Weierstrass method for quaternionic polynomial root-finding

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

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

Information

Keywords:
Quaternionic polynomials,
Root-finding methods,
Weierstrass algorithm.

Original publication:
Mathematical Methods in the Applied Sciences, 41 (2018), 423–437
DOI:10.1002/mma.4623
Wiley Online Library

Abstract

Quaternions, introduced by Hamilton in 1843 as a generalization of complex numbers, have found, in more recent years, a wealth of applications in a number of different areas that motivated the design of efficient methods for numerically approximating the zeros of quaternionic polynomials. In fact, one can find in the literature recent contributions to this subject based on the use of complex techniques, but numerical methods relying on quaternion arithmetic remain scarce. In this paper we propose a Weierstrass-like method for finding simultaneously *all* the zeros of unilateral quaternionic polynomials. The convergence analysis and several numerical examples illustrating the performance of the method are also presented.

1 Introduction

The increasing interest in using quaternions and their applications in areas as diverse as number theory, robotics, virtual reality or image processing (see e.g. [4, 10, 24, 25, 26]), motivated several authors to consider extending well-known (complex) numerical methods, in particular root-finding methods, to the quaternion algebra framework. However, the problem of finding the zeros of quaternionic polynomials turns out to be much more demanding than the analogous problem over the real and complex fields. Niven, in his pioneering work, [23], gave a first extension of the Fundamental Theorem of Algebra for the quaternion context, proving that any quaternionic polynomial of positive degree whose coefficients are located only on one side of the powers must have at least one quaternionic root. In the aforementioned paper, Niven also proposed a method for computing the roots of such polynomials. This algorithm is, however, as stated by Niven a "not very practical" one, due to the need of solving two coupled nonlinear equations for the determination of pairs of real constants. Later, in [34], the authors, by making use of (a complexified version of) the companion matrix of the polynomial, turned the ideas of Niven into what can be considered as the first really usable numerical algorithm.

Nowadays, other quaternionic root-finding algorithms are available that essentially replace the problem of computing the roots of a quaternionic polynomial of degree n, by the problem of determining the roots of a real or complex polynomial of degree $2n$ (usually with multiple roots), relying in this way on algorithms for complex polynomial root-finding (see [5, 17, 35] and the references therein). Several experiments performed by two of the authors of this paper ([8, 22]) have shown the substantial gain in computational effort that can be achieved when using a direct quaternionic approach to this problem.

The Weierstrass method, also known in the literature as the Durand-Kerner method or Dochev method, is one of the most popular iterative methods for obtaining simultaneously approximations to all the roots of a given polynomial with complex coefficients (for a survey on most of the traditional methods for root-finding we refer to [20]). The method was first proposed by Weierstrass in his famous work [37], where a semilocal convergence analysis was also provided and later rediscovered and derived in different ways by Durand [7], Dochev [6], Kerner [18] and Presić [29].

The main purpose of this paper is to present an adaptation of the Weierstrass method to the case of quaternionic polynomials. By making use of the so-called Factor Theorem for quaternions we derive an iterative method that shows fast convergence and robustness with respect to the initial approximations.

The paper is organized as follows: in Section 2 we review some basic results on the algebra of real quaternions and on quaternionic polynomials; Section 3 contains the main results of the paper; after revisiting the classical (complex) Weierstrass method we derive a generalization to the quaternionic case and prove, under some natural assumptions, its quadratic order of convergence; in Section 4 we present several numerical experiments illustrating the results obtained in Section 3; finally, in Section 5 we draw some conclusions and indicate some future work.

2 Basic results on quaternions

In this section we present a brief summary on the main results on the algebra of real quaternions and on the ring of polynomials over the quaternions needed in the sequel.

2.1 The algebra of real quaternions

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 noncommutative product generates the well-known algebra of real quaternions $\mathbb{H}$.

Given a quaternion $q = q_0 + q_1\mathbf{i} + q_2\mathbf{j} + q_3\mathbf{k} \in \mathbb{H}$, its *conjugate* $\overline{q}$ is defined as $\overline{q} = q_0 - q_1\mathbf{i} - q_2\mathbf{j} - q_3\mathbf{k}$; the number q_0 is called the *real part* of q and denoted by $\operatorname{Re} q$ and the *vector part* of q, denoted by $\operatorname{Vec} q$, is given by $\operatorname{Vec} q = q_1\mathbf{i} + q_2\mathbf{j} + q_3\mathbf{k}$; the *norm* of q, $|q|$, is given by $|q| = \sqrt{q\overline{q}} = \sqrt{q_0^2 + q_1^2 + q_2^2 + q_3^2}$; the *inverse* of q (if $q \neq 0$), denoted by q^{-1} is the (unique) quaternion such that $qq^{-1} = q^{-1}q = 1$ and is given by $q^{-1} = \dfrac{\overline{q}}{|q|^2}$.

We say that a quaternion q is *congruent* to a quaternion q', and write $q \sim q'$, if there exists a non-zero quaternion h such that $q' = hqh^{-1}$. This is an equivalence relation in $\mathbb{H}$, partitioning $\mathbb{H}$ in the so-called *congruence classes*. We denote by $[q]$ the congruence class containing a given quaternion q. It can be shown (see, e.g. [38]) that

$$[q] = \{q' \in \mathbb{H} : \operatorname{Re} q = \operatorname{Re} q' \text{ and } |q| = |q'|\}. \tag{1}$$

It follows that $[q]$ reduces to a single element if and only if q is a real number. If $q = q_0+q_1\mathbf{i}+q_2\mathbf{j}+q_3\mathbf{k}$ is not real, its congruence class can be identified with the three-dimensional sphere in the hyperplane $\{(x_0, x_1, x_2, x_3) \in \mathbb{R}^4 : x_0 = q_0\}$, with center $(q_0, 0, 0, 0)$ and radius $\sqrt{q_1^2 + q_2^2 + q_3^2}$.

2.2 Ring of left quaternionic polynomials

Because of the noncommutativity of quaternion multiplication, one can consider different classes of polynomials in one quaternion variable, depending on whether the variable commutes with the polynomial coefficients or not. General polynomials in the indeterminate x are defined as finite sums of noncommutative monomials of the form $a_0xa_1 \dots xa_j$. In this work we restrict our attention to polynomials whose coefficients are located only on the left-hand side of the powers of x, i.e. have the special form

$$P(x) = a_nx^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0, \quad a_i \in \mathbb{H}; \quad i = 0, \dots, n. \tag{2}$$

These polynomials are usually called in the literature *one-sided* or *unilateral* (*left*) polynomials. As usual, if $a_n \neq 0$, we will say that the *degree* of the polynomial $P(x)$ is n and refer to a_n as the leading coefficient of the polynomial. When $a_n = 1$, we say that $P(x)$ is *monic*. If the coefficients a_i in (2) are real, then we say that $P(x)$ is a *real polynomial* and write $P(x) \in \mathbb{R}[x]$.

The set of polynomials of the form (2) is a ring with respect to the operations of addition and multiplication defined as in the commutative case: for any two polynomials $P(x) = \sum_{i=0}^{n} a_i x^i$ and $Q(x) = \sum_{j=0}^{m} b_j x^j$,

$$P(x) + Q(x) := \sum_{k=0}^{\max\{m,n\}} (a_k + b_k) x^k,$$
$$P(x) * Q(x) := \sum_{k=0}^{m+n} \Big(\sum_{j=0}^{k} a_j b_{k-j} \Big) x^k,$$

with the implicit assumption that $a_k = 0$ for $k > n$ and $b_k = 0$ for $k > m$. We will denote this ring of polynomials by $\mathbb{H}[x]$. Naturally, due to the noncommutativity of the quaternionic multiplication, $\mathbb{H}[x]$ is a noncommutative ring. However, if $P(x)$ is a real polynomial, then $P(x)$ commutes with any polynomial in $\mathbb{H}[x]$.

