Newton Method in the Context of Quaternion Analysis

M.I. Falcão

Departamento de Matemática e Aplicações and CMAT
Universidade do Minho, Portugal

Information

Abstract

In this paper we propose a version of Newton method for finding zeros of a quaternion function of a quaternion variable, based on the concept of quaternion radial derivative. Several numerical examples involving elementary functions are presented.

1   Introduction

Since 1928 R. Fueter, one of the founders of quaternion analysis ([1, 2]), tried to develop a function theory to generalize the theory of holomorphic functions of one complex variable, by considering quaternion valued functions of a quaternion variable. Nowadays, this well known and developed theory is recognized as a powerful tool for modeling and solving problems in both theoretical and applied mathematics. For a survey on quaternion analysis and a list of references we refer to the book [3]; a historical perspective of the subject and several applications can be found in [4].

In this work we revisit the classical Newton method for finding roots (or zeros) of a complex function and propose a quaternion analogue, based on the concept and properties of quaternion radially holomorphic functions. We show that for a certain class of functions (including simple polynomials and other elementary functions) this method produces the same sequence as the classical Newton method for vector valued functions. In this way, we can obtain, with less computational effort, local quadratic convergence for a class of quaternion functions.

This idea was already considered by Janovská and Opfer in [5], where the authors formally adapted, for the first time, Newton method for finding roots of Hamilton quaternions, by considering the quaternion equation $x^n - a = 0$. More recently, Kalantari in [6], using algebraic-combinatorial arguments, proposed a Newton method for finding roots of special quaternion polynomials. Working in the framework of quaternion analysis, we can provide a motivation for the techniques used in those works and simultaneously extend the applicability of the method.

The paper is organized as follows. In Section 2 we introduce the basic notations and the results that are needed for our work in Section 3.

Section 3 contains the main results of the paper. Here, by making use of the theory given in Section 2, and after establishing new properties on the radial derivative of a special class of functions, we propose a Newton method in the framework of quaternion analysis.

Finally, in Section 4 several numerical examples illustrating the applicability of the aforementioned methods are presented.

2 Quaternion Analysis

We start by first recalling some basic results concerning Hamilton quaternion algebra $\mathbb{H}$, which can be found in classic books on this subject. For results concerning quaternion analysis we refer to [7, 8, 3].

Let $\{1, \mathbf{i}, \mathbf{j}, \mathbf{k}\}$ be an orthonormal basis of the Euclidean vector space $\mathbb{R}^4$ with a product given according to the multiplication rules

$$\mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = -1, \ \ \mathbf{ij} = -\mathbf{ji} = \mathbf{k}.$$

This non-commutative product generates the well known algebra of real quaternions $\mathbb{H}$. The real vector space $\mathbb{R}^4$ will be embedded in $\mathbb{H}$ by identifying the element $x = (x_0, x_1, x_2, x_3) \in \mathbb{R}^4$ (or the column vector in $\mathbb{R}^{4 \times 1}$, $x = (x_0 \ x_1 \ x_2 \ x_3)^T$) with the element $x = x_0 + x_1 \mathbf{i} + x_2 \mathbf{j} + x_3 \mathbf{k} \in \mathbb{H}$. Throughout this paper, we will also use the symbol $\boldsymbol{x}$ to represent an element in $\mathbb{R}^4$, whenever we need to distinguish the structure of $\mathbb{H}$ from $\mathbb{R}^4$.

The real or *scalar part* of a quaternion $x = x_0 + x_1 \mathbf{i} + x_2 \mathbf{j} + x_3 \mathbf{k}$ is denoted by $\mathrm{Sc}\, x$ and is equal to x_0, the *vector part* of x is $\underline{x} := x_1 \mathbf{i} + x_2 \mathbf{j} + x_3 \mathbf{k}$ and therefore x can be written as $x = x_0 + \underline{x}$. The *conjugate* of x is $\overline{x} := x_0 - x_1 \mathbf{i} - x_2 \mathbf{j} - x_3 \mathbf{k} = x_0 - \underline{x}$. The mapping $x \mapsto \overline{x}$ is called *conjugation* and has the property $\overline{xy} = \overline{y}\,\overline{x}$, for all $x, y \in \mathbb{H}$. The norm $|x|$ of x is defined by $|x|^2 = x\overline{x} = \overline{x}x$ and coincides with the corresponding Euclidean norm of $\boldsymbol{x}$, as a vector in $\mathbb{R}^4$. It follows that $|xy| = |x||y|$ and each non-zero $x \in \mathbb{H}$ has an inverse given by $x^{-1} = \frac{\overline{x}}{|x|^2}$. Moreover, $|x|^{-1} = |x^{-1}|$.

In this work we are going to consider also the representation of the quaternion $x = x_0 + x_1 \mathbf{i} + x_2 \mathbf{j} + x_3 \mathbf{k}$ by means of the real matrix in $\mathbb{R}^{4 \times 4}$

$$L_x := \begin{pmatrix} x_0 & -x_1 & -x_2 & -x_3 \\ x_1 & x_0 & -x_3 & x_2 \\ x_2 & x_3 & x_0 & -x_1 \\ x_3 & -x_2 & x_1 & x_0 \end{pmatrix}. \tag{1}$$

This representation is called *matrix left representation* of x and can be associated with the product of quaternions

$$
\begin{aligned}
xy &= (x_0 + x_1\mathbf{i} + x_2\mathbf{j} + x_3\mathbf{k})(y_0 + y_1\mathbf{i} + y_2\mathbf{j} + y_3\mathbf{k}) \\
&= (x_0y_0 - x_1y_1 - x_2y_2 - x_3y_3) + (x_1y_0 + x_0y_1 - x_3y_2 + x_2y_3)\mathbf{i} \\
&= (x_2y_0 + x_3y_1 + x_0y_2 - x_1y_3)\mathbf{j} + (x_3y_0 - x_2y_1 + x_1y_2 + x_0y_3)\mathbf{k}, \quad (2)
\end{aligned}
$$

through the identification

$$
z = xy \quad \longrightarrow \quad \boldsymbol{z} = L_x\boldsymbol{y}, \tag{3}
$$

where $\boldsymbol{y}$ is the (column) vector in $\mathbb{R}^4$ corresponding to the quaternion y.

Any arbitrary nonreal quaternion x can also be written in the so-called *complex-like* form

$$
x = x_0 + \omega(\underline{x})|\underline{x}|, \tag{4}
$$

where

$$
\omega(\underline{x}) := \frac{\underline{x}}{|\underline{x}|}
$$

belongs to the unit sphere in $\mathbb{R}^3$. Since $\omega(\overline{x}) = \overline{\omega(\underline{x})} = -\omega(\underline{x})$, it follows immediately that $\omega(\underline{x})^2 = -\omega(\underline{x})\overline{\omega(\underline{x})} = -|\omega(\underline{x})|^2 = -1$. In other words, we can consider that ω behaves like the imaginary unit and therefore the complex-like form (4) is similar to the complex form $a + ib$. We use the convention $\omega(\underline{x}) := 0$, for real quaternions x. The following properties play an important role in the present work.

