (A^{*})^{\dagger} A^{\dagger}\right)^{*} = \left(AA^{\dagger}\right)^{*} \\ =& AA^{\dagger} = AA^{*} (A^{*})^{\dagger} A^{\dagger} = B\left[(A^{*})^{\dagger} A^{\dagger}\right], \end{aligned}$$

- the last condition:

$$\begin{aligned} \left(\left[(A^{*})^{\dagger} A^{\dagger}\right] B\right)^{*} &= \left((A^{*})^{\dagger} A^{\dagger} AA^{*}\right)^{*} = \left((A^{*})^{\dagger} A^{*}\right)^{*} \\ &= (A^{*})^{\dagger} A^{*} = (A^{*})^{\dagger} A^{\dagger} AA^{*} = \left[(A^{*})^{\dagger} A^{\dagger}\right] B. \end{aligned}$$

Therefore, $(A^{*})^{\dagger} A^{\dagger} = (AA^{*})^{\dagger} = B^{\dagger}$. Next,

$$\begin{aligned} (AA^{*})^{\dagger} AA^{*} &= (A^{*})^{\dagger} A^{\dagger} AA^{*} = (A^{*})^{\dagger} A^{*} \\ &= \left((A^{*})^{\dagger} A^{*}\right)^{*} = AA^{\dagger} = AA^{*} \left(A^{\dagger}\right)^{*} A^{\dagger} = AA^{*} (AA^{*})^{\dagger}, \end{aligned}$$

which is, $B^{\dagger}B = BB^{\dagger}$. □

Lemma 3.2.2. *Let $A \in \mathbb{H}[x]^{m\times n}$ have the generalized inverse $A^{\dagger}$. Set $B = AA^*$. Then $\left(B^{\dagger}\right)^k = \left(B^k\right)^{\dagger}$ where $k \in \mathbb{N}$.*

Proof. We show that $\left(B^{\dagger}\right)^k$ satisfies the defining relations of the generalized inverse stated in Definition 3.1.3,

- the first condition:

$$\begin{aligned} B^k\left(B^{\dagger}\right)^k B^k =& B^k\left(B^{\dagger}\right)^{k-1}\left(BB^{\dagger}\right)B^{k-1} \\ =& B^k\left(B^{\dagger}\right)^{k-1}B^{k-1} = \cdots = B^k, \end{aligned}$$

- the second condition:

$$\begin{aligned} \left(B^{\dagger}\right)^k B^k\left(B^{\dagger}\right)^k &= \left(B^{\dagger}\right)^{k-1}\left(BB^{\dagger}\right)B^{k-1}\left(B^{\dagger}\right)^k \\ &= \left(B^{\dagger}\right)^{k-1}B^{k-1}\left(B^{\dagger}\right)^k = \cdots = \left(B^{\dagger}\right)^k, \end{aligned}$$

- the third condition:

$$\left(B^k\left(B^{\dagger}\right)^k\right)^* = \left[(B^*)^{\dagger}\right]^k (B^*)^k = \left(B^{\dagger}\right)^k B^k = B^k\left(B^{\dagger}\right)^k,$$

- the last condition:

$$\left(\left(B^{\dagger}\right)^k B^k\right)^* = \left((B^*)^k\right)\left((B^*)^{\dagger}\right)^k = B^k\left(B^{\dagger}\right)^k = \left(B^{\dagger}\right)^k B^k.$$

Therefore, $\left(B^{\dagger}\right)^k = \left(B^k\right)^{\dagger}$ where $k \in \mathbb{N}$. □

Lemma 3.2.3. *(Adaption of (Penrose, 1955)) Let $A \in \mathbb{H}[x]^{m\times n}$, $B \in \mathbb{H}[x]^{p\times q}$ and $C \in \mathbb{H}[x]^{m\times q}$. If $A^{\dagger}$ and $B^{\dagger}$ both exist, then the quaternion polynomial matrix equation $AXB = C$ has a solution in $\mathbb{H}[x]^{n\times p}$ if and only if $AA^{\dagger}CB^{\dagger}B = C$, in which case the general solution is*

$$X = A^{\dagger}CB^{\dagger} + Y - A^{\dagger}AYBB^{\dagger},$$

where $Y \in \mathbb{H}[x]^{n\times p}$ *is arbitrary.*

Proof. Suppose X satisfies $AXB = C$, then

$$C = AXB = AA^{\dagger}AXBB^{\dagger}B = AA^{\dagger}CB^{\dagger}B.$$

Conversely, if $C = AA^{\dagger}CB^{\dagger}B$ then $A^{\dagger}CB^{\dagger}$ is a particular solution of $AXB = C$.

For the general solution, we must solve $AXB = 0$. Any expression of the form $X = Y - A^{\dagger}AYBB^{\dagger}$ satisfies $AXB = 0$, and conversely if $AXB = 0$ then $X = A^{\dagger}CB^{\dagger} + Y - A^{\dagger}AYBB^{\dagger}$. □

Theorem 3.2.4. *(Adaption of (Penrose, 1955)) Let* $A \in \mathbb{H}[x]^{m\times n}$ *have the generalized inverse* $A^{\dagger}$. *For any* $x \in \mathbb{R}$, *set* $B = AA^{*}$. *Suppose the generalized characteristic polynomial of* A *is*

$$g(\lambda) = \lambda^{m} + a_1\lambda^{m-1} + \cdots + a_k\lambda^{m-k} + \cdots + a_{m-1}\lambda + a_m,$$

where $a_i \in \mathbb{R}[x]$. *If* k *is the largest integer such that* $a_k \neq 0$, *then the generalized inverse of* A *is given by*

$$A^{\dagger} = -\frac{1}{a_k}A^{*}\left[B^{k-1} + a_1B^{k-2} + \cdots + a_{k-1}I\right].$$

If $a_i = 0$ *for all* $1 \leq i \leq m$, *then* $A^{\dagger} = 0$.

Proof. By Corollary 3.1.20, we have:

$$0 = B^{m} + a_1B^{m-1} + \cdots + a_kB^{m-k} + \cdots + a_{m-1}B + a_mI_m.$$

If k is the largest integer such that $a_k \neq 0$ and define $B^0 = I$, we may write:

$$B^{m-k}\left(B^{k} + a_1B^{k-1} + \cdots + a_{k-1}B + a_kI\right) = 0.$$

This equation guarantees a solution of the matrix equation $B^{m-k}X = 0$ and hence, by Lemma 3.2.1 and 3.2.3, all solutions are given by:

$$X = Y - \left(B^{m-k}\right)^{\dagger} B^{m-k}Y = Y - B^{\dagger}BY,$$

where $Y \in \mathbb{H}^{m\times m}$ is arbitrary. In particular, there exists Y_1 such that:

$$B^k + a_1B^{k-1} + \cdots + a_{k-1}B + a_kI = Y_1 - B^{\dagger}BY_1 = Y_1 - AA^{\dagger}Y_1$$

Left multiplication of the latter equation by $A^{\dagger}$ gives:

$$A^{\dagger}B^k + a_1A^{\dagger}B^{k-1} + \cdots + a_{k-1}A^{\dagger}B + a_kA^{\dagger} = 0.$$

By Lemma 3.1.7, this can be simplified to:

$$A^*B^{k-1} + a_1A^*B^{k-2} + \cdots + a_{k-1}A^* = -a_kA^{\dagger},$$

which gives:

$$A^{\dagger} \;\; = \;\; -\frac{1}{a_k}A^*\left[B^{k-1} + a_1B^{k-2} + \cdots + a_{k-1}I\right].$$

If $a_i = 0$ for all $1 \le i \le m$, then $A = 0$ and therefore $A^{\dagger} = 0$. □

Lemma 3.2.5. *(Adaption of (Kalman, 1978)) Let $A \in \mathbb{H}\left[x\right]^{m\times n}$ have the generalized inverse $A^{\dagger}$. For any real $x \in \mathbb{R}$, set $B = AA^*$. Let $\lambda_1,\ \cdots,\ \lambda_{m'}$, where $m' \le m$, be all the non-zero eigenvalues of B. Then for $1 \le k \le m$,*

$$tr\left[\left(B^k + a_1B^{k-1} + \cdots + a_{k-1}B\right)\right] = -ka_k,$$

where the a_i arise from the generalized characteristic polynomial of A:

$$g\left(\lambda\right) \;\; = \;\; \lambda^m + a_1\lambda^{m-1} + \cdots + a_k\lambda^{m-k} + \cdots + a_{m-1}\lambda + a_m.$$

Proof. Let $Y = yI$ where $y \in \mathbb{R}$. We can write $g(Y)$ as:

$$\begin{aligned} g(Y) = &(Y - B) \\ &\times \left(Y^{m-1} + (B + a_1 I) Y^{m-2} + \left(B^2 + a_1 B + a_2 I\right) Y^{m-3} + \cdots \right. \\ &\left. + \left(B^{m-1} + a_1 B^{m-2} + \cdots + a_m I\right)\right). \end{aligned}$$

As long as y is not an eigenvalue of B, $(yI - B) = Y - B$ is non-singular, so we can write:

$$\begin{aligned} (Y - B)^{-1} g(Y) = &Y^{m-1} + (B + a_1 I) Y^{m-2} + \left(B^2 + a_1 B + a_2 I\right) Y^{m-3} + \cdots \\ &+ \left(B^{m-1} + a_1 B^{m-2} + \cdots + a_m I\right). \end{aligned}$$

Taking the traces gives:

$$\begin{aligned} \operatorname{tr}\left[(Y - B)^{-1} g(Y)\right] = &m y^{m-1} + \operatorname{tr}\left[(B + a_1 I)\right] y^{m-2} \\ &+ \operatorname{tr}\left[\left(B^2 + a_1 B + a_2 I\right)\right] y^{m-3} + \cdots \\ &+ \operatorname{tr}\left(B^{m-1} + a_1 B^{m-2} + \cdots + a_m I\right). \end{aligned}$$

Let $C = (Y - B)^{-1} g(Y)$. Since $g(Y) = g(yI) = g(y) I$, $C = g(y)(Y - B)^{-1}$. Therefore,

$$\operatorname{tr} C = g(y) \operatorname{tr}\left[(Y - B)^{-1}\right].$$

$\operatorname{tr}\left[(Y - B)^{-1}\right]$ is the sum of the eigenvalue of $(Y - B)^{-1}$. We will show these eigenvalues are the fractions $\frac{1}{y-\lambda_1}, \cdots, \frac{1}{y-\lambda_{m'}}$.