We should also observe that the *evaluation map* at a given quaternion q, defined, for the polynomial $P(x)$ given by (2), by

$$P(q) = a_n q^n + a_{n-1} q^{n-1} + \cdots + a_1 q + a_0,$$

is not a homomorphism from the ring $\mathbb{H}[x]$ into $\mathbb{H}$. In fact, $P(x) = L(x) * R(x)$ does not lead, in general, to $P(q) = L(q)R(q)$.

Remark 1. *Since all the polynomials considered will be in the indeterminate x, we will usually omit the reference to this variable and write simply P when referring to an element $P(x) \in \mathbb{H}[x]$, the expression $P(q)$ being preferably reserved for the evaluation of P at a specific value $q \in \mathbb{H}$.*

We say that a quaternion q is a *zero* of a polynomial P, if $P(q) = 0$, and we use the notation $\mathbf{Z}_P$ to denote the *zero-set* of P, i.e. the set of all the zeros of P. Since this work is concerned with the computation of zeros of polynomials, there is no loss of generality in assuming that the polynomials are monic and we will do so in what follows.

We now review some basic properties of unilateral (left) quaternionic polynomials needed in the sequel.

The next theorem shows a way of evaluating the product of two polynomials at a given quaternion, without explicitly performing their product. The proof of the first two results can be seen in e.g. [19] and the last result is a simple consequence of the definition of the product of polynomials and of the fact that any real number commutes with a quaternion.

Theorem 1. *Let $P = L * R$ with $L, R \in \mathbb{H}[x]$, $q \in \mathbb{H}$ and $h = R(q)$.*

(i) If $h = 0$, then $P(q) = 0$ (i.e. if q is a zero of the right factor R, then q is also a zero of the product P).

(ii) If $h \neq 0$, then
$$P(q) = L(\tilde{q})R(q) \quad \text{with} \quad \tilde{q} = hqh^{-1}. \tag{3}$$
In particular, if q is a zero of P that is not a zero of R, then $\tilde{q}$ is a zero of L.

(iii) If $L \in \mathbb{R}[x]$, then
$$P(q) = R(q)L(q). \tag{4}$$

The following result, first proved by Gordon and Motzkin [14], can also be seen in [19].

Theorem 2 (Factor Theorem). *Let $P \in \mathbb{H}[x]$ and $q \in \mathbb{H}$. Then, q is a zero of P if and only if there exists $Q \in \mathbb{H}[x]$ such that*
$$P(x) = Q(x) * (x - q).$$

In 1941, Niven [23] proved the Fundamental Theorem of Algebra for unilateral quaternionic polynomials, establishing that any non-constant polynomial in $\mathbb{H}[x]$ always has a zero in $\mathbb{H}$. More general results are contained in the following theorem.

Theorem 3. *Let P be a monic polynomial of degree $n\,(n \geq 1)$ in $\mathbb{H}[x]$. Then:*

(i) P admits a factorization into linear factors, i.e. there exist $x_1, \ldots, x_n \in \mathbb{H}$, such that

$$P(x) = (x - x_n) * (x - x_{n-1}) * \cdots * (x - x_1).$$

(ii) For the factor terms x_i referred in (i), we have:

(a) $\mathbf{Z}_P \subseteq \bigcup_{i=1}^{n} [x_i]$.

(b) Each of the congruence classes $[x_i]; i = 1, \ldots, n$, contains (at least) a zero of P.

(iii) 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 $h_i \in \mathbb{H}; i = 1, \ldots, n$, such that

$$y_{\pi(i)} = h_i x_i h_i^{-1}.$$

The first result in the above theorem is an immediate consequence of the Fundamental Theorem of Algebra for quaternionic polynomials and of the Factor Theorem; the proof of the other results can be found in [19] and [35].

Given a polynomial $P(x) = \sum_{k=0}^{n} a_k x^k$, its *conjugate polynomial*, denoted by $\overline{P}(x)$, is given by

$$\overline{P}(x) = \sum_{k=0}^{n} \overline{a_k} x^k.$$

It is very simple to verify that, for all $P, Q \in \mathbb{H}[x]$:

$$\overline{P * Q} = \overline{Q} * \overline{P},$$
$$P * \overline{P} \in \mathbb{R}[x] \text{ and } P * \overline{P} = \overline{P} * P.$$

To each quaternion q, we will associate the following polynomial

$$\mathcal{Q}_q(x) := (x - q) * (x - \overline{q}) = x^2 - 2\,\mathrm{Re}\,q\,x + |q|^2,$$

called the *characteristic polynomial* of q. Since the characteristic polynomial of q only depends on the real part and norm of q and recalling (1), we immediately conclude that $\mathcal{Q}_q = \mathcal{Q}_{q'}$ if and only if $[q] = [q']$. Note that $\mathcal{Q}_q$ is a quadratic polynomial with real coefficients. It can also be shown that the zero-set of $\mathcal{Q}_q$ is the congruence class of q, i.e. $\mathbf{Z}_{\mathcal{Q}_q} = [q]$; see, e.g. [38]. This result already shows that, in what concerns the number of zeros, polynomials in $\mathbb{H}[x]$ can behave very differently from complex polynomials: a polynomial in $\mathbb{H}[x]$ can have an infinite number of zeros. However, as Theorem 3 shows, the zeros of a polynomial of degree n belong to, at most, n congruence classes in $\mathbb{H}$.

The zeros of an unilateral quaternionic polynomial can be of two distinct types, the so-called *isolated zeros* and *spherical zeros*, whose definitions we now recall. Let q be a zero of a given polynomial P. We say that q is an *isolated zero* of P if the congruence class of q contains no other zero of P. If q is not an isolated zero of P, we call it a *spherical zero* of P. Note that, according to the definition, real zeros are always isolated zeros. The next theorem gives conditions under which a nonreal zero is a spherical zero (see e.g. [28]).

Theorem 4. *Let q be a nonreal zero of a given polynomial $P \in \mathbb{H}[x]$. Then, q is a spherical zero of P if and only if any of the following equivalent conditions hold:*

(i) q and $\overline{q}$ are both zeros of P.

(ii) $[q] \subseteq \mathbf{Z}_P$.

(iii) The characteristic polynomial of q, $\mathcal{Q}_q$, is a divisor of P, i.e. there exists a polynomial $Q \in \mathbb{H}[x]$ such that $P = Q * \mathcal{Q}_q$.

Recalling that the congruence classes of nonreal quaternions can be identified with spheres, condition (ii) justifies the choice of the term *spherical* to designate this type of zeros. When q is a spherical zero, we also say that q *generates the sphere of zeros* $[q]$.

3 The Weierstrass method in $\mathbb{H}[x]$

Let P be a complex monic polynomial of degree n with roots $\zeta_1, \dots, \zeta_n$ and let $z_1^{(0)}, \dots, z_n^{(0)}$ be n given distinct numbers. The classical *Weierstrass method* for approximating the roots ζ_i is defined by the iterative scheme:

$$z_i^{(k+1)} = z_i^{(k)} - \frac{P(z_i^{(k)})}{\prod_{\substack{j=1\\ j\neq i}}^{n} (z_i^{(k)} - z_j^{(k)})}; i = 1, \dots, n;\ k = 0, 1, 2, \dots \tag{5}$$

If the roots $\zeta_1, \dots, \zeta_n$ are distinct and $z_1^{(0)}, \dots, z_n^{(0)}$ are sufficiently good initial approximations to these roots, then the method converges at a quadratic rate, as was firstly proven by Dochev [6]. A complete history and an improvement of Dochev's theorem can be found in [32] (see also [31, Sect. 6]). For multiple roots, the method still converges (locally) but the quadratic convergence is lost; see e.g. [11].