Proposition 1 *If $x = x_0 + x_1\mathbf{i} + x_2\mathbf{j} + x_3\mathbf{k}$ and $y = y_0 + y_1\mathbf{i} + y_2\mathbf{j} + y_3\mathbf{k}$ are quaternions, the following statements are equivalent:*

(i) $xy = yx$;

(ii) $x_1y_2 = x_2y_1$ *and* $x_1y_3 = x_3y_1$ *and* $x_2y_3 = x_3y_2$;

(iii) $\omega(\underline{x})\omega(\underline{y}) = \omega(\underline{y})\omega(\underline{x})$;

(iv) $\omega(\underline{x}) = \pm\omega(\underline{y})$.

Proof: The equivalence between the first three statements follows at once from the multiplication rules (cf. (2)). If $\omega(\underline{x}) = \pm\omega(\underline{y})$, clearly $\omega(\underline{x})\omega(\underline{y}) = \mp 1 = \omega(\underline{y})\omega(\underline{x})$ and we conclude that (iv) implies (iii).

Now, we prove that (ii) implies (iv) for the case of nonreal quaternions. If y is nonreal, we can assume that, for example, $y_1 \neq 0$. Using (ii) one can write $x_2 = x_1\frac{y_2}{y_1}$ and $x_3 = x_1\frac{y_3}{y_1}$ and this, in turn, implies that $\underline{x} = x_1\mathbf{i} + x_2\mathbf{j} + x_3\mathbf{k} = (y_1\mathbf{i} + y_2\mathbf{j} + y_3\mathbf{k})\frac{x_1}{y_1}$. This means that $|\underline{x}| = |\underline{y}|\frac{|x_1|}{|y_1|}$ and the result $\omega(\underline{x}) = \frac{\underline{x}}{|\underline{x}|} = \frac{\underline{y}}{|\underline{y}|}\frac{x_1|y_1|}{|x_1|y_1} = \pm\omega(\underline{y})$ follows. $\qquad\square$

Corollary 1 *If x and y are quaternions such that $\omega(\underline{x})$ and $\omega(\underline{y})$ commute then:*

(i) $x\overline{y} = \overline{y}x$;

(ii) $xy^{-1} = y^{-1}x$.

Proof: It is sufficient to note that $\omega(\underline{y}^{-1}) = \omega(\overline{y}) = -\omega(\underline{y})$ and apply the results of Proposition 1. $\qquad\square$

In what follows, we consider complex-like functions $f : \Omega \subset \mathbb{R}^4 \to \mathbb{H} \cong \mathbb{R}^4$ of the form

$$f(x) = u(x_0, |\underline{x}|) + \omega(\underline{x})v(x_0, |\underline{x}|), \qquad (5)$$

where $\underline{x} \neq 0$ and u and v are real valued functions. Continuity, differentiability or integrability are defined coordinate-wisely.

R. Fueter proposed a generalization of complex analyticity to the quaternion case ([1, 2]) which leads to close analogues of several important results from classical complex function theory. In this framework the analogue of holomorphic functions, usually refer to as *monogenic functions*, is obtained as the set of null-solutions of a generalized Cauchy-Riemann system. In addition, Sudbery, in [9], defined a *regular* quaternion valued function by the existence of its quaternion derivative. Unfortunately, contrary to the complex case, neither the canonical quaternion variable x nor any of its nonnegative integer powers x^n are monogenic functions and therefore, the quaternion derivative of x^n does not exist. This means, among other things, that based on the concept of monogenicity one can not generalize Newton method to finding roots of simple quaternion polynomials, i.e. polynomials with quaternion coefficients located on only one side of the powers.

In order to introduce a notion of *regular function* that fits our purpose, we follow [8, 3, 10] and adopt the following usual definition.

Definition 1 Let f be a complex-like function of the form (5) and let $h = h_0 + \omega(\underline{x})h_r \in \mathbb{H}$. Such function f is called *radially holomorphic* or *radially regular* in Ω if

$$\lim_{h \to 0} (f(x + h) - f(x)) h^{-1}$$

exists. In the case of existence, this limit is called the *radial derivative* of f at x and is denoted by $f'(x)$. (This notation will be justified later).

Defining on the set $\mathscr{C}^1(\Omega, \mathbb{H})$ the *radial operators*

$$\partial_{\mathrm{rad}} := \tfrac{1}{2}(\partial_0 - \omega(\underline{x})\partial_r), \qquad \bar{\partial}_{\mathrm{rad}} := \tfrac{1}{2}(\partial_0 + \omega(\underline{x})\partial_r), \qquad (6)$$

where $r := |\underline{x}|$, $\partial_0 := \frac{\partial}{\partial x_0}$ and $\partial_r := \frac{\partial}{\partial r}$, one can prove the following essential results.

Theorem 1 ([10, Corollary 5.2]) *If f is a radially holomorphic function, then*

$$f'(x) = \partial_{rad}f(x).$$

Theorem 2 ([3, Corollary 11.30]) *A function f of the form (5) is radially holomorphic if and only if*

$$\bar{\partial}_{rad}f = 0,$$

which is a Cauchy-Riemann type differential equation

$$\begin{cases} \partial_0 u & = \partial_r v, \\ \partial_0 v & = -\partial_r u. \end{cases}$$

4

We note that when $\bar{\partial}_{\mathrm{rad}} f = 0$, the first radial operator in (6) simplifies to $\partial_{\mathrm{rad}} f = \partial_0 f$ and Theorem 1 together with Theorem 2 imply that if f is a radially holomorphic function, then

$$f'(x) = \partial_0 f(x) = \partial_0 u(x_0, r) + \omega(\underline{x}) \partial_0 v(x_0, r), \tag{7}$$

like in the complex case, which justifies the use of the same notation for the complex derivative of a complex holomorphic function and the radial derivative of a quaternion radially holomorphic function. Next straightforward result supports these ideas.

Proposition 2 *If f and g are radially holomorphic functions of the form*

$$f(x) = u(x_0, |\underline{x}|) + \omega(\underline{x}) v(x_0, |\underline{x}|) \quad \text{and} \quad g(x) = u^*(x_0, |\underline{x}|) + \omega(\underline{x}) v^*(x_0, |\underline{x}|),$$

where u, v, u^* and v^* are real valued functions, then

1. *$f + g$ is radially holomorphic and $(f + g)'(x) = f'(x) + g'(x)$;*

2. *if $\alpha \in \mathbb{R}$, then αf is radially holomorphic and $(\alpha f)'(x) = \alpha f'(x)$;*

3. *fg is radially holomorphic and $(fg)'(x) = f'(x)g(x) + g'(x)f(x)$.*

Example 1 The following functions are radially holomorphic functions:

- monomials of the form x^α, $\alpha \in \mathbb{R}$;

- the exponential function $e^x := \sum_{k=0}^{\infty} \frac{x^k}{k!} = e^{x_0}(\cos r + \omega(\underline{x}) \sin r)$.

