RepositóriUM
Universidade do Minho

CMAT
Centro de Matemática da Universidade do Minho
Campus de Gualtar 4710-057 Braga Portugal
www.cmat.uminho.pt

Universidade do Minho
Escola de Ciências
Centro de Matemática

Modified Quaternion Newton Methods

M. Irene Falcão   Fernando Miranda

CMAT/DMA, Universidade do Minho, Portugal

Information

Keywords:
Quaternion Analysis, Newton methods.

Original publication:
Computational Science and its Applications, Lecture Notes in Computer Science, vol. 8579, pp. 146-161, 2014
DOI: 10.1007/978-3-319-09144-0_11
www.springerlink.com

Abstract

We revisit the quaternion Newton method for computing roots of a class of quaternion valued functions and propose modified algorithms for finding *multiple* roots of simple polynomials. We illustrate the performance of these new methods by presenting several numerical experiments.

1 Introduction

In this work we concentrate on the problem of finding roots of special quaternion polynomials of the form

$$p_n(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0,\ a_i \in \mathbb{H},\ i = 0, \ldots, n,\ a_n \neq 0. \tag{1}$$

Since the work of Niven [18], several authors gave contributions to the problem of finding roots of quaternion valued polynomials (see e.g. [3,4,11,14,16,21,23]), by following different approaches and with different motivations.

Recently a quaternion version of the well known Newton method for finding roots of a class of quaternion functions was proposed in [5]. This work was motivated by [13], where the authors formally adapted, for the first time, Newton method for finding roots of quaternions, i.e. for solving quaternion equations of the form $x^n + a_0 = 0$.

Due to the non-commutativity of quaternion multiplication, the use of root-finding methods involving quaternion iterative functions requires close attention. In particular in the framework of Newton-like methods, left and right quaternion versions have to be considered.

The results of [5] are based on the equivalence between the classical multivariate Newton method and a quaternion version derived by the use of the so-called radial derivative.

In this work we give new insights on the quaternion Newton method, by making the link, under certain conditions, to the complex approach. Quaternion versions of well known variants of the classical Newton method for multiple roots are also derived.

2 Quaternion Analysis Toolbox

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

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

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 $\boldsymbol{x} = (x_0, x_1, x_2, x_3) \in \mathbb{R}^4$ with the element $x = x_0 + x_1\mathbf{i} + x_2\mathbf{j} + x_3\mathbf{k} \in \mathbb{H}$. Thus, throughout this paper, we will not distinguish an element in $\mathbb{R}^4$ and the corresponding quaternion in $\mathbb{H}$, unless we need to stress the context.

The conjugate of x is defined as

$$\bar{x} = x_0 - x_1\mathbf{i} - x_2\mathbf{j} - x_3\mathbf{k}$$

and instead of the real and the imaginary parts we will distinguish between the scalar part of x

$$\mathrm{Sc}\, x := x_0 = \tfrac{1}{2}(x + \bar{x})$$

and the vector part of x

$$\mathrm{Vec}\, x = \underline{x} := x_1\mathbf{i} + x_2\mathbf{j} + x_3\mathbf{k} = \tfrac{1}{2}(x - \bar{x}).$$

When $\mathrm{Sc}\, x = 0$, x is called a pure quaternion. The norm $|x|$ of x is defined by

$$|x|^2 = x\bar{x} = \bar{x}x = x_0^2 + x_1^2 + x_2^2 + x_3^2$$

and it immediately follows that each non-zero $x \in \mathbb{H}$ has an inverse given by

$$x^{-1} = \frac{\bar{x}}{|x|^2}.$$

Quaternions x such that $|x| = 1$ are called unit quaternions. Observe that any arbitrary non-real quaternion x can be written as

$$x = x_0 + \underline{x} = x_0 + \omega(\underline{x})|\underline{x}|\,, \tag{2}$$

where $\omega(\underline{x})$ is the unit quaternion

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

very much like a complex number is written in the form $a+ib$. Moreover, since $\omega(\underline{x})^2 = -1$, one can argue that $\omega(\underline{x})$ behaves like the imaginary unit. In what follows we use the convention $\omega(\underline{x}) := 0$, for real quaternions x. Now, if x and y are quaternions such that $\omega(\underline{x}) = \omega(\underline{y}) =: \boldsymbol{\omega}$, i.e. if $x = a + \boldsymbol{\omega} b$ and $y = c + \boldsymbol{\omega} d$, then all the algebraic operations can be computed as if x and y were complex numbers, in particular,

$$xy = yx = ac - bd + \boldsymbol{\omega}(ad + bc).$$

$$xy^{-1} = y^{-1}x = \frac{ac + bd}{c^2 + d^2} + \boldsymbol{\omega}\frac{bc - ad}{c^2 + d^2}. \tag{3}$$

For all the above reasons, we call (2) the *complex-like* representation of a quaternion x.

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

$$f(x) = f(x_0 + w(\underline{x})r) = u(x_0, r) + \omega(\underline{x})v(x_0, r)\,, \tag{4}$$

where $x_0 = \mathrm{Sc}\, x$, $r := |\underline{x}|$ and u and v are real valued functions. Continuity and differentiability are defined coordinate wise.

In order to prepare next results, we define on the set $\mathscr{C}^1(\Omega, \mathbb{H})$ the so-called *radial operators*

$$\partial_{\mathrm{rad}} := \tfrac{1}{2}(\partial_0 - \boldsymbol{\omega}\partial_r), \qquad \bar{\partial}_{\mathrm{rad}} := \tfrac{1}{2}(\partial_0 + \boldsymbol{\omega}\partial_r), \tag{5}$$

where $\partial_0 := \frac{\partial}{\partial x_0}$ and $\partial_r := \frac{\partial}{\partial r}$.

Definition 1. *Let f be a function of the form (4), $x \in \Omega$ and $h = h_0 + \omega(\underline{x})h_r$, $h_0, h_r \in \mathbb{R}$. Such function f is called radially holomorphic or radially regular in x if*

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

exists. In the case of existence, this limit is called the radial derivative of f at x and is denoted by $f'(x)$.

The following results are well known and play a fundamental role in the present work (see [12,24]).

Proposition 1. *f is radially holomorphic if and only if $\bar{\partial}_{\rm rad} f = 0$.*

Remark 1. Let g be a complex holomorphic function in the complex variable $z = x + iy$ and recall that the complex partial derivatives (also called Wirtinger derivatives)

$$\frac{\partial g}{\partial \bar{z}} = \frac{1}{2}\Big(\frac{\partial g}{\partial x} + i\frac{\partial g}{\partial y}\Big), \qquad \frac{\partial g}{\partial z} = \frac{1}{2}\Big(\frac{\partial g}{\partial x} - i\frac{\partial g}{\partial y}\Big)$$

allow to express the Cauchy-Riemann equations in the form $\frac{\partial g}{\partial \bar{z}} = 0$. In other words, $\bar{\partial}_{\rm rad} f = 0$ is, in fact, a Cauchy-Riemann type differential equations, which can be written as

$$\partial_0 u = \partial_r v, \quad \partial_0 v = -\partial_r u.$$

Proposition 2. *If f is radially holomorphic then $f' = \partial_{\rm rad} f$.*

Remark 2. Since a radially holomorphic function belongs to the kernel of $\bar{\partial}_{\rm rad}$, it follows that, in fact, $f' = \partial_0 f = -\partial_r f$ i.e.