Let ζ be an eigenvalue of $(Y - B)^{-1}$ with corresponding eigenvector Z such that:

$$(Y - B)^{-1} Z = Z\zeta.$$

ζ is real by Lemma 3.1.16, and hence

$$(Y-B)\,Z = Z\frac{1}{\zeta}$$

$$\Rightarrow$$

$$BZ = Z\left(y-\frac{1}{\zeta}\right).$$

Therefore, $y-\frac{1}{\zeta}=\lambda_i \Rightarrow \zeta=\frac{1}{y-\lambda_i}$ for some $1\le i\le m'$.

Since $g(y) = (y-\lambda_1)(y-\lambda_2)\cdots(y-\lambda_{m'})$, we have that $g'(y) = g(y)\left(\frac{1}{y-\lambda_1}+\cdots+\frac{1}{y-\lambda_{m'}}\right)$ and

$$\mathrm{tr}C = g'(y)$$

The derivative of g is also equal to:

$$g'(y) = my^{m-1}+a_1(m-1)\,y^{m-2}+\cdots+a_{m-1}.$$

Therefore,

$$\begin{aligned} my^{m-1}+a_1(m-1)\,y^{m-2}+\cdots+a_{m-1} =& my^{m-1}+\mathrm{tr}\ [(B+a_1I)]\,y^{m-2}+\cdots \\ &+\mathrm{tr}\ \left(B^{m-1}+a_1B^{m-2}+\cdots+a_mI\right) \end{aligned}$$

Comparing the coefficient of y^{m-k-1} on both sides, we obtain

$$a_k(m-k) = \mathrm{tr}\left[\left(B^k+a_1B^{k-1}+\cdots+a_{k-1}B+a_kI\right)\right]$$

$$\Rightarrow$$

$$a_k(m-k) = \mathrm{tr}\left[\left(B^k+a_1B^{k-1}+\cdots+a_{k-1}B\right)\right]+\mathrm{tr}a_kI$$

$$\Rightarrow$$

$$-ka_k = \mathrm{tr}\left[\left(B^k+a_1B^{k-1}+\cdots+a_{k-1}B\right)\right]$$

□

Lemma 3.2.6. *(Adaption of (Decell, 1965; Faddeeva, 1959)) Let $A\in\mathbb{H}[x]^{m\times n}$ have* *the generalized inverse $A^{\dagger}$. For any $x \in \mathbb{R}$, set $B = AA^*$. Suppose the generalized characteristic polynomial of A is*

$$g(\lambda) = \lambda^m + a_1\lambda^{m-1} + \cdots + a_k\lambda^{m-k} + \cdots + a_{m-1}\lambda + a_m,$$

where $a_i \in \mathbb{R}[x]$. Define $a_0 = 1$. If p is the largest integer such that $a_p \neq 0$ and we construct the sequence $A_0, \cdots, A_p$ as follows:

$$\begin{array}{lll} A_0 = 0 & -1 = q_0 & B_0 = I \\ \vdots & \vdots & \vdots \\ A_{p-1} = AA^*B_{p-2} & \dfrac{trA_{p-1}}{p-1} = q_{p-1} & B_{p-1} = A_{p-1} - q_{p-1}I \\ A_p = AA^*B_{p-1} & \dfrac{trA_p}{p} = q_p & B_p = A_p - q_pI \end{array}$$

then $q_i(x) = -a_i(x)$, $i = 0, \cdots, p$.

Proof. We will show the desired result by mathematical induction.

For $p \geq i \geq 0$, let P_i be the statement:

$$q_i(x) = -a_i(x).$$

P_0 is the statement:

$$q_0 = -a_0,$$

which is true by definition.

Suppose P_i holds for all $0 \leq i \leq k-1$, it remains to show that P_k holds. We have

$$\begin{aligned}
A_k =& BB_{k-1} \\
=& B\left(A_{k-1} - q_{k-1}I\right) \\
=& B\left(B\left(A_{k-2} - q_{k-2}I\right) - q_{k-1}I\right) \\
=& \cdots \\
=& B^k - q_1 B^{k-1} - q_2 B^{k-2} - \cdots - q_{k-1}B \\
=& B^k + a_1 B^{k-1} + \cdots + a_{k-1}B.
\end{aligned}$$

Therefore,

$$\text{tr } A_k \quad = \quad \text{tr } \left[\left(B^k + a_1 B^{k-1} + \cdots + a_{k-1}B\right)\right],$$

which by Lemma 3.2.5 is equal to $-ka_k$. So $q_k = \frac{\text{tr } A_k}{k} = -a_k$.

Thus, P_k holds.

Therefore, by the principle of mathematical induction, P_i holds for all $p \geq i \geq 0$. That is, $q_i = -a_i$, $i = 0, \cdots, p$. $\square$

3.3. Generalized Inversion by Interpolation

In this section we present a method to obtain the generalized inverse of a quaternion polynomial matrix by interpolation at real number data points.

Lemma 3.3.1. *(See (Lam, 2001) for more details) An element $r \in \mathbb{H}$ is a root of a nonzero polynomial $f \in \mathbb{H}[x]$ iff $x - r$ is a right divisor of f. The set of polynomials in $\mathbb{H}[x]$ having r as a root is the left ideal $\mathbb{H}[x] \cdot (x - r)$.*

Proof. By direct calculation. □

Lemma 3.3.2. *(See (Lam, 2001) for more details) Let f, g and $h \in \mathbb{H}[x]$, $f = gh$ and $r \in \mathbb{H}$ be such that $\beta = h(r) \neq 0$. Then*

$$f(r) = g\left(\beta r \beta^{-1}\right) h(r).$$

In particular, if r is a root of f but not of h, then $\beta r \beta^{-1}$ is a root of g.

Proof. Let $g = \sum_{i=0}^{n} a_i x^i$. Then $f = \left(\sum_{i=0}^{n} a_i x^i\right) h = \sum_{i=0}^{n} a_i h x^i$. Therefore,

$$\begin{aligned} f(r) &= \sum_{i=0}^{n} a_i h(r) r^i = \sum_{i=0}^{n} a_i \beta r^i \\ &= \sum_{i=0}^{n} a_i \beta r^i \beta^{-1} \beta = \sum_{i=0}^{n} a_i \left(\beta r \beta^{-1}\right)^i \beta \\ &= \left(\sum_{i=0}^{n} a_i \left(\beta r \beta^{-1}\right)^i\right) \beta = g\left(\beta r \beta^{-1}\right) h(r). \end{aligned}$$

The last conclusion follows since $\mathbb{H}$ has no zero-divisors. □

Theorem 3.3.3. *(Adaption of (Gordon and Motzkin, 1965)) Let $f \in \mathbb{H}[x]$ be of degree n. Then the roots of f lie in at most n conjugacy classes of $\mathbb{H}$.*

Proof. We will show the desired result by mathematical induction.

For $n \geq 1$, let P_n be the statement:

Let $f \in \mathbb{H}[x]$ be of degree n. Then the roots of f lie in at most n conjugacy classes.

P_1 is the statement:

Let $f \in \mathbb{H}[x]$ be of degree 1. Then the root of f lies in at most 1 conjugacy class.

P_1 is trivially true.

Suppose P_n holds for all $1 \leq n \leq k-1$, it remains to show that P_k holds. Let $f \in \mathbb{H}[x]$ be of degree k with roots $c \neq d$. By Lemma 3.3.1, we can write

$$f = g \cdot (x - c),$$

where g has a root that is, by Lemma 3.3.2, conjugate to d. Invoking the inductive hypothesis, d lies in at most $k-1$ conjugacy classes. So, the roots of f lie in at most k conjugacy classes. Thus, P_k holds. Therefore, by the principle of mathematical induction, P_n holds for all $n \geq 1$. □

Theorem 3.3.4. *(Adaption of (Bray and Whaples, 1983)) Let $c_1, \cdots, c_n$ be n pairwise non-conjugate elements of $\mathbb{H}$. Then there is a unique polynomial $g_n \in \mathbb{H}[x]$, monic of degree n, such that $g_n(c_1) = \cdots = g_n(c_n) = 0$. Moreover, $c_1, \cdots, c_n$ are the only roots of g_n in $\mathbb{H}$.*

Proof. We first show the existence of g_n for all $n \geq 1$ by mathematical induction.

For $n \geq 1$, let P_n be the statement:

Let $c_1, \cdots, c_n \in \mathbb{H}$ be pairwise non-conjugate. Then there is a polynomial $g_n \in \mathbb{H}[x]$, monic of degree n, such that $g_n(c_1) = \cdots = g_n(c_n) = 0$. Moreover, $c_1, \cdots, c_n$ are the only roots of g in $\mathbb{H}$.

P_1 is the statement:

Let $c_1 \in \mathbb{H}$. Then there is a polynomial $g_1 \in \mathbb{H}[x]$, monic of degree 1, such that $g_1(c_1) = 0$. Moreover, c_1 is the only root of g in $\mathbb{H}$.

P_1 is trivially true as $g = x - c_1$.

Suppose P_n holds for all $1 \leq n \leq k-1$, it remains to show that P_k holds. Let $c_1, \cdots, c_k \in \mathbb{H}$ be pairwise non-conjugate. Invoking the inductive hypothesis, there exists a monic polynomial g_{k-1} of degree $k-1$ with $c_2, \cdots, c_k$ as its only roots, that is, $g_{k-1}(c_1) \neq 0$. Construct g_k as follows,

$$g_k = \left(x - g_{k-1}(c_1)\, c_1 g_{k-1}(c_1)^{-1}\right) \cdot g_{k-1}.$$

By Lemma 3.3.2, $g_k(c_1) = 0$. Thus, P_k holds.

Therefore, by the principle of mathematical induction, for all $n \geq 1$, P_n holds.