Formula (5) is performed in *parallel mode* and is often called the *total-step* mode. The convergence of the method can be accelerated by using a different variant that makes use of the most recent updated approximations to the roots as soon as they are available, as follows:

$$z_i^{(k+1)} = z_i^{(k)} - \frac{P(z_i^{(k)})}{\prod_{j=1}^{i-1} (z_i^{(k)} - z_j^{(k+1)}) \prod_{j=i+1}^{n} (z_i^{(k)} - z_j^{(k)})}; i = 1, \dots, n;\ k = 0, 1, 2, \dots$$

The above variant of the Weierstrass method is usually referred to as the *serial*, *sequential* or *single-step* mode (see [1] and references therein).

3.1 A quaternionic Weierstrass-like scheme

Our purpose is to adapt the idea of the Weierstrass method to the computation of the zeros of quaternionic polynomials. So let P be a given monic polynomial of degree n in $\mathbb{H}[x]$. Corresponding to the assumption imposed in the complex case to guarantee the quadratic convergence of the method – i.e. that the zeros of the polynomial are simple – we will now assume that the polynomial P has n distinct isolated roots. By analogy with the complex case, in this situation, we will still say that P has only *simple roots.* As stated in the previous section, P can be factorized in the form

$$P(x) = (x - x_n) * (x - x_{n-1}) * \cdots * (x - x_1), \tag{6}$$

with the factor terms $x_i \in \mathbb{H}$. For simplicity, we introduce the following convenient notation, which we borrow and adapt from [12],

$$\prod_{i=k}^{m}{}^{*} (x - \alpha_i) := (x - \alpha_m) * (x - \alpha_{m-1}) * \cdots * (x - \alpha_k).$$

Remark 2. *Note that the order of the factors, due to the noncommutativity of the product in $\mathbb{H}[x]$, is important. We also adopt the convention that*

$$\prod_{i=k}^{m}{}^{*} (x - \alpha_i) := 1, \quad \textit{whenever}\ \ k > m.$$

We first present a simple lemma, relating the roots of P with the quaternions involved in any of its factorizations.

Lemma 1. *Let P be a (monic) polynomial of degree n in $\mathbb{H}[x]$ with simple roots and let (6) be one of its factorizations. Then:*

(i) The congruence classes of the elements $x_j; j = 1, \dots, n$, in (6) are distinct.

(ii) The roots $\zeta_1,\dots,\zeta_n$ of P can be obtained from the quaternions $x_1,\dots,x_n$ as follows:

$$\zeta_i=\overline{R_i}(x_i)\,x_i\,\big(\overline{R_i}(x_i)\big)^{-1};\ i=1,2,\dots,n, \tag{7}$$

where R_i are the polynomials given by

$$R_i:=\prod_{j=1}^{i-1}{}^{*}(x-x_j). \tag{8}$$

Proof. The fact that the congruence classes $[x_j]; j=1,\dots,n$, are distinct is an immediate consequence of the results in Theorem 3 and of the assumption that P has only simple roots, i.e. it has exactly n isolated roots. The proof that the roots of P are given by (7) is a simple adaptation of the proof of [19, Proposition 16.3]. □

Following the idea of the Weierstrass method in its sequential version, we will now show how to obtain sequences converging, at a quadratic rate, to the factor terms in (6) of a given polynomial P. Then, we will show how these sequences can be used to estimate the zeros of P.

Theorem 5. *Let P be a polynomial of degree n in $\mathbb{H}[x]$ with simple roots and, for $i=1,\dots,n$; $k=0,1,2,\dots$, let*

$$z_i^{(k+1)}=z_i^{(k)}-\left(\overline{\mathcal{L}_i^{(k)}}*P*\overline{\mathcal{R}_i^{(k)}}\right)(z_i^{(k)})\,\left(\mathcal{Q}_i^{(k)}(z_i^{(k)})\right)^{-1}, \tag{9}$$

where

$$\mathcal{L}_i^{(k)}(x):=\prod_{j=i+1}^{n}{}^{*}\,(x-z_j^{(k)}), \tag{10}$$

$$\mathcal{R}_i^{(k)}(x):=\prod_{j=1}^{i-1}{}^{*}\,(x-z_j^{(k+1)}) \tag{11}$$

and

$$\mathcal{Q}_i^{(k)}(x):=\prod_{j=1}^{i-1}\mathcal{Q}_{z_j^{(k+1)}}(x)\prod_{j=i+1}^{n}\mathcal{Q}_{z_j^{(k)}}(x), \tag{12}$$

with $\mathcal{Q}_q$ denoting the characteristic polynomial of q. If the initial approximations $z_i^{(0)}$ are sufficiently close to the factor terms x_i in a factorization of P in the form (6), then the sequences $\{z_i^{(k)}\}$ converge quadratically to x_i.

Proof. Let $z_i^{(k)}$ be approximations to x_i with errors $\varepsilon_i^{(k)}$, i.e.

$$\varepsilon_i^{(k)}:=x_i-z_i^{(k)};\ i=1,\dots,n, \tag{13}$$

and let

$$\varepsilon^{(k)}:=\max_i|\varepsilon_i^{(k)}|.$$

We assume that $\varepsilon^{(k)}$ is *small enough*, i.e. that $z_i^{(k)}$ are *sufficiently good* approximations to x_i. We want to show that the next iterates $z_i^{(k+1)}$ are approximations to x_i with errors $\varepsilon_i^{(k+1)}$ such that

$$\varepsilon_i^{(k+1)}=\mathcal{O}\big((\varepsilon^{(k)})^2\big).$$

We will do this by induction on i. For simplicity, we will omit the iteration superscript (k), writing simply z_i for $z_i^{(k)}$, ε_i for $\varepsilon_i^{(k)}$, $\mathcal{L}_i$ for $\mathcal{L}_i^{(k)}$ etc. and will replace the superscript $(k+1)$ by a tilde symbol, using $\tilde{z}_i$ for $z_i^{(k+1)}$, $\tilde{\varepsilon}_i$ for $\varepsilon_i^{(k+1)}$, etc.

Step 1: We first prove that the result is true for $i=1$, i.e. that we have $\tilde{\varepsilon}_1=\mathcal{O}(\varepsilon^2)$.

By making use of (13), we can rewrite the polynomial $P(x)$ as

$$
\begin{aligned}
P(x) &= \prod_{j=1}^{n}{}^{*}(x - x_j) = \prod_{j=2}^{n}{}^{*}(x - z_j - \varepsilon_j) * (x - z_1 - \varepsilon_1) \\
&= \Big(\prod_{j=2}^{n}{}^{*}(x - z_j) + \mathscr{E}_1(x)\Big) * (x - z_1 - \varepsilon_1),
\end{aligned}
$$

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

$$-(x - z_n) * \cdots * (x - z_{j-1}) * \varepsilon_j * (x - z_{j+1}) * \cdots * (x - z_2),$$

$(j = 2, \ldots, n)$ with terms with $*$-products involving at least two ε_j's. By using the definition (10) of the polynomial $\mathcal{L}_1$, we can write $P(x)$ in the following form

$$
\begin{aligned}
P(x) &= \Big(\mathcal{L}_1(x) + \mathscr{E}_1(x)\Big) * (x - z_1 - \varepsilon_1) \\
&= \mathcal{L}_1(x) * (x - z_1 - \varepsilon_1) + \mathscr{E}_1(x) * (x - z_1 - \varepsilon_1).
\end{aligned}
$$

Let $\overline{\mathcal{L}}_1$ be the conjugate of $\mathcal{L}_1$, and note that $\overline{\mathcal{L}}_1 * \mathcal{L}_1$ is precisely the real polynomial $\boldsymbol{\mathcal{Q}}_1$ defined by (12). Hence, if we multiply $P(x)$ on the left by $\overline{\mathcal{L}}_1$ and evaluate the resulting polynomial at the point $x = z_1$, we obtain, recalling the results (3) and (4) in Theorem 1,

$$\big(\overline{\mathcal{L}}_1 * P\big)(z_1) = -\varepsilon_1 \boldsymbol{\mathcal{Q}}_1(z_1) - \Big(\overline{\mathcal{L}}_1 * \mathscr{E}_1\Big)(\hat{z}_1)\, \varepsilon_1,$$

where $\hat{z}_1 = \varepsilon_1 z_1 \varepsilon_1^{-1}$. Observing that 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 recalling the definition of $\mathscr{E}_1$, it is easily seen that we have