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

which is similar to the complex case. Moreover it follows immediately that u and v are harmonic functions and therefore f' is also radially holomorphic.

Proposition 3. *If f and g are radially holomorphic functions of the form (4) then*

1. *$\alpha f + \beta g$, with $\alpha, \beta \in \mathbb{R}$, is radially holomorphic and*

$$(\alpha f + \beta g)'(x) = \alpha f'(x) + \beta g'(x);$$

2. *fg is radially holomorphic and*

$$(fg)'(x) = f'(x)g(x) + f(x)g'(x);$$

3. *$\frac{1}{f}$ is radially holomorphic and*

$$(\tfrac{1}{f})'(x) = -f'(x)f(x)^{-2} = -f(x)^{-2}f'(x), \qquad (f(x) \neq 0).$$

Example 1. The function

$$f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0,\ a_i \in \mathbb{R},\ i = 0, \ldots, n$$

is radially holomorphic and $f'(x) = na_n x^{n-1} + (n-1)a_{n-1}x^{n-2} + \cdots + a_1$.

Remark 3. The construction of radially regular functions goes back to the work [7] of R. Fueter, one of the founders of quaternion analysis. This class of functions is also related to the standard intrinsic functions studied by Rinehart [22] and later on by Cullen [2]. Recently, following the approach of this last work, a new theory of regular functions of one quaternion variable, the so-called slice regular functions, has been introduced by Gentili and Struppa [9,10] and is now very well developed (see [1] and the references therein).

3 Remarks on the Zeros of Quaternion Polynomials

It is a well known fact that the algebraic as well as the geometric properties of complex holomorphic functions are not the same for their generalizations in quaternion analysis. In particular, because of the non-commutativity of quaternion multiplication, one can consider different classes of polynomials in one quaternion variable, depending on whether a variable commutes with polynomial coefficients or not. General polynomials are defined as finite sums of non-commutative monomials of the form $a_0xa_1xa_2\ldots xa_j$.

In this work we consider only polynomials whose coefficients are located only on the left-hand side of the powers, i.e. they have the special form (1). These polynomials are usually called in the literature, *simple* or *one-sided* polynomials. If the coefficients a_i in (1) are real, then we say that p_n is a real polynomial. As usual, when $a_n = 1$ the polynomial is called monic.

In this section we review some basic properties of simple quaternion polynomials needed in the sequel.

Definition 2. *Two quaternions x and y are called equivalent, written $x \sim y$, if $x = h^{-1}yh$, for some $h \in \mathbb{H}$. The equivalence class of x, denoted by $[x]$, is the set*

$$[x] = \{y \in \mathbb{H} : y \sim x\}.$$

It is easy to see that, in fact,

$$[x] = \big\{y \in \mathbb{H} : \operatorname{Sc} x = \operatorname{Sc} y \text{ and } |\underline{x}| = |\underline{y}|\big\}. \tag{6}$$

Definition 3. *Let Z_{p_n} denote the zero-set of a simple polynomial p_n. A non-real root z^* of p_n is called a spherical root (or one says that z^* generates a spherical root) if $[z^*] \subset Z_{p_n}$. In this case $[z^*]$ is called a sphere of zeros for p_n. A root z^* of p_n is called isolated if either z^* is real or it does not generate a spherical root.*

We note that any non-real quaternion $z^* = z_0^* + \underline{z}^*$ is a root of the real quadratic polynomial

$$q_2(x) = x^2 - 2z_0^*x + |z^*|^2$$

and, taking into account (6), it is clear that z^* is, in fact, a spherical root of q_2. The following property can be used in order to identify a spherical root.

Proposition 4 ([21, Corollary 2.1]). *A non-real quaternion z^* generates a spherical root of a simple polynomial p_n if and only if $\overline{z^*} \in Z_{p_n}$.*

The following result characterizes the zero-set Z_{p_n} of a simple polynomial p_n, by saying that its zeros fall in two classes.

Proposition 5 ([21, Theorem 6]). *The zero-set of a simple polynomial p_n consists of r isolated roots and s spheres of zeros with $r + 2s \leq n$.*

Example 2. The zero-sets of the following simple polynomials

$$p_4(x) = (x - \mathbf{i} + \mathbf{j})(x^2 + 1)(x - 1), \quad q_2(x) = x^2 - (\mathbf{i} + \mathbf{j})x + \mathbf{k} \quad \text{and} \quad s_2(x) = x^2 - 2\mathbf{i},$$

are, respectively

$$Z_{p_4} = \{1, \mathbf{i} - \mathbf{j}\} \cup [\mathbf{i}], \quad Z_{q_2} = \{\mathbf{j}\} \quad \text{and} \quad Z_{s_2} = \{-1 - \mathbf{i}, 1 + \mathbf{i}\}.$$

Example 2 illustrates some typical features of the zeros of simple polynomials that we need to be aware of. In particular, denoting by $z_1 = -1 - \mathbf{i}$ and $z_2 = 1 + \mathbf{i}$ the roots of s_2, observe that $(x - z_1)(x - z_2) = x^2 - x\mathbf{i} + \mathbf{i}x - 2\mathbf{i} \neq s_2(x)$ and $(x - z_2)(x - z_1) = x^2 + x\mathbf{i} - \mathbf{i}x - 2\mathbf{i} \neq s_2(x)$.

Nevertheless, when p_n is a real polynomial, the zeros of p_n can be obtained by considering p_n as a polynomial over $\mathbb{C}$.

Proposition 6. *If p_n is a real monic polynomial and the zero-set of p_n in $\mathbb{C}$ is*

$$Z_{p_n}^{\mathbb{C}} = \{z_1, \dots, z_r, \zeta_1, \dots, \zeta_s, \overline{\zeta_1}, \dots, \overline{\zeta_s}\},$$

where $z_1, \dots, z_r$ are real numbers and $\zeta_1, \dots, \zeta_s$ are non-real complex numbers, then the zero-set of p_n in $\mathbb{H}$ is

$$Z_{p_n}^{\mathbb{H}} = \{z_1, \dots, z_r\} \cup [\zeta_1] \cup \dots \cup [\zeta_s].$$

Moreover, there exist positive integers $m_1, \dots, m_r, n_1, \dots, n_s$, such that

$$p_n(x) = (x - z_1)^{m_1} \dots (x - z_r)^{m_r} (x^2 - 2x \operatorname{Sc} \zeta_1 + |\zeta_1|^2)^{n_1} \dots (x^2 - 2x \operatorname{Sc} \zeta_s + |\zeta_s|^2)^{n_s},$$

with $\sum_{i=1}^{r} m_i + 2 \sum_{i=1}^{s} n_i = n$.