We next show that g_n is unique for all $n \geq 1$. For some fixed n, let $g' \neq g_n$ be monic of degree n, such that $g'(c_1) = \cdots = g'(c_n) = 0$ too. Then $\deg(g_n - g') \leq n-1$ but $g_n - g'$ has roots $c_1, \cdots, c_n$ which lie in n conjugacy classes of $\mathbb{H}$. This contradicts Theorem 3.3.3. Therefore, g_n is unique for all $n \geq 1$. □

Lemma 3.3.5. *Let $c_1, \cdots, c_{n+1} \in \mathbb{H}$ be pairwise non-conjugate and let $d_1, \cdots, d_{n+1} \in \mathbb{H}$. Then there is a unique lowest degree polynomial $f \in \mathbb{H}[x]$, of degree $p \leq n$, such that $f(c_i) = d_i$ for all $1 \leq i \leq n+1$.*

Proof. For any $1 \leq s' \leq n+1$, let $S = \{1, \cdots, n+1\} \setminus \{s'\}$. By Theorem 3.3.4, for any $s \in S$, we can find unique $h_S \in \mathbb{H}[x]$, monic of degree n, such that $h_S(c_s) = 0$ and that $\{c_s \mid s \in S\}$ are the only roots of h_S in $\mathbb{H}$. So, $h_S(c_{s'}) \neq 0$ and therefore we can construct a polynomial g_S of degree n such that $g_S(c_\alpha) = \begin{cases} 0 & \alpha \in S, \\ 1 & \alpha = s', \end{cases}$ as below:

$$g_S = h_S(c_{s'})^{-1} h_S.$$

We can then construct the polynomial f, of degree at most n, such that $f(c_i) = d_i$

for all $1 \leq i \leq n+1$ as below:

$$f = \sum_{s'=1}^{n+1} d_{s'} g_S.$$

We next show that f is unique. Suppose we have $f' \in \mathbb{H}[x]$ of degree $p' \leq p \leq n$, such that $f' \neq f$ and that $f'(c_i) = d_i$ for all $1 \leq i \leq n+1$ too. Then $f - f' \neq 0$ is of degree at most $p \leq n$. But $f - f'$ has roots $c_1, \cdots, c_{n+1}$ which lie in $n+1$ conjugacy classes of $\mathbb{H}$. This contradicts Theorem 3.3.3. Therefore, f is unique. □

Definition 3.3.6. The degree of a given quaternion polynomial matrix $A \in \mathbb{H}[x]^{m \times n}$ is $\deg A = \max\{\deg(A_{ij}) \mid 1 \leq i \leq m, 1 \leq j \leq n\}$.

Lemma 3.3.7. *Let $A \in \mathbb{H}[x]^{m \times n}$ have the generalized inverse $A^\dagger$. Then* $\deg A^\dagger \leq$ *$(2m-1)\deg A$.*

Proof. By Theorem 3.2.4, $\deg A^\dagger \leq \deg\left(A^* (AA^*)^{m-2}\right) \leq \deg(A^{2m-1}) \leq (2m-1)\deg A$. □

Remark 3.3.8. The upper bound mentioned in Lemma 3.3.7 is rarely achieved in practice, because some elements of AA^* (and other matrices) do not have maximal degree.

Lemma 3.3.9. *Let $c_1, \cdots, c_{k+1} \in \mathbb{H}$ be pairwise non-conjugate and let $A_1, \cdots, A_{k+1} \in \mathbb{H}^{n \times m}$. Then there is a unique lowest degree matrix $A \in \mathbb{H}[x]^{n \times m}$ of degree $p \leq k$, such that $A(c_i) = A_i$ for all $1 \leq i \leq k+1$.*

Proof. For any $1 \leq n_1 \leq n$ and $1 \leq m_1 \leq m$, by Lemma 3.3.5, there is a lowest degree polynomial $A_{n_1 m_1}$ determined by the values $c_1, \cdots, c_{k+1}$ and $(A_1)_{n_1 m_1}, \cdots, (A_{k+1})_{n_1 m_1}$. In fact, for any $1 \leq s' \leq k+1$, let $S = \{1, \cdots, k+1\} \setminus \{s'\}$. Then

$$A_{n_1 m_1} = \sum_{s'=1}^{k+1} (A_{s'})_{n_1 m_1} g_S,$$

where $g_S(c_\alpha) = \begin{cases} 0 & \alpha \in S, \\ 1 & \alpha = s' \end{cases}$. Since n_1 and m_1 are chosen randomly, a lowest degree matrix A that satisfies $A(c_i) = A_i$ for all $1 \leq i \leq k+1$ is determined by $A = \sum_{s'=1}^{k+1} A_{s'} g_S$.

Next we show that A is unique. Suppose $A' \neq A$ of degree $p' \leq p$ also satisfies $A'(c_i) = A_i$ for all $1 \leq i \leq k+1$. Then for some $1 \leq n_2 \leq n$ and $1 \leq m_2 \leq m$, $(A - A')_{n_2 m_2} \neq 0$. But $(A - A')_{n_2 m_2}$, of degree at most $p \leq k$, has roots $c_1, \cdots, c_{k+1}$ which lie in $k+1$ conjugacy classes of $\mathbb{H}$. This contradicts Theorem 3.3.3. Therefore, A is unique. □

Let $A \in \mathbb{H}[x]^{m\times n}$ satisfy conditions of Theorem 3.1.15. For any $x \in \mathbb{R}$, set $B = AA^*$. Let p be the largest integer such that $a_p \neq 0$. We construct the sequence $A_0, \cdots, A_p$ as follows:

$$\begin{array}{ccc} A_0 = 0 & -1 = q_0 & B_0 = I \\ \vdots & \vdots & \vdots \\ A_{p-1} = AA^* B_{p-2} & \dfrac{\mathrm{tr} A_{p-1}}{p-1} = q_{p-1} & B_{p-1} = A_{p-1} - q_{p-1} I \\ A_p = AA^* B_{p-1} & \dfrac{\mathrm{tr} A_p}{p} = q_p & B_p = A_p - q_p I \end{array}$$

Theorem 3.3.10. *In the above setting, let $k = (2m-1)\deg A$ and $c_1, \cdots, c_{k+1} \in \mathbb{R}$ be $k+1$ distinct real numbers such that $q_p(c_{s'}) \neq 0$ for any $1 \leq s' \leq k+1$. Let $S = \{1, \cdots, k+1\} \setminus \{s'\}$. Then*

$$A^\dagger = \sum_{s'=1}^{k+1} A(c_{s'})^\dagger g_S$$

where

$$A(c_{s'})^\dagger = \frac{1}{q_p(c_{s'})} A(c_{s'})^* \left[B(c_{s'})^{p-1} - q_1(c_{s'}) B(c_{s'})^{p-2} - \cdots - q_{p-1}(c_{s'}) I \right]$$

and

$$g_S(c_\alpha) = \begin{cases} 0 & \alpha \in S, \\ 1 & \alpha = s'. \end{cases}$$

Proof. By Theorem 3.2.4, Lemma 3.2.6 and Lemma 3.3.9. $\square$

Example 3.3.11. Find, by interpolation, the generalized inverse of the quaternion polynomial matrix

$$A = \begin{pmatrix} 14x+14+76\mathbf{i}+70\mathbf{j}+56\mathbf{k} & 56-28\mathbf{i}-70\mathbf{j}+70\mathbf{k} & 28\mathbf{j}-56\mathbf{k} & 14x-56-8\mathbf{i}-14\mathbf{j}-56\mathbf{k} \\ -2x-2-43\mathbf{i}-10\mathbf{j}-8\mathbf{k} & -8+4\mathbf{i}+10\mathbf{j}-10\mathbf{k} & -4\mathbf{j}+8\mathbf{k} & -2x+8-31\mathbf{i}+2\mathbf{j}+8\mathbf{k} \\ -3x-3+3\mathbf{i}-15\mathbf{j}-12\mathbf{k} & -12+6\mathbf{i}+15\mathbf{j}-15\mathbf{k} & -6\mathbf{j}+12\mathbf{k} & -3x+12+21\mathbf{i}+3\mathbf{j}+12\mathbf{k} \\ -4x-4+4\mathbf{i}-20\mathbf{j}-16\mathbf{k} & -16+8\mathbf{i}+20\mathbf{j}-20\mathbf{k} & -8\mathbf{j}+16\mathbf{k} & -4x+16+28\mathbf{i}+4\mathbf{j}+16\mathbf{k} \end{pmatrix}$$

The following steps have been implemented into a single procedure in Maple. For details, see the "LFI" procedure of §A.2.

As mentioned in Remark 3.3.8, here we only need two data points because direct calculation shows that $\deg A^\dagger = 2$. Let $c_1 = 0$ and $c_2 = 1$. We have:

$$A(c_1) = \begin{pmatrix} 14+76\mathbf{i}+70\mathbf{j}+56\mathbf{k} & 56-28\mathbf{i}-70\mathbf{j}+70\mathbf{k} & 28\mathbf{j}-56\mathbf{k} & -56-8\mathbf{i}-14\mathbf{j}-56\mathbf{k} \\ -2-43\mathbf{i}-10\mathbf{j}-8\mathbf{k} & -8+4\mathbf{i}+10\mathbf{j}-10\mathbf{k} & -4\mathbf{j}+8\mathbf{k} & 8-31\mathbf{i}+2\mathbf{j}+8\mathbf{k} \\ -3+3\mathbf{i}-15\mathbf{j}-12\mathbf{k} & -12+6\mathbf{i}+15\mathbf{j}-15\mathbf{k} & -6\mathbf{j}+12\mathbf{k} & 12+21\mathbf{i}+3\mathbf{j}+12\mathbf{k} \\ -4+4\mathbf{i}-20\mathbf{j}-16\mathbf{k} & -16+8\mathbf{i}+20\mathbf{j}-20\mathbf{k} & -8\mathbf{j}+16\mathbf{k} & 16+28\mathbf{i}+4\mathbf{j}+16\mathbf{k} \end{pmatrix}$$

and

$$A(c_2) = \begin{pmatrix} 28+76\mathbf{i}+70\mathbf{j}+56\mathbf{k} & 56-28\mathbf{i}-70\mathbf{j}+70\mathbf{k} & 28\mathbf{j}-56\mathbf{k} & -42-8\mathbf{i}-14\mathbf{j}-56\mathbf{k} \\ -4-43\mathbf{i}-10\mathbf{j}-8\mathbf{k} & -8+4\mathbf{i}+10\mathbf{j}-10\mathbf{k} & -4\mathbf{j}+8\mathbf{k} & 6-31\mathbf{i}+2\mathbf{j}+8\mathbf{k} \\ -6+3\mathbf{i}-15\mathbf{j}-12\mathbf{k} & -12+6\mathbf{i}+15\mathbf{j}-15\mathbf{k} & -6\mathbf{j}+12\mathbf{k} & 9+21\mathbf{i}+3\mathbf{j}+12\mathbf{k} \\ -8+4\mathbf{i}-20\mathbf{j}-16\mathbf{k} & -16+8\mathbf{i}+20\mathbf{j}-20\mathbf{k} & -8\mathbf{j}+16\mathbf{k} & 12+28\mathbf{i}+4\mathbf{j}+16\mathbf{k} \end{pmatrix}.$$