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

and therefore

$$\big(\overline{\mathcal{L}}_1 * P\big)(z_1) = -\varepsilon_1 \boldsymbol{\mathcal{Q}}_1(z_1) + \mathcal{O}(\varepsilon^2).$$

Since we are assuming that the congruence classes $[x_j]$ are distinct, then, for sufficiently small ε, $|\boldsymbol{\mathcal{Q}}_1(z_1)|$ is bounded away from zero and so, by multiplying both sides of the above equality on the right by $\big(\boldsymbol{\mathcal{Q}}_1(z_1)\big)^{-1}$, we obtain

$$\big(\overline{\mathcal{L}}_1 * P\big)\,(z_1)\ \big(\boldsymbol{\mathcal{Q}}_1(z_1)\big)^{-1} = -\varepsilon_1 + \mathcal{O}(\varepsilon^2),$$

or, in other words (cf. (13)),

$$x_1 = z_1 - \big(\overline{\mathcal{L}}_1 * P\big)\,(z_1)\,\big(\boldsymbol{\mathcal{Q}}_1(z_1)\big)^{-1} + \mathcal{O}(\varepsilon^2),$$

which means that the next approximation to x_1

$$\tilde{z}_1 = z_1 - (\overline{\mathcal{L}}_1 * P)(z_1)\,\big(\boldsymbol{\mathcal{Q}}_1(z_1)\big)^{-1}$$

is such that

$$\tilde{\varepsilon}_1 = x_1 - \tilde{z}_1 = \mathcal{O}(\varepsilon^2).$$

Step i: We now assume that, for $j = 1, \ldots, i-1$, $\tilde{z}_j$ approximates x_j with an error $\tilde{\varepsilon}_j$ such that $\tilde{\varepsilon}_j = \mathcal{O}(\varepsilon^2)$ and prove that $\tilde{z}_i$ is also an $\mathcal{O}(\varepsilon^2)$ approximation to x_i.

Using the polynomials

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

we can write

$$L_i(x) = \prod_{j=i+1}^{n} {}^{*}\,(x - z_j - \varepsilon_j) = \prod_{j=i+1}^{n} {}^{*}\,(x - z_j) + \mathscr{E}_i(x) = \mathcal{L}_i(x) + \mathscr{E}_i(x) \tag{14}$$

and

$$R_i(x) = \prod_{j=1}^{i-1} {}^{*}(x - \tilde{z}_j - \tilde{\varepsilon}_j) = \prod_{j=1}^{i-1} {}^{*}(x - \tilde{z}_j) + \tilde{\mathscr{E}}_i(x) = \mathcal{R}_i(x) + \tilde{\mathscr{E}}_i(x), \tag{15}$$

where $\mathscr{E}_i$ and $\tilde{\mathscr{E}}_i$ are remainder polynomials defined in an analogous manner to $\mathscr{E}_1$, with the obvious modifications. Note that $\mathscr{E}_i$ is a sum of terms, all of which involve at least the product by a ε_j $(j \in \{i+1, \ldots, n\})$ and $\tilde{\mathscr{E}}_i$ a sum of terms, all of which involve at least the product by an $\tilde{\varepsilon}_j$ $(j \in \{1, \ldots, i-1\})$. Therefore

$$\mathscr{E}_i(\alpha) = \mathcal{O}(\varepsilon) \quad \text{and} \quad \tilde{\mathscr{E}}_i(\alpha) = \mathcal{O}(\varepsilon^2),\ \forall \alpha \in \mathcal{D}.$$

Hence the polynomial P can be written as

$$P(x) = L_i(x) * (x - x_i) * R_i(x) = \big(\mathcal{L}_i(x) + \mathscr{E}_i(x)\big) * (x - z_i - \varepsilon_i) * \big(\mathcal{R}_i(x) + \tilde{\mathscr{E}}_i(x)\big).$$

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

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

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

$$\begin{aligned}\big(\overline{\mathcal{L}}_i * P * \overline{\mathcal{R}}_i\big)(z_i) &= -\varepsilon_i \boldsymbol{\mathcal{Q}}_i(z_i) - \big(\overline{\mathcal{L}}_i * \mathcal{R}_i * \overline{\mathcal{R}}_i * \mathscr{E}_i\big)(\hat{z}_i)\varepsilon_i + \mathcal{O}(\varepsilon^2)\\ &= -\varepsilon_i \boldsymbol{\mathcal{Q}}_i(z_i) + \mathcal{O}(\varepsilon^2),\end{aligned} \tag{16}$$

where $\hat{z}_i = \varepsilon_i z_i \varepsilon_i^{-1}$. Multiplying by $(\boldsymbol{\mathcal{Q}}_i(z_i))^{-1}$ on the right and observing, once more, that $|\boldsymbol{\mathcal{Q}}_i(z_i)|$ is bounded away from zero, we obtain

$$\big(\overline{\mathcal{L}}_i * P * \overline{\mathcal{R}}_i\big)\,(z_i)\ (\boldsymbol{\mathcal{Q}}_i(z_i))^{-1} = -\varepsilon_i + \mathcal{O}(\varepsilon^2)$$

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

$$\big(\overline{\mathcal{L}}_i * P * \overline{\mathcal{R}}_i\big)\,(z_i)\ (\boldsymbol{\mathcal{Q}}_i(z_i))^{-1} = z_i - x_i + \mathcal{O}(\varepsilon^2)$$

showing that

$$\tilde{z}_i = z_i - \big(\overline{\mathcal{L}}_i * P * \overline{\mathcal{R}}_i\big)\,(z_i)\ (\boldsymbol{\mathcal{Q}}_i(z_i))^{-1}$$

is an $\mathcal{O}(\varepsilon^2)$ approximation to x_i, which is precisely the result that we wanted to establish. □

Remark 3. *We should observe that, for each $i = 1, \ldots, n$, formula (9) for the computation of the approximation $z_i^{(k+1)}$ to x_i involves the polynomials $\mathcal{R}_i^{(k)}$ and $\boldsymbol{\mathcal{Q}}_i^{(k)}$ which make use of the already computed $z_1^{(k+1)}, \ldots, z_{i-1}^{(k+1)}$, i.e. the method here described can be seen as a generalization of the sequential version of the Weierstrass method. A careful analysis of the proof, namely the deduction of formula (16), shows that the use of the updated $z_j^{(k+1)}; j = 1, \ldots, i-1$, when computing $z_i^{(k+1)}$, is essential for establishing the quadratic order of convergence of the method.*

We now show how, with some additional little effort, one can use the iterative scheme (9)–(12) to produce, not only the factor terms, but also the roots of the polynomial.

Theorem 6. *Let P be a monic polynomial of degree n in $\mathbb{H}[x]$ with simple roots and let $\{z_i^{(k)}\}$ be the sequences defined by the Weierstrass iterative scheme (9)–(12) under the assumptions of Theorem 5. Finally, let $\{\zeta_i^{(k)}\}$ be the sequences defined by*

$$\zeta_i^{(k+1)} := \overline{\mathcal{R}_i^{(k)}}(z_i^{(k+1)})\, z_i^{(k+1)} \left(\overline{\mathcal{R}_i^{(k)}}(z_i^{(k+1)})\right)^{-1};\ k = 0, 1, 2, \ldots, \tag{17}$$

where $\mathcal{R}_i^{(k)}$ are the polynomials given by (11). Then, $\{\zeta_1^{(k)}\}, \ldots, \{\zeta_n^{(k)}\}$ converge quadratically to the roots of P.

