A Two-step Quaternionic Root-finding Method

M. Irene Falcão[a,c] Fernando Miranda[a,c] Ricardo Severino[a,c] M. Joana Soares[b]

[a] CMAT - Centre of Mathematics, University of Minho, Portugal
[b] NIPE - Economic Policies Research Unit, University of Minho, Portugal
[c] Department of Mathematics, University of Minho, Portugal

Information

Keywords:
Quaternions · Zeros · Weierstrass method

Original publication:
Lecture Notes in Computer Science, vol. 14104 pp. 708–721, 2023
DOI: 10.1007/978-3-031-37105-9_47
https://link.springer.com

Abstract

In this paper we present a new method for determining simultaneously all the simple roots of a quaternionic polynomial. The proposed algorithm is a two-step iterative Weierstrass-like method and has cubic order of convergence. We also illustrate a variation of the method which combines the new scheme with a recently proposed deflation procedure for the case of polynomials with spherical roots.

1 Introduction

In this paper we focus on the problem of approximating the zeros of polynomials of the form

$$P(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0,\ a_n \neq 0, \tag{1}$$

where the coefficients a_k are quaternions.

Newton-like methods, based on quaternion arithmetic, have been considered in the past [6, 7, 14] to obtain approximations to the zeros of special functions. One important issue in the framework of Newton or any similar derivative-based method is related to the notions of regularity of a quaternionic function and its derivatives, which in turn restrict the application of this class of methods to a certain class of quaternionic functions (see [6] for details).

One of the most frequently used methods for simultaneous approximation of all simple polynomial zeros is the Weierstrass method [22], also known in the literature as the Durand-Kerner method [3] or Dochev method [2]. This is a free-derivative method relying on the factorization of the polynomial which makes its extension to the quaternion setting possible. Such generalization was derived in [5], where it was also proved that, as in the classical case, the method has quadratic order of convergence for the simple roots of a polynomial.

More recently [9], an approach combining a deflation procedure with the Weierstrass method allowed to obtain approximations also to the non-isolated zeros of P. In this work, we suggest an improvement of the quaternionic Weierstrass method, by using a two-step strategy and prove that this new procedure has cubic order of convergence. We also illustrate, by examples, that this new method can be combined with the aforementioned deflation process to obtain the isolated and non-isolated zeros of P.

The paper is organized as follows: in Sect. 2 we recall some results concerning the ring of quaternionic polynomials, fundamental throughout the paper. Section 3 contains the main result of the paper: a two-step Weierstrass method which we prove to have, under certain assumptions, cubic order of convergence. In Sect. 4 we illustrate the performance of the method by considering some examples and computing the corresponding computational order of convergence. We also apply, in Examples 3 and 4, the technique described in [9], to obtain successfully both the isolated and non-isolated zeros of the polynomial under consideration. The paper ends with some remarks and conclusions.

2 Basic Definitions and Results

We start by first recalling some aspects of the algebra of quaternions $\mathbb{H}$ needed for this work; for more details on this algebra, we refer to [13, 15, 23]. Here we will adopt the following notation: a quaternion x is an element of the noncommutative division algebra $\mathbb{H}$ of the form $x = x_0 + x_1\mathbf{i} + x_2\mathbf{j} + x_3\mathbf{k},\ x_i \in \mathbb{R}$, where the imaginary units $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$ satisfy the multiplication rules

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

In analogy with the complex case, we define the real part of x, $\mathrm{Re}(x) := x_0$, the conjugate of x, $\bar{x} := x_0 - \mathbf{i}x_1 - \mathbf{j}x_2 - \mathbf{k}x_3$ and the norm of x, $|x| := \sqrt{x\bar{x}} = \sqrt{\bar{x}x}$. Any quaternion $x \neq 0$ is invertible and its inverse is given by $x^{-1} = \frac{\bar{x}}{|x|^2}$.

On $\mathbb{H}$, the relation $q \sim q'$ if $\mathrm{Re}\, q = \mathrm{Re}\, q'$ and $|q| = |q'|$, is an equivalence relation and, as usual, $[q] := \{q' \in \mathbb{H} : q \sim q'\}$ denotes the equivalence class of q.

In this work we consider polynomials P of the form (1), i.e., polynomials whose coefficients a_k are quaternions located only on the left-hand side of the powers; similar results could be derived by considering the coefficients on the right.

The set of polynomials of the form (1), with the addition and multiplication defined as in the commutative case, is a ring, usually denoted by $\mathbb{H}[x]$ and called the ring of (left) one-sided polynomials.

We introduce now some definitions and results concerning $\mathbb{H}[x]$, which will play an important role in the sequel (see [11, 15] for other details). We mainly follow the notions and notations of [5, 9].