By the algorithm stated in Lemma 3.2.6, we calculate and obtain:

$$A(c_1)^\dagger = A(0)^\dagger = \frac{1}{230175} \times$$
$$\begin{pmatrix} 140-560\mathbf{i}-228\mathbf{j}-342\mathbf{k} & 355+1730\mathbf{i}-96\mathbf{j}+81\mathbf{k} & -255-870\mathbf{i}+126\mathbf{j}+54\mathbf{k} & -340-1160\mathbf{i}+168\mathbf{j}+72\mathbf{k} \\ 276+88\mathbf{i}+426\mathbf{j}-382\mathbf{k} & 282+416\mathbf{i}-93\mathbf{j}-149\mathbf{k} & -252-276\mathbf{i}-72\mathbf{j}+204\mathbf{k} & -336-368\mathbf{i}-96\mathbf{j}+272\mathbf{k} \\ 32+16\mathbf{i}-176\mathbf{j}+292\mathbf{k} & -176-88\mathbf{i}+68\mathbf{j}+194\mathbf{k} & 96+48\mathbf{i}+12\mathbf{j}-204\mathbf{k} & 128+64\mathbf{i}+16\mathbf{j}-272\mathbf{k} \\ -140-122\mathbf{i}+228\mathbf{j}+342\mathbf{k} & -355+2021\mathbf{i}+96\mathbf{j}-81\mathbf{k} & 255-1176\mathbf{i}-126\mathbf{j}-54\mathbf{k} & 340-1568\mathbf{i}-168\mathbf{j}-72\mathbf{k} \end{pmatrix}$$

and

$$A\left(c_2\right)^{\dagger} = A\left(1\right)^{\dagger} = \frac{1}{230175}\times$$
$$\begin{pmatrix} 152-550\mathbf{i}-244\mathbf{j}-330\mathbf{k} & 289+1675\mathbf{i}-8\mathbf{j}+15\mathbf{k} & -219-840\mathbf{i}+78\mathbf{j}+90\mathbf{k} & -292-1120\mathbf{i}+104\mathbf{j}+120\mathbf{k} \\ 268+104\mathbf{i}+406\mathbf{j}-402\mathbf{k} & 326+328\mathbf{i}+17\mathbf{j}-39\mathbf{k} & -276-228\mathbf{i}-132\mathbf{j}+144\mathbf{k} & -368-304\mathbf{i}-176\mathbf{j}+192\mathbf{k} \\ 32+16\mathbf{i}-160\mathbf{j}+300\mathbf{k} & -176-88\mathbf{i}-20\mathbf{j}+150\mathbf{k} & 96+48\mathbf{i}+60\mathbf{j}-180\mathbf{k} & 128+64\mathbf{i}+80\mathbf{j}-240\mathbf{k} \\ -152-132\mathbf{i}+244\mathbf{j}+330\mathbf{k} & -289+2076\mathbf{i}+8\mathbf{j}-15\mathbf{k} & 219-1206\mathbf{i}-78\mathbf{j}-90\mathbf{k} & 292-1608\mathbf{i}-104\mathbf{j}-120\mathbf{k} \end{pmatrix}.$$

By Theorem 3.3.10, we have:

$$\begin{aligned} A^{\dagger} &= \sum_{s'=1}^{2} A\left(c_{s'}\right)^{\dagger} g_S = A\left(0\right)^{\dagger}\left(1-x\right) + A\left(1\right)^{\dagger} x \\ &= \frac{1}{230175}\times \end{aligned}$$
$$\left(\begin{matrix} \left(12+10\mathbf{i}-16\mathbf{j}+12\mathbf{k}\right)x+140-560\mathbf{i}-228\mathbf{j}-342\mathbf{k} & \left(-66-55\mathbf{i}+88\mathbf{j}-66\mathbf{k}\right)x+355+1730\mathbf{i}-96\mathbf{j}+81\mathbf{k} \\ \left(-8+16\mathbf{i}-20\mathbf{j}-20\mathbf{k}\right)x+276+88\mathbf{i}+426\mathbf{j}-382\mathbf{k} & \left(44-88\mathbf{i}+110\mathbf{j}+110\mathbf{k}\right)x+282+416\mathbf{i}-93\mathbf{j}-149\mathbf{k} \\ \left(16\mathbf{j}+8\mathbf{k}\right)x+32+16\mathbf{i}-176\mathbf{j}+292\mathbf{k} & \left(-88\mathbf{j}-44\mathbf{k}\right)x-176-88\mathbf{i}+68\mathbf{j}+194\mathbf{k} \\ \left(-12-10\mathbf{i}+16\mathbf{j}-12\mathbf{k}\right)x-140-122\mathbf{i}+228\mathbf{j}-342\mathbf{k} & \left(66+55\mathbf{i}-88\mathbf{j}+66\mathbf{k}\right)x-355+2021\mathbf{i}+96\mathbf{j}-81\mathbf{k} \end{matrix}\right.$$
$$\left.\begin{matrix} \left(36+30\mathbf{i}-48\mathbf{j}+36\mathbf{k}\right)x-255-870\mathbf{i}+126\mathbf{j}+54\mathbf{k} & \left(48+40\mathbf{i}-64\mathbf{j}+48\mathbf{k}\right)x-340-1160\mathbf{i}+168\mathbf{j}+72\mathbf{k} \\ \left(-24+48\mathbf{i}-60\mathbf{j}-60\mathbf{k}\right)x-252-276\mathbf{i}-72\mathbf{j}+204\mathbf{k} & \left(-32+64\mathbf{i}-80\mathbf{j}+80\mathbf{k}\right)x-366-368\mathbf{i}-96\mathbf{j}+272\mathbf{k} \\ \left(48\mathbf{j}+24\mathbf{k}\right)x+96+48\mathbf{i}+12\mathbf{j}-204\mathbf{k} & \left(64\mathbf{j}+32\mathbf{k}\right)x+128+64\mathbf{i}+16\mathbf{j}-272\mathbf{k} \\ \left(-36-30\mathbf{i}+48\mathbf{j}-36\mathbf{k}\right)x+255-1176\mathbf{i}-126\mathbf{j}-54\mathbf{k} & \left(-48-40\mathbf{i}+64\mathbf{j}-48\mathbf{k}\right)x+340-1568\mathbf{i}-168\mathbf{j}-72\mathbf{k} \end{matrix}\right).$$

Direct calculation shows that $A^{\dagger}$ satisfies the defining relations of the generalized inverse stated in Definition 3.1.3.

A. Maple Codes

A.1. Overview

At the early stage of my graduate studies, I realized the need of utilizing Maple when trying to solve quaternion quadratic equations. There are existing Maple packages dedicated to quaternions (Carter, 2007; Harder, 2010), but not quaternion polynomials or quaternion polynomial matrices. Therefore, I started writing my own Maple module. During my two years of graduate studies, I gradually extended this module to have over thirty procedures. These procedures enable the user to do basic quaternion and quaternion polynomial matrix operations such as additions and multiplications; to do Lagrange and Newton interpolations of a list of quaternion pairs; to obtain the generalized inverse of a quaternion polynomial matrix; to solve multiple quaternion equations and quaternion least norm problems; and so on. In §A2 the codes of the procedures are provided with descriptions and examples.

A.2. The Codes

```
Q := module()

option package;

export M, `*`, `^`, eval, Qdef, Qrand, QPrand, Mrand,

PMrand, Qreal, Qimag, Qconj, Mconj, Qnorm, Qinv, AXXB,

sortcollect, ZeroP, QMFNorm, PRotate, LagrangeP, AXBc,

NewtonP, QAdjointMatrix, LFI, NoSim1, NoSim2, Meval, AXB,

SemiSim, CSSim, LN1, LN2;

local QMultiplyScalar, QMatrixMultiply, QScalarMultiply;

M := proc(a,b)

#M is the multiplication operator for quaternions and quaternion polynomials.

local a1 := Qreal(a), b1 := coeff(a, I),

c1 := coeff(a, J), d1 := coeff(a, K),

a2 := Qreal(b), b2 := coeff(b, I),

c2 := coeff(b, J), d2 := coeff(b, K);

return

sortcollect (a1*a2 - b1*b2 - c1*c2 - d1*d2

+ (a1*b2 + b1*a2 + c1*d2 - d1*c2)*I

+ (a1*c2 - b1*d2 + c1*a2 + d1*b2)*J

+ (a1*d2 + b1*c2 - c1*b2 + d1*a2)*K, x);

end proc;
```

```
QMatrixMultiply := proc(A::Matrix, B::Matrix)

#Local procedure for quaternion matrix multiplication.

local nrowsA := LinearAlgebra:-RowDimension(A),

ncolsA := LinearAlgebra:-ColumnDimension(A),

ncolsB := LinearAlgebra:-ColumnDimension(B),

AB := Matrix(1..nrowsA, 1..ncolsB), i, j;

for i from 1 to nrowsA do

for j from 1 to ncolsB do

AB(i, j) := simplify(add(M(A(i, k), B(k, j)), k = 1 .. ncolsA));

end do;

end do;

end proc;

QMultiplyScalar := proc(A::Matrix, B)

#Local procedure to multiply a quaternion matrix by a quaternion scalar.

local nrowsA := LinearAlgebra:-RowDimension(A),

ncolsA := LinearAlgebra:-ColumnDimension(A),

AB := Matrix(1..nrowsA, 1..ncolsA), i, j;

for i from 1 to nrowsA do

for j from 1 to ncolsA do

AB(i, j) := simplify(M(A(i, j), B));

end do;

end do;
```

```
end proc;

QScalarMultiply := proc(A, B::Matrix)

#Local procedure to multiply a quaternion scalar by a quaternion matrix.

local nrowsB := LinearAlgebra:-RowDimension(B),

ncolsB := LinearAlgebra:-ColumnDimension(B),

AB := Matrix(1..nrowsB, 1..ncolsB), i, j;

for i from 1 to nrowsB do

for j from 1 to ncolsB do

AB(i, j) := simplify(M(A, B(i, j)));

end do;

end do;

end proc;

`*` := proc(m, n)

#Newly defined multiplication operator for quaternions related elements.

options remember;

if type(m, Matrix) and type(n, Matrix) then

QMatrixMultiply(m, n);

elif type(m, Matrix) then

QMultiplyScalar(m, n);

elif type(n, Matrix) then

QScalarMultiply(m, n);

else M(m, n);
```

end if;

end proc;