Proof. We start by first recalling that the roots ζ_i; $i = 1, \ldots, n$, of P are related to x_i in (6) through (7), i.e.

$$\zeta_i = \overline{R_i}(x_i)\, x_i \left(\overline{R_i}(x_i)\right)^{-1};\ i = 1, \ldots, n,$$

with R_i defined by (8).

Next, denote by $\varepsilon_i^{(k+1)}$ the errors in the approximations $z_i^{(k+1)}$ to x_i and let $\varepsilon^{(k+1)} := \max_i |\varepsilon_i^{(k+1)}|$. We will show that

$$\zeta_i^{(k+1)} = \zeta_i + \mathcal{O}(\varepsilon^{(k+1)}).$$

This, conjugated with the results of Theorem 5, will prove the assertion of the theorem.

Similarly to what we did in the proof of Theorem 5, we will simply write $\tilde{z}_i$ for $z_i^{(k+1)}$, $\tilde{\varepsilon}_i$ for $\varepsilon_i^{(k+1)}$, $\tilde{\varepsilon}$ for $\varepsilon^{(k+1)}$, $\tilde{\zeta}_i$ for $\zeta_i^{(k+1)}$ and $\mathcal{R}_i$ for $\mathcal{R}_i^{(k)}$. Taking into account that the polynomials R_i in (8) are exactly the same polynomials presented in (15), we can write

$$\overline{R_i}(x) = \overline{\mathcal{R}_i}(x) + \overline{\tilde{\mathscr{E}}}_i(x)$$

and therefore

$$\overline{R_i}(x_i) = \overline{\mathcal{R}_i}(x_i) + \mathcal{O}(\tilde{\varepsilon}).$$

Expressing $\overline{\mathcal{R}_i}(x)$ in the expanded form $\sum_{j=1}^{i-1} \bar{r}_j x^j$, it follows at once that

$$\overline{R_i}(x_i) = \overline{\mathcal{R}_i}(\tilde{z}_i + \tilde{\varepsilon}_i) + \mathcal{O}(\tilde{\varepsilon}) = \overline{\mathcal{R}_i}(\tilde{z}_i) + \mathcal{O}(\tilde{\varepsilon}). \tag{18}$$

Combining the fact that both $|\overline{R_i}(x_i)|$ and $|\overline{\mathcal{R}_i}(\tilde{z}_i|$ are bounded away from zero with the result (18), we can conclude that

$$\left(\overline{R_i}(x_i)\right)^{-1} = \left(\overline{\mathcal{R}_i}(\tilde{z}_i)\right)^{-1} + \mathcal{O}(\tilde{\varepsilon}). \tag{19}$$

Finally, result (7) together with (18), (19) and the assumption (17) gives

$$\begin{aligned}
\zeta_i &= \overline{R_i}(x_i)\, x_i \left(\overline{R_i}(x_i)\right)^{-1} \\
&= \left(\overline{\mathcal{R}_i}(\tilde{z}_i) + \mathcal{O}(\tilde{\varepsilon})\right)\left(\tilde{z}_i + \tilde{\varepsilon}_i\right)\left(\left(\overline{\mathcal{R}_i}(\tilde{z}_i)\right)^{-1} + \mathcal{O}(\tilde{\varepsilon})\right) \\
&= \overline{\mathcal{R}_i}(\tilde{z}_i)\, \tilde{z}_i \left(\overline{\mathcal{R}_i}(\tilde{z}_i)\right)^{-1} + \mathcal{O}(\tilde{\varepsilon}) \\
&= \tilde{\zeta}_i + \mathcal{O}(\tilde{\varepsilon}),
\end{aligned}$$

which is precisely the result we want to prove. □

3.2 Computational details

We now summarize the proposed algorithm for computing the roots of a given quaternionic unilateral polynomial P of degree n and make some practical comments regarding its implementation.

Quaternionic-Weierstrass algorithm

INPUT:

- polynomial coefficients
- initial values $z_i^{(0)}$
- error tolerances $\varepsilon_1, \varepsilon_2$
- maximum number of iterations $kmax$

1. Set $\zeta_i^{(0)} = z_i^{(0)}$
2. For $k = 1, 2, \dots$ until Stopping Criterion is true
   (a) Compute $z_i^{(k)}$, by means of (9)-(12).
   (b) Compute $\zeta_i^{(k)}$, by means of (17) and (11).

   Stopping Criterion: $\left(\max_i |\zeta_i^{(k)} - \zeta_i^{(k-1)}| < \varepsilon_1 \text{ and } \max_i |P(\zeta_i^{(k)})| < \varepsilon_2\right)$ or $k = kmax$.

OUPUT: Factors $\tilde{x}_i = z_i^{(k)}$ and roots $\tilde{\zeta}_i = \zeta_i^{(k)}$.

Choice of initial approximations

In the classical case, Weierstrass method seems in practice to converge from nearly all starting points (see [20] and the references therein for details). The numerical experiments that we have conducted also show the robustness of the quaternionic version of the method in what concerns the choice of initial approximations. In any case, there are some aspects that should be taken into account.

First, for formula (9) to be meaningful, a first requirement one has to have in mind when choosing the initial approximations $z_1^{(0)}, \dots, z_n^{(0)}$ is that all of them belong to distinct congruence classes. This does not necessarily guarantee that, in the course of the computations, two approximations do not fall into the same congruence class, although this is very unlikely to happen. In such a case, a small perturbation of the initial guesses should be sufficient to regain convergence.

Second, it is, naturally, convenient to select the initial approximations from a region where the x_i in any factorization of the polynomial P are known to lie. Since the x_i and the roots ζ_i of P have the same norm, bounds on $|\zeta_i|$ are also valid for $|x_i|$. Moreover, since $P(x) * \overline{P}(x)$ is a real polynomial[1], whose roots r_i also have the same norm as the roots ζ_i of P (this is an immediate consequence of Theorem 2 in [35] and the characterization of the congruence classes given by (1)), one can use any known result on bounds on (complex) polynomial roots to obtain a region from where the initial approximations should be selected.

Non simple zeros

The proof of Theorem 5 was done under the assumption that the roots $\zeta_1, \dots, \zeta_n$ of the polynomial (6) are simple, i.e. that $[\zeta_i] \neq [\zeta_j]$ for all $i \neq j$, or, equivalently, $[x_i] \neq [x_j]$ (cf. Lemma 1). When $[x_i] = [x_j]$ for some $i \neq j$, the characterization of the zero-set of the polynomial can be done taking into account the following two results.

Lemma 2. *If $x_1, x_2 \in \mathbb{H}$ and $h = \overline{x}_2 - x_1$, then*

$$(x - x_2) * (x - x_1) = \begin{cases} (x - h^{-1}x_1 h) * (x - h^{-1}x_2 h), & \text{if } h \neq 0 \\ (x - x_1) * (x - x_2), & \text{if } h = 0. \end{cases}$$

Proof. The result follows by simple manipulation; see also [35] for a different, but equivalent result. □

[1] The use of the polynomial $P(x) * \overline{P}(x)$ goes back to the work of Niven [23].

Lemma 3. *Consider a quadratic polynomial factorized in the form*

$$P(x) = (x - x_2) * (x - x_1),$$

where $x_1, x_2 \in \mathbb{H} \setminus \mathbb{R}$ *and* $[x_1] = [x_2]$.

(i) If $x_1 \neq \overline{x}_2$*, then the only zero of* P *is* x_1.