The problem of giving an appropriate notion of multiplicity of a root of a quaternion polynomial is not a trivial task. However, for real polynomials and taking into account last proposition, it seems natural to call the numbers m_i the multiplicities of the real roots x_i and say that $[\zeta_j]$ are spheres of zeros of order n_j. We point out that, similar to the complex case, the derivatives of p_n satisfy

$$p_n^{(k)}(z_i) = 0, \; k = 0, \dots, m_i - 1 \quad \text{and} \quad p_n^{(m_i)}(z_i) \neq 0,$$
$$p_n^{(k)}(x) = 0, \; k = 0, \dots, n_j - 1 \quad \text{and} \quad p_n^{(n_j)}(x) \neq 0, \; \forall x \in [\zeta_j].$$

Finally we would like to call reader's attention to the work [15] of Janovská and Opfer where they give a result connecting the zeros of a simple polynomial p_n with those of a certain real polynomial of degree $2n$. This strategy was already used in the pioneer work of Niven [18] and in, among others, [21].

4 Quaternion Newton Method

Newton methods in quaternion context were formally adapted for the first time by Janovská and Opfer in [13], where the authors solved equations of the form $x^n - a = 0$, $a \in \mathbb{H}$. Later on, Kalantari in [17], using algebraic-combinatorial arguments, proposed a Newton method for finding roots of special quaternion polynomials.

Here we follow the ideas in [5], where the equivalence between the classical multivariate Newton method (4D-NM) and the quaternion Newton methods ($\mathbb{H}$-NM$_l$ and $\mathbb{H}$-NM$_r$) for radially holomorphic functions was established and reads as follows.

Proposition 7 ([5, Theorem 4]). *Let $f(x) = \sum_{i=0}^{s} \alpha_i f_i(x)$ be a function defined on the set $\mathscr{C}^1(\Omega, \mathbb{H})$ such that f_i, $i = 0, \dots, s$, are radially holomorphic functions in Ω and α_i are quaternions not simultaneously zero. If z^* is a root of f such that $J\boldsymbol{f}(\boldsymbol{z}^*)$ is nonsingular and $J\boldsymbol{f}$ 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*

$$\mathbb{H}\text{-NM}_r: \quad z_{k+1} = z_k - f(z_k) \Big(\sum_{i=0}^{s} \alpha_i f_i'(z_k) \Big)^{-1}, \qquad z_0 = c; \tag{7}$$

$$\mathbb{H}\text{-NM}_l: \quad z_{k+1} = z_k - \Big(\sum_{i=0}^{s} \alpha_i f_i'(z_k) \Big)^{-1} f(z_k), \qquad z_0 = c; \tag{8}$$

$$\text{4D-NM}: \quad \boldsymbol{z}_{k+1} = \boldsymbol{z}_k - (J\boldsymbol{f}(\boldsymbol{z}_k))^{-1} \boldsymbol{f}(\boldsymbol{z}_k), \qquad \boldsymbol{z}_0 = \boldsymbol{c} \tag{9}$$

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

Here, for the sake of clarity, we have used the bold symbol $\boldsymbol{f}$ to denote the vector valued function defined in $\mathbb{R}^4$ corresponding to the quaternion valued function f as well as for the other elements in Proposition 7.

In this paper we provide new insights on the $\mathbb{H}$-NM (7)-(8) and propose modified versions prepared to deal with the case of multiple roots.

We recall here the well known fact that the 4D-NM for solving the equation $\boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{0}$ has only linear convergence in the neighborhood of a root $\boldsymbol{z}^*$ such that $\det(J\boldsymbol{f}(\boldsymbol{z}^*)) = 0$ (see e.g. [19], [25]). When z^* is a multiple isolated root of the polynomial p_n or generates a sphere of zeros of order greater than one, in the sense of the definition introduced in last section, then $\det(J\boldsymbol{p}_n(\boldsymbol{z}^*)) = 0$ and a modification of the 4D-NM and consequently of the $\mathbb{H}$-NM$_l$ and/or $\mathbb{H}$-NM$_r$ is required in order to gain the second order convergence.

When the multiplicity m of a root is known, a classical way to re-establish the order of convergence of the real/complex Newton method (NM) for solving $f(x) = 0$ is to consider the iterative process

$$z_{k+1} = z_k - m\frac{f(z_k)}{f'(z_k)} \tag{10}$$

or, alternatively, to apply NM to the function

$$g(x) = \frac{f(x)}{f'(x)}, \tag{11}$$

since this function has the same roots as f and all of them are simple.

We point out that the use of (10) in its quaternion version is straightforward. To adapt (11) to the quaternion context we note the following:

1. If f is a quaternion radially holomorphic function, then
$$u(x) := f(x)(f'(x))^{-1} = (f'(x))^{-1}f(x)$$
is also radially holomorphic (see (3) and Proposition 3) and, therefore, Proposition 7 applies.
2. The same reasoning can be used when the function f is of the general form considered in the assumptions of Proposition 7. In fact, if f is of the form $f(x) = \sum_{i=0}^{s} \alpha_i f_i(x)$, where f_i are radially holomorphic functions and the coefficients α_i commute pairwise, then one can prove (using arguments similar to those in [5]) that
$$u(x) := \sum_{i=0}^{s} \alpha_i f_i(x)\Big(\sum_{i=0}^{s} \alpha_i f_i'(x)\Big)^{-1} = \Big(\sum_{i=0}^{s} \alpha_i f_i'(x)\Big)^{-1}\sum_{i=0}^{s} \alpha_i f_i(x). \tag{12}$$
Moreover, u can be written as $u(x) = \sum_{i=0}^{s} \beta_i u_i(x)$, for some radially holomorphic functions u_i and some coefficients $\beta_i \in \mathbb{H}$ commuting pairwise. Therefore Proposition 7 is valid for the function u.
3. It is clear that u has the same roots as f and they are all simple. Therefore, by Proposition 7 the sequence
$$z_{k+1} = z_k - u(z_k)\Big(\sum_{i=0}^{s} \beta_i u_i'(z_k)\Big)^{-1} \tag{13}$$
converges quadratically to z^*, when $c = z_0$ is sufficiently close to z^*.

We underline that the use of (10) requires the explicit knowledge of the multiplicity m of the root, whereas the use of (11) just requires the knowledge of the existence of a non-simple root. When this is not the case, i.e. if we are not aware of the existence of multiple roots, one can consider adaptive Newton methods (ANM) for estimating, in each iteration k, the value of m to be used in (10). For such purpose we compute approximations to m based on the following known results:

$$m \approx m_k := \frac{z_k - z_{k-1}}{g(z_k) - g(z_{k-1})}, \qquad m \approx \widetilde{m}_k := \frac{\log|f(z_k)|}{\log|g(z_k)|}, \tag{14}$$

where g is the function defined in (11). In this work we consider natural adaptations of (14) to derive three ANM. The algorithm for determining both z^* and m can be written as follows.