‘^‘ := proc(m, n::integer)

#Newly defined exponential operator to include quaternion raised to positive integer powers.

options remember;

local i, p := 1;

if type(m, freeof({I, J, K})) = false and n > 1 then

for i from 1 to n do

p := M(p, m);

end do;

return p;

else m^n;

end if;

end proc;

Qdef := proc(a, b, c, d)

#Defines a quaternion $a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k}$.

a + b*I + c*J + d*K;

end proc:

Qrand := proc(m::integer, n::integer)

#Generates a random quaternion with all integer coordinates between m and n.

Qdef(rand(m..n)(), rand(m..n)(), rand(m..n)(), rand(m..n)());

end proc:

QPrand := proc (m,n,k)

#Generates a random quaternion polynomial of degee k, with coefficients $Qrand\,(m, n)$.

local j := k;

return convert([seq(Qrand(m, n)*x^(j - i), i = 0 .. k)], ‘+‘)

end proc;

Mrand := proc(m, n, p, q)

#Generates a random $p \times q$ Matrix with quaternion entries $Qrand\,(m, n)$.

local M := Matrix(1 .. p, 1 .. q), i, j;

for i from 1 to p do

for j from 1 to q do

M(i, j) := Qrand(m, n);

end do;

end do;

end proc:

PMrand := proc(m, n, k1, k2, p, q)

#Generates a random $p \times q$ Matrix with random quaternion polynomial entries.

local M := Matrix(1 .. p, 1 .. q), i, j;

for i from 1 to p do

for j from 1 to q do

M(i, j) := QPrand(m, n, rand(k1..k2)());

```
end do;

end do;

end proc:

Qreal := proc(a)

#Calculates the real part of a quaternion.

subs(I = 0, J = 0, K = 0, a);

end proc:

Qimag := proc(a)

#Calculates the imaginary part of a quaternion.

a - Qreal(a);

end proc:

Qconj := proc(f)

#Calculates the conjugate of a quaternion or a quaternion polynomial.

sortcollect(Qreal(f) - Qimag(f), x);

end proc:

Mconj := proc(A)

#Calculates the conjugate of a quaternion or quaternion polynomial matrix.

local nrowsA:=LinearAlgebra:-RowDimension(A),

ncolsA:=LinearAlgebra:-ColumnDimension(A),

AC := Matrix(1..nrowsA, 1..ncolsA), i, j;

for i from 1 to nrowsA do

for j from 1 to ncolsA do
```

AC(i, j) := Qconj(A(i, j))

end do;

end do;

end proc:

sortcollect := proc (A, a)

#Returns A as a polynomial in standard form of variable a.

return sort(collect(A, a), a)

end proc;

eval := proc(N, n)

#Calculates $N(n)$ where $N(x)$ is a polynomial of x.

options remember;

local i, f := 1, m := N;

if has(N, x) = false then

expand(N);

elif is(expand(N),‘+‘) = true then

map(eval,expand(N),n);

else for i from 1 to degree(N, x) do

f := M(f, n);

end do;

M(coeff(N, x, degree(N, x)), f);

end if;

end proc;

```
Meval := proc(N, n)

#Calculates N (n) where N (x) is a quaternion polynomial matrix of x.

options remember;

local i, j,

m := Matrix(LinearAlgebra:-RowDimension(N),

LinearAlgebra:-ColumnDimension(N));

for i from 1 to LinearAlgebra:-RowDimension(m) do

for j from 1 to LinearAlgebra:-ColumnDimension(m) do

m(i, j) := eval(N(i, j), n);

end do;

end do;

end proc;

PRotate := proc(m::integer)

#Generates a block matrix.

local M := Matrix(1 .. 2*m, 1 .. 2*m), i;

for i from 1 to m do

M(i, i + m) := -1;

end do;

for i from m + 1 to 2*m do

M(i, i - m) := 1;

end do;

end proc;
```

```
Qnorm := proc(a)

#Calculates the 2 norm of a quaternion.

(Qreal(a)^2 + coeff(a, I)^2 + coeff(a, J)^2 + coeff(a, K)^2)^(1/2);

end proc:

Qinv := proc(a)

#Calculates the inverse of a quaternion or a quaternion polynomial.

sortcollect(Qreal(a) - Qimag(a), x)/sortcollect

((Qreal(a)^2 + coeff(a,I)^2 + coeff(a, J)^2 + coeff(a, K)^2), x)

end proc:

NoSim1 := proc(A::list)

#Ensures that a list of quaternion is ready for zero polynomial interpolation.

options remember;

local i, j, k, B := A, n := nops(B);

for i from 2 to n do

for j from 1 to i - 1 do

if Qreal(B[j]) = Qreal(B[i]) and Qnorm(B[j]) = Qnorm(B[i]) then

for k from j + 1 to i - 1 do

if Qreal(B[k]) = Qreal(B[i]) and Qnorm(B[k]) = Qnorm(B[i]) then

B := subsop(i = z, B); next;

end if;

end do;

end if;
```

```
end do;

end do;

return remove(t -> t = z, B);

end proc;

ZeroP := proc(B::list)

#Calculates an zero polynomial for a list of quaternions.

options remember;

local A := [op({op(NoSim1(B))})], i, n := nops(A), f;

f[1] := x - op(1, A);

for i from 2 to n do

f[i] := M((x - M(M(eval((f[i - 1]), op(i, A)), op(i, A)),

Qinv(eval((f[i - 1]), op(i, A))))), f[i - 1]);

end do;

return sortcollect(f[n], x);

end proc;

NoSim2 := proc(A::listlist)

#Ensures that a list of quaternion pairs is ready for Lagrange Interpolation.

options remember;

local i, j, k, B := A, C, n1 := nops(B), n2;

for i from 2 to n1 do

for j from 1 to i - 1 do

if A[j][1] = A[i][1] then
```

```
B := subsop([i,1] = z, B);
end if;
end do;
end do;
C := remove(t -> t[1] = z, B); n2 := nops(C);
for i from 3 to n2 do
for j from 1 to i-2 do
if Qreal(C[j][1]) = Qreal(C[i][1]) and Qnorm(C[j][1]) = Qnorm(C[i][1]) then
for k from j + 1 to i - 1 do
if Qreal(C[k][1]) = Qreal(C[i][1]) and Qnorm(C[k][1]) = Qnorm(C[i][1]) then
C := subsop([i, 1] = z, C); next;
end if;
end do;
end if;
end do;
end do;
return remove(t -> t[1] = z, C);
end proc;
LagrangeP := proc(B::listlist)
#Calculates the Lagrange polynomial for a list of quaternion pairs.
local A := NoSim2(B), n := nops(A), b, c, i, j, f, g;
options remember;
```

```
for i from 1 to n do

b[i] := op(1, op(i, A));

c[i] := op(2, op(i, A));

end do;

for i from 1 to n do

f[i] := ZeroP([seq(b[j], j in {seq(1..n)} minus {i})]);

end do;

for i from 1 to n do

g[i] := M(Qinv(eval(f[i], b[i])), f[i]);

end do;

return sortcollect(add(M(c[i], g[i]), i = 1..n), x);

end proc:

QAdjointMatrix := proc(A::Matrix)

#Calculates the adjoint matrix of a quaternion matrix.

local nrowsA := LinearAlgebra:-RowDimension(A),

ncolsA := LinearAlgebra:-ColumnDimension(A),

A1 := Matrix(1..nrowsA, 1..ncolsA), A1C := Matrix(1..nrowsA, 1..ncolsA),

A2 := Matrix(1..nrowsA, 1..ncolsA), A2C := Matrix(1..nrowsA, 1..ncolsA),

AM := Matrix(1..2*nrowsA, 1..2*ncolsA), i, j;

for i from 1 to nrowsA do

for j from 1 to ncolsA do

A1(i, j) := A(i, j) - coeff(A(i, j), J)*J - coeff(A(i, j), K)*K;
```

```
A1C(i, j) := Qconj(A1(i, j));

A2(i, j) := coeff(A(i, j), J) + coeff(A(i, j), K)*I;

A2C(i,j) := Qconj(A2(i, j));

AM(i, j) := A1(i, j);

end do;

end do;

for i from 1 to nrowsA do

for j from ncolsA + 1 to 2*ncolsA do

AM(i, j) := A2(i, j - ncolsA);

end do;

end do;

for i from nrowsA + 1 to 2*nrowsA do

for j from 1 to ncolsA do

AM(i, j) := -A2C(i - nrowsA, j);

end do;

end do;

for i from nrowsA + 1 to 2*nrowsA do

for j from ncolsA + 1 to 2*ncolsA do

AM(i, j) := A1C(i - nrowsA, j - ncolsA);

end do;

end do;

for i from 1 to 2*nrowsA do
```

```
for j from 1 to 2*ncolsA do

AM(i, j) := subs(I = ii, AM(i, j));

end do;

end do; return AM;

end proc;

QMFNorm := proc(A::Matrix)

#Calculates the frobinius norm of a quaternion matrix.

local nrowsA := LinearAlgebra:-RowDimension(A),

ncolsA := LinearAlgebra:-ColumnDimension(A), i, j, N := 0;

for i from 1 to nrowsA do

for j from 1 to ncolsA do

N := N + coeff(A[i, j], I)^2 + coeff(A[i, j], J)^2

+ coeff(A[i, j], K)^2 + (Qreal(A[i, j]))^2;

end do;

end do;

return (expand(N))^(1/2);

end proc;

NewtonP := proc(B::listlist)

#Calculates the Newton polynomials and the divided difference for a list of quater-
nion pairs.

local g, i, j, d, b, c, A := NoSim2(B), n := nops(A);

options remember;
```

```
for i from 1 to n do
b[i] := op(1, op(i, A)); c[i] := op(2, op(i, A));
end do;
d[0] := c[1]; g[0] := 1;
for i from 1 to n - 1 do
g[i] := ZeroP([seq(b[j], j = 1..i)]);
d[i] := M(c[i + 1] - c[1] - add(M(d[j], eval(g[j], b[i + 1])), j = 1..i - 1),
Qinv(eval(g[i], b[i + 1])));
end do;
return [seq(g[i], i = 0..n - 1)], [seq(d[i], i = 0..n - 1)];
end proc;
LFI := proc (A::Matrix)
#Calculates the generalized inverse of a quaternion polynomial matrix.
local nrowsA := LinearAlgebra:-RowDimension(A),
H := LinearAlgebra:-Transpose(Mconj(A)),
P := QMatrixMultiply(A, H),
a, B, AA, i, j, k := 0, LFI; a[0] := 1;
B[0] := LinearAlgebra:-IdentityMatrix(nrowsA);
for i from 1 to nrowsA do
AA[i] := QMatrixMultiply(P, B[i - 1]);
a[i] := -LinearAlgebra:-Trace(AA[i])/i;
B[i] := AA[i] + QScalarMultiply(a[i], B[0]);
```

```
end do;
for j from 1 to nrowsA do
if a[j] <> 0 then
k := j;
end if;
end do; if k = 0 then
LFI := 0;
else
LFI := [(-1/a[k]), QMatrixMultiply(H, B[k - 1])];
end if;
end proc:
AXXB:= proc (A, B, C)
#Calculates the solutions of x^2 + Ax + xB + C = 0.
local a0 := Qreal(A), a1 := coeff(A, I),
a2 := coeff(A, J), a3 := coeff(A, K),
b0 := Qreal(B), b1 := coeff(B, I),
b2 := coeff(B, J), b3 := coeff(B, K),
c := ((a0 + b0)/2)^2 - M((A + B), ((a0 + b0)/2)) + C,
c0 := Qreal(c), c1 := coeff(c, I),
c2 := coeff(c, J),c3 := coeff(c, K),
one := -4*(a1 + b1)*x^3 + 4*(a2*b3 - a3*b2 - c1)*x^2
+ (b1 - a1)*(a1^2 + a2^2 + a3^2 - b1^2 - b2^2 - b3^2)*x
```

```
+ 2*(c2*(b3 - a3) + c3*(a2 - b2))*x

+ (b1 - a1)*(c1*(a1 - b1) + c2*(a2 - b2) + c3*(a3 - b3)),

two := -4*(a2 + b2)*x^3 + 4*(a3*b1 - a1*b3 - c2)*x^2

+ (b2 - a2)*(a1^2 + a2^2 + a3^2 - b1^2 - b2^2 - b3^2)*x

+ 2*(c1*(a3 - b3) + c3*(b1 - a1))*x

+ (b2 - a2)*(c1*(a1 - b1) + c2*(a2 - b2) + c3*(a3 - b3)),

three := -4*(a3 + b3)*x^3 + 4*(a1*b2 - a2*b1 - c3)*x^2

+ (b3 - a3)*(a1^2 + a2^2 + a3^2 - b1^2 - b2^2 - b3^2)*x

+ 2*(c1*(b2 - a2) + c2*(a1 - b1))*x

+ (b3 - a3)*(c1*(a1 - b1) + c2*(a2 - b2) + c3*(a3 - b3)),

d := 8*x^3 + 2*((a1 - b1)^2 + (a2 - b2)^2 + (a3 - b3)^2)*x,

f := 16*x^3

+ 8*(a1^2 + a2^2 + a3^2 + b1^2 + b2^2 + b3^2 + 2*c0)*x^2

+ 4*c0*((a1 - b1)^2 + (a2 - b2)^2 + (a3 - b3)^2)*x

+ (a1^2 + a2^2 + a3^2 - b1^2 - b2^2 - b3^2)^2*x

+ (8*(a2*b3 - a3*b2)*c1 - 4*c1^2)*x

+ (8*(a3*b1 - a1*b3)*c2 - 4*c2^2)*x

+ (8*(a1*b2 - a2*b1)*c3 - 4*c3^2)*x

- (a1*c1 + a2*c2 + a3*c3 - b1*c1 - b2*c2 - b3*c3)^2,

Z := evalf(op(select(t -> evalf(t) > 0,

map(Re + Im*ii, simplify(evalc([solve(f,x)])))))));

return map(t ->
```

```
(t + subs(x = t,one/d)*I

+ subs(x = t,two/d)*J + subs(x = t,three/d)*K - (a0 + b0)/2),

{sqrt(Z), -sqrt(Z)});

end proc:

AXBc:= proc (A, B, c)
```