A quaternion q is a zero or a root of P, if $P(q) = 0$, being the evaluation of P at q defined as $P(q) := a_n q^n + a_{n-1}q^{n-1} + \cdots + a_1 q + a_0$. We use the notation $\mathbf{Z}_P$ to represent the set of all the zeros of P. A zero q is called an isolated zero of P, if $[q]$ contains no other zeros of P, otherwise the zero is called a spherical zero of P; in this last case all the elements of $[q]$ are zeros of P (we point out that $[q]$, $q \in \mathbb{H} \setminus \mathbb{R}$, can be identified with the three-dimensional sphere in the hyperplane $\{(x_0, x, y, z) \in \mathbb{R}^4 : x_0 = q_0\}$, with center $(q_0, 0, 0, 0)$ and radius $\sqrt{q_1^2 + q_2^2 + q_3^2}$).

The conjugate of P, denoted by $\overline{P}$, is obtained by conjugating the coefficients of P; the characteristic polynomial of a quaternion q is the real polynomial

$$\Psi_q(x) := (x - q)(x - \overline{q}) = x^2 - 2\,\mathrm{Re}(q)\,x + |q|^2. \tag{2}$$

Concerning the zero-structure and the factorization of polynomials in $\mathbb{H}[x]$ we recall the following results (see e.g. [1, 11, 15, 18] for the proofs) essential for next section.

Result 1. *Consider the factorization of a polynomial $P \in \mathbb{H}[x]$ in the form $P(x) = L(x)R(x)$ with $L, R \in \mathbb{H}[x]$.*

1. *If $q \in \mathbb{H}$ is a zero of the right factor R, then q is a zero of the product P.*

2. *When q is a zero of P which is not a zero of R, we have*

$$P(q) = L(\tilde{q})R(q), \tag{3}$$

where $\tilde{q} := R(q)qR(q)^{-1}$ is a zero of L;

3. *If $L \in \mathbb{R}[x]$, then*

$$P(q) = R(q)L(q). \tag{4}$$

Result 2. *Let P be a monic polynomial of degree $n\,(n \geq 1)$ in $\mathbb{H}[x]$. Then,*

1. *P admits a factorization into linear factors*

$$P(x) = (x - x_n)(x - x_{n-1}) \cdots (x - x_1), \tag{5}$$

   being the quaternions $x_1, \ldots, x_n$ called factor terms of P;

2. *$\mathbf{Z}_P \subseteq \bigcup_{i=1}^{n} [x_i]$ and each of the equivalence classes $[x_i]; i = 1, \ldots, n$, contains (at least) a zero of P;*

3. *If $P(x) = (x - y_n)(x - y_{n-1}) \cdots (x - y_1)$ is another factorization of P into linear factors, then there exists a permutation π of $(1, 2, \ldots, n)$ and $h_i \in \mathbb{H}$ such that $y_{\pi(i)} = h_i x_i h_i^{-1}; i = 1, \ldots, n$.*

Result 3. *Let P be a monic polynomial of degree n in $\mathbb{H}[x]$ with n isolated roots and let (5) be one of its factorizations.*

1. *The equivalence classes of the factor terms $x_1, \ldots, x_n$ in (5) are distinct;*

2. *Consider the polynomials*

$$R_i := \prod_{j=1}^{i-1} (x - \bar{x}_j).$$

   The relation between the roots $\zeta_1, \ldots, \zeta_n$ and the factor terms $x_1, \ldots, x_n$ of P is the following:

$$\zeta_i = R_i(x_i)\, x_i \left(R_i(x_i)\right)^{-1} \quad \text{and} \quad x_i = \bar{R}_i(\zeta_i)\, \zeta_i \left(\bar{R}_i(\zeta_i)\right)^{-1}, \tag{6}$$

   for $i = 1, \ldots, n$.

3 A Two-Step Weierstrass Method

From now on, we assume, for simplicity, that the polynomial P in (1) is monic, i.e., $a_n = 1$.

In the classical case, i.e., when the coefficients of P are complex, the popular Weierstrass method can be written as

$$\tilde{z}_i = z_i - W_i(z_i);\ \ i = 1, \ldots, n,$$

where the so-called *Weierstrass correction* W_i is the rational function

$$W_i(x) = \frac{P(x)}{\prod\limits_{\substack{j=1 \\ j \neq i}}^{n} (x - z_j)}.$$

For simplicity, we use $\tilde{z}_i$ and z_i to denote, respectively, the $k+1$ and the k iterates of the method. If all the zeros of P are simple,[1] and we start with sufficiently close approximations to the roots, this method has quadratic order of convergence [2].

In [5] the authors of this paper extended the Weierstrass method in its sequential version to quaternionic context, by considering the scheme

$$\tilde{z}_i = z_i - \mathcal{P}_i(z_i)\ \left(\mathcal{Q}_i(z_i)\right)^{-1};\ \ i = 1, \ldots, n,$$

where $\mathcal{P}_i(x) = \mathcal{L}_i(x) P(x) \mathcal{R}_i(x)$, with

$$\mathcal{L}_i(x) = \prod_{j=i+1}^{n} \left(x - \overline{z_j}\right), \quad \mathcal{R}_i(x) = \prod_{j=1}^{i-1} \left(x - \overline{\tilde{z}_j}\right) \tag{7}$$

[1] The zeros are all simple if they are all distinct and isolated.

and (cf. (2))

$$\mathcal{Q}_i(x) = \prod_{j=1}^{i-1} \Psi_{\tilde{z}_j}(x) \prod_{j=i+1}^{n} \Psi_{z_j}(x).$$

They also showed that, under certain conditions, the method converges quadratically to the factor terms x_i of P.