Modified Quaternion Newton Methods:

1. Choose a value of z_0, m, kmax and a tolerance ε
2. For $k = 1, 2, \ldots$
   i. Compute z_k by means of the appropriated quaternion version of (10)

ii. `Compute` m `by choosing one of the estimates`

$$
\begin{aligned}
m_1 &= \max\left\{1, \frac{|z_k - z_{k-1}|}{|u(z_k) - u(z_{k-1})|}\right\} && \% \ \mathbb{H} - \mathrm{ANM}_1 \\
m_2 &= \max\left\{1, \frac{\log|f(z_k)|}{\log|u(z_k)|}\right\} && \% \ \mathbb{H} - \mathrm{ANM}_2 \\
m_3 &= \mathrm{round}(m_2) && \% \ \mathbb{H} - \mathrm{ANM}_3
\end{aligned}
$$

where u is defined in (12)

3. `Repeat Step 2 until` k `is such that` $|z_k - z_{k-1}| < \varepsilon$ `or` $k =$ `kmax`

Next section contains several numerical results illustrating the performance of these methods. In particular, we present estimates for the computational order of convergence ρ of each method based on the use of

$$
\rho \approx \rho_{k+1} := \frac{\log\left(|z^* - z_{k+1}|/|z^* - z_k|\right)}{\log\left(|z^* - z_k|/|z^* - z_{k-1}|\right)},
$$

where z_{k-1}, z_k, z_{k+1} are three consecutive iterations close to z^* (see e.g. [20]).

It is known from the literature that the order of convergence of the classical ANM_1 is $\frac{1+\sqrt{5}}{2} \approx 1.62$, whereas the convergence of ANM_2 is essentially linear. For details and comments on the order of convergence of these modified methods we refer to [25] and [8]. The numerical experiments reported in next section confirm a similar computational order of convergence of the adaptative quaternion versions $\mathbb{H}$-ANM.

5 Numerical Examples

The numerical experiments reported is this section were obtained by the use of a package [6] designed by the authors of this paper with the purpose of endowing the Mathematica `Quaternions` package with the ability of operating 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 modified quaternion Newton methods proposed in Sect. 4, we consider, as in [5], a function $N(c)$ which gives the number of iterations required for each process 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 z_k such that $|z_k - z_{k-1}| < \varepsilon$ or $k = \mathrm{kmax}$.

We have considered different initial guesses c, by choosing points in special regions $\Omega := \Omega(x, y) \subset \mathbb{R}^4$ and show density plots of N as a function of x and y. In all figures, the white regions correspond to a choice of $c \in \Omega$ for which the method under consideration did not reach the level of precision ε with kmax iterations.

Example 3. We consider as a first example the real polynomial

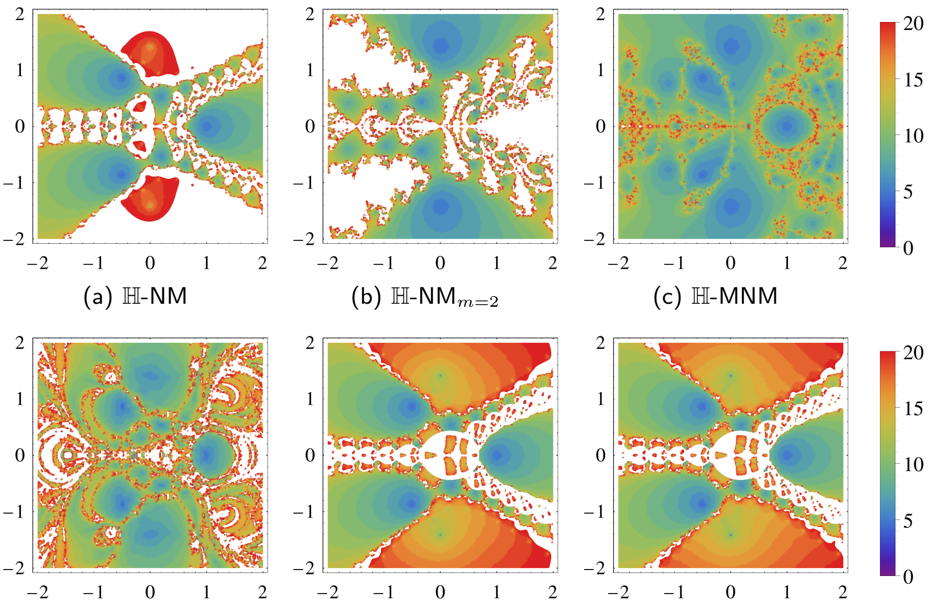
$$
p_7(x) = (x^3 - 1)(x^2 + 2)^2
$$

which has the real isolated root $z_1 = 1$ and two spherical roots generated by $\zeta_1 = -\frac{1}{2} + \frac{\sqrt{3}}{2}\mathbf{i}$ and $\zeta_2 = \sqrt{2}\mathbf{i}$. According to the definition introduced in Sect. 3, z_1 is a root of multiplicity $m_1 = 1$, $[\zeta_1]$ is a sphere of zeros of order $n_1 = 1$ and $[\zeta_2]$ is a sphere of zeros of order $n_2 = 2$. Furthermore, $m_1 + 2(n_1 + n_2) = 7$, as expected (cf. Proposition 6).

Since p_7 is radially holomorphic (see Example 1) the 4D-Newton method is equivalent to its quaternion versions (see Proposition 7) and all the considerations concerning the modified Newton methods introduced in the previous section are valid. This is precisely what is illustrated in Fig. 1 which contains density plots of the functions $N(c)$ associated to the methods (from left to right, from top to bottom),

$\mathbb{H}$-NM	–	classical quaternion Newton method (see (7)),
$\mathbb{H}$-NM$_{m=2}$	–	modified quaternion Newton method, with $m = 2$ in (10),
$\mathbb{H}$-MNM	–	modified quaternion Newton method (see (11)),
$\mathbb{H}$-ANM$_1$	–	adaptative quaternion Newton method,
$\mathbb{H}$-ANM$_2$	–	adaptative quaternion Newton method,
$\mathbb{H}$-ANM$_3$	–	adaptative quaternion Newton method,

when c is chosen in the region $\Omega_1 = \{(x, y, 0, 0) \in \mathbb{R}^4 : -2 \leq x \leq 2,\ -2 \leq y \leq 2\}$.