(ii) If $x_1 = \overline{x}_2$*, then* x_1 *generates the sphere of zeros* $[x_1]$*, i.e.* x_1 *is a spherical zero.*

Proof. See e.g. [26]. □

The problem of finding a natural definition of multiplicity for zeros of quaternionic polynomials is a rather complicated task and as a consequence one can find in the literature different (not always equivalent, see [9]) concepts of multiplicity [2, 3, 13, 26, 36]. In the case (i) above, we will say that x_1 is a root with (*isolated*) *multiplicity* equal to two. For example, the polynomials $(x+1-\mathbf{i})*(x+1+\mathbf{k})$ and $(x+1+\mathbf{k})*(x+1+\mathbf{k})$ both have $\zeta = -1 - \mathbf{k}$ as a root with multiplicity two.

Returning to the case of a general polynomial of degree n of the form (6), we consider, for simplicity, that $[x_i]$ and $[x_j]$ are the only non-distinct congruent classes. Using Lemma 2, we can freely move the factors $(x - x_i)$ and $(x - x_j)$ to the right of the factorization without changing the set of congruence classes so that in the new factorization

$$P(x) = (x - y_n) * (x - y_{n-1}) * \cdots * (x - y_2) * (x - y_1) = Q(x) * (x - y_2) * (x - y_1)$$

we have $[y_1] = [y_2]$. Observe that all the roots of the $n-2$ degree polynomial Q are simple and, therefore, the complete characterization of the roots of P can be done applying Lemma 3 to the quadratic polynomial $(x - y_2) * (x - y_1)$.

We considered the application of the quaternionic Weierstrass method to several examples of polynomials having double (isolated) or spherical roots and, in all the cases, we have observed the following: when y_1 is a double isolated root ($y_1 \neq \overline{y}_2$), the behavior is analogous to the one observed in the classical case, i.e. the rate of convergence drops to one; on the other hand, if y_1 is a spherical root, the iterative scheme produces two distinct roots ζ_1 and ζ_2 belonging to the congruence class $[y_1]$ and still shows a quadratic order of convergence.

4 Numerical examples

In this section we present several examples illustrating the performance of the quaternionic Weierstrass method introduced in Section 3.

All the numerical experiments here reported were obtained by the use of the Mathematica add-on application `QuaternionAnalysis` [21] designed by two of the authors of this paper for symbolic manipulation of quaternion valued functions. A collection of new functions, including an implementation of the Weierstrass method described in this paper, has been recently developed to endow the aforementioned package with the ability to perform operations in the noncommutative ring of polynomials $\mathbb{H}[x]$.

Example 1. *Our first test example is a polynomial which fulfills the assumptions of Theorem 6. In fact, it is easy to see that the polynomial*

$$P(x) = (x + 2\mathbf{i}) * (x + 1 + \mathbf{k}) * (x - 2) * (x - 1) * (x - 2 + \mathbf{j}) * (x - 1 + \mathbf{i}), \tag{20}$$

has only simple roots, namely

$$\zeta_1 = 1 - \mathbf{i}, \qquad \zeta_2 = 1, \qquad \zeta_3 = -1 - \tfrac{29}{39}\mathbf{i} + \tfrac{14}{39}\mathbf{j} - \tfrac{22}{39}\mathbf{k},$$
$$\zeta_4 = 2, \qquad \zeta_5 = -\tfrac{224}{113}\mathbf{i} - \tfrac{30}{113}\mathbf{k}, \qquad \zeta_6 = 2 - \tfrac{2}{3}\mathbf{i} - \tfrac{1}{3}\mathbf{j} + \tfrac{2}{3}\mathbf{k}.$$

Since, in this case, the polynomial roots ζ_i *are known exactly, we replace the stopping criterion based on the incremental size of the iterations by the following one:*

$$\epsilon^{(k)} := \max_i \{\epsilon_i^{(k)}\} < \varepsilon_1, \ \textit{with}\ \epsilon_i^{(k)} := |\zeta_i^{(k)} - \zeta_{\pi_k(i)}|,$$

Table 1: Quaternionic Weierstrass method for Example 1

k	$\epsilon_1^{(k)}$	$\epsilon_2^{(k)}$	$\epsilon_3^{(k)}$	$\epsilon_4^{(k)}$	$\epsilon_5^{(k)}$	$\epsilon_6^{(k)}$	$\rho^{(k)}$
0	2.7 e−1	6.0 e−2	4.5 e−1	2.0 e−2	8.0 e−2	3.3 e−1	–
1	9.3 e−2	1.8 e−2	7.7 e−2	7.2 e−3	6.0 e−2	3.1 e−2	1.36
2	7.9 e−3	1.9 e−3	5.7 e−3	5.6 e−4	5.3 e−3	1.1 e−3	2.03
3	6.0 e−5	2.0 e−5	4.0 e−5	4.5 e−6	9.2 e−6	3.1 e−7	2.17
4	2.4 e−9	1.5 e−9	3.3 e−9	3.7 e−10	2.0 e−9	2.2 e−13	2.04
5	1.5 e−17	5.3 e−18	1.6 e−17	6.1 e−19	7.5 e−18	8.1 e−26	2.06

where π_k is an appropriate permutation of $\{1,\ldots,6\}$. Here, we considered $\varepsilon_1 = \varepsilon_2 = 10^{-16}$ and chose initial approximations so that $\epsilon^{(0)} \le 0.5$.

The Weierstrass method applied to the extended form of P produced, after 5 iterations, the following approximations to the factor terms (with 15 decimal places[2])

$$
\begin{aligned}
x_1^{(5)} &= 1.(0) - 1.(0)\mathbf{i}\\
x_2^{(5)} &= 1.(0)\\
x_3^{(5)} &= -1.(0) - 0.545454545454545\mathbf{i} - 0.181818181818182\mathbf{j} - 0.818181818181818\mathbf{k}\\
x_4^{(5)} &= 2.(0)\\
x_5^{(5)} &= -1.587878787878788\mathbf{i} - 0.911515151515152\mathbf{j} + 0.804848484848485\mathbf{k}\\
x_6^{(5)} &= 2.(0) + 0.133333333333333\mathbf{i} + 0.093333333333333\mathbf{j} - 0.986666666666667\mathbf{k}
\end{aligned}
$$

corresponding to the approximate roots

$$
\begin{aligned}
\zeta_1^{(5)} &= 1.(0) - 1.(0)\mathbf{i}\\
\zeta_2^{(5)} &= 1.(0)\\
\zeta_3^{(5)} &= -1.(0) - 0.743589743589744\mathbf{i} + 0.358974358974359\mathbf{j} - 0.564102564102564\mathbf{k}\\
\zeta_4^{(5)} &= 2.(0)\\
\zeta_5^{(5)} &= -1.982300884955752\mathbf{i} - 0.265486725663717\mathbf{k}\\
\zeta_6^{(5)} &= 2.(0) - 0.666666666666667\mathbf{i} - 0.333333333333333\mathbf{j} + 0.666666666666667\mathbf{k}
\end{aligned}
$$

It is interesting to observe that the approximations $x_i^{(5)}$ to the factor terms lead to a factorization of P different from (20), but of course in line with Theorem 3 (iii).

Table 1 contains the relevant information concerning the errors in the successive approximations $\zeta_i^{(k)}$ ($k = 0,\ldots,5$, $i = 1,\ldots,6$) to the roots ζ_i of P. Estimates ρ for the computational local order of convergence of the method, based on the use of (see e.g. [15] for details)

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

were also computed and are included in the last column of the table.

To illustrate Remark 3 we have also implemented the parallel version of Weierstrass method. In this case, using the same initial guesses, 9 iterations were required to achieve the same precision. The results presented in Table 2 clearly indicate the deterioration of the speed of convergence of this version of the method.

[2] The notation (0) after the decimal point represents a sequence of 15 zeros.