In addition, $(x^\alpha)' = \alpha x^{\alpha-1}$ and $(e^x)' = e^x$.

We point out that usually when $\alpha \notin \mathbb{R}$, αf can not be written in the complex-like form (5), since, in general, $\alpha f(x) \neq f(x)\alpha$. For more details about radially holomorphic functions as well as other examples, see e.g. [8, 3, 10]. For a recent survey on elementary functions in the context of quaternion analysis, see [11].

3 Newton method: from $\mathbb{C}$ to $\mathbb{H}$

The well known Newton method for finding a zero x^* of a holomorphic complex function f of one complex variable, consists on approximating x^* by means of the iterative process

$$P_1: \ z_{k+1} = z_k - \frac{f(z_k)}{f'(z_k)}, \ k = 0, 1, \ldots, \tag{8}$$

with z_0 sufficiently close to x^* and $f'(z_k) \neq 0$.

Writing $f(z) = f(x + iy) = u(x, y) + iv(x, y)$, where u and v are real valued functions and identifying the complex number $z = x + iy \in \mathbb{C}$ with the vector $\boldsymbol{z} = (x, y) \in \mathbb{R}^2$, we can interpret $f(z) = 0$ as a system of two equations in two real unknowns, namely $\boldsymbol{f}(\boldsymbol{z}) = (u(\boldsymbol{z}), v(\boldsymbol{z})) = (0, 0)$ and apply the classical 2D-Newton method

$$P_2: \ \boldsymbol{z}_{k+1} = \boldsymbol{z}_k - (J\boldsymbol{f}(\boldsymbol{z}_k))^{-1} f(\boldsymbol{z}_k), \ k = 0, 1, \ldots, \tag{9}$$

provided that the Jacobian matrix $J\boldsymbol{f}(\boldsymbol{z}_k)$ is nonsingular.

We recall that, if f is holomorphic, due to the Cauchy-Riemann equations, we can write

$$J\boldsymbol{f} = \begin{pmatrix} u_x & -v_x \\ v_x & u_x \end{pmatrix} \qquad \text{and} \qquad J\boldsymbol{f}^{-1} = \frac{1}{u_x^2 + v_x^2} \begin{pmatrix} u_x & v_x \\ -v_x & u_x \end{pmatrix},$$

where $u_x := \partial u/\partial x$, $u_y := \partial u/\partial y$, etc. Moreover, the determinant of the Jacobian matrix satisfies

$$|J\boldsymbol{f}(\boldsymbol{x})| = |f'(x)|^2. \tag{10}$$

Therefore, we can conclude, after simplifications, the well known fact that, in such cases, the 2D-Newton method (9) is identical to the complex Newton method (8), i.e. P_1 and P_2 produce the same sequence.

The problem concerning finding zeros of a quaternion function is, as expected, more difficult than the root finding problem in the complex plane. In contrast to the complex case, the zeros of a quaternion function are not necessarily isolated, and its range is not necessarily open (see e.g. [9]). Since the well known work of Niven [12], several authors gave contributions to this subject, in particular in connection to the study of zeros of quaternion polynomials [13, 14, 5, 15, 16, 17] (see also [18, 19] and [6, Section 4] and the references therein).

In this section, we are going to show that the use of the radial derivative leads to an iterative process analogue to the usual 4D-Newton process which can be extended to a class of functions *wider* than the class of radially holomorphic functions.

Let f be a radially holomorphic complex-like function f of the form (5), i.e.

$$f(x) = f(x_0 + \underline{x}) = u(x_0, r) + \omega(\underline{x})v(x_0, r),$$

where $r = |\underline{x}| = \sqrt{x_1^2 + x_2^2 + x_3^2}$, with radial derivative given by (7), i.e.

$$f'(x) = \partial_0 u(x_0, r) + \omega(\underline{x})\partial_0 v(x_0, r).$$

Consider an iterative process of the form

$$P_1: \qquad z_{k+1} = z_k - f(z_k)(f'(z_k))^{-1}, \ k = 0, 1, \ldots,$$
$$z_0 = c,$$

where $c \in \mathbb{H}$ and $f'(z_k) \neq 0$, for all $k = 0, 1, \ldots$. Observe that due to the structure of f, this process produces exactly the same sequence as

$$P_2: \qquad \tilde{z}_{k+1} = \tilde{z}_k - (f'(\tilde{z}_k))^{-1}f(\tilde{z}_k), \ k = 0, 1, \ldots,$$
$$\tilde{z}_0 = c,$$

with $f'(\tilde{z}_k) \neq 0$, since the use of Proposition 1 and Corollary 1 allows to conclude that f and $(f')^{-1}$ commute. Our objective now is to compare these two equivalent processes with the classical Newton method for vector valued functions

$$P_3: \qquad \boldsymbol{z}_{k+1} = \boldsymbol{z}_k - (J\boldsymbol{f}(\boldsymbol{z}_k))^{-1}\boldsymbol{f}(\boldsymbol{z}_k), \ k = 0, 1, \ldots,$$
$$\boldsymbol{z}_0 = \boldsymbol{c},$$

where $\boldsymbol{c} \in \mathbb{R}^4$ is the vector corresponding to the quaternion c in processes P_1 and P_2, $\boldsymbol{f}$ is the vector valued function associated with the quaternion valued function f, i.e.

$$\boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{f}(x_0, x_1, x_2, x_3) = \left(u(x_0, r), \frac{x_1}{r} v(x_0, r), \frac{x_2}{r} v(x_0, r), \frac{x_3}{r} v(x_0, r) \right) \quad (11)$$

and $J\boldsymbol{f}(\boldsymbol{z}_k)$ is the Jacobian matrix of $\boldsymbol{f}$ at $\boldsymbol{z}_k$ which we assume is nonsingular.

Next results reveal new important and interesting relationships between the radial derivative of a function f and the derivative of $\boldsymbol{f}$.