(a) $\mathbb{H}$-NM (b) $\mathbb{H}$-NM$_{m=2}$ (c) $\mathbb{H}$-MNM

(d) $\mathbb{H}$-ANM$_1$ (e) $\mathbb{H}$-ANM$_2$ (f) $\mathbb{H}$-ANM$_3$

Fig. 1. $N(c)$ for Example 3 with $c \in \Omega_1$, $\varepsilon = 10^{-6}$ and kmax$= 20$.

We point out that, if $c \in \Omega_1$, the sequences produced by each of the aforementioned methods lie also on Ω_1 and therefore it makes sense to consider the basins of attraction of the roots

$$r_1 = 1,\ r_2 = -\tfrac{1}{2} + \tfrac{\sqrt{3}}{2}\mathbf{i},\ r_3 = -\tfrac{1}{2} - \tfrac{\sqrt{3}}{2}\mathbf{i},\ r_4 = \sqrt{2}\mathbf{i},\ r_5 = -\sqrt{2}\mathbf{i},$$

with $r_2, r_3 \in [\zeta_1] \cap \Omega_1$ and $r_4, r_5 \in [\zeta_2] \cap \Omega_1$, with respect to the iterative functions associated with each method. 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, $i = 1, \ldots, 5$ (see Fig. 2). In addition to illustrate the methods performance, this example aims to call the attention to the relation between the complex Newton methods and the quaternion ones. In fact, we can reproduce all the figures presented in Figs. 1-2 by considering p_7 as a polynomial in $\mathbb{C}$, where the order of convergence of the Newton method and its variants is well studied. More precisely, if $x_k = a_k + ib_k$ denotes the sequence produced by the a complex Newton method, converging to the complex root $x^* = a + ib$ of p_7, it is easy to see that the sequence produced by the corresponding quaternion Newton method with initial guess $c = a_0 + \omega b_0$, where ω is any unit pure quaternion, converges to the quaternion root $\zeta \in [x^*]$ such that $\zeta = a + \omega b$. As a matter of fact, one can prove that the converse can also be established, as far as we adjust the definition of $\omega(\underline{c})$. The proof of this result is beyond the scope of the paper.

When $c \in \Omega_2 = \{(0, x, y, 0) \in \mathbb{R}^4 : -2 \leq x \leq 2,\ -2 \leq y \leq 2\}$, the situation is rather different since here the sequences produced by the Newton methods lie on $\Omega_2^* = \{(a, b, c, 0) \in \mathbb{R}^4\}$ and not on Ω_2. In this case when we observe convergence to values generating spherical roots, the behavior of the $\mathbb{H}$-NM and variants is clear: if c is the initial guess, then the Newton sequence converges to the root $r \in [\zeta_1] \cap \Omega_2^*$ (or

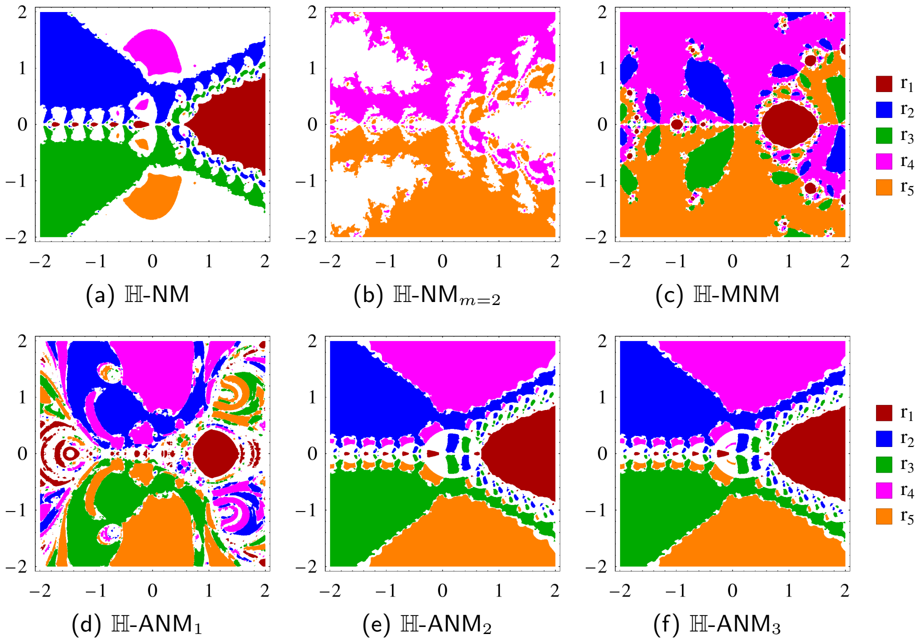


Fig. 2. Basins of attraction of the roots $r_i,\ i = 1, \dots, 5$, for Example 3, with $c \in \Omega_1$.

$r \in [\zeta_2] \cap \Omega_2^*$) such that $\omega(\underline{r}) = \pm\omega(\underline{c})$. The vector part of several sequences obtained by the use of the ℍ-NM for different choices of the initial guess (marked with the symbol $\circ$) are illustrated in Fig. 3. Figure 4 contains the basins of attraction of each spherical root. We remark that in this domain we can not observe convergence to the real root $r_1 = 1$. To check the effectiveness of the modified Newton methods proposed in last section, we compare them with the ℍ-NM$_r$ (7) (which is equivalent, as already mentioned, to the ℍ-NM$_l$). Table 1 contains estimates, for different choices of the initial guess, to the multiplicity m of the root and to the order of convergence ρ corresponding to the last three iterations computed with the stopping criterion $\varepsilon = 10^{-12}$. Numerical computations needed to produce all the tables presented in this paper have been carried out in Mathematica environment with the precision increased to 512 significant digits.

Example 4. Considerer now the polynomial

$$p_4(x) = (x - \mathbf{i} - \mathbf{j})(x^3 + 2x).$$

Since $\mathbf{i} + \mathbf{j} \in [\sqrt{2}\mathbf{i}]$, the zero-set of p_4 is $Z_{p_4} = \{0\} \cup [\sqrt{2}\mathbf{i}] = \{0\} \cup [\mathbf{i} + \mathbf{j}]$. Although we have only defined the multiplicity of a root for real polynomials, it seems natural, in the context of Newton method, to consider the root $\mathbf{i} + \mathbf{j}$ as a "double" root even if it generates a sphere of simple zeros. We point out that the polynomial $q_2(x) = x^2 + 2$ is such that $q_2(\mathbf{i} + \mathbf{j}) = q_2(-\mathbf{i} - \mathbf{j}) = 0$, but $(x - \mathbf{i} - \mathbf{j})(x + \mathbf{i} + \mathbf{j}) = x^2 - (\mathbf{i} + \mathbf{j})x - x(\mathbf{i} + \mathbf{j}) + 2 \neq q_2(x)$. In fact, one-sided polynomials can not have non spherical roots of

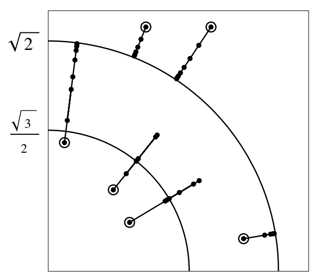


Fig. 3. Convergence to spherical roots in Ω_2 - Example 3.