One can find in the literature several simultaneous methods based on the Weierstrass corrections with higher order of convergence (see e.g. [20]), which usually depend on the derivatives of the polynomial. If we are looking for a higher order free derivative method, the use of multi-step methods can be a solution. The well-known two-step Newton method [21],

$$\begin{cases} y_i = z_i - \frac{P(z_i)}{P'(z_i)} \\ \tilde{z}_i = y_i - \frac{P(y_i)}{P'(y_i)} \end{cases}$$

which has a fourth order convergence (see also [16]), can be easily adapted to quaternionic context by replacing first $\frac{P(x)}{P'(x)}$ by $W_i(x)$. In the complex case there are more competitive methods available (see e.g. [19]), but on the contrary, to the best of our knowledge, the quaternionic method that we are going to introduce, based on this strategy, is the one with the highest order of convergence.

Next theorem contains the main result of the paper.

Theorem 1. *Let P be a monic polynomial of degree n in $\mathbb{H}[x]$ with n isolated distinct roots and, for $i = 1, \ldots, n$; $k = 0, 1, 2, \ldots$, let*

$$\begin{cases} y_i^{(k)} = z_i^{(k)} - \mathcal{P}_i^{(k)}(z_i^{(k)}) \left(\mathcal{Q}_i^{(k)}(z_i^{(k)})\right)^{-1} \\ \\ z_i^{(k+1)} = y_i^{(k)} - \mathcal{P}_i^{(k)}(y_i^{(k)}) \left(\mathcal{Q}_i^{(k)}(y_i^{(k)})\right)^{-1} \end{cases} \tag{8}$$

where $\mathcal{P}_i^{(k)}(x) = \left(\mathcal{L}_i^{(k)}(x) P(x) \mathcal{R}_i^{(k)}(x)\right)$ with

$$\mathcal{L}_i^{(k)}(x) := \prod_{j=i+1}^{n} \left(x - \overline{z_j^{(k)}}\right), \tag{9}$$

$$\mathcal{R}_i^{(k)}(x) := \prod_{j=1}^{i-1} \left(x - \overline{z_j^{(k+1)}}\right) \tag{10}$$

and

$$\mathcal{Q}_i^{(k)}(x) := \prod_{j=1}^{i-1} \Psi_{z_j^{(k+1)}}(x) \prod_{j=i+1}^{n} \Psi_{z_j^{(k)}}(x). \tag{11}$$

If the initial approximations $z_i^{(0)}$ are sufficiently close to the factor terms x_i in a factorization of P in the form (5), then the sequences $\{z_i^{(k)}\}$ converge to x_i with cubic order of convergence.

Proof. The proof is an adaptation of the proof in [5], now for the case where each iteration involves two steps. In order to make the document complete, we have included all details.

For simplicity of notation, we write the scheme (8) in the form

$$\begin{cases} y_i = z_i - \mathcal{P}_i(z_i) \left(\mathcal{Q}_i(z_i)\right)^{-1} \\ \\ \tilde{z}_i = y_i - \mathcal{P}_i(y_i) \left(\mathcal{Q}_i(y_i)\right)^{-1} \end{cases}$$

omitting all the superscripts corresponding to the iteration number.

Assume that z_i are approximations to x_i with errors ε_i, i.e.,

$$\varepsilon_i := x_i - z_i, \; i = 1, \dots, n, \tag{12}$$

and denote by ε the maximum error, i.e., $\varepsilon := \max_i |\varepsilon_i|$. We assume that ε is *small enough*, i.e., that z_i are *sufficiently good* approximations to x_i. Since, in each iteration, the first step corresponds to the classical quaternionic Weierstrass method, we known that

$$\varepsilon_i' := x_i - y_i, \; i = 1, \dots, n,$$

are such that

$$\varepsilon' := \max_i |\varepsilon_i| = \mathcal{O}\big(\varepsilon^2\big). \tag{13}$$

We now prove, by complete induction on i, that the second step iterates $\tilde{z}_i$ are approximations to x_i with errors $\tilde{\varepsilon}_i$ such that

$$\tilde{\varepsilon}_i = \mathcal{O}\big(\varepsilon^3\big).$$

Base case: We need to prove that $\tilde{\varepsilon}_1 = \mathcal{O}(\varepsilon^3)$.