Proposition 3 *Let $\boldsymbol{f}$ be the vector valued function* (11) *associated with a radially holomorphic complex-like function f of the form* (5). *The Jacobian determinant of $\boldsymbol{f}$ is related to the derivative of f by means of*

$$|J\boldsymbol{f}(\boldsymbol{x})| = |f'(x)|^2 \frac{v(x_0, r)^2}{r^2}. \quad (12)$$

Proof: The Jacobian matrix of $\boldsymbol{f}$ can be written as

$$J\boldsymbol{f}(\boldsymbol{x}) = \begin{pmatrix} A & -B(x_1) & -B(x_2) & -B(x_3) \\ B(x_1) & D(x_1, x_2, x_3) & C(x_1, x_2) & C(x_1, x_3) \\ B(x_2) & C(x_1, x_2) & D(x_2, x_3, x_1) & C(x_2, x_3) \\ B(x_3) & C(x_1, x_3) & C(x_2, x_3) & D(x_3, x_1, x_2) \end{pmatrix},$$

where

$$A = \partial_0 u(x_0, r);$$
$$B(x) = \frac{x}{r} \, \partial_0 v(x_0, r);$$
$$C(x, y) = \frac{xy}{r^2} \, \partial_0 u(x_0, r) - \frac{xy}{r^3} \, v(x_0, r);$$
$$D(x, y, z) = \frac{x^2}{r^2} \, \partial_0 u(x_0, r) + \frac{y^2 + z^2}{r^3} \, v(x_0, r).$$

The result follows by cumbersome, but straightforward calculations. $\qquad \square$

Remark 1 *We underline the fact that, when $x \in \mathbb{H}$, $J\boldsymbol{f}(\boldsymbol{x})$ can be singular while $f'(x) \neq 0$. This situation is quite different from the complex case where the equations $|J\boldsymbol{f}(\boldsymbol{x})| = 0$ and $f'(x) = 0$ are equivalent, for $x \in \mathbb{C}$ (cf.* (10)).

Proposition 4 *If f is a radially holomorphic quaternion valued function and $a \in \mathbb{H}$ is such that $\omega(\underline{x})$ commutes with $\omega(\underline{a})$ then*

$$J\boldsymbol{f}(\boldsymbol{x})\boldsymbol{a} = af'(x) = f'(x)a. \quad (13)$$

Proof: If $a = a_0 + a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}$, we can use Proposition 3 to write

$$J\boldsymbol{f}(\boldsymbol{x})\boldsymbol{a} = \begin{pmatrix} a_0 \partial_0 u - \frac{a_1 x_1 + a_2 x_2 + a_3 x_3}{r} \partial_0 v \\ \frac{a_1(x_2^2 + x_3^2) - a_2 x_1 x_2 - a_3 x_1 x_3}{r^3} v + \frac{a_1 x_1^2 + a_2 x_1 x_2 + a_3 x_1 x_3}{r^2} \partial_0 u + \frac{a_0 x_1}{r} \partial_0 v \\ \frac{a_2(x_1^2 + x_3^2) - a_1 x_1 x_2 - a_3 x_2 x_3}{r^3} v + \frac{a_2 x_2^2 + a_1 x_1 x_2 + a_3 x_2 x_3}{r^2} \partial_0 u + \frac{a_0 x_2}{r} \partial_0 v \\ \frac{a_3(x_2^2 + x_1^2) - a_2 x_3 x_2 - a_1 x_1 x_3}{r^3} v + \frac{a_3 x_3^2 + a_2 x_3 x_2 + a_1 x_1 x_3}{r^2} \partial_0 u + \frac{a_0 x_3}{r} \partial_0 v \end{pmatrix},$$

where all functions are considered at (x_0, r). Since $\omega(\underline{x})$ commutes with $\omega(\underline{a})$, we have from Proposition 1 that

$$x_1 a_2 = x_2 a_1, \qquad x_1 a_3 = x_3 a_1, \qquad x_2 a_3 = x_3 a_3 \tag{14}$$

and therefore

$$J\boldsymbol{f}(\boldsymbol{x})\boldsymbol{a} = \begin{pmatrix} a_0 \partial_0 u - \frac{a_1 x_1 + a_2 x_2 + a_3 x_3}{r} \partial_0 v \\ a_1 \partial_0 u + \frac{a_0 x_1}{r} \partial_0 v \\ a_2 \partial_0 u + \frac{a_0 x_2}{r} \partial_0 v \\ a_3 \partial_0 u + \frac{a_0 x_3}{r} \partial_0 v \end{pmatrix}.$$

On the other hand, from (1) and (3), we have

$$af'(x) = \begin{pmatrix} a_0 & -a_1 & -a_2 & -a_3 \\ a_1 & a_0 & -a_3 & a_2 \\ a_2 & a_3 & a_0 & -a_1 \\ a_3 & -a_2 & a_1 & a_0 \end{pmatrix} \begin{pmatrix} \partial_0 u \\ \frac{x_1}{r} \partial_0 v \\ \frac{x_2}{r} \partial_0 v \\ \frac{x_3}{r} \partial_0 v \end{pmatrix}$$

$$= \begin{pmatrix} a_0 \partial_0 u - \frac{a_1 x_1 + a_2 x_2 + a_3 x_3}{r} \partial_0 v \\ a_1 \partial_0 u + \frac{a_0 x_1 - a_3 x_2 + a_2 x_3}{r} \partial_0 v \\ a_2 \partial_0 u + \frac{a_3 x_1 + a_0 x_2 - a_1 x_3}{r} \partial_0 v \\ a_3 \partial_0 u + \frac{-a_2 x_1 + a_1 x_2 + a_0 x_3}{r} \partial_0 v \end{pmatrix}$$

and the first part of the result is proved, using (14). The second part of (13) comes once more from Proposition 1. $\square$

We can also establish the link between the Jacobian matrix and the corresponding determinant of the quaternion functions f and $\alpha f + \beta$, $\alpha, \beta \in \mathbb{H}$.

Proposition 5 *If f and g are quaternion valued functions defined on the set $\mathscr{C}^1(\Omega, \mathbb{H})$ such that $g(x) = \alpha f(x) + \beta$, $\alpha, \beta \in \mathbb{H}$ and $\alpha \neq 0$, then*

$$J\boldsymbol{g}(\boldsymbol{x}) = L_\alpha J\boldsymbol{f}(\boldsymbol{x}), \tag{15}$$

where L_α is the matrix left representation (1) of the quaternion α. Therefore

$$|J\boldsymbol{g}(\boldsymbol{x})| = |\alpha|^4 |J\boldsymbol{f}(\boldsymbol{x})|. \tag{16}$$

Proof: Writing $f(x) = f_0(x) + f_1(x)\mathbf{i} + f_2(x)\mathbf{j} + f_3(x)\mathbf{k}$, $\alpha = \alpha_0 + \alpha_1\mathbf{i} + \alpha_2\mathbf{j} + \alpha_3\mathbf{k}$ and $\beta = \beta_0 + \beta_1\mathbf{i} + \beta_2\mathbf{j} + \beta_3\mathbf{k}$ and recalling (1) it follows that

$$\boldsymbol{g}(\boldsymbol{x}) = L_\alpha \boldsymbol{f}(\boldsymbol{x}) + \boldsymbol{\beta} = \begin{pmatrix} \alpha_0 f_0(x) - \alpha_1 f_1(x) - \alpha_2 f_2(x) - \alpha_3 f_3(x) + \beta_0 \\ \alpha_1 f_0(x) + \alpha_0 f_1(x) - \alpha_3 f_2(x) + \alpha_2 f_3(x) + \beta_1 \\ \alpha_2 f_0(x) + \alpha_3 f_1(x) + \alpha_0 f_2(x) - \alpha_1 f_3(x) + \beta_2 \\ \alpha_3 f_0(x) - \alpha_2 f_1(x) + \alpha_1 f_2(x) + \alpha_0 f_3(x) + \beta_3 \end{pmatrix}.$$

Straightforward calculations lead to result (15). Result (16) follows from the fact that $\det L_\alpha = |\alpha|^4$. $\square$

We are now in position to prove one of our main results.