Table 2: Parallel version of Weierstrass method for Example 1

k	$\epsilon_1^{(k)}$	$\epsilon_2^{(k)}$	$\epsilon_3^{(k)}$	$\epsilon_4^{(k)}$	$\epsilon_5^{(k)}$	$\epsilon_6^{(k)}$	$\rho^{(k)}$
0	2.7 e−1	6.0 e−2	4.5 e−1	2.0 e−2	8.0 e−2	3.3 e−1	–
1	9.3 e−2	2.3 e−2	8.6 e−2	5.9 e−3	5.3 e−3	9.8 e−2	3.26
2	1.7 e−2	4.4 e−3	2.8 e−2	2.0 e−3	2.0 e−2	3.3 e−2	1.47
3	1.0 e−3	4.3 e−4	3.1 e−3	3.1 e−4	2.9 e−3	4.2 e−3	1.60
4	5.0 e−5	1.9 e−5	1.5 e−4	1.2 e−5	2.4 e−4	3.0 e−4	1.48
5	2.2 e−7	1.1 e−7	2.1 e−6	1.0 e−7	7.0 e−6	7.9 e−6	1.44
6	4.5 e−10	2.2 e−10	5.6 e−8	2.5 e−10	6.0 e−7	1.9 e−6	1.12
7	2.5 e−13	1.2 e−13	9.4 e−11	2.0 e−13	1.2 e−8	8.6 e−9	1.39
8	1.6 e−18	1.6 e−18	1.6 e−15	3.7 e−18	2.4 e−12	3.0 e−12	1.45
9	2.0 e−25	1.7 e−25	7.9 e−21	5.5 e−25	1.5 e−17	2.3 e−17	1.44

Our next examples concern situations where the polynomials under consideration have zeros which are not simple.

Example 2. *The polynomial*

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

has, apart from the isolated zeros $-\mathbf{i}+\mathbf{k}$ and $1-\mathbf{k}$, a whole sphere of zeros, $[\,\mathbf{i}\,]$. In this case, since the spherical roots have the same real part and modulus, we replaced the stopping criterion used in the previous example by the following one:

$$\epsilon^{(k)} = \max\{\epsilon_R^{(k)}, \epsilon_N^{(k)}\} < 10^{-16},$$

where

$$\epsilon_R^{(k)} := \max_i\{|\operatorname{Re}(\zeta_i^{(k)}) - \operatorname{Re}(\zeta_{\pi_k(i)})|\} \quad \textit{and} \quad \epsilon_N^{(k)} := \max_i\{\big||\zeta_i^{(k)}| - |\zeta_{\pi_k(i)}|\big|\}.$$

Starting with an initial guess chosen so that $\epsilon^{(0)} \leq 0.15$, we obtained, after 5 iterations, the following approximations:

$$\begin{aligned}
\zeta_1^{(5)} &= 0.099934477851162\mathbf{i} - 0.917198737816235\mathbf{j} - 0.385693629043728\mathbf{k}\\
\zeta_2^{(5)} &= -0.799427021998164\mathbf{i} - 0.519295977566198\mathbf{j} - 0.302073044449043\mathbf{k}\\
\zeta_3^{(5)} &= 1.(0) - 1.(0)\mathbf{j}\\
\zeta_4^{(5)} &= -1.(0)\mathbf{i} + 1.(0)\mathbf{k}
\end{aligned}$$

The spherical root can be identified at once by observing that, up to the required precision, we have $[\zeta_1^{(5)}] = [\zeta_2^{(5)}]$, since $\operatorname{Re}\zeta_1^{(5)} = \operatorname{Re}\zeta_2^{(5)} = 0$ and $|\zeta_1^{(5)}| = |\zeta_2^{(5)}| = 1$.

The numerical details related to this example are displayed in Table 3. Here the numerical computations have been conducted with the precision increased to 512 significant digits.

As we can observe from Table 3, the quaternionic Weierstrass method works, produces all the roots simultaneously with machine precision and exhibits quadratic order of convergence. As expected, for the case of the spherical root, we obtain convergence to two distinct members of the sphere of zeros.

Example 3. *In our last example we address the problem of using Weierstrass method in cases where the polynomial under consideration has multiple (isolated) roots. The polynomials*

$$P(x) = (x-\mathbf{i}) * (x+1+\mathbf{k}) * (x+1+\mathbf{k}) \textit{ and } Q(x) = (x-\mathbf{i}) * (x+1-\mathbf{i}) * (x+1+\mathbf{k})$$

Table 3: Weierstrass method for spherical roots - Example 2

k	$\epsilon_1^{(k)}$	$\epsilon_2^{(k)}$	$\epsilon_3^{(k)}$	$\epsilon_4^{(k)}$	$\rho^{(k)}$
0	$1.3\,\mathrm{e}{-2}$	$7.1\,\mathrm{e}{-2}$	$7.6\,\mathrm{e}{-2}$	$1.3\,\mathrm{e}{-1}$	–
1	$6.3\,\mathrm{e}{-3}$	$4.5\,\mathrm{e}{-2}$	$6.1\,\mathrm{e}{-3}$	$1.1\,\mathrm{e}{-2}$	1.52
2	$1.2\,\mathrm{e}{-4}$	$9.8\,\mathrm{e}{-4}$	$1.4\,\mathrm{e}{-3}$	$9.7\,\mathrm{e}{-5}$	2.11
3	$9.6\,\mathrm{e}{-8}$	$1.1\,\mathrm{e}{-6}$	$2.6\,\mathrm{e}{-6}$	$1.8\,\mathrm{e}{-8}$	1.96
4	$9.2\,\mathrm{e}{-12}$	$6.1\,\mathrm{e}{-11}$	$1.6\,\mathrm{e}{-11}$	$1.0\,\mathrm{e}{-15}$	1.83
5	$1.4\,\mathrm{e}{-22}$	$9.7\,\mathrm{e}{-22}$	$4.9\,\mathrm{e}{-21}$	$8.0\,\mathrm{e}{-31}$	1.99

Table 4: Weierstrass method for double roots - Example 3

P		Q	
$\epsilon^{(k)}$	$\rho^{(k)}$	$\epsilon^{(k)}$	$\rho^{(k)}$
$6.8\,\mathrm{e}{-10}$	1.05	$1.9\,\mathrm{e}{-13}$	1.03
$2.6\,\mathrm{e}{-10}$	1.04	$7.3\,\mathrm{e}{-14}$	1.03
$1.0\,\mathrm{e}{-10}$	1.04	$2.8\,\mathrm{e}{-14}$	1.03
$3.9\,\mathrm{e}{-11}$	1.04	$1.1\,\mathrm{e}{-14}$	1.03
$1.5\,\mathrm{e}{-11}$	1.04	$4.1\,\mathrm{e}{-15}$	1.03
$5.9\,\mathrm{e}{-12}$	1.04	$1.6\,\mathrm{e}{-15}$	1.03
$2.3\,\mathrm{e}{-12}$	1.04	$6.1\,\mathrm{e}{-16}$	1.03
$8.8\,\mathrm{e}{-13}$	1.04	$2.4\,\mathrm{e}{-16}$	1.03
$3.4\,\mathrm{e}{-13}$	1.03	$9.1\,\mathrm{e}{-17}$	1.03

have one nonreal root with multiplicity one and $-1-\mathbf{k}$ as a double root. The approximations to the roots of P obtained by the use of the quaternionic Weierstrass method are

$$
\begin{aligned}
\zeta_1 &= -1.(0) - 1.(0)\mathbf{k} \\
\zeta_2 &= -0.230769230769231\mathbf{i} - 0.307692307692308\mathbf{j} - 0.923076923076923\mathbf{k} \\
\zeta_3 &= -1.(0) - 1.(0)\mathbf{k}
\end{aligned}
$$

while, for the roots of Q, we obtained

$$
\begin{aligned}
\zeta_1 &= 0.333333333333333\mathbf{i} - 0.666666666666667\mathbf{j} - 0.666666666666667\mathbf{k} \\
\zeta_2 &= -1.(0) - 1.(0)\mathbf{k} \\
\zeta_3 &= -1.(0) - 1.(0)\mathbf{k}
\end{aligned}
$$

As we can observe from Table 4, the behavior of the quaternionic Weierstrass method is very similar to that one observed for the classical complex case, where the rate of convergence is linear. This table shows $\epsilon^{(k)}$ for the last 9 iterations of the method together with $\rho^{(k)}$ for both polynomials.