Observe that the polynomial P can be written, by the use of (12), as

$$\begin{aligned} P(x) &= \prod_{j=1}^{n} (x - x_{n-j+1}) = \prod_{j=1}^{n-1} (x - z_{n-j+1} - \varepsilon_{n-j+1})(x - y_1 - \varepsilon_1') \\ &= \Big(\prod_{j=1}^{n-1} (x - z_{n-j+1}) + \mathscr{E}_1(x) \Big)(x - y_1 - \varepsilon_1'), \end{aligned}$$

where $\mathscr{E}_1(x)$ designates a remainder polynomial consisting of a sum of $n-1$ terms of the form

$$-(x - z_n)(x - z_{n-1}) \dots (x - z_{j-1}) \varepsilon_j (x - z_{j+1}) \dots (x - z_2), \; j = 2, \dots, n,$$

with terms consisting of products involving at least two ε_j's. We may assume that we are working in a bounded domain $\mathcal{D}$ of $\mathbb{H}$ (a sufficiently large disk containing all z_i) and therefore, we have

$$\mathscr{E}_1(\alpha) = \mathcal{O}(\varepsilon), \; \forall \alpha \in \mathcal{D}. \tag{14}$$

Taking into account the definition (7) of the polynomial $\mathcal{L}_1$, P can be written as

$$P(x) = \Big(\overline{\mathcal{L}}_1(x) + \mathscr{E}_1(x) \Big)(x - y_1 - \varepsilon_1') = \overline{\mathcal{L}}_1(x)(x - y_1 - \varepsilon_1') + \mathscr{E}_1(x)(x - y_1 - \varepsilon_1').$$

If we multiply P on the left by $\mathcal{L}_1$ and evaluate the resulting polynomial at the point $x = y_1$, we obtain, recalling the results (3) and (4) in Result 1 and the definition (11) of $\mathcal{Q}_1$,

$$\big(\mathcal{L}_1 P\big)(y_1) = -\varepsilon_1' \mathcal{Q}_1(y_1) - \Big(\mathcal{L}_1 \mathscr{E}_1\Big)(\hat{z}_1)\, \varepsilon_1',$$

where $\hat{z}_1 = \varepsilon_1' z_1 (\varepsilon_1')^{-1}$. Using now (14) we can write

$$\big(\mathcal{L}_1 P\big)(y_1) = -\varepsilon_1' \mathcal{Q}_1(y_1) + \mathcal{O}(\varepsilon \varepsilon_1').$$

Since we are assuming that the equivalence classes $[x_j]$ are distinct then, for sufficiently small ε, $|\mathcal{Q}_1(y_1)|$ is bounded away from zero and so, by multiplying both sides of the above equality on the right by $\big(\mathcal{Q}_1(y_1)\big)^{-1}$, we obtain

$$\left(\mathcal{L}_1 P\right)(y_1)\; \left(\mathcal{Q}_1(y_1)\right)^{-1} = -\varepsilon_1' + \mathcal{O}(\varepsilon \varepsilon_1'),$$

which means by (8), that

$$x_1 = y_1 - \left(\mathcal{L}_1 P\right)(z_1) \left(\mathcal{Q}_1(y_1)\right)^{-1} + \mathcal{O}(\varepsilon \varepsilon_1').$$

Finally, we may conclude from (13) that the next approximation to x_1

$$\tilde{z}_1 = y_1 - (\mathcal{L}_1 P)(y_1)\left(\mathcal{Q}_1(y_1)\right)^{-1}$$

is such that

$$\tilde{\varepsilon}_1 = x_1 - \tilde{z}_1 = \mathcal{O}(\varepsilon\varepsilon_1') = \mathcal{O}(\varepsilon^3).$$

Induction Step: We now prove that $\tilde{z}_i$ approximates x_i with an error $\mathcal{O}(\varepsilon^3)$, assuming that, for $j = 1, \dots, i-1$, $\tilde{z}_j$ are $\mathcal{O}(\varepsilon^3)$ approximations to x_j.

Define the polynomials

$$L_i(x) = \prod_{j=i+1}^{n} (x - \bar{x}_j) \quad \text{and} \quad R_i(x) = \prod_{j=1}^{i-1} (x - \bar{x}_j)$$

which can be written as

$$L_i(x) = \prod_{j=i+1}^{n} (x - \bar{z}_j - \bar{\varepsilon}_j) = \prod_{j=i+1}^{n} (x - \bar{z}_j) + \mathscr{E}_i(x) = \mathcal{L}_i(x) + \mathscr{E}_i(x)$$

and

$$R_i(x) = \prod_{j=1}^{i-1} (x - \tilde{z}_j - \tilde{\varepsilon}_j) = \prod_{j=1}^{i-1} (x - \tilde{z}_j) + \tilde{\mathscr{E}}_i(x) = \mathcal{R}_i(x) + \tilde{\mathscr{E}}_i(x),$$

where $\mathscr{E}_i$ and $\tilde{\mathscr{E}}_i$ are remainder polynomials defined similarly to $\mathscr{E}_1$, with the appropriate modifications. Since $\mathscr{E}_i$ is a sum of terms, all of which involve at least the product by a $\bar{\varepsilon}_j$ ($j \in \{i+1, \dots, n\}$), we conclude that $\mathscr{E}_i(\alpha) = \mathcal{O}(\varepsilon)$. On the other hand $\tilde{\mathscr{E}}_i$ is a sum of terms involving at least the product by an $\bar{\tilde{\varepsilon}}_j$ ($j \in \{1, \dots, i-1\}$), which means that we can write, using the induction hypothesis,

$$\tilde{\mathscr{E}}_i(\alpha) = \mathcal{O}(\varepsilon^3), \ \forall \alpha \in \mathcal{D}.$$