Theorem 3 *Let* $g(x) = \alpha f(x) + \beta$ *be a function defined on the set* $\mathscr{C}^1(\Omega, \mathbb{H})$ *such that f is a radially holomorphic function in* Ω, $\alpha, \beta \in \mathbb{H}$ *and* $\alpha \neq 0$. *Formulas*

$$
\begin{array}{lll}
z_{k+1} = \mathcal{P}_1(z_k) & \tilde{z}_{k+1} = \mathcal{P}_2(\tilde{z}_k) & \boldsymbol{z}_{k+1} = \mathcal{P}_3(\boldsymbol{z}_k) \\
z_0 = c & \tilde{z}_0 = c & \boldsymbol{z}_0 = \boldsymbol{c}
\end{array}
\qquad (17)
$$

where $c \in \mathbb{H}$ and $\mathcal{P}_i$, $i = 1, 2, 3$ *are the iterative functions*

$$
\mathcal{P}_1(z) = z - g(z)(\alpha f'(z))^{-1}, \qquad (18)
$$

$$
\mathcal{P}_2(z) = z - (\alpha f'(z))^{-1} g(z), \qquad (19)
$$

$$
\mathcal{P}_3(\boldsymbol{z}) = \boldsymbol{z} - (J\boldsymbol{g}(\boldsymbol{z}))^{-1}\boldsymbol{g}(\boldsymbol{z}) \qquad (20)
$$

produce the same sequence $(k = 0, 1, \dots)$, if $\omega(\underline{c})$ commutes with $\omega(\underline{\alpha})$ and $\omega(\underline{\beta})$ and $J\boldsymbol{f}(\boldsymbol{z}_k)$ *is nonsingular.*

Proof: Observe that under the assumption that $\omega(\underline{c})$ commutes with $\omega(\underline{\alpha})$ and $\omega(\underline{\beta})$ we can conclude that α and β commute, as well as $\bar{\alpha}$ and β (see Proposition 1 and Corollary 1). In addition, using the fact that f is a radially holomorphic function of the form (5), we can also establish the commutativity between $f(c)$ and $\overline{f'(c)}$, α and $f'(c)$, f and $\overline{\alpha}$ and, finally, β and $\overline{f'(c)}$. Therefore

$$
z_1 = c - (af(c) + \beta)(af'(c))^{-1} = c - \frac{1}{|af'(c)|^2}(af(c) + \beta)(\overline{af'(c)})
$$

$$
= c - \frac{1}{|af'(c)|^2}(af(c) + \beta)(\overline{f'(c)\alpha})
$$

$$
= c - \frac{1}{|a|^2|f'(c)|^2}\left(af(c)\overline{\alpha}\,\overline{f'(c)} + \beta\overline{\alpha}\,\overline{f'(c)}\right)
$$

$$
= c - \frac{f(c)\overline{f'(c)}}{|f'(c)|^2} - \frac{\beta\overline{\alpha}}{|\alpha^2|}\frac{\overline{f'(c)}}{|f'(c)|^2}
$$

and

$$
\tilde{z}_1 = c - (af'(c))^{-1}(af(c) + \beta) = c - \frac{1}{|af'(c)|^2}(\overline{af'(c)})(af(c) + \beta)
$$

$$
= c - \frac{1}{|a|^2|f'(c)|^2}\left(\overline{f'(c)}|\alpha|^2 f(c) + \overline{f'(c)}\overline{\alpha}\beta\right)
$$

$$
= c - \frac{f(c)\overline{f'(c)}}{|f'(c)|^2} - \frac{\beta\overline{\alpha}}{|\alpha^2|}\frac{\overline{f'(c)}}{|f'(c)|^2}.
$$

This proves that if $\omega(\underline{c})$ commutes with $\omega(\underline{\alpha})$ and $\omega(\underline{\beta})$, then $z_1 = \tilde{z}_1$. Moreover, using similar arguments we can conclude that $\omega(\underline{z_1})$ commutes with $\omega(\underline{\alpha})$ and $\omega(\underline{\beta})$ and therefore $z_2 = \tilde{z}_2$. Using induction one can, in fact, prove that

$$
z_k = \tilde{z}_k, \ k = 1, 2, \dots.
$$

We proceed now by examining formula (20). We point out that, by (16), $J\boldsymbol{g}(\boldsymbol{x}_k)$ is nonsingular if and only if $J\boldsymbol{f}(\boldsymbol{x}_k)$ is nonsingular. In addition, using Proposition 5 we obtain

$$
(J\boldsymbol{g}(\boldsymbol{c}))^{-1}\boldsymbol{g}(\boldsymbol{c}) = (L_\alpha J\boldsymbol{f}(\boldsymbol{c}))^{-1}(L_\alpha \boldsymbol{f}(\boldsymbol{c}) + \boldsymbol{\beta})
$$

$$
= (J\boldsymbol{f}(\boldsymbol{c}))^{-1}L_\alpha^{-1}(L_\alpha \boldsymbol{f}(\boldsymbol{c}) + \boldsymbol{\beta})
$$

$$
= (J\boldsymbol{f}(\boldsymbol{c}))^{-1}\boldsymbol{f}(\boldsymbol{c}) + (J\boldsymbol{f}(\boldsymbol{c}))^{-1}L_\alpha^{-1}\boldsymbol{\beta}.
$$

Since $c = c_0 + \omega(\underline{c})|\underline{c}|$ commutes with $f(c) = u(c_0, |\underline{c}|) + \omega(\underline{c})v(c_0, |\underline{c}|)$ and also, under the assumptions of the theorem, with α and β and therefore with $\alpha^{-1}\beta$ (see Proposition 1 and Corollary 1), it follows from Proposition 4 that

$$\begin{aligned}
(J\boldsymbol{g}(\boldsymbol{c}))^{-1}\boldsymbol{g}(\boldsymbol{c}) &= (f'(c))^{-1}f(c) + (f'(c))^{-1}\alpha^{-1}\beta \\
&= (f'(c))^{-1}\alpha^{-1}\alpha f(c) + (\alpha f'(c))^{-1}\beta \\
&= (\alpha f'(c))^{-1}(\alpha f(c) + \beta) = (\alpha f'(c))^{-1}g(c).
\end{aligned}$$

This proves that $\boldsymbol{z}_1 = \tilde{z}_1$. Repeating the same arguments we can prove by induction that $\boldsymbol{z}_k = \tilde{z}_k, \; k = 1, 2, \ldots,$ $\qquad\square$

When α and β are real numbers, the function g in Theorem 3 is radially holomorphic (see Proposition 2) and the result reads as follows:

Corollary 2 *If f is a radially holomorphic function defined on the set $\mathscr{C}^1(\Omega, \mathbb{H})$, then the processes (17) with the iterative functions*

$$\mathcal{P}_1(z) = z - f(z)(f'(z))^{-1}, \tag{21}$$

$$\mathcal{P}_2(z) = z - (f'(z))^{-1}f(z), \tag{22}$$

$$\mathcal{P}_3(\boldsymbol{z}) = \boldsymbol{z} - (J\boldsymbol{f}(z))^{-1}\boldsymbol{f}(\boldsymbol{z}), \tag{23}$$

produce the same sequence, for all $c \in \mathbb{H}$, if $J(\boldsymbol{f}(\boldsymbol{z}_k))$ is nonsingular.