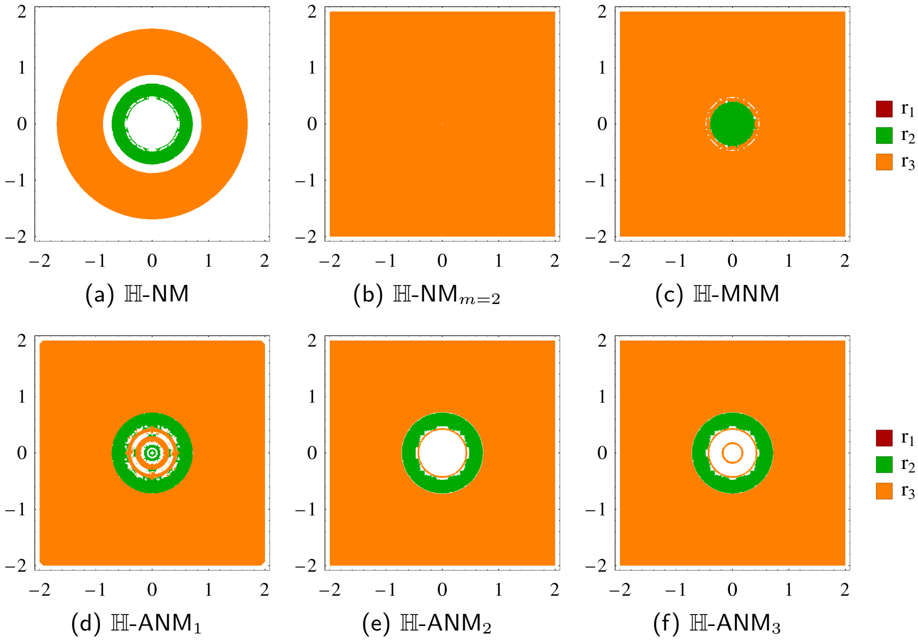
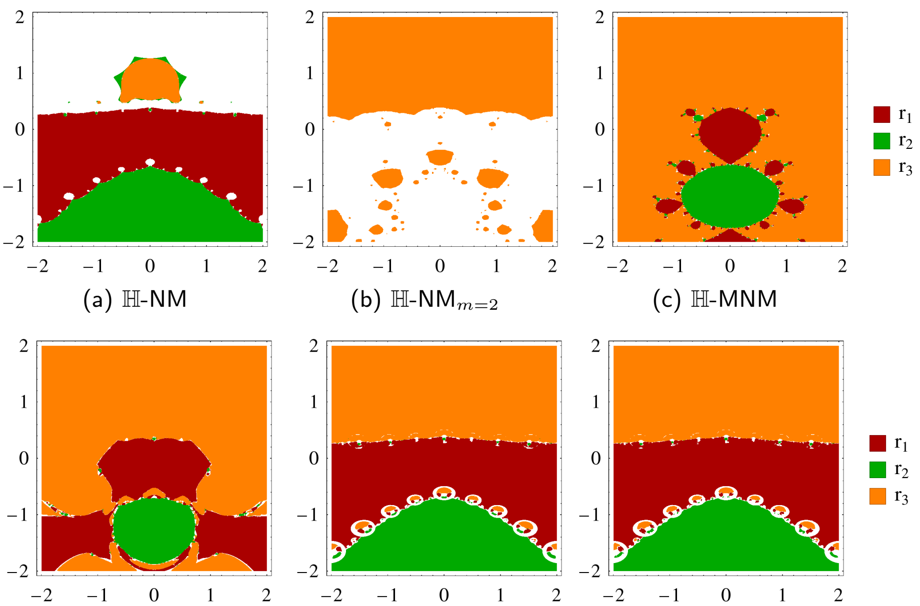


Fig. 4. Basins of attraction of the roots r_i, $i = 1, 2, 3$ for Example 3 with $c \in \Omega_2$

Table 1. ℍ-NM versus its variants for Example 3. COR - convergence to other root; ND - not defined ($10|\rho_n - \rho_{n-1}| > \min\{\rho_n, \rho_{n-1}\}$ as in [20]).

c	root	ℍ-NM		ℍ-NM $m=2$		ℍ-MNM		ℍ-ANM$_1$			ℍ-ANM$_2$			ℍ-ANM$_3$		
		k	ρ_k	k	ρ_k	k	ρ_k	k	m_k	ρ_k	k	m_k	ρ_k	k	m_k	ρ_k
		41	1.00	10	2.00	9	2.00	14	2.00	1.56	22	1.81	1.04	16	2	ND
$1.1+1.3\mathbf{i}$	$\sqrt{2}\mathbf{i}$	42	1.00	11	2.00	10	2.00	15	2.00	1.58	23	1.83	1.04	17	2	2.00
		43	1.00	12	2.00	11	2.00	16	2.00	1.63	24	1.84	1.04	18	2	2.00
		6	2.00	14		7	2.00	7	1.00	2.71	6	1.00	2.00	6	1	2.00
$-0.7-0.6\mathbf{i}$	$-\frac{1}{2}+\frac{\sqrt{3}}{2}\mathbf{i}$	7	2.00	15	COR	8	2.00	8	1.00	1.20	7	1.00	2.00	7	1	2.00
		8	2.00	16		9	2.00	9	1.00	2.66	8	1.00	2.00	8	1	2.00
		36	1.00	5	2.00	5	2.00	7	2.00	1.62	17	1.81	1.04	11	2	ND
$0.2\mathbf{i}-0.7\mathbf{j}+\mathbf{k}$	$\frac{2\mathbf{i}-7\mathbf{j}+10\mathbf{k}}{3\sqrt{\frac{17}{2}}}$	37	1.00	6	2.00	6	2.00	8	2.00	1.62	18	1.82	1.04	12	2	2.00
		38	1.00	7	2.00	7	2.00	9	2.00	1.62	19	1.84	1.04	13	2	2.00
		6	2.00	40		9	2.00	8	1.00	1.03	6	1.00	2.00	6	1	2.00
$1.2+0.3\mathbf{i}$	1	7	2.00	41	COR	10	2.00	9	1.00	2.09	7	1.00	2.00	7	1	2.00
		8	2.00	42		11	2.00	10	1.00	2.41	8	1.00	2.00	8	1	2.00



(a) ℍ-NM (b) ℍ-NM$_{m=2}$ (c) ℍ-MNM

(d) ℍ-ANM$_1$ (e) ℍ-ANM$_2$ (f) ℍ-ANM$_3$

Fig. 5. Basins of attraction of the roots r_i for Example 4 with $c \in \Omega_3$.

multiplicity $m > 1$, in the usual sense, rather than the real ones. Figure 5 illustrates the features of the roots of the above polynomial. The real root is denoted by r_1 and, to distinguish the multiplicity of the root $\mathbf{i}+\mathbf{j}$ as a "double" root from the multiplicity of $[\mathbf{i}+\mathbf{j}]$, we have assigned the label r_3 to the root $\mathbf{i}+\mathbf{j}$ and r_2 to the other roots in $[\mathbf{i}+\mathbf{j}]$. In this way, the basin of attraction of the spherical root $[\mathbf{i}+\mathbf{j}]$ corresponds to the basin of attraction of r_2 and r_3.