Therefore the polynomial P can be written as

$$P(x) = \bar{L}_i(x)(x - x_i)\bar{R}_i(x) = \big(\bar{\mathcal{L}}_i(x) + \bar{\mathscr{E}}_i(x)\big)(x - y_i - \varepsilon_i')\big(\bar{\mathcal{R}}_i(x) + \bar{\tilde{\mathscr{E}}}_i(x)\big).$$

Multiplying both sides of the last equality on the left by $\mathcal{L}_i$ and on the right by $\mathcal{R}_i$ and evaluating at $x = y_i$, we obtain

$$\begin{aligned}\big(\mathcal{L}_i P \mathcal{R}_i\big)(y_i) = &\Big(\mathcal{L}_i \overline{\mathcal{L}}_i \overline{\mathcal{R}}_i \mathcal{R}_i (x - y_i - \varepsilon_i')\Big)(y_i) + \Big(\mathcal{L}_i \overline{\mathcal{R}}_i \mathcal{R}_i \bar{\mathscr{E}}_i (x - y_i - \varepsilon_i')\Big)(y_i) \\ &+ \Big(\mathcal{L}_i \overline{\mathcal{L}}_i (x - y_i - \varepsilon_i') \bar{\tilde{\mathscr{E}}}_i \mathcal{R}_i\Big)(y_i) + \Big(\mathcal{L}_i \mathscr{E}_i (x - y_i - \varepsilon_i') \bar{\tilde{\mathscr{E}}}_i \mathcal{R}_i\Big)(y_i),\end{aligned}$$

where we made use of the result that, since $\mathcal{R}_i \overline{\mathcal{R}}_i$ is a real polynomial, it commutes with any other polynomial. Observing that $\mathcal{L}_i \overline{\mathcal{L}}_i \overline{\mathcal{R}}_i \mathcal{R}_i$ is the real polynomial $\mathcal{Q}_i$, using again the results (3) and (4) in Result 1 and having in mind the form of the remainder polynomials $\mathscr{E}_i$ and $\tilde{\mathscr{E}}_i$, we can write

$$\begin{aligned}\big(\mathcal{L}_i P \mathcal{R}_i\big)(y_i) &= -\varepsilon_i \mathcal{Q}_i(y_i) - \big(\mathcal{L}_i \overline{\mathcal{R}}_i \mathcal{R}_i \mathscr{E}_i\big)(\hat{y}_i)\varepsilon_i' + \mathcal{O}(\varepsilon^3) \\ &= -\varepsilon_i' \mathcal{Q}_i(y_i) + \mathcal{O}(\varepsilon\varepsilon_i') + \mathcal{O}(\varepsilon^3), \end{aligned} \tag{15}$$

where $\hat{y}_i = \varepsilon_i' y_i (\varepsilon_i')^{-1}$. Since $|\mathcal{Q}_i(y_i)|$ is bounded away from zero, multiplying (15) on the right by $\left(\mathcal{Q}_i(z_i)\right)^{-1}$ leads to

$$\left(\mathcal{L}_i P \mathcal{R}_i\right)(y_i) \ \left(\mathcal{Q}_i(y_i)\right)^{-1} = -\varepsilon_i + \mathcal{O}(\varepsilon\varepsilon_i') + \mathcal{O}(\varepsilon^3)$$

or, equivalently, recalling the definition of the errors ε_i',

$$\left(\mathcal{L}_i P \mathcal{R}_i\right)(y_i) \ \left(\mathcal{Q}_i(y_i)\right)^{-1} = y_i - x_i + \mathcal{O}(\varepsilon^3).$$

This proves that

$$\tilde{z}_i = y_i - \left(\mathcal{L}_i P \mathcal{R}_i\right)(y_i) \ \left(\mathcal{Q}_i(y_i)\right)^{-1}$$

is an $\mathcal{O}(\varepsilon^3)$ approximation to x_i. $\square$

Remark 1. The use of the Weierstrass method in its sequential version is essential to reach the cubic order of convergence, as a careful analysis of the proof reveals; for more details see [5, Remark 3]. This is the reason why the final order of convergence of this quaternionic two-step method is three instead of four, as in the complex case.

Remark 2. We point out that, for each $\tilde{z}_i$, the scheme performs four polynomial evaluation per iteration, which corresponds to the number of evaluations required by two iterations of the one-step Weierstrass method. However the number of operations involved in each iteration is substantially less, since the demanding process of constructing the polynomials $\mathcal{P}_i$ and $\mathcal{Q}_i$ is done just one time per iteration. Details on evaluation schemes of polynomials with quaternion floating point coefficients from the complexity and stability point of view can be obtained in [4].

As in the case of the classical quaternionic Weierstrass method, the iterative scheme (8) can produce, not only the factor terms, but also the roots of the polynomial. Using the relations (6) between the roots and factor terms of a polynomial and the arguments of the proof of the classical case [5, Theorem 6], the following result can be easily obtained.