We observe that the equivalence of the processes (21)-(23) does not depend on the choice of the initial value c, exactly as in the case of the complex Newton method (8) and the 2D-Newton method (9). This comes from the fact that for radially holomorphic functions, $f(x)$ always commutes with x and with $(f'(x))^{-1}$.

Once the equivalence of the process $\mathcal{P}_3$ and $\mathcal{P}_1$ (and $\mathcal{P}_2$) is established, the local quadratic convergence of the method (18) (and (19)) can be established, provided that the initial guess $\boldsymbol{c}$ is chosen sufficiently close to the root $\boldsymbol{z}^*$ of $\boldsymbol{f}$ (see [20, 21], for the quadratic convergence of the Newton method for systems of nonlinear equations). This idea of proving the equivalence of the processes and, as a consequence, of establishing the local quadratic convergence of the Newton methods goes back to [5], where the authors have considered the roots of the particular polynomial $x^n - a$. The general form of this idea can now be stated easily as follows:

Theorem 4 *Let $g(x) = \sum_{i=0}^{s} \alpha_i g_i(x)$ be a function defined on the set $\mathscr{C}^1(\Omega, \mathbb{H})$ such that $g_i, \; i = 0, \ldots, s$ are radially holomorphic functions in Ω and α_i are quaternions not simultaneously zero. If z^* is a root of g such that $Jg(\boldsymbol{z}^*)$ is nonsingular and Jg is Lipschitz continuous on a neighborhood of $\boldsymbol{z}^*$, then for all $c \in \mathbb{H}$ sufficiently close to z^*, such that $\omega(\underline{c})$ commutes with all $\omega(\underline{\alpha_i})$, the Newton processes (17) with the iterative functions*

$$\mathcal{P}_1(z) = z - g(z)\Big(\sum_{i=0}^{s} \alpha_i f_i'(z)\Big)^{-1}, \tag{24}$$

$$\mathcal{P}_2(z) = z - \Big(\sum_{i=0}^{s} \alpha_i f_i'(z)\Big)^{-1}g(z), \tag{25}$$

$$\mathcal{P}_3(\boldsymbol{z}) = \boldsymbol{z} - (J\boldsymbol{g}(\boldsymbol{z}))^{-1}\boldsymbol{g}(\boldsymbol{z}), \tag{26}$$

produce the same sequence, which converges quadratically to z^*.

4 Numerical Examples

We report now several experiments obtained by the use of the package [22] which was design with the purpose of endowing the Mathematica standard package `Quaternions` with the ability to perform operations on symbolic expressions involving quaternion valued functions.

All simulations have been performed in Mathematica 9.0 (64-bit) on a computer with Intel Xeon E5607 4C 2.26GHz/1066Mhz/8MB processors and 64GB of RAM.

In order to illustrate and compare the behavior of the quaternion Newton methods proposed in this work with the classical 4D-Newton method, we consider a function $N(c, i)$ which gives the number of iterations required for each process (17) with iterative functions $\mathcal{P}_i$, $i = 1, 2, 3$ to converge, within a certain precision, to one of the solutions of the problem under consideration, using c as initial guess. The stopping criteria used is based on the incremental sizes and number of iterations, i.e. the iterative process stops whenever it produces an approximation x_k such that $|x_k - x_{k-1}| < 10^{-12}$ or $k = 50$.

We have considered different initial guesses c, by choosing points in several special regions $\Omega := \Omega(x, y) \subset \mathbb{R}^4$ and plot density plots of N as a function of x and y.

Example 2 Consider the function $g(x) = x^2 - 1 - \mathbf{i}$, where x is a quaternion of the form $x_0 + x_1\mathbf{i} + x_2\mathbf{j} + x_3\mathbf{k}$. This function can be written as $g(x) = f(x) - \mathbf{i}$, where $f(x) = x^2 - 1$ is a radially holomorphic function (see Example 1). The equation $g(x) = 0$ has two known solutions, namely

$$r_1 = \sqrt{\frac{1 + \sqrt{2}}{2}} + \sqrt{\frac{1 - \sqrt{2}}{2}}\mathbf{i} \quad \text{and} \quad r_2 = -\sqrt{\frac{1 + \sqrt{2}}{2}} - \sqrt{\frac{1 - \sqrt{2}}{2}}\mathbf{i} \qquad (27)$$

Moreover, $|J\boldsymbol{f}(\boldsymbol{x})| = 16x_0^2|x|^2$ and $f'(x) = 2x$. We have considered different initial guesses c, by choosing points in each of the following regions:

$$\Omega_1 = \{(x, y, 0, 0) : \ x, y \in [-2, 2]\},$$
$$\Omega_2 = \{(0, x, -x, y) : \ x, y \in [-2, 2]\},$$
$$\Omega_3 = \{(0, x, y, y) : \ x, y \in [-2, 2]\}.$$

Concerning the numerical results, we observe the following facts:

— Theorem 3 holds with $\alpha = 1$ and $\beta = -\mathbf{i}$, provided that the initial guess c commutes with $\mathbf{i}$ and $|J\boldsymbol{f}(\boldsymbol{z}_k)| \neq 0$, $k = 0, 1, \ldots$. In other words, if

$$c = a + b\mathbf{i}, \ a, b \in \mathbb{R} \qquad (28)$$

and

$$\operatorname{Sc} z_k \neq 0 \qquad (29)$$

then $\mathcal{P}_1(z_k) = \mathcal{P}_2(z_k) = \mathcal{P}_3(\boldsymbol{z}_k)$, $k = 1, 2, \ldots$.

— When $c \in \Omega_1$, (28) is fulfilled and $\mathcal{P}_i^k(c) \in \Omega_1$, $k = 1, 2, \ldots$. Figure 1(a) illustrates the behavior of Newton methods, under the assumption (29). When c is chosen sufficiently close to one of the roots r_1 or r_2, we observe, as expected,

fast convergence. In fact, since $\Omega_1 \cong \mathbb{C}$, the quaternion methods (18) and (19) reduce to the complex Newton method (8), once we identify the quaternion valued function g with the complex valued function $x^2 - 1 - i$.