#Calculates the solutions of $x^2 + AxB + c = 0$.

```
local a1 := coeff(A, I), a2 := coeff(A, J), a3 := coeff(A, K),

b1 := coeff(B, I), b2 := coeff(B, J), b3 := coeff(B, K),

one := x*(a3*b2 - a2*b3),

two := x*(a1*b3 - a3*b1),

three := x*(a2*b1 - a1*b2),

a := (a1^2 + a2^2 + a3^2)*(b1^2 + b2^2 + b3^2),

b := a1*b1 + a2*b2 + a3*b3,

d := 2*x + b,

f := x^4 + (c - (3/4)*a)*x^2

+ (1/4)*b*(4*c - a)*x + (1/4)*c*b^2,

sol := fsolve(f, x, real);

return map(t -> (t

+ subs(x = t,one/d)*I + subs(x = t,two/d)*J + subs(x = t,three/d)*K),

{sol});

end proc:

AXB:= proc (A, B)
```

```
#Calculates the solutions of x^2 + AxB = 0.

local a0 := Qreal(A), a1 := coeff(A, I),

a2 := coeff(A, J), a3 := coeff(A, K),

b0 := Qreal(B), b1 := coeff(B, I),

b2 := coeff(B, J), b3 := coeff(B, K),

m11 := a0*b0 - a1*b1 + a2*b2 + a3*b3,

m12 := a0*b3 - a1*b2 - a2*b1 - a3*b0,

m13 := -a0*b2 - a1*b3 + a2*b0 - a3*b1,

m21 := -a0*b3 - a1*b2 - a2*b1 + a3*b0,

m22 := a0*b0 + a1*b1 - a2*b2 + a3*b3,

m23 := a0*b1 - a1*b0 - a2*b3 - a3*b2,

m31 := a0*b2 - a1*b3 - a2*b0 - a3*b1,

m32 := -a0*b1 + a1*b0 - a2*b3 - a3*b2,

m33 := a0*b0 + a1*b1 + a2*b2 - a3*b3,

d := 8*x^3 + 4*(m11 + m22 + m33)*x^2

+ 2*(m11*m22 + m11*m33 - m12*m21 - m13*m31 - m23*m32 + m22*m33)*x

+ m11*m22*m33 - m11*m23*m32 - m12*m21*m33

+ m13*m21*m32 + m12*m23*m31 - m13*m22*m31,

n1 := -a0*b1 - a1*b0 - a2*b3 + a3*b2,

n2 := -a0*b2 + a1*b3 - a2*b0 - a3*b1,

n3 := -a0*b3 - a1*b2 + a2*b1 - a3*b0,

one := 4*n1*x^3 + 2*(n1*(m22 + m33) - n2*m12 - n3*m13)*x^2
```

```
+ n1*(m22*m33 - m23*m32)*x

+ n2*(m13*m32 - m12*m33)*x

+ n3*(m12*m23 - m13*m22)*x,

two := 4*n2*x^3 + 2*(n2*(m11 + m33) - n1*m21 - n3*m23)*x^2

+ n1*(m23*m31 - m21*m33)*x

+ n2*(m11*m33 - m13*m31)*x

+ n3*(m13*m21 - m11*m23)*x,

three := 4*n3*x^3 + 2*(n3*(m11 + m22) - n1*m31 - n2*m32)*x^2

+ n1*(m21*m32 - m22*m31)*x

+ n2*(m12*m31 - m11*m32)*x

+ n3*(m11*m22 - m12*m21)*x,

ab := (a0^2 + a1^2 + a2^2 + a3^2)*(b0^2 + b1^2 + b2^2 + b3^2),

h := x^3 + 2*a0*b0*x^2 + ((4/3)*a0^2*b0^2

- (1/12)*(a0^2 - 3*a1^2 - 3*a2^2 - 3*a3^2)

*(b0^2 - 3*b1^2 - 3*b2^2 - 3*b3^2))*x

+ (1/4)*ab*(a0*b0 - a1*b1 - a2*b2 - a3*b3),

g := x^4 + 2*a0*b0*x^3

+ (1/2)*(4*a0^2*b0^2

- (a1^2 + a2^2 + a3^2 - a0^2)*(b1^2 + b2^2 + b3^2 - b0^2))*x^2

+ (1/2)*a0*b0*ab*x + (1/16)*ab^2,

solg := fsolve(g, x, real),

solh := fsolve(h, x, real);
```

if a0 = 0 and b0 = 0 then

return map(t -> (t

+ subs(x = t,one/d)*I + subs(x = t,two/d)*J + subs(x = t,three/d)*K),

{solh});

else

return map(t -> (t

+ subs(x = t, one/d)*I + subs(x = t, two/d)*J + subs(x = t, three/d)*K),

{solg, solh});

end if;

end proc:

SemiSim:= proc (B, A)

#Calculates (x, y) such that $\begin{cases} yAx = & B \\ xBy = & A \end{cases}.$

local a0 := Qreal(A), a1 := coeff(A, I), a2 := coeff(A, J),

a3 := coeff(A, K), b0 := Qreal(B), b1 := coeff(B, I),

b2 := coeff(B, J), b3 := coeff(B, K), Z;

if Qnorm(A) <> Qnorm(B) then

return "No Solution"

elif abs(a0) <> abs(b0) then

return "No Solution"

else Z := LinearAlgebra:-LinearSolve(

Matrix([[a0^2 - b0^2, -a0*a3 - b0*b3, a0*a2 + b0*b2, a0*a1 - b0*b1, 0], [a0*a3 + b0*b3, a0^2 - b0^2, -a0*a1 - b0*b1, a0*a2 - b0*b2, 0],

[-a0*a2 - b0*b2, a0*a1 + b0*b1, a0^2 - b0^2, a0*a3 - b0*b3, 0],

[-a0*a1 + b0*b1, -a0*a2 + b0*b2, -a0*a3 + b0*b3, a0^2 - b0^2, 0]]),

free = 't');

end if;

return sort(Z[4] + Z[1]*I + Z[2]*J + Z[3]*K, [I, J, K]) and

simplify(M(B, M(Qinv(Z[4] + Z[1]*I + Z[2]*J + Z[3]*K),Qinv(A))));

end proc:

CSSim:= proc (B, A)

#Calculates (x, y) such that $\begin{cases} \overline{y}Ax = & B \\ \overline{x}By = & A \end{cases}.$

#Calculates x,y such that

local a0 := Qreal(A), a1 := coeff(A, I), a2 := coeff(A, J),

a3 := coeff(A, K), b0 := Qreal(B), b1 := coeff(B, I),

b2 := coeff(B, J), b3 := coeff(B, K), Z;

if Qnorm(A) <> Qnorm(B) then

return "No Solution"

else Z := LinearAlgebra:-LinearSolve(

Matrix([[0, a0*b3 + a3*b0, -a0*b2 - b0*a2, a2*b3 - a3*b2, 0],

[-a0*b3 - a3*b0, 0, a0*b1 + a1*b0, -a1*b3 + a3*b1, 0],

[a0*b2 + a2*b0, -a0*b1 - a1*b0, 0, a1*b2 - a2*b1, 0],

[-a2*b3 + a3*b2, a1*b3 - a3*b1, -a1*b2 + a2*b1, 0, 0]]),

```
free = 't');
end if;
return simplify(Z[4] + Z[1]*I + Z[2]*J + Z[3]*K) and
simplify(M(M(Qinv(B),Qinv(Qconj(Z[4] + Z[1]*I + Z[2]*J + Z[3]*K))),A))
end proc:
LN1:= proc (A, B, C)
```

#Calculates the solution of $\min_{x\in\mathbb{H}} |ax - xb - c|$.

```
local a0 := Qreal(A), a1 := coeff(A, I), a2 := coeff(A, J),
a3 := coeff(A, K), b0 := Qreal(B), b1 := coeff(B, I),
b2 := coeff(B, J), b3 := coeff(B, K), c0 := Qreal(C),
c1 := coeff(C, I), c2 := coeff(C, J), c3 := coeff(C, K), Z;
if Qnorm(A) = Qnorm(B) and Qreal(A) = Qreal(B) then
Z := LinearAlgebra:-MatrixMatrixMultiply(LinearAlgebra:-MatrixInverse(
Matrix([[a0 - b0, -a3 - b3, a2 + b2, a1 - b1],
[a3 + b3, a0 - b0, -a1 - b1, a2 - b2],
[-a2 - b2, a1 + b1, a0 - b0, a3 - b3],
[-a1 + b1, -a2 + b2, -a3 + b3, a0 - b0]]), method = pseudo),
Matrix(4, 1, {(1, 1) = c1, (2, 1) = c2, (3, 1) = c3, (4, 1) = c0}));
else
Z := LinearAlgebra:-LinearSolve(
Matrix([[a0 - b0, -a3 - b3, a2 + b2, a1 - b1, c1],
[a3 + b3, a0 - b0, -a1 - b1, a2 - b2, c2],
```

```
[-a2 - b2, a1 + b1, a0 - b0, a3 - b3, c3],
[-a1 + b1, -a2 + b2, -a3 + b3, a0 - b0, c0]]), free = 't');
end if;
return simplify(Z[4] + Z[1]*I + Z[2]*J + Z[3]*K);
end proc:
LN2:= proc (A,B,C)
```

#Calculates the solution of $\min_{x\in\mathbb{H}} |ax - \bar{x}b - c|$.

```
local a0 := Qreal(A), a1 := coeff(A, I), a2 := coeff(A, J),
a3 := coeff(A, K), b0 := Qreal(B), b1 := coeff(B, I),
b2 := coeff(B, J), b3 := coeff(B, K), c0 := Qreal(C),
c1 := coeff(C, I), c2 := coeff(C, J), c3 := coeff(C, K), Z;
if Qnorm(A) = Qnorm(B) and Qreal(A) = Qreal(B) then
Z := LinearAlgebra:-MatrixMatrixMultiply(LinearAlgebra:-MatrixInverse(
Matrix([[a0 + b0, -a3 + b3, a2 - b2, a1 - b1],
[a3 - b3, a0 + b0, -a1 + b1, a2 - b2],
[-a2 + b2, a1 - b1, a0 + b0, a3 - b3],
[-a1 - b1, -a2 - b2, -a3 - b3, a0 - b0]]), method = pseudo),
Matrix(4, 1, {(1, 1) = c1, (2, 1) = c2, (3, 1) = c3, (4, 1) = c0}));
else
Z := LinearAlgebra:-LinearSolve(
Matrix([[a0 + b0, -a3 + b3, a2 - b2, a1 - b1, c1],
[a3 - b3, a0 + b0, -a1 + b1, a2 - b2, c2],
```

```
[-a2 + b2, a1 - b1, a0 + b0, a3 - b3, c3],

[-a1 - b1, -a2 - b2, -a3 - b3, a0 - b0, c0]]), free = 't');

end if;

return simplify(Z[4] + Z[1]*I + Z[2]*J + Z[3]*K);

end proc:

end module:
```

Examples

```
> a:=Qdef(1,2,3,4);
```

$$a := 1 + 2I + 3J + 4K \quad (1.1)$$

```
> b:=Qrand(-10,10);
```

$$b := 4 + 2I - 5J - 9K \quad (1.2)$$

```
> Qreal(a);
```

$$1 \quad (1.3)$$

```
> Qimag(b);
```

$$2I - 5J - 9K \quad (1.4)$$

```
> c:=a+b;
```

$$c := 5 + 4I - 2J - 5K \quad (1.5)$$

```
> Qnorm(c);
```

$$\sqrt{70} \quad (1.6)$$

```
> d:=a*b;
```

$$d := 51 + 3I + 33J - 9K \quad (1.7)$$

```
> e:=Qconj(a);
```

$$e := 1 - 2I - 3J - 4K \quad (1.8)$$

```
> f:=Qinv(b);
```

$$f := \frac{2}{63} - \frac{1}{63}I + \frac{5}{126}J + \frac{1}{14}K \quad (1.9)$$

```
> A:=LagrangeP([[a,b],[c,d],[e,f]]);
```

$$A := \left(\frac{3553}{14616}J - \frac{137}{2436}I + \frac{3011}{2520}K - \frac{3419}{146160}\right)x^2 + \left(-\frac{1067}{7308}I - \frac{18757}{24360} - \frac{1751}{7308}J - \frac{107479}{36540}K\right)x - \frac{3191}{7308}I + \frac{8342}{1827}J + \frac{19447}{609}K + \frac{31145}{14616} \quad (1.10)$$

```
> is(eval(A,a)=b);
```

$$true \quad (1.11)$$

```
> B:=NewtonP([[a,b],[c,d],[e,f]]);
```

$$B := \left[1, x - 1 - 2I - 3J - 4K, x^2 + \left(-6 - \frac{92}{63}I + \frac{136}{63}J + \frac{16}{63}K\right)x + \frac{270}{7} + \frac{220}{63}J + \frac{100}{63}I + \frac{220}{9}K\right], \left[4 + 2I - 5J - 9K, 2I + 3J + 4K, \frac{3553}{14616}J - \frac{137}{2436}I + \frac{3011}{2520}K - \frac{3419}{146160}\right] \quad (1.12)$$

```
> C:=zip((x,y)->x*y,B[2],B[1]);
```

$$C := \left[4 + 2I - 5J - 9K, (2I + 3J + 4K)x + 29 - 2I - 3J - 4K, \left(\frac{3553}{14616}J - \frac{137}{2436}I + \frac{3011}{2520}K - \frac{3419}{146160}\right)x^2 + \left(-\frac{15683}{7308}I - \frac{18757}{24360} - \frac{23675}{7308}J - \frac{253639}{36540}K\right)x + \frac{27364}{609}K - \frac{451183}{14616} - \frac{3191}{7308}I + \frac{22958}{1827}J\right] \quad (1.13)$$

```
> is(add(i, i in C)=A);
```

$$(1.14)$$

$$true \quad (1.14)$$

```
> AXXB(1-2*I-2*J-2*K,1+I-3*J+K,-1-3*I+J-2*K);
```

$$\{-2.408361978 + 0.8684696550\,I + 1.654437636\,J - 1.925435601\,K,\ 0.4083619780 + 0.7998322759\,I + 0.004052709407\,J + 1.143297118\,K\} \quad (1.15)$$

```
> AXBc(-I-3*J+2*K,-2*I+J-K,21);
```