Theorem 2. *Let P be a monic polynomial of degree n in $\mathbb{H}[x]$ with n isolated distinct roots and let $\{z_i^{(k)}\}$ be the sequences defined by the two-step Weierstrass iterative scheme (8)–(11) under the assumptions of Theorem 1. Finally, let $\{\zeta_i^{(k)}\}$ be the sequences defined by*

$$\zeta_i^{(k+1)} := \mathcal{R}_i^{(k)}(z_i^{(k+1)})\, z_i^{(k+1)} \left(\mathcal{R}_i^{(k)}(z_i^{(k+1)})\right)^{-1};\ k = 0, 1, 2, \ldots, \tag{16}$$

where $\mathcal{R}_i^{(k)}$ are the polynomials given by (10). Then, $\{\zeta_1^{(k)}\}, \ldots, \{\zeta_n^{(k)}\}$ converge to the roots of P with cubic order of convergence.

Remark 3. Observe that the polynomials $\mathcal{R}_i$ in (16) used to obtain the roots are the same polynomials presented in (8) to obtain the factor terms.

4 Numerical Examples

We illustrate the performance of the two-step quaternionic Weierstrass method (8)-(11) by considering several examples.

For the first two experiments we have used the Mathematica add-on application `QuaternionAnalysis` [17] specially designed for symbolic manipulation of quaternion valued functions together with the collection of functions `QPolynomial` [8, 10] for solving polynomial problems in $\mathbb{H}[x]$.

To evaluate the quality of the approximations produced by the numerical scheme, all the examples were constructed so that the exact solution ζ is known. In this way, the error $\varepsilon^{(k)}$ in each iteration k is computed as

$$\varepsilon^{(k)} = \max_i \{|\zeta_i - z_i^{(k)}|\},$$

where $z_i^{(k)}$ is given by (8). To obtain estimates for ρ, the local order of convergence of the method, we used the following computational estimate (see e.g. [12] for details)

$$\rho \approx \rho^{(k)} := \frac{\log \varepsilon^{(k)}}{\log \varepsilon^{(k-1)}}.$$

We point out that, in some cases, we had to take advantages of the fact that the Mathematica system allows to carry out the numerical computations using arbitrary precision arithmetic.

Example 1. Consider the polynomial

$$P(x) = x^3 + (3 + 3\mathbf{i} + 3\mathbf{j} + 5\mathbf{k})x^2 + (-3 + \mathbf{i} - 3\mathbf{j} + 17\mathbf{k})x + 2 - 16\mathbf{i} - 6\mathbf{j} + 8\mathbf{k}.$$

This polynomial was constructed, with the help of (6), so that its factor terms and roots are, respectively,

$$x_1 = -2 - \mathbf{j} - \mathbf{k}, \quad x_2 = -1 - 2\mathbf{i} - 3\mathbf{j} - 4\mathbf{k} \quad \text{and} \quad x_3 = -\mathbf{i} + \mathbf{j}$$

Table 1: Results for Example 1

QWM			2QWM		
k	$\varepsilon^{(k)}$	$\rho^{(k)}$	k	$\varepsilon^{(k)}$	$\rho^{(k)}$
8	7.6×10^{-2}	$--$	4	1.9×10^{-1}	$--$
9	1.9×10^{-3}	2.44	5	1.3×10^{-3}	4.08
10	7.1×10^{-7}	2.26	6	6.1×10^{-10}	3.17
11	3.9×10^{-14}	2.18	7	6.0×10^{-29}	3.06
12	7.7×10^{-29}	2.09	8	1.5×10^{-85}	3.01

Table 2: Results for Example 2

QWM			2QWM		
k	$\varepsilon^{(k)}$	$\rho^{(k)}$	k	$\varepsilon^{(k)}$	$\rho^{(k)}$
18	9.1×10^{-3}	$--$	11	2.5×10^{-1}	$--$
19	1.9×10^{-4}	1.82	12	1.6×10^{-2}	3.06
20	5.4×10^{-8}	1.95	13	6.1×10^{-6}	2.90
21	5.5×10^{-16}	1.96	14	1.3×10^{-16}	3.05
22	1.3×10^{-29}	2.03	15	2.2×10^{-48}	3.00

and

$$\zeta_1 = -2 - \mathbf{j} - \mathbf{k}, \quad \zeta_2 = -1 - \frac{27}{23}\mathbf{i} - \frac{76}{23}\mathbf{j} - \frac{94}{23}\mathbf{k} \quad \text{and} \quad \zeta_3 = \frac{8}{27}\mathbf{i} + \frac{35}{27}\mathbf{j} + \frac{13}{27}\mathbf{k}.$$

Starting with the initial approximation $z^{(0)} = (1, 2, 1+\mathbf{i}+\mathbf{j})$, we reached the precision 10^{-16} after 12 iterations of the quaternionic Weierstrass method (QWM) and just 7 iterations of the two-step quaternionic Weierstrass method (2QWM). Table 1 contains the results concerning the computational order of convergence for both methods. These results agree, as expected, with the conclusions of Theorem 1 (P fulfills its assumptions).