The numerical results (see Table 2) confirm the idea that on the hyperplane $\Omega_3 = \{(x, y, y, 0) \in \mathbb{R}^4\}$, where Proposition 7 is valid, the root $\mathbf{i}+\mathbf{j}$ behaves as double. On the other hand, on $\Omega_4 = \{(x, y, 0, 0) \in \mathbb{R}^4\}$, the conditions of Proposition 7 are not fulfilled, since quaternions of the form $c = x + y\mathbf{i} \in \Omega_4$ do not commute with $\mathbf{i}+\mathbf{j}$. As a consequence, we have to consider right and left versions of each method.

Table 2. ℍ-NM versus its variants for Example 4 and $c \in \Omega_3$. NC - no convergence.

c	root	ℍ-NM		ℍ-NM $m=2$		ℍ-MNM		ℍ-ANM$_1$			ℍ-ANM$_2$			ℍ-ANM$_3$		
		k	ρ_k	k	ρ_k	k	ρ_k	k	m_k	ρ_k	k	m_k	ρ_k	k	m_k	ρ_k
$1+\mathbf{i}+\mathbf{j}$	$\mathbf{i}+\mathbf{j}$	40	1.00	6	2.00	6	2.00	8	2.00	1.61	17	1.88	1.04	11	2	2.00
		41	1.00	7	2.00	7	2.00	9	2.00	1.62	18	1.90	1.04	12	2	2.00
		42	1.00	8	2.00	8	2.00	10	2.00	1.62	19	1.91	1.04	13	2	2.00
$0.2-0.1\mathbf{i}-0.1\mathbf{j}$	0	4	1.99	NC		4	2.03	5	1.00	2.17	4	1.00	1.99	4	1	1.99
		5	2.00			5	2.00	6	1.00	2.00	5	1.00	2.00	5	1	2.00
		6	2.00			6	2.00	7	1.00	2.00	6	1.00	2.00	6	1	2.00
$0.5-\mathbf{i}-\mathbf{j}$	$-\mathbf{i}-\mathbf{j}$	7	2.00	NC		7	2.00	7	1.00	1.00	7	1.00	2.00	7	1	2.00
		8	2.00			8	2.00	8	1.00	2.00	8	1.00	2.00	8	1	2.00
		9	2.00			9	2.00	9	1.00	2.50	9	1.00	2.00	9	1	2.00
$0.1+0.5\mathbf{i}+0.5\mathbf{j}$	$\mathbf{i}+\mathbf{j}$	40	1.00	7	2.00	7	2.00	11	2.00	1.56	17	1.89	1.04	11	2	2.00
		41	1.00	8	2.00	8	2.00	12	2.00	1.52	18	1.90	1.04	12	2	2.00
		42	1.00	9	2.00	9	2.00	13	2.00	1.65	19	1.91	1.03	13	2	2.00

The performances of both versions are illustrated in Figs. 6 and 7. We can see on Table 3 that the right quaternion Newton methods seem to converge slower than their left versions. In addition it was not

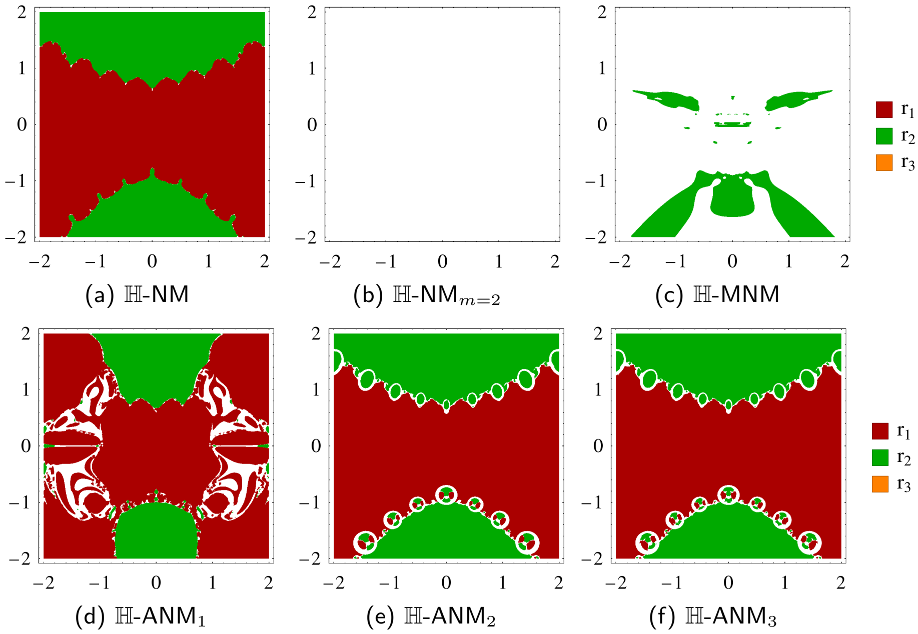


Fig. 6. Basins of attraction of the roots r_i for Example 4 with $c \in \Omega_4$ - left versions.

possible to observe convergence to the real root using any right Newton methods or variants, whilst the left versions seem not to converge to $\mathbf{i}+\mathbf{j}$. For both right and left versions, the three ℍ-ANM produced

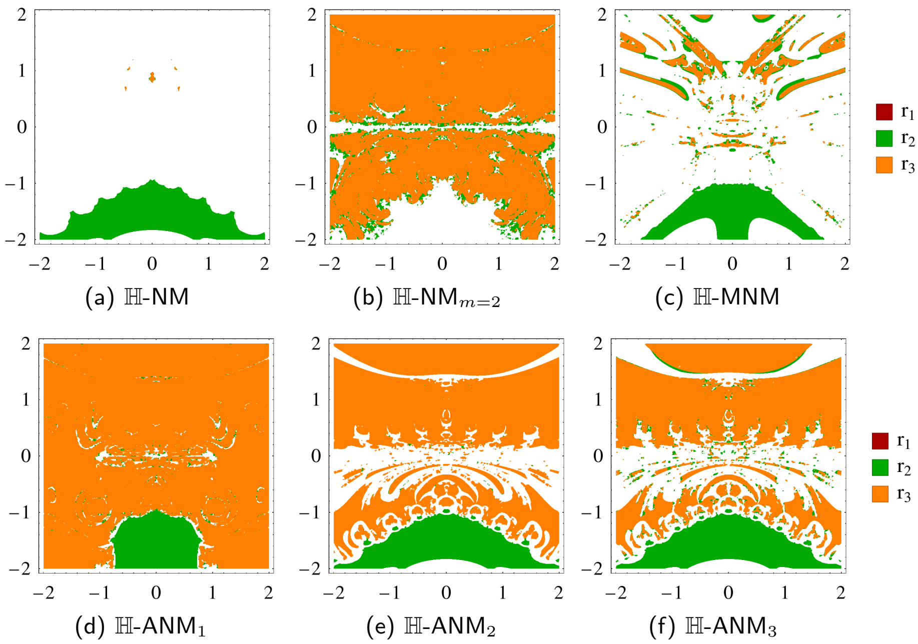


Fig. 7. Basins of attraction of the roots r_i for Example 4 with $c \in \Omega_4$ - right versions.