$$\{-6.390863387 - 0.4049533666\,I + 2.024766833\,J + 2.834673566\,K,\ -1.075576667 - 0.2088030771\,I + 1.044015386\,J + 1.461621540\,K,\ 1.075576667 + 1.267103601\,I - 6.335518005\,J - 8.869725207\,K,\ 6.390863387 - 0.6533471582\,I + 3.266735791\,J + 4.573430107\,K\} \quad (1.16)$$

```
> AXB(-19+15*I-J+11*K,9+10*I-4*J+10*K);
```

$$\{443.0282687 + 62.30337605\,I - 7.351923515\,J + 100.3304128\,K\} \quad (1.17)$$

```
> subs(t[2]=2,t[3]=3, SemiSim(2+4*I-3*J+13*K,2+5*I+12*J-5*K));
```

$$24\,I + 2\,J + 3\,K - 11 \ \textbf{and}\ -\frac{11}{710} - \frac{12}{355}\,I - \frac{1}{355}\,J - \frac{3}{710}\,K \quad (1.18)$$

```
> subs(t[1]=1,t[3]=3, CSSim(17+I-24*J-7*K,1-25*I+8*J+15*K));
```

$$-\frac{19}{2} + I + 13\,J + 3\,K \ \textbf{and}\ \frac{2}{1077} + \frac{56}{1077}\,I - \frac{4}{359}\,J - \frac{32}{1077}\,K \quad (1.19)$$

```
> LN1(5-10*I-5*J+2*K,3-4*I-4*J-8*K,-9-2*I+10*J-2*K);
```

$$-\frac{3364}{2905} + \frac{128}{415}\,I - \frac{1073}{2905}\,J + \frac{2372}{2905}\,K \quad (1.20)$$

```
> LN2(6-8*I+J+5*K,6+I+5*J-8*K,-3+I+J-5*K)[1];
```

$$-\frac{39}{205} - \frac{119}{7380}\,I + \frac{1133}{8610}\,J - \frac{2519}{17220}\,K \quad (1.21)$$

```
> final := Matrix(4, 4, {(1, 1) = 14*x+14+76*I+70*J+56*K, (1,
  2) = 56-28*I-70*J+70*K, (1, 3) = 28*J-56*K, (1, 4) = 14*x-56
  -8*I-14*J-56*K, (2, 1) = -2*x-2-43*I-10*J-8*K, (2, 2) = -8+4*
  I+10*J-10*K, (2, 3) = -4*J+8*K, (2, 4) = -2*x+8-31*I+2*J+8*K,
  (3, 1) = -3*x-3+3*I-15*J-12*K, (3, 2) = -12+6*I+15*J-15*K,
  (3, 3) = -6*J+12*K, (3, 4) = -3*x+12+21*I+3*J+12*K, (4, 1) =
  -4*x-4+4*I-20*J-16*K, (4, 2) = -16+8*I+20*J-20*K, (4, 3) =
  -8*J+16*K, (4, 4) = -4*x+16+28*I+4*J+16*K});
```

$$final := [[14\,x + 14 + 76\,I + 70\,J + 56\,K,\ 56 - 28\,I - 70\,J + 70\,K,\ 28\,J - 56\,K,\ 14\,x - 56 - 8\,I - 14\,J - 56\,K],$$
$$[-2\,x - 2 - 43\,I - 10\,J - 8\,K,\ -8 + 4\,I + 10\,J - 10\,K,\ -4\,J + 8\,K,\ -2\,x + 8 - 31\,I + 2\,J + 8\,K],$$
$$[-3\,x - 3 + 3\,I - 15\,J - 12\,K,\ -12 + 6\,I + 15\,J - 15\,K,\ -6\,J + 12\,K,\ -3\,x + 12 + 21\,I + 3\,J + 12\,K],$$
$$[-4\,x - 4 + 4\,I - 20\,J - 16\,K,\ -16 + 8\,I + 20\,J - 20\,K,\ -8\,J + 16\,K,\ -4\,x + 16 + 28\,I + 4\,J + 16\,K]] \quad (1.22)$$

```
> FI:=LFI(final):
> final*FI*final-final;
```

(1.23)

$$\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \tag{1.23}$$

```
> FI*final*FI-FI;
```

$$\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \tag{1.24}$$

```
> Transpose(Mconj(FI*final))-FI*final;
```

$$\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \tag{1.25}$$

```
> Transpose(Mconj(final*FI))-final*FI;
```

$$\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \tag{1.26}$$

Bibliography

Adler, S., 1995. Quaternionic quantum mechanics and quantum fields. International Series of Monographs on Physics. Oxford University Press, USA.

Au-Yeung, Y., 2003. An explicit solution for the quaternionic equation $x^2 + bx + xc + d = 0$. Southeast Asian Bull. Math. 26 (5), 717–724.

Baker, A., 1999. Right eigenvalues for quaternionic matrices: a topological approach. Linear Algebra Appl. 286 (1-3), 303–309.

Bjõrck, A., 1996. Numerical methods for least squares problems. Society for Industrial and Applied Mathematics.

Bray, U., Whaples, G., 1983. Polynomials with coefficients from a division ring. Canad. J. Math. 35 (3), 509–515.

Brenner, J. L., 1951. Matrices of quaternions. Pacific J. Math. 1, 329–335.

Carter, M., March 2007. Quaternions. Maple package. Maplesoft.
URL `http://www.maplesoft.com/applications/view.aspx?SID=4886`

Cayley, A., 1894. On coordinates versus quaternions. Proc. Roy. Soc. Edinburgh 20, 271–275.

Churchill, L., 1990. Quaternion analysis in physics. East Carolina University.

Decell, H., 1965. An application of the Cayley-Hamilton theorem to generalized matrix inversion. SIAM Rev. 7, 526–528.

Derbyshire, J., 2006. Unknown quantity: a real and imaginary history of algebra. Joseph Henry Press.

Eilenberg, S., Niven, I., 1944. The "fundamental theorem of algebra" for quaternions. Bull. Amer. Math. Soc. 50, 246–248.

Faddeeva, V., 1959. Computational methods of linear algebra. Dover books on advanced mathematics. Dover Publications.

Finkelstein, D., Jauch, J., Schiminovich, S., Speiser, D., 1962. Foundations of quaternion quantum mechanics. Journal of Mathematical Physics 3 (2), 207–220.

Gibbs, J., Bumstead, H., Van Name, R., Longley, W., 1902. The collected works of J. Willard Gibbs. No. v. 2 in The Collected Works of J. Willard Gibbs. Longman, Green and Co.

Gordon, B., Motzkin, T. S., 1965. On the zeros of polynomials over division rings. Trans. Amer. Math. Soc. 116, 218–226.

Hanson, A., 2005. Visualizing quaternions. The Morgan Kaufmann Series In Interactive 3D Technology. Morgan Kaufmann.

Harder, D., September 2010. Quaternions. Maple package. Maplesoft.
URL `http://www.maplesoft.com/applications/view.aspx?SID=96897`

Hastings, C., 2010. Biographical Memoir of Josiah Willard Gibbs, 1839-1903: By Charles S. Hastings. BiblioBazaar.

Horn, R., Johnson, C., 1990. Matrix analysis. Cambridge University Press.

Huang, L., So, W., 2002. Quadratic formulas for quaternions. Appl. Math. Lett. 15 (5), 533–540.

Johnson, R. E., 1944. On the equation $\chi\alpha = \gamma\chi + \beta$ over an algebraic division ring. Bull. Amer. Math. Soc. 50, 202–207.

Kalman, D., 1978. An extension of the remainder theorem. Delta 8, 77–80.

Karampetakis, N. P., 1997. Computation of the generalized inverse of a polynomial matrix and applications. Linear Algebra Appl. 252, 35–60.

Kuipers, J., 2002. Quaternions and rotation sequences: a primer with applications to orbits, aerospace and virtual reality. Princeton University Press.

Lam, T., 2001. A first course in noncommutative rings. Graduate Texts in Mathematics (131). Springer.

Lee, H. C., 1949. Eigenvalues and canonical forms of matrices with quaternion coefficients. Proc. Roy. Irish Acad. Sect. A. 52, 253–260.

Maxwell, J., 1873. Quaternions. Nature 9, 137–138.

McAulay, A., 1892. Quaternions as a practical instrument of physical research. Philosophical Magazine 33, 477–495.

Morandi, P., 1996. Field and Galois theory. Graduate Texts in Mathematics. Springer.

Muir, T., 2003. A treatise on the theory of determinants. Dover Phoenix Editions. Dover Publications.

Nahin, P., 2002. Oliver Heaviside: the life, work, and times of an electrical genius of the Victorian age. Oliver Heaviside. Johns Hopkins University Press.

Niven, I., 1941. Equations in quaternions. Amer. Math. Monthly 48, 654–661.

Niven, I., 1942. The roots of a quaternion. Amer. Math. Monthly 49, 386–388.

Penrose, R., 1955. A generalized inverse for matrices. Proc. Cambridge Philos. Soc. 51, 406–413.

Puystjens, R., de Smet, H., 1980. The Moore-Penrose inverse for matrices over skew polynomial rings. Ring Theory Antwerp, Lecture Notes in Mathematics 825, 94–103.

Smith, C., Wise, N., 1989. Energy and empire: a biographical study of Lord Kelvin. Cambridge University Press.

Stanimirović, P., Petković, M., 2006. Computing generalized inverse of polynomial matrices by interpolation. Appl. Math. Comput. 172 (1), 508–523.

Stanimirović, P., Tasić, M., Krtolica, P., Karampetakis, N., 2007. Generalized inversion by interpolation. Filomat 21 (1), 67–86.

Tait, P., 1890. Vii. On the importance of quaternions in physics. Philosophical Magazine Series 5 29 (176), 84–97.

Tian, Y., 2004. The equations $ax - xb = c$, $ax - \overline{x}b = c$ and $\overline{x}ax = b$ in quaternions. Southeast Asian Bull. Math. 28 (2), 343–362.

Tian, Y., 2010. Solving two pairs of quadratic equations in quaternions. Adv. Appl. Clifford Algebr. 20 (1), 185–193.

van der Waerden, B., 1985. A history of algebra: from al-Khwārizmī to Emmy Noether. Springer.

Zhang, F., 1997. Quaternions and matrices of quaternions. Linear Algebra and its Applications 251 (0), 21–57.