Example 2. We consider now the polynomial borrowed from [5]:

$$P(x) = (x + 2\mathbf{i})(x + 1 + \mathbf{k})(x - 2)(x - 1)(x - 2 + \mathbf{j})(x - 1 + \mathbf{i}),$$

whose roots are

$$\begin{array}{lll} \zeta_1 = 1 - \mathbf{i}, & \zeta_2 = 1, & \zeta_3 = -1 - \frac{29}{39}\mathbf{i} + \frac{14}{39}\mathbf{j} - \frac{22}{39}\mathbf{k}, \\ \zeta_4 = 2, & \zeta_5 = -\frac{224}{113}\mathbf{i} - \frac{30}{113}\mathbf{k}, & \zeta_6 = 2 - \frac{2}{3}\mathbf{i} - \frac{1}{3}\mathbf{j} + \frac{2}{3}\mathbf{k}. \end{array}$$

We used as initial approximation $z^{(0)} = (\frac{1}{2}, \frac{3}{2} - \mathbf{j}, \frac{3}{2} + \mathbf{i} - \mathbf{j} + \mathbf{k}, \frac{3}{2} + \mathbf{i} - \mathbf{j}, -\frac{1}{2}, -1 - 2\mathbf{i})$ and reached the precision 10^{-16} after 22 iterations of the quaternionic Weierstrass method and 15 iterations of the two-step quaternionic Weierstrass method. The details about this example are presented in Table 2.

Example 3. Our next example concerns a polynomial P with a spherical zero, i.e., P does not fulfill the assumptions of Theorem 1.

Table 3: QWM and 2QMW for Example 3

Roots	Type	Error (QWM)	Error(2QWM)
$1-\mathbf{j}$	Isolated	2.5×10^{-16}	2.4×10^{-12}
$-\mathbf{i}+\mathbf{k}$	Isolated	2.0×10^{-15}	3.6×10^{-13}
$[i]$	Spherical	7.7×10^{-9}	4.5×10^{-8}

Table 4: Modified QWM and 2QWM for Example 3

Roots	Method	Type	Error	k
$1-\mathbf{j}$	Weierstrass	Isolated	7.4×10^{-17}	11
$-\mathbf{i}+\mathbf{k}$	Weierstrass	Isolated	2.3×10^{-18}	11
$1-\mathbf{j}$	2 step Weierstrass	Isolated	1.6×10^{-16}	6
$-\mathbf{i}+\mathbf{k}$	2 step Weierstrass	Isolated	2.0×10^{-16}	6
$[\mathbf{i}]$	Deflation	Spherical	0	–

Recently, we have proposed a deflation algorithm [9] to be used together with the quaternionic Weierstrass method which allows to obtain quadratic order of convergence for isolated and spherical roots without requiring higher order precision. The same technique can be used now for the two-step method. The first part of the method concerns the determination of the spherical roots, while the second one consists of applying the Weierstrass method to a deflate polynomial.

In this example we revisit the first example of [9], where the polynomial

$$P(x) = x^4 + (-1+\mathbf{i})x^3 + (2-\mathbf{i}+\mathbf{j}+\mathbf{k})x^2 + (-1+\mathbf{i})x + 1 - \mathbf{i} + \mathbf{j} + \mathbf{k},$$

was considered. This polynomial has the isolated zeros $-\mathbf{i}+\mathbf{k}$ and $1-\mathbf{j}$ and the sphere of zeros, $[\,\mathbf{i}\,]$.

The numerical computations have been performed, as in the aforementioned work, in the Matlab system with double floating point arithmetic.

Starting with the initial guess $z^{(0)} = (1, -2, 0.5\mathbf{i}, 1+\mathbf{i})$, we obtained the results presented in Table 3, without using the deflate strategy. The results of Table 3 can be easily explained if we take into account that in the proof of Theorem 1 we assume that $|\mathcal{Q}_i(y_i)|$ is bounded away from zero, which is not the case when we have two factor terms "almost" in the same equivalence class. If arbitrary precision arithmetic is not available, the faster the method, the more quickly this effect is expected to be observed.

Applying now the Weierstrass algorithm to a 2nd degree deflate polynomial with the initial approximation $z^{(0)} = (1, 1+\mathbf{i})$, we obtain, after 11 iterations of QWM and 6 iterations of the 2QWM, the results presented in Table 4.

Example 4. Consider now the 9th degree polynomial

$$Q(x) = P(x)(x^2+4)(x^3+9x),$$

where P is the polynomial of Example 3. Apart from the same two isolated roots $-\mathbf{i}+\mathbf{k}$, $1-\mathbf{j}$ and the spherical zero $[\mathbf{i}]$, this polynomial has also 1 as isolated zero and $[2\mathbf{i}]$ and $[3\mathbf{i}]$ as spherical zeros.

The one- and two-step Weierstrass algorithm applied to the 3rd degree polynomial, obtained by the deflation procedure, with the initial approximation $z^{(0)} = (-1, 2, 1+\mathbf{i})$, produce, after 12 iterations of QWM and 9 iterations of the 2QWM, the results presented in Table 5.