5 Final Remarks

In this paper we proposed a generalization to the quaternionic context of the well-known Weierstrass method for approximating *all* zeros of a polynomial simultaneously. Becuase of the structure of the zero-set of a quaternionic polynomial, the claim that the method we have proposed produces *all* the zeros simultaneously, requires an additional explanation. Assuming the convergence of the method to the roots $\zeta_1, \ldots, \zeta_n$ of a polynomial P of degree n, it is easy to identify $\mathbf{Z}_P$, once we test if each element of $\{\zeta_1, \ldots, \zeta_n\}$ is an isolated or a spherical zero of P (cf. Theorem 4).

The quaternionic Weierstrass algorithm is entirely based on quaternionic arithmetic and shows fast convergence for simple and spherical roots. We proved the quadratic convergence of the sequential iterative scheme, under the assumptions that all the roots of the polynomial are simple, and presented numerical examples supporting this fact. In [8], it was proved that the same rate of convergence can be achieved by quaternion versions of Newton's method for the so-called radially holomorphic functions [16, p. 234]. None of the polynomials presented in this section are radially holomorphic or are in the less restrictive conditions of [8, Theorem 4]. As far as we are aware, the method proposed in this paper is the first numerical method entirely based on quaternionic arithmetic for which we can observe theoretical and experimental results for general unilateral quaternionic polynomials.

Several authors have described conditions for the safe convergence of the classical method depending only on the initial approximations. The history of this problem can be found in [30, 31] (see also [33] and [27]). This is a very interesting question that we intend to address in the near future, in the quaternionic case.

One can find in the literature several modifications to the classical Weierstrass method which improve the speed of convergence to multiple roots (see e.g. [11]). It is also in our plans of research to consider adaptations of such strategies.

Acknowledgments

Research at CMAT was financed by Portuguese Funds through FCT - Fundação para a Ciência e a Tecnologia, within the Project UID/MAT/00013/2013. Research at NIPE was carried out within the funding with COMPETE reference number POCI-01-0145-FEDER-006683 (UID/ECO/03182/2013), with the FCT/MEC's (Fundação para a Ciência e a Tecnologia, I.P.) financial support through national funding and by the ERDF through the Operational Programme on "Competitiveness and Internationalization - COMPETE 2020" under the PT2020 Partnership Agreement.

References

[1] G. Alefeld and J. Herzberger. On the convergence speed of some algorithms for the simultaneous approximation of polynomial roots. *SIAM Journal on Numerical Analysis*, 11(2):237–243, 1974.

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

[3] V. Bolotnikov. Zeros, factorizations and least common multiples of quaternion polynomials. Journal of Algebra and its Applications. 2017;16(10) 1750181:1-23

[4] F. Brackx, E. Hitzer, and S. Sangwine. History of quaternion and Clifford-Fourier transforms and wavelets. In E. Hitzer and S. Sangwine, editors, *Quaternion and Clifford Fourier transforms and wavelets*, volume 27 of *Trends in Mathematics*, pages XI–XXVII. Springer Basel AG, 2013.

[5] S. De Leo, G. Ducati, and V. Leonardi. Zeros of unilateral quaternionic polynomials. *Electron. J. Linear Algebra*, 15:297–313, 2006.

[6] K. Dočev. 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.

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

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

[9] M. I. Falcão, F. Miranda, R. Severino, and M. J. Soares. Quaternionic polynomials with multiple zeros: A numerical point of view. *AIP Conference Proceedings*, 1798(1):020099, 2017.

[10] R. T. Farouki, G. Gentili, C. Giannelli, A. Sestini, and C. Stoppato. A comprehensive characterization of the set of polynomial curves with rational rotation-minimizing frames. *Adv. Comput. Math.*, 43(1):1–24, 2017.

[11] P. Fraigniaud. The Durand-Kerner polynomials roots-finding method in case of multiple roots. *BIT*, 31(1):112–123, 1991.

[12] G. Gentili, C. Stoppato, and D. C. Struppa. *Regular Functions of a Quaternionic Variable*. Springer Monographs in Mathematics. Springer-Verlag Berlin Heidelberg, 2013.

[13] G. Gentili and D. C. Struppa. On the multiplicity of zeroes of polynomials with quaternionic coefficients. *Milan J. Math.*, 76:15–25, 2008.

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

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

[16] K. Gürlebeck, K. Habetha, and W. Sprößig. *Holomorphic functions in the plane and n-dimensional space*. Birkhäuser Verlag, Basel, 2008. Translated from the 2006 German original.

[17] D. Janovská and G. Opfer. A note on the computation of all zeros of simple quaternionic polynomials. *SIAM J. Numer. Anal.*, 48(1):244–256, 2010.

[18] I. O. Kerner. Ein Gesamtschrittverfahren zur Berechnung der Nullstellen von Polynomen. *Numer. Math.*, 8:290–294, 1966.

[19] T.-Y. Lam. *A first course in noncommutative rings*. Graduate Texts in Mathematics. Springer-Verlag, New York, 1991.

[20] J. M. McNamee. *Numerical methods for roots of polynomials*, volume 14 of *Part I*. Elsevier B. V., Amsterdam, 2007.

[21] F. Miranda and M. I. Falcão. QuaternionAnalysis Package User's Guide. http://w3.math.uminho.pt/QuaternionAnalysis, 2014.

[22] F. Miranda and M. I. Falcão. Modified quaternion Newton methods. *Lecture Notes in Comput. Sci.*, 8579(1):146–161, 2014.

[23] I. Niven. Equations in quaternions. *Amer. Math. Monthly*, 48:654–661, 1941.

[24] R. Pereira and P. Rocha. On the determinant of quaternionic polynomial matrices and its application to system stability. *Math Methods Appl Sci.*, 31(1):99–122, 2008.

[25] R. Pereira, P. Rocha, and P. Vettori. Algebraic tools for the study of quaternionic behavioral systems. *Linear Algebra Appl.*, 400(1-3):121–140, 2005.

[26] R. Pereira and P. Vettori. Stability of quaternionic linear systems. *IEEE Trans. Automat. Control*, 51(3):518–523, 2006.

[27] M. Petković. Point estimation of root finding methods. *Lecture Notes in Mathematics*, 1933:1–210, 2008.

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

[29] S. B. Prešić. Un procédé itératif pour la factorisation des polynômes. *C. R. Acad. Sci. Paris Sèr. A*, 262:862–863, 1966.

[30] P. Proinov. A new semilocal convergence theorem for the Weierstrass method from data at one point. *Comptes Rendus de L'Academie Bulgare des Sciences*, 59:131–136, 2006.

[31] P. Proinov. General convergence theorems for iterative processes and applications to the weierstrass root-finding method. *Journal of Complexity*, 33:118–144, 2016.

[32] P. Proinov and M. Petkova. Convergence of the Weierstrass method for simultaneous approximation of polynomial zeros. *Comptes Rendus de L'Academie Bulgare des Sciences*, 66(6):809–818, 2013.

[33] P. Proinov and M. Petkova. A new semilocal convergence theorem for the Weierstrass method for finding zeros of a polynomial simultaneously. *Journal of Complexity*, 30(3):366–380, 2014.

[34] R. Serôdio, E. Pereira, and J. Vitória. Computing the zeros of quaternion polynomials. *Comput. Math. Appl.*, 42(8-9):1229–1237, 2001.

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

[36] N. Topuridze. On roots of quaternion polynomials. *J. Math. Sci.*, 160(6):843–855, 2009.

[37] K. Weierstrass. 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*, volume II. S.1085-1101. Berlin, 1891.

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