1 as the estimate of m. We do not compare the values of the (right and left) computational order of convergence since most of the times the right and left version of a method do not converge to the same root.

Table 3. Left and right versions of $\mathbb{H}$-NM versus its variants for Example 4 and $c \in \Omega_4$. NC - no convergence; $*$ - convergence to a different root in $[\mathbf{i}+\mathbf{j}]$.

c	$\mathbf{i}+\mathbf{j}+2\mathbf{k}$		$-\mathbf{j}+2\mathbf{k}$		$1+2\mathbf{i}$	
	Left	Right	Left	Right	Left	Right
$x*$	$[\mathbf{i}+\mathbf{j}]$	$\mathbf{i}+\mathbf{j}$	$[\mathbf{i}+\mathbf{j}]$	$\mathbf{i}+\mathbf{j}$	$[\mathbf{i}+\mathbf{j}]$	$\mathbf{i}+\mathbf{j}$
$\mathbb{H}$-NM	9	46	8	53	10	46
$\mathbb{H}$-NM$_{m=2}$	NC	25	NC	29	NC	29
$\mathbb{H}$-MNM	10*	49	8*	88	14*	50
$\mathbb{H}$-ANM1	12*	26	11*	26	12*	26
$\mathbb{H}$-ANM2	9	27	8	32	10	28
$\mathbb{H}$-ANM3	9	31	8	36	10	31

Further investigation on the behavior of the $\mathbb{H}$-NM and variants, whenever the assumptions of Proposition 7 are not met, is needed in order to justify the experimental results.

Acknowledgment The research was partially supported by the Research Centre of Mathematics of the University of Minho with the Portuguese Funds from the "Fundação para a Ciência e a Tecnologia", through the Project PEstOE/MAT/UI0013/2014.

References

1. Colombo, F., Sabadini, I., Struppa, D.C.: Noncommutative functional calculus, Progress in Mathematics, vol. 289. Birkhäuser/Springer Basel AG, Basel (2011), Theory and applications of slice hyperholomorphic functions
2. Cullen, C.G.: An integral theorem for analytic intrinsic functions on quaternions. Duke Math. J. 32, 139–148 (1965)
3. De Leo, S., Ducati, G., Leonardi, V.: Zeros of unilateral quaternionic polynomials. Electron. J. Linear Algebra 15, 297–313 (2006)
4. Eilenberg, S., Niven, I.: The "fundamental theorem of algebra" for quaternions. Bull. Amer. Math. Soc. 50, 246–248 (1944)
5. Falcão, M.I.: Newton method in quaternion context. Appl. Math. Comput. 236, 458–470 (2014)
6. Falcão, M.I., Miranda, F.: Quaternions: A Mathematica package for quaternionic analysis. In: Murgante, B., Gervasi, O., Iglesias, A., Taniar, D., Apduhan, B. (eds.) Computational Science and Its Applications - ICCSA 2011, Lecture Notes in Computer Science, vol. 6784, pp. 200–214. Springer-Verlag Berlin Heidelberg (2011)
7. Fueter, R.: Die funktionentheorie der differetialgleichungen $\Delta u = 0$ und $\Delta\Delta u = 0$ mit vier reellen variablen. Comm. Math. Helv. (7), 307–330 (1934-35)
8. Galántai, A., Hegedűs, C.J.: A study of accelerated Newton methods for multiple polynomial roots. Numer. Algorithms 54(2), 219–243 (2010)
9. Gentili, G., Struppa, D.C.: A new approach to Cullen-regular functions of a quaternionic variable. C. R. Math. Acad. Sci. Paris 342(10), 741–744 (2006)
10. Gentili, G., Struppa, D.C.: A new theory of regular functions of a quaternionic variable. Adv. Math. 216(1), 279–301 (2007)
11. Gentili, G., Struppa, D.C., Vlacci, F.: The fundamental theorem of algebra for Hamilton and Cayley numbers. Math. Z. 259(4), 895–902 (2008)
12. Gürlebeck, K., Habetha, K., Sprößig, W.: Holomorphic functions in the plane and n-dimensional space. Birkhäuser Verlag, Basel (2008)
13. Janovská, D., Opfer, G.: Computing quaternionic roots by Newton's method. Electron. Trans. Numer. Anal. 26, 82–102 (2007)
14. Janovská, D., Opfer, G.: The classification and the computation of the zeros of quaternionic, two-sided polynomials. Numer. Math. 115(1), 81–100 (2010)
15. Janovská, D., Opfer, G.: A note on the computation of all zeros of simple quaternionic polynomials. SIAM J. Numer. Anal. 48(1), 244–256 (2010)
16. Jia, Z., Cheng, X., Zhao, M.: A new method for roots of monic quaternionic quadratic polynomial. Comput. Math. Appl. 58(9), 1852–1858 (2009)
17. Kalantari, B.: Algorithms for quaternion polynomial root-finding. J. Complexity 29(3-4), 302–322 (2013)
18. Niven, I.: Equations in quaternions. Amer. Math. Monthly 48, 654–661 (1941)

19. Ostrowski, A.M.: Solution of equations and systems of equations. Pure and Applied Mathematics, Vol. IX. Academic Press, New York-London (1960)
20. Özban, A.Y.: Some new variants of Newton's method. Appl. Math. Lett. 17(6), 677–682 (2004)
21. Pogorui, A., Shapiro, M.: On the structure of the set of zeros of quaternionic polynomials. Complex Var. Theory Appl. 49(6), 379–389 (2004)
22. Rinehart, R.F.: Elements of a theory of intrinsic functions on algebras. Duke Math. J 27, 1–19 (1960)
23. Serôdio, R., Siu, L.S.: Zeros of quaternion polynomials. Appl. Math. Lett. 14(2), 237–239 (2001)
24. Sprößig, W.: On operators and elementary functions in Clifford analysis. Journal for Analysis and its Applications 18(2), 349–360 (1999)
25. Traub, J.F.: Iterative Methods for the Solution of Equations. Prentice-Hall, Inc. Englewood Cliffs, NJ (1964)