Table 5: Modified QWM and 2QWM for Example 4

Roots	Method	Type	Error	k
$1.000 - 1.000\mathbf{j}$	Weierstrass	Isolated	3.3×10^{-15}	12
$-1.000\mathbf{i} + 1.000\mathbf{k}$	Weierstrass	Isolated	1.7×10^{-15}	12
1.000	Weierstrass	Isolated	2.7×10^{-15}	12
$1.000 - 1.000\mathbf{j}$	2 step Weierstrass	Isolated	3.4×10^{-15}	9
$-1.000\mathbf{i} + 1.000\mathbf{k}$	2 step Weierstrass	Isolated	1.7×10^{-15}	9
1.000	2 step Weierstrass	Isolated	3.0×10^{-15}	9
$[\mathbf{i}]$	Deflation	Spherical	1.1×10^{-15}	–
$[2\mathbf{i}]$	Deflation	Spherical	1.4×10^{-15}	–
$[3\mathbf{i}]$	Deflation	Spherical	1.9×10^{-15}	–

5 Conclusions

We have derived a two-step method based on the Weierstrass method for computing the roots of a quaternionic polynomial and have proved its cubic order of convergence, under the assumptions that all the roots are isolated and distinct (and the initial guesses are sufficiently "good"). A modified version of the two-step method was also considered allowing to overcome the issues associated to spherical roots.

We hope it is possible to modify the scheme in order to improve its efficiency, in particular, in what concerns the number of evaluation required in each iteration. We intend to focus on this aspect in the near future.

Acknowledgments

Research at CMAT was partially financed by Portuguese funds through FCT - Fundação para a Ciência e a Tecnologia, within the Projects UIDB/00013/2020 and UIDP/00013/2020. Research at NIPE has been financed by FCT, within the Project UIDB/03182/2020.

References

[1] Beck, B.: Sur les équations polynomiales dans les quaternions. Enseign. Math. 25, 193–201 (1979)

[2] Dočev, K.: A variant of Newton's method for the simultaneous approximation of all roots of an algebraic equation. Fiz. Mat. Spis. Bŭlgar. Akad. Nauk. 5 (38), 136–139 (1962)

[3] Durand, É.: Solutions Numériques des Equations Algébriques. Tome I: Equations du type F(x); Racines d'un Polynôme. Masson et Cie (1960)

[4] Falcão, M.I., Miranda, F., Severino, R., Soares, M.J.: Evaluation schemes in the ring of quaternionic polynomials. BIT Numerical Mathematics 58(1), 51–72 (2018)

[5] Falcão, M.I., Miranda, F., Severino, R., Soares, M.J.: Weierstrass method for quaternionic polynomial root-finding. Mathematical Methods in the Applied Sciences 41(1), 423–437 (2018)

[6] Falcão, M.I.: Newton method in the context of quaternion analysis. Appl. Math. Comput. 236, 458–470 (2014)

[7] Falcão, M.I., Miranda, F.: Quaternions: A Mathematica package for quaternionic analysis. Lecture Notes in Comput. Sci. 6784, 200–214 (2011)

[8] Falcão, M.I., Miranda, F., Severino, R., Soares, M.J.: Mathematica tools for quaternionic polynomials. Lecture Notes in Comput. Sci. 10405, 394–408 (2017)

[9] Falcão, M.I., Miranda, F., Severino, R., Soares, M.J.: A modified quaternionic Weierstrass method. Lecture Notes in Comput. Sci. 13377, 407–419 (2022)

[10] Falcão, M.I., Miranda, F., Severino, R., Soares, M.J.: Computational aspects of quaternionic polynomials - Part II: Root-finding methods. The Mathematica Journal (2018)

[11] Gordon, B., Motzkin, T.: On the zeros of polynomials over division rings I. Trans. Amer. Math. Soc. 116, 218–226 (1965)

[12] Grau-Sánchez, M., Noguera, M., Grau, À., Herrero, J.R.: On new computational local orders of convergence. Appl. Math. Lett. 25(12), 2023–2030 (2012)

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

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

[15] Lam, T.Y.: A first course in noncommutative rings. Graduate Texts in Mathematics, Springer-Verlag, New York (1991)

[16] Magreñán Ruiz, A.A., Argyros, I.K.: Two-step newton methods. Journal of Complexity 30(4), 533–553 (2014)

[17] Miranda, F., Falcão, M.I.: QuaternionAnalysis Package User's Guide. `http://w3.math.uminho.pt/QuaternionAnalysis` (2014)

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

[19] Petković, I., Herceg, D.: Computer methodologies for comparison of computational efficiency of simultaneous methods for finding polynomial zeros. Journal of Computational and Applied Mathematics 368, 112513 (2020)

[20] Sakurai, T., Petković, M.: On some simultaneous methods based on Weierstrass' correction. Journal of Computational and Applied Mathematics 72(2), 275–291 (1996)

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

[22] Weierstrass, K.: Neuer Beweis des Satzes, dass jede ganze rationale Function einer Veränderlichen dargestellt werden kann als ein Product aus linearen Functionen derselben Veränderlichen. In: Sitzungsberichte der Königlich Preussischen Akademie der Wissenschaften zu Berlin, vol. II. S.1085-1101. Berlin (1891)

[23] Zhang, F.: Quaternions and matrices of quaternions. Linear Algebra Appl. 251, 21–57 (1997)