$-$ The situation is rather different when one choose $c \in \Omega_2$. Here, the equivalence of the processes does not take place, because the assumptions (28)-(29) are no longer true, which means that the iterative function $\mathcal{P}_3$ associated to the classical 4D-Newton method can not be used at all. On the other hand, the iterative functions $\mathcal{P}_1$ and $\mathcal{P}_2$ can be used whenever $f'(\mathrm{x}_k) = 2\mathrm{x}_k \neq 0$. Figure 1(b) illustrates this situation. The iteration sequences are different, but both methods converge, albeit slower than the previous case.

$-$ When $c \in \Omega_3$, assumption (28) is, in general, not true and therefore the three methods don't produce the same iteration sequence. This fact is well reported in Figure 1(c), where we can also see that the classical 4D-Newton method converges, in general, faster (less iterations are needed) than the other two.

Example 3 Consider the function $g(x) = x^3 - x^2\mathbf{j} - x + \mathbf{j}$, which can be written as $g(x) = g_0(x) - g_1(x)\mathbf{j}$, where $g_0(x) = x^3 - x$ and $g_1(x) = x^2 - 1$ are radially holomorphic functions. The quaternion polynomial g has three distinct roots, namely
$$r_1 = 1, \quad r_2 = -1, \quad r_3 = \mathbf{j}. \tag{30}$$
The aim of this example is to analyze the behavior of the iterative functions (24)-(26), for the case $\alpha_1 = 1$, $\alpha_2 = -\mathbf{j}$. Figure 2 contains density plots of the functions $N(c, i)$ associated to each function $\mathcal{P}_i$ $(i = 1, 2, 3)$, when c is chosen in one of the following regions:

$$\Omega_1 = \{(x, y, 0, 0) : \ x, y \in [-2, 2]\},$$
$$\Omega_2 = \{(0, x, y, 0) : \ x, y \in [-2, 2]\},$$
$$\Omega_3 = \{(x, 0, y, 0) : \ x, y \in [-2, 2]\}.$$

Concerning the numerical results, we point out the following:

$-$ In Ω_1 the processes (17) produce different sequences, which converge to one of the roots (30), but the three methods don't converge always to the same root. In fact, starting from $c = 1.31 + 2\mathbf{i}$, process associated to $\mathcal{P}_1$ converges to the root r_3, process associated to $\mathcal{P}_2$ converges to the root r_1 while the third process converges to the root r_2, as Table 1 shows.

$-$ For the sake of better visibility, we also present the basins of attraction of the roots (30) with respect to the iterative function $\mathcal{P}_i$ (see Figure 3). The color code used is the following: choosing any initial guess c in the region corresponding to color i, causes the process to converge to the root r_i; $i = 1, 2, 3$.

$-$ It is easy to see that $\mathcal{P}_i(x) \in \Omega_2$, for all $x \in \Omega_2$ and therefore it is not possible to obtain convergence to any of the real roots of h. Thus the three methods produce three different sequences (Corollary 4 does not apply) all converging to the nonreal root $r_3 = \mathbf{j}$, for any choice of $c \in \Omega_2$.

$-$ In Ω_3, Theorem 4 applies and we obtain the same sequence for all the processes.

(a) $N(c, 1) = N(c, 2) = N(c, 3)$, $c \in \Omega_1$

(b) $N(c, i)$, $c \in \Omega_2$ and $i = 1, 2$ (from the left to the right)

(c) $N(c, i)$, $c \in \Omega_3$ and $i = 1, 2, 3$ (from the left to the right)

Figure 1: Newton methods for Example 2

13

Table 1: Newton iterations for Example 3, with $c = 1.31 + 2\mathbf{i}$

Iterative function $\mathcal{P}_1$

$z_0 = 1.31 + 2\mathbf{i}$
$z_1 = 0.908706241411 + 1.387337773146\mathbf{i} + 0.439339998353\mathbf{j} - 0.212237448464\mathbf{k}$
$\vdots$
$z_6 = 0.005549336281 - 0.004999963424\mathbf{i} + 1.005002523097\mathbf{j} + 0.008395675752\mathbf{k}$
$\vdots$
$z_9 = 1.\mathbf{k}$

Iterative function $\mathcal{P}_2$

$\tilde{z}_0 = 1.31 + 2\mathbf{i}$
$\tilde{z}_1 = 0.908706241411 + 1.225324453708\mathbf{i} - 0.050277419391\mathbf{j} + 0.103775488306\mathbf{k}$
$\vdots$
$\tilde{z}_6 = 1.002873960381 + 0.006924303843\mathbf{i} + 0.019152349106\mathbf{j} + 0.013086632643\mathbf{k}$
$\vdots$
$\tilde{z}_9 = 1.$

Iterative function $\mathcal{P}_3$

$\boldsymbol{z}_0 = (0.31, 2, 0, 0)$
$\boldsymbol{z}_1 = (0.725569466462, 0.897281973274, -2.725798320733, -0.088716357547)$
$\vdots$
$\boldsymbol{z}_9 = (-0.921970127170, 0.2445632666550, -0.251209684043, 0.463755091394)$
$\vdots$
$\boldsymbol{z}_{15} = (-1., 0., 0., 0.)$

(a) $N(c, i)$, $c \in \Omega_1$ and $i = 1, 2, 3$ (from the left to the right)

(b) $N(c, i)$, $c \in \Omega_2$ and $i = 1, 2, 3$ (from the left to the right)

(c) $N(c, 1) = N(c, 2) = N(c, 3)$, $c \in \Omega_3$

Figure 2: Newton methods for Example 3

(a) Iterative functions $\mathcal{P}_i$: $c \in \Omega_1$ and $i = 1, 2, 3$ (from the left to the right)

(b) $\mathcal{P}_1 \equiv \mathcal{P}_2 \equiv \mathcal{P}_3$: $c \in \Omega_3$

Figure 3: Basins of attraction of the roots of Example 3 with respect to the iterative functions $\mathcal{P}_i$

Table 2: Comparison of the computational efforts between the three versions of Newton method for Example 3

	Ω_1	Ω_2	Ω_3
$\mathcal{P}_1$	5.89	1.55	5.86
$\mathcal{P}_2$	5.48	1.49	6.57
$\mathcal{P}_3$	15.36	8.91	18.16

Figure 4: Representation of some zeros of the function considered in Example 4

− Table 2 shows the CPU time (in minutes) spent by each of the processes (24)-(26) to produce the images presented in Figure 2.

Example 4 Consider now the radially elementary function

$$f(x) = \exp(x) - x = e^{x_0} \cos r - x_0 + \omega(\underline{x})(e^{x_0} \sin r - r),$$

where $r = |\underline{x}|$ (see Example 1). There are an infinite number of nonisolated roots $r = r_0 + \underline{r}$ of f, as illustrated in Figure 4 where a plot of r_0 versus $|\underline{r}|$ is presented.

In this case, we can apply Corollary 2 to conclude that processes corresponding to (21)-(23) are equivalent, provided that $J\boldsymbol{f}(\boldsymbol{z}_k)$ is nonsingular. From (12) it follows that

$$|J\boldsymbol{f}(\boldsymbol{x})| = 0 \Leftrightarrow f'(x) = e^x - 1 = 0 \quad \text{or} \quad v(x_0, r) = e^{x_0} \sin r - r = 0.$$

This means that for all pure quaternions x such that $|\underline{x}| = 2k\pi, \ k \in \mathbb{Z}$, the three processes break down, since $f'(x) = 0$, whereas process $\mathcal{P}_3$ also terminates whenever $e^{x_0} \sin r = r$. We illustrate the performance of the three methods by considering two particular domains:

$$\Omega_1 = \{(x, y, 0, 0) : \ x \in [-5, 11], y \in [-8, 8]\},$$

$$\Omega_2 = \{(0, x, y, 0) : \ x, y \in [-8, 8]\}.$$

It is visible, from Figure 5 the influence of the points for which Newton methods break down. These points correspond to the white regions of the figure and

17

(a) $N(c, i)$, $c \in \Omega_1$ and $i = 1, 3$ (from the left to the right)

(b) $N(c, i)$, $c \in \Omega_2$ and $i = 1, 3$ (from the left to the right)

Figure 5: Newton methods for Example 4

we can see that the classical 4D-Newton method, corresponding to the right-hand side of Figures 5(a) and 5(b), is, as expected, extremely affected by this phenomenon. The quaternion versions of Newton method constitute, for this example, an effective alternative to the classical version, not only because they involve a lower computational cost, but also because they produce a significantly lower number of points for which the Newton method stops (see also Figure 6, where the basins of attraction of the roots $0.318132 \pm 1.33724\mathbf{i}$ and $2.0622 \pm 7.58863\mathbf{i}$, with respect to the iterative function $\mathcal{P}_1(\equiv \mathcal{P}_2)$ are presented).

5 Final Remarks

We have shown, by deriving new relations between the quaternion radial derivative of a radially holomorphic function and the derivative of the associated vector function, that the quaternion versions of the Newton method for finding roots of a class of functions are, under certain conditions, equivalent to the classical Newton method for vector functions. Since these class of functions includes simples polynomials, Theorem 4 can be seen as a quaternion analysis argument to justify the results of [5].

The numerical results presented in last section (and some more produced during this work with the same purpose) indicate that the quaternion Newton methods involve a lower computational cost than the classical Newton method,

Figure 6: Basins of attraction of some roots of Example 4, with respect to the iterative function $\mathcal{P}_1 (\equiv \mathcal{P}_2)$, for $c \in \Omega_1$.

since inverting a matrix involves more operations than inverting a quaternion. Moreover, Proposition 3 shows that

$$\{ \boldsymbol{x} \in \mathbb{R}^4 : |\boldsymbol{J}\boldsymbol{f}(\boldsymbol{x})| = 0 \} \supset \{ x \in \mathbb{H} : f'(x)) = 0 \}$$

and therefore, the quaternion Newton methods produce a significantly lower number of points for which the method stops, than the classical one. The numerical experiments also show evidences of local convergence (even quadratically), in some cases where the assumptions of Theorem 4 are not fulfilled. Forthcoming work will be on the behavior of the methods in such cases.

References

[1] R. Fueter, Die funktionentheorie der differetialgleichungen $\Delta u = 0$ und $\Delta\Delta u = 0$ mit vier reellen variablen, Comm. Math. Helv. (7) (1934-35) 307–330.

[2] R. Fueter, Über die analytische Darstellung der regulären Funktionen einer Quaternionenvariablen, Comment. Math. Helv. 8 (1) (1935) 371–378.

[3] K. Gürlebeck, K. Habetha, W. Sprößig, Holomorphic functions in the plane and n-dimensional space, Birkhäuser Verlag, Basel, 2008.

[4] H. R. Malonek, Quaternions in applied sciences. A historical perspective of a mathematical concept, 17th Inter. Conf. on the Appl. of Computer Science and Mathematics on Architecture and Civil Engineering, Weimar, 2003.

[5] D. Janovská, G. Opfer, Computing quaternionic roots by Newton's method, Electron. Trans. Numer. Anal. 26 (2007) 82–102.

[6] B. Kalantari, Algorithms for quaternion polynomial root-finding, J. Complexity 29 (3-4) (2013) 302–322.

[7] F. Brackx, R. Delanghe, F. Sommen, Clifford analysis, Pitman, Boston-London-Melbourne, 1982.

[8] K. Gürlebeck, W. Sprößig, Quaternionic and Cliford calculus for physicists and engineers, John Wiley & Sons, 1997.

[9] A. Sudbery, Quaternionic analysis, Math. Proc. Camb. Phil. Soc. 85 (1979) 199–225.

[10] W. Sprößig, On operators and elementary functions in Clifford analysis, Journal for Analysis and its Applications 18 (2) (1999) 349–360.

[11] S. Georgiev, J. Morais, W. Sprössig, New aspects on elementary functions in the context of quaternionic analysis, Cubo 14 (1) (2012) 93–110.

[12] I. Niven, Equations in quaternions, Amer. Math. Monthly 48 (1941) 654–661.

[13] S. De Leo, G. Ducati, V. Leonardi, Zeros of unilateral quaternionic polynomials, Electron. J. Linear Algebra 15 (2006) 297–313.

[14] G. Gentili, D. C. Struppa, F. Vlacci, The fundamental theorem of algebra for Hamilton and Cayley numbers, Math. Z. 259 (4) (2008) 895–902.

[15] Z. Jia, X. Cheng, M. Zhao, A new method for roots of monic quaternionic quadratic polynomial, Comput. Math. Appl. 58 (9) (2009) 1852–1858.

[16] A. Pogorui, M. Shapiro, On the structure of the set of zeros of quaternionic polynomials, Complex Var. Theory Appl. 49 (6) (2004) 379–389.

[17] R. Serôdio, L.-S. Siu, Zeros of quaternion polynomials, Appl. Math. Lett. 14 (2) (2001) 237–239.

[18] D. Janovská, G. Opfer, The classification and the computation of the zeros of quaternionic, two-sided polynomials, Numer. Math. 115 (1) (2010) 81–100.

[19] D. Janovská, G. Opfer, Zero points of general quaternionic polynomials, AIP Conference Proceedings, 1558, 554 (2013).

[20] J. M. Ortega, W. C. Rheinboldt, Iterative solution of nonlinear equations in several variables, Academic Press, New York, 1970.

[21] J. F. Traub, Iterative Methods for the Solution of Equations, Prentice-Hall, Inc. Englewood Cliffs, NJ, 1964.

[22] M. I. Falcão, F. Miranda, Quaternions: A Mathematica package for quaternionic analysis, in: B. Murgante, O. Gervasi, A. Iglesias, D. Taniar, B. Apduhan (Eds.), Computational Science and Its Applications - ICCSA 2011, Vol. 6784 of Lecture Notes in Computer Science, Springer-Verlag Berlin Heidelberg, 2011, pp. 200–214.